16

Money and the Trade Cycle

16.1 INTRODUCTION

In the years following 1870 there was no change in monetary economics comparable with the changes which took place in the theory of value and distribution, though Jevons, Menger and Walras all made important contributions.[1] Jevons' contribution was primarily statistical, developing index numbers and using them to analyse the effects of the Californian gold discoveries. Walras integrated monetary theory into his general equilibrium model in a way not matched by his contemporaries. Menger, in complete contrast, analysed money as an institution which emerged as a natural result of men's economizing behaviour. As regards monetary policy, both Jevons and Walras were concerned with schemes to stabilize prices, Jevons proposing his tabular standard of value, Walras his *billon d'argent régulateur*. There was, however, within these contributions, no new theme comparable with the new ideas which emerged in the theory of value and distribution. There was no significant dividing line between monetary economics before and after 1870.

There was, however, one problem which dominated discussions of monetary policy during the period: that of fluctuating prices. Fluctuating prices were of concern not only to Jevons and Walras, but to all the major contributors to the period's monetary economics. Wicksell, Marshall and Fisher, for example, all produced schemes designed to produce greater price stability than was afforded by the gold standard. This concern carried through into the decades leading up to 1930, when price fluctuations were used by many economists to explain the trade cycle. There was, however, no break with classical monetary economics. Indeed, it could be argued that the key to much of the period's monetary economics is Wicksell's cumulative process, itself a rediscovery of Thornton's theory of interest and prices.[2]

This continuity extends to discussions of the trade cycle. Marshall, for example, took his analysis of the trade cycle from Overstone.[3] There were, however, important statistical contributions dating from the 1860s, which made the nature of the trade cycle much clearer. Jevons, through analysing time series data, established the existence of a regular cycle of 10 to 11 years' duration.[4] Probably more important was the work of Juglar who, though he also analysed time series data on the cycle, contributed the terminology which came to be used to describe the various phases of the cycle. Unlike Jevons, Juglar saw the cycle as fundamentally independent of outside events, such as harvests and wars, seeing cycles as arising "out of the

behaviour, the activities, and above all the saving habits of the population, and the way they employ the capital and credit available".[5] He saw depression as a response to the preceding prosperity, rather than as brought about by outside forces. In common with Mitchell, Juglar was more concerned to explain how cycles develop, in particular their financial aspects, rather than to provide a theoretical explanation of what caused what.[6]

16.2 MONETARY THEORY BEFORE 1930

Wicksell

Wicksell's main contributions to monetary theory came in *Interest and Prices* (1898), a book whose concern was to restate and defend the quantity theory of money, the classic statement of which Wicksell found in Ricardo.[7] In developing the quantity theory he was influenced above all by two things: (1) the capital theory of Jevons and Böhm-Bawerk, which had laid the foundations for a real theory of the rate of interest; and (2) Tooke's objections to the quantity theory, which comprised both a number of puzzles (e.g. why do interest rates rise with inflation and fall with depression?) and a number of suggestions, such as his statement that only incomes determine prices. Wicksell argued that although the quantity theory was the only theory with any claim to scientific importance, it was incomplete, in that it failed to give any explanation of *how* prices changed. His solution involved arguing, in denial of Say's Law, that prices change because supply is not equal to demand.

The first step was to define the *normal, or natural rate of interest*, the rate of interest at which the supply of savings equals the demand for loan capital.[8] This will be roughly equal to the expected yield on newly-created capital. In a simple credit economy, in which savers lend directly to borrowers, competition between borrowers would ensure that the rate of interest on loans equalled this natural rate of interest. In an organized credit economy, however, the situation is more complicated, for banks can grant loans in excess of the amount of savings deposited with them, for the money they lend will be returned to them as deposits. In a "pure credit" economy, where the only form of money is bank deposits backed by loans, there is no limit to the amount of credit that can be created in this way. It is only when banks have to hold reserves of metallic money that this process of credit creation has limits.

In such a credit economy the crucial factor is the *money rate of interest*, the rate of interest charged on loans. Consider a pure credit economy. If the money rate of interest is less than the natural rate, borrowing will exceed saving, with the result that demand for goods will exceed supply, and the price level will rise. Similarly, prices will fall if the money rate is above the natural rate. Price stability requires that the money rate of interest equal the natural rate. Thus Wicksell argued that in a pure credit economy the money supply was completely elastic, being able to sustain any equilibrium price

level, changes in this price level depending on the banking system's interest rate policy. The quantity theory and Say's Law held in equilibrium, but not in disequilibrium. Similar arguments would be true of an economy in which money comprised both credit money and metallic currency, but here there were limits to the ability of the banking system to change the money supply. In an expansion, for example, a shortage of reserves might force the banking system to raise the money rate of interest, bringing the expansion to an end.

To explain cyclical variations in interest rates (Tooke's puzzle mentioned above) Wicksell assumed that the natural rate of interest fluctuated, with the banking system changing the money rate of interest only with a lag. He thus had a view of the cycle in which increased optimism, or an increase in the rate of technical progress, could raise the natural rate of interest. Because the market rate lagged behind this, borrowing would exceed saving, credit would expand, and the excess of demand over supply would raise prices. In recessions this process would be reversed.

Wicksell thus supplied a theory of a credit economy, based on the ability of the banking system to cause divergences between the natural and the money rates of interest. Though developed independently, this had much in common with Thornton's theory. This monetary theory was then combined with a real theory of the trade cycle to provide an explanation of interest rates and the price level over the cycle. These ideas, in particular the *cumulative process* whereby a divergence between the money and natural rates of interest produces, not a once-for-all rise in the price level, but continuous inflation, influenced most twentieth century monetary theory, though the degree of Wicksell's influence varied enormously. Prior to the 1930s, for example, Wicksell's influence on English speaking economics was only indirect, coming in particular through Fisher, Cassel and Mises; whereas in the 1930s, due to the work of Hayek and Keynes, Wicksellian ideas were widely discussed. In Sweden itself, a definite Wicksellian school emerged.

Fisher

The second major contribution on which twentieth century monetary economics rests is that of Irving Fisher. Though contributing to the subject throughout the inter-war period, his major contributions date from before 1914: *Appreciation and Interest* (1896), *The Rate of Interest* (1907) and *The Purchasing Power of Money* (1911). Of his interest theory,[9] the aspect most relevant for monetary economics is his emphasis on the distinction between real and nominal interest rates.[10] As for his theory of money itself, this was, like Wicksell's, expressed in terms of the quantity theory, Fisher being responsible for the most widely used version of the equation of exchange,

$$MV = PT,$$

where M is the quantity of money, V its velocity of circulation, P the price

level, and T the volume of transactions. When bank deposits were brought in as well as currency (it was still currency that was thought of as money) the equation became,

$$MV + M'V' = PT,$$

M' and V' being the quantity and velocity of circulation of bank deposits. In this version of the quantity theory the emphasis was on the circulation of the means of payment, something Fisher took from Newcomb, whose *Principles of Political Economy* (1885) had contained a section on "the societary circulation", in which the equation of exchange was stated. This emphasis on the circulation of money used to finance transactions led Fisher to define T so as to include all transactions: not simply income, but transactions in intermediate goods and financial assets.[11]

In using the equation of exchange Fisher distinguished carefully between the "permanent or ultimate effects" of a change in, say, the quantity of money, and "temporary effects during periods of transition".[12] Permanent effects were obtained using a simple quantity theory: changes in the quantity of money would ultimately produce equiproportionate changes in the price level; the interest rate, velocity and the volume of transactions settling down at their normal values. During transition periods, however, monetary changes would produce changes in *all* the variables in the equation of exchange. The major reason for these temporary effects was, according to Fisher, the failure of the interest rate to respond sufficiently rapidly to price changes. For example, if M increased, this would increase the inflation rate, but if the money interest rate did not respond fully, the real interest rate would fall, and business would be stimulated. The volume of bank lending would increase, raising M' relative to M, and velocities would increase, causing prices to rise still further. Prices would continue to rise so long as the interest rate lagged behind its normal rate. During this inflation output would expand excessively, for prices "have to be pushed up, so to speak, by increased purchases".[13] Eventually, however, a rise in the interest rate will bring this process to an end.

Several points need to be made about Fisher's theory. (1) Fisher used this theory of transition periods to provide a monetary theory of the cycle, claiming that when taken in conjunction with maladjustments in the rate of interest, it was monetary factors that were the main cause of crises.[14] He argued that "overconsumption" and "overinvestment", in terms of which other economists were explaining crises,[15] arose because of monetary causes: "people spend more than they can afford [because] they are relying on the dollar as a stable unit when as a matter of fact its purchasing power is rapidly falling".[16] (2) He laid much more stress than did many of his contemporaries on the level of indebtedness relative to the changes in capital values brought about by changing prices and interest rates. This emphasis is clearest in his later "Debt-deflation theory of great depressions" (1933), in which he explained the great depression as resulting from the effects of deflation on an economy in which the level of indebtedness had risen

excessively. (3) Though Fisher stressed variations in the quantity of money, where Wicksell emphasized changes in the natural rate of interest, his analysis of transition periods had much in common with Wicksell's cumulative process. However, Fisher chose to stress the long run in a way Wicksell did not. For example, in *The Purchasing Power of Money* he wrote,

> So far as I can discover, *except to a limited extent during transition periods, or during a passing season* (e.g. the fall) there is no truth whatsoever in the idea that the price level is an independent cause of changes in any of the magnitudes, M, M', V, V', or the Q's.[17]

This emphasis on the long run was reinforced by his illustration, comparing the effects of a monetary change to the movement of cars in a train:

> The peculiar effects during transition periods are analogous to the peculiar effects in starting or stopping a train of cars. Normally the caboose keeps exact pace with the locomotive, but when the train is starting or stopping this relationship is modified by the gradual transmission of the effects through the intervening cars.[18]

The Cambridge school

Marshall's successors at Cambridge developed his cash balance approach to the quantity theory, according to which people desire to hold a certain fraction, k, of their resources, R, as money, M, with P being the *value of money* (i.e. the reciprocal of the price level as we are used to thinking of it). This is of course equivalent to Fisher's formula, with resources, or income, substituted for transactions. Despite this, however, the difference in formulation mattered, for it put greater stress on psychological factors and individual decisions. Thus all the major advocates of the "Cambridge equation" were led to stress changes in expectations and confidence as an important cause of changes in the value of money.

This emphasis is perhaps clearest in Pigou, who argued that people had two uses for resources, consumption apart: to hold money for the convenience and security it gives; or to use resources for the production of commodities, in order to obtain a profit. Resources will be allocated between these two uses so as to equate the marginal utility of investing (dependent on "the expected fruitfulness of industrial activity") with the marginal utility of holding money (dependent on the pattern of income receipts, the availability of other means of settling debts, and expectations of price changes).[19] These expectations were liable to change, this being the main factor explaining changes in the value of money.

The Cambridge version of the equation of exchange was thus regarded as a framework within which to analyse various effects on the value of money, not as the expression of a rigid quantity theory. According to Pigou,

> The quantity theory is often defended and opposed as though it were a definite set of propositions that must either be true or false. But in fact the formulas employed in the exposition of that theory are merely devices for enabling us to bring together in an orderly way the principal causes by which the value of money is determined.[20]

Yet the theory was regarded as more than a truism, for, in Keynes' words, it "flows from the fact that money as such has no utility except what is derived from its exchange value, that is to say from the utility of things which it can buy".[21] This was the justification for the assertion that demand for money was demand for a certain quantity of real cash balances, and for asserting that *ceteris paribus*,[22] the demand curve for money was a rectangular hyperbola.[23] However, even this was not consistently asserted. For example, Keynes argued that in the short run changes in M and P might cause changes in kR.[24]

As with Fisher's quantity theory, the Cambridge version was developed to allow for bank deposits, this being done in Pigou's equation,

$$P = (kR/M)[c + h(1 - c)],$$

where M, this time, is the quantity of legal tender money (what we might call high-powered, or base money), c is the fraction of money balances held as legal tender money and h is the fraction of their deposits that banks hold as reserves of legal tender money.[25] As with the other versions of the Cambridge equation, this provided a framework within which causes of price level changes could be analysed.

The evils of fluctuating prices were stressed both by Marshall and his followers, a variety of schemes being proposed to deal with the problem. Marshall emphasized the uncertainty introduced by price fluctuations, and the element of speculation this introduced into business life. He integrated this into a theory of the trade cycle:

> The consequence of this uncertainty is that, when prices are likely to rise, people rush to borrow money and buy goods, and this helps prices to rise; business is inflated ...; those working on borrowed money pay back less real value than they borrowed.

Because nominal wages are fixed,

> the employer pays smaller real salaries and wages than usual, at the very time when his profits are largest in other ways, and is thus prompted to over-estimate his strength, and engage in ventures which he will not be able to pull through after the tide turns.[26]

When credit is shaken, and prices fall, the process is reversed. Employees are unwilling to let money wages fall, so falling prices raise real wages and workers become unemployed as businesses are shut down. Firms reduce production in order to improve the market for their own goods, but in doing so they reduce demand in other markets. Thus fluctuations in the standard of value, according to Marshall, are always "either flurrying up business activity to an unwholesome fervour, or else closing factories and workshops by the thousand".[27]

Marshall's remedy for this situation was twofold. Firstly, he revived the idea of a tabular standard of value, whereby contracts could be made in terms of a unit of fixed purchasing power. This would reduce the significance of fluctuations in the value of money. Secondly, he advocated

basing the currency on both gold and silver, a unit of currency corresponding to a fixed quantity of gold *plus* a fixed quantity of silver. The value of such a currency would fluctuate with the average of the values of gold and silver, fluctuating by less than either metal alone.

Whereas for Marshall money was thought of primarily as metal, this was not the case for his successors. Though Keynes and Pigou both shared his view as to the effects of price fluctuations,[28] together with his desire to stabilize the price level, they attached greater importance to credit money, and their remedies varied accordingly, both of them stressing the management of the money supply. This stress on the need for management was particularly marked in the case of Keynes, whose emphasis on it, though going back to his *Indian Currency and Finance* (1913), was most clearly stated in his *Tract on Monetary Reform* (1923).[29]

To understand the *Tract on Monetary Reform* we need to consider the other side of the value of money: the exchange rate, which was explained in terms of purchasing power parity, the elements of which, although the name was coined by Cassel,[30] can be found in Marshall, and, before him, Ricardo. Purchasing power parity, in the Cambridge version, determined the equilibrium value of the exchange rate. If the exchange rate were too high, for example, equilibrium required that the price level fall sufficiently for purchasing power parity to be restored. In such circumstances, Keynes argued in the *Tract*, the government had to decide between devaluation and deflation. This involved management: a *decision* as to the appropriate price level. Given that capitalism could not be efficient, and might not even survive, without a stable price level, Keynes had no hesitation in advocating devaluation rather than deflation:

> For these grave causes we must free ourselves from the deep distrust which exists against allowing the regulation of the standard of value to the subject of *deliberate decision*. We can no longer leave it in the category of ... matters which are settled by natural causes, or are the resultant of the separate actions of many individuals acting independently, or require a revolution to change them.[31]

16.3 BUSINESS CYCLE THEORY BEFORE 1910

Before 1900[32]

In the period before 1900 there was little sustained analysis of the trade cycle. As was the case with Marshall and Wicksell, most of the leading economists were more concerned with the formulation and development of equilibrium theory. Discussion of the trade cycle was, apart from brief treatments, such as those considered above, confined to others. During the period, however, a variety of ideas were put forward, many of these anticipating ideas found in later writing. Nasse, for example, in 1879, attributed booms to inventions, associating particular booms with particular industries. In the period after the crisis of 1873, Price ascribed crises and depressions to over-consumption's destroying more wealth than was

produced, something that might happen as a result of excessive investment in fixed capital, high wages, or high government consumption. Similar was the explanation of Guyot who, in 1892, argued that over-investment in fixed capital could produce a shortage of circulating capital. Set against these "real" theories were ones stressing psychological factors and credit. We have already mentioned Marshall. Similarly Bagehot, referring to "a great many stupid people" having "a great deal of stupid money",[33] argued that credit gets extended during the optimism of a boom, but that this over-optimism will eventually be discovered, whereupon the structure of confidence and credit will collapse. In his account elements of multiplier and accelerator processes can even be discerned.

In addition to all these theories were those which attacked Say's Law: the under-consumptionist theories. An American, Hawley (1882) reached such a theory via the route of Mill's stagnation thesis, of the tendency of the rate of profits to a minimum.[34] He argued that there was a tendency towards over-accumulation, especially when profits were high during a boom. Another American, Crocker, in the 1880s reached similar conclusions through attacking Mill's doctrine of the impossibility of general over-production.[35] The main performance here, however, was that of Hobson. In his first book, *The Physiology of Industry* (1889)[36] he claimed that production was not equal to the maximum set by factor supply, being kept below this by excessive saving. Although Hobson was taken by Keynes as a precursor, this analysis, although containing many insights into the macroeconomics of saving and investment, was not altogether free from the Smithian perspective: saving was equated with investment, and differences between a barter and a monetary economy were not recognized.

Tugan Baranovsky

Despite this variety of earlier writings on the trade cycle, the modern literature stems, above all, from one work: Tugan-Baranovsky's *Industrial Crises in England* (1894). The background to this work was the debate amongst Russian Marxists in the 1890s over the question of whether Russian capitalism was in a position to create a market sufficient for its own development, an issue with important political implications.[37] Tugan-Baranovsky's position was that capitalism could expand indefinitely, demand being sustained by increased production of capital goods, accumulation being an end in itself. Contrary to Marx, but in agreement with Lenin, Tugan-Baranovsky argued that though capitalism would suffer periodic crises, these would not cause its collapse.

In this context Tugan-Baranovsky viewed cycles as an integral part of the process of capitalist development, explicitly rejecting "exogenous" explanations of the cycle, such as those of Jevons and Juglar. Cycles were connected with the persistent tendency, in a capitalist economy, towards over-production of capital goods, something which could, in a monetary economy, lead to general over-production. Whilst he denies Say's Law, recognizing the possibility of general over-production, Tugan-Baranovsky

argued that it was the behaviour of investment that was crucial. He found empirical evidence for this in fluctuations in iron production, iron being used above all, Tugan-Baranovsky claimed, in the production of capital goods.

Why, however, should the accumulation of fixed capital not proceed at a steady rate? Tugan-Baranovsky's answer was that, because so much of national income goes to the capitalist class, capital is not accumulated in the right proportions. In the upswing capitalists draw on accumulated funds to accumulate capital as fast as they dare. But the result is that eventually the accumulated funds become exhausted, and the interest rate rises. Because too many capital goods have been produced investment then falls, and a depression begins. In the following depression, loanable capital accumulates, for savings continue at a fairly steady rate, but before this can be transformed into productive capital it needs to be appropriately distributed amongst the various branches of production, something that need not occur, given the anarchy of an individualistic, competitive economy. So loanable capital accumulates until the pressure of funds seeking investment outlets is sufficient to overcome industry's resistance, and it begins to be transformed into fixed capital. As expansion in one area tends to spread throughout the economy, new demand is rapidly created and the economy enters a new phase of prosperity.[38]

Spiethoff

Spiethoff's contributions, dating from the year after the German edition of *Industrial Crises in England*, were strongly influenced by Tugan-Baranovsky's work, as well as by the German historical school.[39] Like Tugan-Baranovsky, he rejected Say's Law as inappropriate to a monetary economy, but where Tugan-Baranovsky explained fluctuations in investment in terms of alternating shortages and gluts of loanable capital, Spiethoff introduced other features. Firstly, he explained the boom in terms of innovations, or the discovery of overseas markets, either of which could raise profitability in some particular sector, thus starting the upswing. Secondly, he explained the crisis in terms of limited opportunities for investment: during the boom, output of consumption goods will lag behind investment, so consumption goods prices stay high, keeping profits high. However, when the new investment eventually starts to result in increased production of consumption goods, prices must eventually fall, even though cartels may manage to maintain prices for a while. Investment will have to be curtailed, because once the new plant is installed, there remains only the task of maintenance and replacement.

These influences, operating on investment through profitability and the incentive to invest, were seen by Spiethoff as complementary to Tugan-Baranovsky's explanations in terms of the supply of capital. In a boom the process of expansion was limited not only by limits to the demand for real capital, but also by shortages of capital. These shortages, occuring at the peak of the cycle, were very much real shortages, for it was shortages of real

capital goods which underlay the tightened conditions in the capital market. The wrong type of goods had been produced. Thus monetary and credit policies would do nothing to prevent the crisis.

1900–1910

The work of Tugan-Baranovsky and Spiethoff stimulated much work on business cycles, and in the first decade of the twentieth century a number of important contributions were made. The first of these concerns the acceleration principle, the theory that a relatively small increase in the demand for consumption goods can produce a much larger increase in the demand for investment goods. The essence of the idea was first put forward by Carver (1903), but the credit for developing it and integrating it into a theory of the cycle is due to Aftalion (1910).[40]

The key to Aftalion's theory of the cycle was the idea that wants could become saturated, something he explained in an Austrian manner, in terms of diminishing marginal utilities. The stock of capital goods had to be adjusted to the demand for consumer goods, something that was difficult, for three reasons: (1) the acceleration principle meant that even small fluctuations in demand for consumer goods could produce large fluctuations in the demand for capital goods; (2) the long period required for the construction of capital goods; and (3) the durability of capital goods. Thus in a boom projects would be started without initially producing any goods to satisfy demand; when they started to produce, demand would become saturated, and investment would fall. Depression would follow, and this would last until a sufficient number of capital goods had worn out for there to be a shortage of capital goods relative to demand. The gestation lag in new investment projects, and the durability of capital, were thus important in determining the length of the cycle. Aftalion thus found the amplitude and the timing of the cycle to be inherent in the techniques of production.

Another important contribution was that of Schumpeter, who emphasized the role of innovations in the process of economic development. Schumpeter's argument was that innovations inevitably occur in waves: when an entrepreneur innovates, others follow, for the task of following an innovator is easier than that of first making the discovery. The appearance of an innovation moves the economy out of equilibrium, creating new opportunities for profit, which are gradually exploited. Depression then follows as the economy settles down to a new equilibrium. Boom and slump are, for Schumpeter, essential aspects of the process through which equilibrium is re-established after a wave of innovations.

As regards the monetary aspects of the cycle, an important contribution was that of Johannsen (1908) who, though not alone in arguing the under-consumptionist case, was the most original. His originality was twofold: a perceptive analysis of the relationship between saving and investment; and a statement of the "multiplying principle" linking consumption to investment. Johannsen distinguished carefully between saving, the act of refraining from consumption, and investment, the purchase of

capital goods. It is only when savings are invested in new wealth (what Johannsen calls the "capitalistic form" of savings) that aggregate demand will be unaffected by the level of saving. If savings are hoarded, or are used to purchase already-existing property ("impair savings") this will reduce demand and hence profits, something akin to Keynes' theory in the *Treatise on Money*. Johannsen's version of the multiplier, too, is reminiscent of Keynes: a fall in one group's spending lowers the incomes of another group, which in turn reduces its spending, thus reducing the incomes of a third group, and so on.

16.4 BUSINESS CYCLE THEORY, 1910–1930

If the previous period was the one in which the main explanations of the business cycle were first put forward, virtually all the materials for a comprehensive theory being available by 1910, it was the following two decades that constituted the heyday of business cycle theory. Between 1910 and 1930 many studies of the cycle were published, the major ones being by Mitchell (1913, 1927), Cassel (1918a), Hawtrey (1913, 1919), Robertson (1915, 1926) and Pigou (1912, 1927). From the 1930s the question of the trade cycle became, in part due to Keynesian influence, to a certain extent a theme subsidiary to the theory of employment.

Mitchell

Arguably the most important of these writers on the cycle was, at the time at least, Mitchell, for it was his work which popularized the notion of the business cycle, not merely amongst economists but amongst businessmen and politicians.[41] In addition, the statistical work on the business cycle undertaken both by Mitchell and his students, marks him out from other economists working on the problem of cycles.

Mitchell's emphasis was on the business cycle as an integral part of business life, its rhythmic pattern being evident not merely in one or two major aggregates, but permeating all economic activity. He saw his task as investigating the nature of these cycles – the nature of the process whereby prosperity is transformed into depression, and depression into prosperity – and to do this he accumulated and brought together an ever-increasing wealth of statistical data. This task of building up a comprehensive picture was started in Mitchell's *Business Cycles* (1913), continuing in his work with the National Bureau for Economic Research, an organization Mitchell was instrumental in founding, becoming its director from its foundation in 1920 to 1945. It was through the National Bureau that his *Business Cycles: the Problem and its Setting* (1927), and *Measuring Business Cycles* (1946) were produced.

This statistical work on business cycles, together with the other statistical work associated with it, such as estimates of national income, and studies of its distribution, constitute in their own right a chapter in the history of

economics. Our main concern here is with the implications of Mitchell's work for the theory of the business cycle. Mitchell has been portrayed as an advocate of measurement without theory, but this is far from the case.[42] There is a sense in which it is truer to see Mitchell as synthesizing the various theories of the cycle. He argued that we need working hypotheses to guide the selection and analysis of data, this being provided, in his work in business cycles, by a survey of current theories.[43] In surveying these theories Mitchell found a place for virtually all the then current theories, seeing them as differing in emphasis rather than in principle.[44] It is possible to view Mitchell as being too uncritical when he claims that the problem is not that there are too few acceptable theories, but that virtually every theory appears justifiable.

Where Mitchell considered earlier writers to have gone wrong was in isolating one particular factor, seeing that as the true cause of the cycle. He argued that because the cycle was an extremely complex phenomenon, comprising numerous actions and reactions, it was more profitable to concentrate on trying to understand the nature of the process as a whole, rather than on trying to single out any one cause as fundamental. This was a task which, according to Mitchell, required statistical analysis, not deductive theory.[45]

However, just as it would be wrong to see Mitchell as a pure empiricist, so too would it be wrong to see him simply as an eclectic in his theory. There was, though it was deliberately not formulated as a formal theory, a theoretical framework underlying his approach, one very much influenced by Veblen.[46] Mitchell saw the business cycle as rooted in a particular institutional structure – that of a system of inter-related, large-scale organizations, the object of which was to make money. The interdependence of all prices was described in almost Walrasian terms. There was substitution between goods on the demand side; prices and costs were linked; competition equalized the rate of profit; and security prices were linked to profits.

> At whatever point analysis may begin, tracing the interlocking links of the price chain, to that point will it come round again if it proceeds far enough. ... Thus all prices in a business economy are continually influencing one another. To account for any one item in the system, one must invoke the whole.[47]

Despite this, however, his assumptions were not Walrasian, for he claimed that the notion of equilibrium was inappropriate for describing an economy, for any economy would always be in motion. Furthermore, he did not accept the hypothesis of perfect competition: not only might there be imperfections of competition (Mitchell was never very clear as to exactly what he assumed here), but also prices and profits were to a substantial extent determined by institutional factors. Many prices were rigid because of contracts and conventions. The institution, however, to which Mitchell attached most importance was that of money. In contrast with the Walrasian, or even the Marshallian system, which in other ways had something in common with Mitchell's, relative prices, and hence profits, depended as much on credit conditions as on real factors.

Business cycles emerged naturally from this framework. Revival emerges naturally from depression as profits rise, something which occurs as a matter of course through costs falling relative to output prices, falling interest rates and the increased availability of finance through the banking system. Profits rise and an expansion develops, until it in turn generates forces which disrupt it: costs rise relative to selling prices, and credit becomes scarce, and hence expensive, or it might be due to other factors, such as errors in businessmen's expectations. Depression begins and the cycle is completed.

For Mitchell, therefore, the business cycle was seen as inherent in an economy organized around making money. The choice of the term "business cycle" as opposed, for example, to the term "industrial fluctuations" favoured by Robertson and Pigou, was no accident: on the one hand it echoed Veblen's distinction[48] between the business, or pecuniary, aspect of economic activity, and the "industrial", or technological, aspect; and on the other hand, it stressed the rhythmic, or cyclical nature of variations in the level of activity.

However, although this framework proved fruitful for empirical research, Mitchell and the National Bureau filling out much of the statistical detail, Mitchell stopped short at explaining *how* this process worked itself out. Nowhere, for example, does he explain *why* prices move relative to costs in the way they must do if his theory is to hold. Despite Mitchell's impressive statistical work, and despite his bringing together many theories of the cycle, we have to look elsewhere to find progress in the theory of the business cycle.

Cassel

Although written in 1914, Cassel's work on the cycle, published as Book IV of *The Theory of Social Economy*, was published in 1918. Like Mitchell, he dealt with cycles rather than crises, seeing them as the product of a specific epoch in economic history:[49] and he used statistical evidence to support his theory. Beyond this, however, he has little in common with Mitchell. Whereas Mitchell stressed the pervasiveness of the cycle, Cassel drew from his statistics the conclusion that the cycle was a phenomenon associated primarily with fixed capital formation, not with consumption, the latter fluctuating little over the cycle. Furthermore, where Mitchell was eclectic in his use of previous theories, Cassel built on Spiethoff's over-investment theory. Where he differed from Spiethoff was in seeing a greater role for monetary factors.

At the start of the upswing, according to Cassel, profits will be high relative to wages, and banks will be lending at too low a rate of interest. These two effects combine to stimulate the production of fixed capital. As the boom proceeds, the proportion of production devoted to capital formation rises, whilst the proportion devoted to savings does not. A shortage of capital develops, and interest rates rise. At the same time wages rise relative to profits, causing demand for investment goods to fall.

Important too is the accelerator, transforming a fall in the growth rate of consumption demand into a fall in the level of investment. In the downswing which follows, these forces are reversed, savings increasing in relation to the production of capital goods, and interest rates falling. At the trough, however, the link between monetary factors and investment is much weaker than at the peak, a conclusion Cassel found to be supported by statistical evidence.

In this theory Cassel explained the role of lags in preventing the economy from being in equilibrium continuously, these including lags in the response of investment to changes interest rates; the reaction of interest rates to changes in investment; and the time taken between the start of an investment project and its completion. In addition, the economy is periodically subject to disturbances which renew the cyclical activity, such as advances in technology (e.g. railways, electricity) or the opening up of new countries. His main difference with Spiethoff was in arguing that the shortage of capital at the crisis was a shortage of *monetary* savings, not a shortage of real capital goods. Banking policy was thus seen as important.

Mises

A theory which laid even greater stress on monetary factors in the cycle was that of Mises (1913). Using a model which in other respects had much in common with Cassel's, Mises brought in Wicksell's cumulative process, whereby prices change in response to any discrepancy between the natural and money rates of interest, to analyse the monetary aspects of the cycle. According to Mises it was the failure of the banking system to keep the money rate of interest equal to the natural rate, together with an ideology, prevalent amongst businessmen and politicians, in favour of low interest rates, which was responsible for the persistent tendency of the economy to expand excessively, producing crises and hence the cycle[50] This view of the cycle is considered below, as it was later taken up by Hayek.

Hawtrey

Whilst Cassel and Mises both emphasized the role of money in the cycle, neither gave it so prominent a place as did Hawtrey, whose most important works during this period were *Good and Bad Trade* (1913) and *Currency and Credit* (1919), the latter in particular being extremely influential in the 1920s.[51] Hawtrey was strongly influenced by Marshall, but differed from him in attaching more importance to money: he argued that although they were correct in claiming that money was not the same thing as wealth, orthodox economists had failed to recognize that money was nevertheless "a most potent factor in economic organization".[52] Whilst Cassel and Mises allowed for the influence of real as well as monetary factors in the cycle, Hawtrey stressed that it was only the latter that really mattered.

The basis of his theory was a distinctive view of the relationship between money and income (*nominal* spending), one which goes beyond the

conventional quantity theory relationship. When a *new* credit is created, Hawtrey argued, through a bank granting a new loan, it will be spent, this expenditure constituting income for its recipients. This income will raise spending, which in turn will generate further income, a process which will carry on until the purchasing power in circulation is returned to the bank to pay off the original loan. It is thus crucial to Hawtrey's theory that when he refers to a given stock of money (or the margin of unspent purchasing power, as he calls it) he is envisaging a situation in which new credit is continually being created, this being counterbalanced by the repayment of older loans. This means that for Hawtrey an increase in the quantity of money occurs when the flow of new credit exceeds the flow of purchasing power being returned to the banks to cancel old loans. Changes in the money supply are thus *directly* linked to changes in the flow of income.

This link between money and income is central to Hawtrey's theory, but it needs to be supplemented by two things: an explanation of why changes in nominal spending produce changes in employment; and an account of Hawtrey's view of the role of interest rates.

Hawtrey saw a strong link between the short term interest rate and the supply of bank credit, the crucial role being played by *dealers*. It is dealers from whom producers receive their orders, and because they hold large stocks of goods, financed by a volume of credit which is large in relation to their turnover, they are very sensitive to changes in the rate of interest. Thus a rise in the short term rate of interest makes dealers wish to reduce their inventories, causing them to reduce their orders from producers. Production falls and unemployment rises. At the same time, because dealers hold lower inventories, demand for credit falls, and the money supply is reduced. Thus a rise in the short term rate of interest would have a strong effect on both money and employment.

The reason why this process leads to a fall in output and unemployment is that prices take time to adjust. In particular, wages are sticky. Thus when, in response to a fall in spending following a monetary contraction, prices and wages fall, interest rates can fall again, monetary demand and the level of employment being restored to their former levels. Money wage rigidity was, for Hawtrey, the key to the link between money and employment.

Three factors in particular are brought in to explain why cycles will emerge in this setting. The first is the inherent instability which Hawtrey sees in the system of bank credit. High profits cause lending to rise, and this in turn raises profits still further, and so on. The second is the lag which occurs between the increase in the volume of spending and the increase in the demand for money which follows it. When new credit is granted, it is used, by dealers, to purchase goods from producers, these being transactions which are settled by cheque rather than with currency. It is as the increase in purchasing power spreads to wage-earners that the demand for currency increases, this occurring only some time after the initial increase in credit. Finally, there is the lag involved in the production process. This is important because Hawtrey assumes that producers require a continual stream of new credit throughout the production process; which means that

once bankers are committed to supporting a project they will be committed, morally if not legally, to continue to advance credit until the process is complete. The outcome of these three factors is that the banking system is unable to respond appropriately. For example, in an expansion, by the time rising demand for currency causes banks to lose reserves and they raise interest rates, the level of demand will already have expanded too far. In addition, despite raising interest rates, banks will be unable to cut back their lending immediately, due to their commitments to producers. Thus monetary expansion will typically be excessive, this necessitating a contraction as banks attempt to restore their reserves to an appropriate level. The downswing is, however, likely to be carried too far, and a cycle is the result.

Hawtrey worked out his theory not only for an isolated economy, but also for an open one, both with fixed and flexible exchange rates (a gold standard and an inconvertible paper currency). He incorporated into it the Wicksellian natural rate of interest and Fisher's theory of the real interest rate. Although Hawtrey did use his theory to analyse the effects of real disturbances, this was only in order to argue that the important influences on the level of activity were monetary in origin: without a change in the money supply there could not be any significant fluctuations in the flow of purchasing power, and hence the scope for fluctuations was very limited.

Hawtrey's analysis, in addition to its influence in the 1920s, is important in the development of Keynesian economics. Although the ideas are not developed so thoroughly as in Keynes' *General Theory*, the multiplier and the view that changes in demand have a direct effect on the volume of output are both present in *Good and Bad Trade.*

Robertson

In contrast to Hawtrey, Robertson, in *A Study of Industrial Fluctuations* (1915) analysed the cycle in terms of real factors: in addition to following Jevons in giving considerable attention to crop cycles and the influence of agriculture on trade, Robertson provided an account, independently of Schumpeter, of the stimulating effect of innovations, and he made use of Aftalion's over-investment theory. Innovations could start an upswing through their effects on costs and demand. During the upswing it would take time before investment led to increased output, and so one possible reason for the downturn, when this came, was that a surplus of particular capital goods might emerge. Shortage of saving provided an alternative reason. Robertson's achievement, in distilling these conclusions out of a welter of facts, has been described by Hicks as involving "almost miraculous insight".[53]

Very different was Roberton's later *Banking Policy and the Price Level* (1926), a book in which he cooperated closely with Keynes,[54] in which he analysed more thoroughly than anyone before him the process of saving and investment. In the key chapter in this book, entitled "The different types of saving", Robertson distinguishes between a multitude of types of saving. His use of the term "lacking", as being more neutral than the more traditional "abstinence" or "waiting", was one of the reasons why his

terminology appears very strange. Whilst Robertson's terminology was far too complicated to appeal to other economists,[55] it was important in that it defined saving and investment (the supply of and demand for lacking) in such a way that they could be unequal. The reason for this was the possibility of forced saving, something which could be brought about through the banking system. Suppose the banking system expands the supply of credit through lending to investors: investors will be able to purchase a larger share of output, doing so at the expense of consumers.

In applying these concepts to the cycle, Robertson retained the idea that there could be real causes of the business cycle, describing fluctuations due to such causes as "appropriate". When real costs or real demand changed, it was quite appropriate for output to change. However, in addition to these fluctuations, there were fluctuations in output that were caused by price fluctuations and the behaviour of the banking system. It was the duty of banking policy to prevent this. However, and this is where Robertson's analysis of saving and investment was so important, the task of the banking system was not so simple as it might sound, for credit creation had two aspects. On the one hand, there was the effect of banking policy on the price level, but on the other there was its effect on forced saving. In addition to the task of assisting price stability, the banking system had the task of ensuring that an appropriate volume of savings, forced if not voluntary, was available for investment. These two tasks could conflict, Robertson arguing that under some circumstances the banking system should abandon the goal of price stability, in order to achieve an adequate supply of savings.

Pigou

We close our discussion of the period up to 1930 with a discussion of Pigou, whose *Industrial Fluctuations* (1927) can be regarded as the best attempt at a synthesis of alternative theories. Pigou's first analysis of the cycle was provided in *Wealth and Welfare* (1912), where it arose naturally out of his aggregative approach to welfare economics. After investigating the factors determining the size and distribution of the national dividend, Pigou turned to the question of its variability, and hence the variability of employment. He argued that the causes of fluctuations were too closely bound up with the general body of economic activity to permit an isolated treatment of them. It was from the second edition of *The Economics of Welfare* (1924) that Pigou dropped this arrangement, assigning discussion of fluctuations to a separate volume.

There were two main features of the framework within which Pigou brought together much of the previous literature on the cycle. The first was his systematic use of the distinction between two problems: that of the initiating impulses disturbing the economic system; and that of the conditions under which such impulses operate to cause fluctuations in economic activity. Pigou was not the first to draw this distinction between what came to be known as the impulse and propagation problems, credit for this being due to Wicksell,[56] but Pigou was influential in bringing it into general use. The second aspect of Pigou's approach was that, like Mitchell, he stressed

expectations of industrial profit as the mechanism through which various impulses affected the economy. Thus although Pigou can, in some respects, be thought of as having a "psychological" explanation of the cycle, his framework was sufficiently elastic to allow for a large variety of influences: autonomous real or monetary factors, just as much as psychological factors, could influence expected profits and hence economic activity.

Pigou's method was very much in the Marshallian tradition. Like Marshall's work, *Industrial Fluctuations* comprised a blend of theory and empirical analysis. More important, however, was the fact that Pigou's theoretical framework was Marshallian. It deals with a competitive, though not perfectly competitive, world; one in which the changing economic equilibrium can be analysed in terms of the elasticities of industry demand and supply curves.

In view of developments in the 1930s and beyond, two features of *Industrial Fluctuations* must be emphasized. The first of these is Pigou's stress on the role of fluctuations in demand as the main cause of fluctuations in employment. In analysing the propagation of demand from one sector to another, Pigou is describing what is essentially a multiplier process: what is missing as compared with later theories is not the idea so much as the *technical* device of the multiplier.[57] It was due to technical limitations that Pigou reached the conclusion that it was not possible to estimate quantitatively the effects on economic activity of an increase in demand.

The second feature of Pigou's work is his emphasis on the labour market. Although in many ways in the tradition of earlier writing on the cycle, *Industrial Fluctuations* is in other ways much closer to modern theory, in which the theory of employment is primary. The key concept here is Pigou's elasticity of supply of floating capital, for it is this which determines the extent to which an increase in profit expectations, caused by whatever initiating impulse, will affect unemployment. Through this elasticity, the reduction in unemployment resulting from an increase in expected profits will depend on the elasticity of supply of credit (and hence on the banking system), and on the extent of any changes in prices and wages (which affect the volume of real resources that a given supply of finance will command). Pigou supplied a variety of reasons why prices and wages would be sticky, failing to equate the supply of and demand for labour. It would be wrong, however, to conclude that Pigou saw wage stickiness as the fundamental problem. If demand were severely depressed, for example, wages might have to fall very low, or even become negative, for full employment to be achieved: aggregate demand, rather than wage stickiness, could well be the problem when there was unemployment.[58]

16.5 THE THEORY OF MONEY AND EMPLOYMENT, 1930–1936

The 1930s opened with two influential works, appearing almost simultaneously – Hayek's *Prices and Production* and Keynes' *Treatise on Money* –

works which dominated discussion of money and the cycle for the next few years. Though both their conclusions and the routes by which they reached these were very different, the ideas in both these books can be traced back to Wicksell, of whose theory they were logical extensions. Given this, together with the influence of Wicksell on Swedish economists working on similar problems, it could be argued that the first half of the 1930s was when Wicksell's influence was at its height.[59]

Hayek

Although Hayek's ideas became neglected after the success of the *General Theory*, they were at the time highly regarded, and the centre of much controversy. Schumpeter refers to the "sweeping success" of *Prices and Production*, a success not equalled by that of any theoretical book at a comparably abstract level.[60] Hayek's key argument was that the trade cycle arose because the organization of the banking system made it difficult to avoid periods of excessive monetary expansion. When, as must inevitably happen, such monetary expansion is brought under control, the result will be depression, the severity of the depression being directly proportional to the length of the previous boom.

The framework Hayek used to reach this conclusion comprised two main elements: Wicksell's cumulative process, and Böhm-Bawerk's theory of capital.[61] During the period of expansion, which might owe its origins to real causes, excessive monetary expansion would lead to forced saving, lowering interest rates, and raising the prices of producers' goods relative to those of consumers' goods.[62] Production of producers' goods would thus rise, but because Hayek assumed an initial state of full employment,[63] the only way in which the capital stock could be increased was through a lengthening of the period of production. However, when monetary expansion ceased, forced saving would also cease, total savings returning to their previous level, thus raising the rate of interest. The price of consumption goods would then rise relative to that of producers' goods. This rise in the interest rate, and the associated shift in relative prices, makes the longer processes of production that were started during the boom unprofitable, and they will be shut down, thus releasing labour. Shorter production processes will be expanded, but this will take time as stocks of all the appropriate intermediate goods have to be built up, and during the time it takes to do this there will be unemployment.

Unemployment occurs because the employment of labour requires capital goods, and the capital goods released when the longer, capital intensive processes are terminated are unsuitable to be used in the shorter processes that need to be started up. Demand for labour is thus, in the short run, very inelastic, and unemployment is the inevitable result of a shortening of the period of production. Monetary expansion will not prevent it: if investment were to be increased, this would merely postpone the required adjustment, making the crisis, when it eventually came, worse; if consumption were to be increased, this too would raise the price of consumption

goods relative to that of producer goods, making still shorter processes profitable, thus exacerbating the situation. The only way the situation could, according to Hayek, be restored, is through an increase in the level of voluntary saving sufficient to make the longer processes of production, introduced during the expansion, profitable.

Keynes' treatise on money

Very different was the *Treatise on Money*, in which a Wicksellian construction was used to argue that monetary expansion could cure unemployment through raising the price level. To understand the argument Keynes was using, we have to consider what he called his *fundamental equations*. The starting point for these was Keynes' idiosyncratic system of national income accounting, in which earnings (E) were defined as including only the *normal* earnings of factors of production, excluding "windfall" profits or losses (Q). Defining the value of output as a whole as Y, the value of investment as I and the value of consumption as C, we have

$$Y = C + I$$
$$Y = E + Q.^{64}$$

Defining savings as *earnings* less consumption,

$$S = E - C,$$

we obtain

$$I = S + Q.$$

That is, investment is equal to savings plus windfall profits, which may be positive or negative. If we now define real output, y, and the price level, P, such that $Y = Py$, we can derive the following equation for the price level:

$$P = (E/y) + (I - S)/y.$$

This shows the price level to be determined by two terms: normal earnings per unit of output, plus the difference between investment and savings (i.e. windfall profits) per unit of output.

The fundamental equations enabled Keynes to distinguish between two types of inflation: *income inflation*, where normal earnings per unit of output were increasing, and *profit inflation*, arising because the actions of the banking system caused investment to exceed saving. The significance of this distinction was that it was only in a profit inflation that entrepreneurs would have an incentive to increase production. It led also to the so-called "widow's cruse" (sic) theory of distribution, for if entrepreneurs were to increase their spending on investment, they would find that this increased their prices and hence their profits: profits were like the widow's cruse,[65] in that however much they were spent, they would never be exhausted.

In using the fundamental equations the crucial factor was the rate of interest. The rate of interest depended on supply and demand for money, where the latter comprised two categories: *cash deposits*, deposits needed to

finance transactions; and *savings deposits*. Demand for cash deposits was explained by Keynes in terms of a quantity theory, but demand for savings deposits was determined very differently, depending on the "bearishness" of the public. The behaviour of the rate of interest thus depended on the willingness of the banking system to respond to the changing demands of the public, these depending on investors' expectations.

A monetary expansion would thus act on the economy in the following way. (1) Given the level of "bearishness", it would lower the rate of interest. (2) This would raise the level of investment relative to savings. (3) The rise in investment relative to savings would raise prices relative to earnings, increasing windfall profits. (4) Finally, the rise in profits would cause entrepreneurs to expand production and increase employment. Monetary expansion could thus be used to lower unemployment, this operating through its effect on prices and profits.

The Swedish contribution

In the early 1930s Wicksellian ideas were also being developed in Sweden. Whether or not they can justifiably be called a "Stockholm School"[66] the economists concerned, Lindahl, Myrdal, Ohlin, Hammarskjold and Lundberg, shared a Wicksellian heritage, and they built upon it to produce works often considered to anticipate the *General Theory* in important respects.[67] Myrdal, for example, claimed that Keynes' *General Theory* constituted, for Swedish economists, an important development along a familiar line of thought, not a revolutionary breakthrough[68] Despite this, however, there were important differences between their approach and that of their British contemporaries.[69]

The first of their contributions was that of Lindahl, in particular his essay "The rate of interest and the price level" (1930). In this Lindahl, after arguing that the quantity theory, though remaining a significant part of the theory of value, failed to provide a "satisfactory and generally valid" explanation of changes in the value of money. Such a theory, he argued, had to be sought through extending the general theory of price to the problem of price relations between periods.[70] To provide such an explanation, Lindahl constructed a dynamic period analysis, in which a "period" had two characteristics: during it, the factors directly influencing prices are taken as constant; and prices are in equilibrium such that supply equals demand.[71] The price level within each period is explained in terms of the relation

$$E(1 - s) = PQ,$$

where E denotes money income, s the proportion saved, and P and Q the price and quantity of consumption goods. Using this framework Lindahl could then analyse the effects, under different circumstances, of a change in the rate of interest. Starting from an initial situation of full employment, he obtained Wicksellian conclusions, but when he started from a situation of unemployment he found that the rise in incomes caused by the fall in the

rate of interest would raise consumption, thus increasing employment, not only in the investment goods sector, but also in the consumption goods sector.[72] Lindahl's emphasis, however, was not on this so much as on the fact that the price level would rise by less in such a situation than in a situation of full employment.[73]

Two years later there appeared the first version of what was to be the most important Swedish contribution: Myrdal's *Monetary Equilibrium*,[74] a book which developed Lindahl's interpretation of Wicksell's cumulative process. Myrdal described his main contribution as being "to include anticipations in the monetary system".[75] He did this through introducing the distinction between the *ex ante* and *ex post* values of a variable. Though he did not provide a full dynamic analysis, the fact that he was, like Lindahl, thinking in terms of a dynamic period analysis was crucial: *ex ante* quantities were defined in terms of the action planned at the start of a period; *ex post* quantities were those measured at the end of the period.[76] This distinction was important only in a dynamic economy, for in a stationary state it would be unimportant whether a period was viewed from the beginning or the end.

This approach led Myrdal to start from the position that,

> In the *ex post* calculus there is, ..., by necessity an exact balance between the invested waiting and the value of gross investment. Looking forward there is no such balance except under certain conditions which remain to be ascertained. In the *ex ante* calculus it is a question not of realized results but of the anticipations, calculations and plans driving the dynamic process forward. [If investment or saving changes] there must be a tendency *ex ante* to a disparity. The real problem to be solved in monetary theory is: How does this tendency to disparity in the saving–investment equation develop into *ex post* balance?[77]

Along with his emphasis on the importance of the period implicit in monetary analysis, Myrdal considered the concepts of *ex ante* and *ex post* to have been his chief contribution.

Using these concepts Myrdal was able to proceed to his "immanent criticism" of Wicksell, reformulating the latter's concept of monetary equilibrium. After replacing Wicksell's natural rate of interest with an anticipated, *ex ante*, yield on investment, he concluded that monetary equilibrium required a profit rate sufficient to stimulate "just the amount of total investment which can be taken care of by the available capital disposal".[78] Though retaining what Myrdal described as the "fundamental part" of Wicksell's theory, he thus abandoned Wicksell's zero profit condition (equality of the natural and money rates of interest). Central to Myrdal's own equilibrium condition was that it was defined *ex ante*. He also parted company with Wicksell when considering the relationship between monetary equilibrium and price stability, for he argued that monetary equilibrium was compatible with any inflation rate.[79] However, because some prices are stickier than others, monetary policy needed, according to Myrdal, to aim at "adapting the flexible prices to the absolute level of the sticky ones".[80]

It is in considering sticky prices that Myrdal derived results which paralleled those of the *General Theory*. He considered a tightening of credit, this lowering capital values, reducing profit margins and lowering investment. This will cause incomes to fall, both in the investment goods and the consumer goods sectors, unemployment rising in both due to wage stickiness. When incomes fall in this way, Myrdal argues, consumption will fall by "significantly less" than total income, savings being reduced despite the rise in the rate of interest[81] He thus sees the possibility that this fall in saving will serve to maintain monetary equilibrium. However, Myrdal is still seeing price changes as the motive force behind the process, for he argues that if consumption goods prices do not fall, the process will come to an end.[82] After this, Myrdal goes on to assert the following, Keynesian sounding, proposition:

> If the forces maintaining consumption are strong enough and if the reaction of total real investment to a shrinking profit margin is sufficiently small, then the effects of the credit policy will be neutralized.[83]

We thus have here two familiar ideas: that monetary equilibrium (savings equal to investment) is possible at a variety of levels of employment; and that monetary policy will, under certain circumstances, be ineffective. Notice, however, that Myrdal's emphasis is throughout, as befits his objective of providing an "immanent criticism" of Wicksell, on the implications of his analysis for the price level rather than for the level of employment.

Although Lindahl and Myrdal made what were probably the most important of the Swedish contributions to the monetary theory, they were far from alone. Others worth mentioning are Hammarskjold, Ohlin and Lundberg. Hammarskjold's contribution was twofold. He not only provided a formal, algebraic exposition of period analysis, but also stressed windfall profits as the link between successive periods. In this he was strongly influenced by Keynes' *Treatise on Money*. Ohlin, as was the case with his work on international trade, stressed quantity adjustments as part of the mechanism whereby equilibrium was brought about. Finally, it was Lundberg (1937) who analysed a process in which there is continual disequilibrium, in the sense that expectations are not fulfilled in each period.

The multiplier

The view that an increase in spending in one sector of the economy might, through its effect on incomes, raise spending in other sectors, was widely held in the late 1920s. Not only was it present, albeit without being given a prominent role, in theoretical writings on the trade cycle, from Hawtrey's *Good and Bad Trade* to Pigou's *Industrial Fluctuations*, but it was also important in discussions of public works expenditure.[84] What was missing from these discussions was the technical device of the multiplier, whereby the secondary effects on employment were determined by the marginal propensity to consume.[85]

One source of the multiplier was Hawtrey who, in a paper prepared for the MacMillan Committee[86] at the beginning of 1931, produced a numerical example relating an increase in investment to an increase in output. In his example, 40% of income was saved, and Hawtrey argued that a £5m increase in investment would produce a £12.5m increase in output.[87] The other, and more well known, source of the multiplier was Kahn's article "On the relation of home investment to unemployment" (1931). In this he calculated the amount of secondary employment that would result from employing an extra man on road building. The ratio of secondary to primary employment was determined by the distribution of income, and the fractions of profits and wages spent on home-produced consumption goods.[88] Meade's contribution to the theory of the multiplier was to show that at the end of the multiplier process, the increase in savings would, ignoring changes in imports and exports, equal the initial increase in investment. In other words, an increase in investment could, under some circumstances, generate enough savings to finance itself.

Because of the way the multiplier is often presented nowadays, it is important to emphasize that Kahn's multiplier was *not* a theory of aggregate demand. It was simply a relationship between the primary and secondary employment effects of an increase in public investment. Similarly, in discussing the relationship between saving and investment, the emphasis was on (1) the reduction in unemployment benefits, and (2) the fall in the balance of trade that would result from a rise in investment. The multiplier was not derived, as in modern theory, from a saving–investment equilibrium condition.

Keynes' General Theory

Immediately after its publication it became clear that there were serious defects in the central theoretical chapters of Keynes' *Treatise on Money*: in particular in those dealing with the "fundamental equations" and the dynamics of the price level. To a certain extent Keynes was aware of the book's inadequacies,[89] but far more important was the barrage of criticism to which the book was immediately subjected. Some criticisms could be dismissed, but there were many that could not be. In particular, it was argued that the fundamental equations were merely truisms, explaining nothing; that there were problems with the units in which Keynes measured quantities; and, above all, that he failed to deal adequately with changes in output. This point had been made by Hawtrey who, even before the book's publication, had criticized Keynes for not recognizing that a fall in demand might, even without any fall in prices, lead to a fall in output.[90] More important, however, was the sustained investigation of this issue by a group of younger Cambridge economists (the so-called "circus", of which Robinson, Meade and Kahn were the most prominent members) which met regularly to discuss the *Treatise*.[91] Kahn's article on the multiplier had been available before the *Treatise* was published, but was too late to have any influence on it. It was as a result of his trying to combine his analysis with

that of the *Treatise* that Kahn reached the conclusion that the fundamental equations applied in their full simplicity only at full employment. It was Kahn who pointed out the error, in the *Treatise*, of ascribing the price levels of investment and consumption goods to completely different causes. It was out of this process of re-examining the *Treatise on Money* that a theory of output began to emerge, albeit one expressed in the same language.[92]

The gist of Keynes' ideas at this time is contained in a series of lectures he gave in Chicago in June 1931.[93] As in the *Treatise*, he spoke of booms and slumps as being determined by excesses or deficiencies of investment over saving, but though still attaching great importance to the role of profits, he emphasized changes in employment more than in the *Treatise*. He came close to the idea of an unemployment equilibrium:

> Now there is a reason for expecting an equilibrium point to be reached. A given deficiency of investment causes a given decline in profit. A given decline in profit causes a given decline in output. Unless there is a constantly increasing deficiency of investment, there is eventually reached, therefore, a sufficiently low level of output which represents a spurious kind of equilibrium.[94]

Although this is not the *General Theory* it is a significant step towards it, the final steps being taken in the next couple of years.[95]

These origins explain some of the puzzling features of the *General Theory*.[96] Because Keynes had little to add to much of the discussion of the monetary system contained in the *Treatise*, the *General Theory*'s treatment of money is fairly rudimentary. More importantly, however, the criticism to which the fundamental equations were subjected explains why Keynes, in the *General Theory*, avoided any measurement of either levels of output or price levels. Everything in the *General Theory* is measured either in terms of quantities of money, or in terms of employment. Thus, for example, the aggregate supply curve relates not price to output, but revenue to employment.

The crucial part of the *General Theory*, as is evident even from a glance at the table of contents, is the theory of aggregate demand. This theory will be familiar, but it is important to point out that there is more to this than the idea of the multiplier. Three main propositions can be distinguished: (1) that an increase in expenditure will produce "multiplier" effects; (2) that there may be equality of saving and investment at a variety of levels of income; and (3) that changes in the level of output provide a mechanism through which, independently of changes in profitability, savings and investment can be brought into equilibrium. Kahn had provided only the first. By the time of his Chicago lectures, quoted above, Keynes had reached the second proposition, though he had still not integrated this with the idea of the multiplier. For the *General Theory* all three propositions were necessary. There is a radical difference between understanding the dynamic multiplier and seeing how this can be used as the basis for a theory of aggregate demand. Keynes achieved the latter; Kahn did not.

The theory of the consumption function, however, can explain the level of demand only given the level of investment. Keynes' explanation of the

latter, the inducement to invest, comprised two parts: the marginal efficiency of capital, and the theory of liquidity preference. The marginal efficiency of capital, the discount rate at which prospective returns from investment equal the cost of the investment, turned out to be the same as Fisher's rate of return over cost.[97] To understand Keynes' use of it, however, we need to consider his approach to the problem of expectations. This was to divide expectations into long and short term expectations.[98] The former are concerned with the proceeds an entrepreneur can expect to derive from the output he is producing; the latter with what he can expect to earn through future production if he purchases an item of capital equipment. Short term expectations are capable of being checked frequently in the light of experience, and so Keynes argues that it is reasonable to neglect them, looking simply at realized proceeds. Long term expectations, however, are a different matter, being based on much more precarious information in which investors have limited confidence. To a large extent, therefore, long term expectations are based on conventions, these being liable to change arbitrarily, and for no apparent reason. This means that the marginal efficiency of capital was, for Keynes, something very volatile, liable to change arbitrarily as investors changed their opinions. This contrasted with consumption decisions, which, being based on short term expectations, could be portrayed as a function of realized income.

Expectations were also important in Keynes' theory of the rate of interest, in which the rate of interest depended on liquidity preference. Demand for money depended on expectations of future interest rates as much as on current interest rates. More important, however, than the details of Keynes' treatment is the fact that it was, as was the treatment of demand for money in the *Treatise*, demand for the *stock* of money which determined the rate of interest. Demand for money was being seen as part of the theory of portfolio choice.[99]

In presenting all the material discussed so far, the main part of the *General Theory*, Keynes made the preliminary assumption that the money wage rate was constant, many variables being measured in wage units, an assumption not relaxed until chapter 19. Despite its coming so late in the *General Theory*, however, this chapter is of particular importance in view of the role of money wage rigidity in previous theories of unemployment. Keynes' argument about money wage changes is the following. (1) Given the propensity to consume, the marginal efficiency of capital and the rate of interest, a reduction in money wage rates can have no effect on employment. If entrepreneurs were to respond to a fall in money wage rates by increasing their production, they would find that demand would, unless the marginal propensity to consume were equal to unity, fail to grow by as much as supply. The additional output would remain unsold, so firms would not have an incentive to produce it. (2) Alternatively, a reduction in money wage rates might affect demand through the propensity to consume (through affecting the distribution of income), through the marginal efficiency of capital (through affecting the ratio of current costs to expected proceeds), or through the rate of interest (through raising the real value of

the money supply). It is through the last of these that Keynes argues that wage reductions must operate. This means that wage reductions have effects exactly equivalent to those of an appropriate increase in the money supply. Two conclusions can be drawn from this. (a) For exactly the same reason that an increase in the money supply may fail to raise employment, a reduction in money wage rates may be ineffective.(b) Any increase in employment which could be brought about by a wage reduction could, much more easily, be brought about by an increase in the quantity of money. (3) Finally, having shown that an increase in the money supply can mimic the effects of a reduction in wage rates, Keynes then argues that money wages *ought not* to be flexible, for changes in the money wage rate cause unfair changes in the distribution of income, changes in the burden of debt, and instability in the price level. Thus, not only does Keynes not blame unemployment on wage rigidity, but he argues that it is better for money wage rates to be inflexible.

16.6 CONCLUSIONS

The period covered by this chapter was an extremely fertile one as regards theories of money and the cycle. Although the quantity theory remained, right up to the time of the *General Theory*, the framework within which such issues were discussed, it was the vehicle for much thought on short run problems. Wicksell's cumulative process and Fisher's transition periods, contain much worthwhile analysis of the relationship of monetary changes and changes in output and the price level. Though they have been expounded separately in this chapter, theories of money and the cycle were closely linked, even by economists for whom monetary factors were not primary. Thus although there was no separate subject of "macroeconomics", the issues we now consider under that heading were far from neglected. Despite the revolutionary claims made in the *General Theory*, therefore, it is very misleading to view Keynes as having created the subject of macroeconomics almost single-handed.