



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Monetising the Classical Equations: a theory of circulation

Author(s): Edward J. Nell

Source: *Cambridge Journal of Economics*, March 2004, Vol. 28, No. 2 (March 2004), pp. 173-203

Published by: Oxford University Press

Stable URL: <https://www.jstor.org/stable/23602123>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Oxford University Press is collaborating with JSTOR to digitize, preserve and extend access to *Cambridge Journal of Economics*

JSTOR

Monetising the Classical Equations: a theory of circulation

Edward J. Nell*

The Classical Equations describe output and income in real terms. To use them to analyse aggregate demand, the transactions they describe must be ‘monetised’. A sum of money equal to the wage bill of the capital goods sector can be shown to be necessary and sufficient to carry out all transactions, in a process of circulation which also defines an expression for velocity. When money has intrinsic value, the quantity approach may hold in the short run but, in the long run, money will be endogenous. In these conditions, the rate of interest will be determined by the supply and demand for reserves, but when money is purely nominal, only a minimum rate will be fixed, and the rate of interest will have to be pegged. The Appendix develops the Classical Equations and shows that they define an invariable unit of account.

Key words: Classical Equations, Circulation, Endogenous money, Rate of interest, Reserves

JEL classifications: O21, O23, O31, 311

Introduction

The Classical Equations describe prices, wages and profits, on the one hand, and quantities, consumption and investment, on the other. The latter equations are written as the dual to the former. Taken as a whole, the two sets of equations describe production and exchange, the distribution of income and the final use of output, customarily leaving a degree of freedom. Behaviour based on varying motivations, expectations and psychology, including rational choice, may be added to close the model. The equations may be written in more or less aggregate form; here they will be considered as describing two sectors of activity—the production of means of production—capital goods—and the production of means of consumption or consumer goods. (In each of these sectors, there may be many specific goods and services.) Two categories of income are paid out, wages and profits, which result in two kinds of spending, consumer expenditure and investment spending. Consumer

Manuscript received 1 September 2000; final version received 27 March 2002.

Address for correspondence: New School for Social Research, The Graduate Faculty, Department of Economics, 65 Fifth Ave, New York 10011, USA; email: EJJNell@aol.com

*New School for Social Research. This paper has been in preparation for a long time, and thanks are due to many friends and colleagues for comments on earlier drafts. A few deserve special mention, however. Warren Mosler commented in detail on several drafts. Stephanie Bell, Ghislain Deleplace, Duncan Foley, Marc Lavoie, Fabio Petri, Louis-Phillipe Rochon, Mario Seccarecia and Randy Wray provided especially important insights, criticisms or suggestions. Leanne Ussher drew the diagrams. The Center for Full Employment and Price Stability at the University of Missouri—Kansas City, and its Director, Mat Forstater, provided financial support. Thanks to all; the remaining errors, alas, are mine.

Cambridge Journal of Economics, Vol. 28, No. 2, © Cambridge Political Economy Society 2004; all rights reserved

spending supports households, enabling them to supply labour, while investment spending buys replacements to maintain, and new capital goods to expand, the system.

The Classical Equations are written in barter form; that is, the exchanges described or implied are swaps of goods for goods, while payments of income distribute command over goods in terms of goods. Prices are expressed in terms of some arbitrarily chosen *numéraire*. No attempt is made to say how the transactions will be carried out, or in what order, or by means of what medium.

One reason for this neglect may lie in the famous ‘Classical Dichotomy’, which holds that to express these transactions in money will not affect the nature or working of the system. In this view, money is invented or adopted because it represents a gain in efficiency, compared with barter (Patinkin, 1965; Goodhart, 2000). A monetary system can be expected to lower transactions costs, and lead to greater efficiency in trade, so there will be some reduction in some of the coefficients. But otherwise there will be no effects. Real forces and relationships will determine real values; while monetary or nominal values will be determined by the monetary system. In this view, neglecting money makes no essential difference.

For some purposes, it is surely reasonable to set monetary questions to one side. But when applied issues arise, the Classical Dichotomy cannot simply be assumed. Most data in modern economies are gathered in monetary form. The Classical Equations are abstract but need not be idealised, and could help us understand the world. But they cannot be applied until we explore the implications of carrying out transactions in monetary form—and when we do we shall find that monetary institutions have costs and affect profits. Worse, in dealing with effective demand the Classical Dichotomy is seriously misleading. Yet the Classical Equations do provide a full account of production and distribution—just what is needed for studying adjustment and effective demand. So we need an account of how the transactions described by the Classical Equations can be ‘monetised’, or to put it in the old-fashioned way, how money, paid as income, can circulate the entire output of the system. This account, like the Classical Equations, should be cast in terms of technology, class and institutions, leaving behaviour open.

1. Circulation and the value of money

1.1 *The idea of monetary circulation*

In mainstream textbooks, the rate of interest and the price level (the value of money) are determined by the supply and demand for ‘real balances’, interacting with savings and investment. The demand for real balances is the amount of money in real terms that households and businesses wish to *hold*, that is, to keep idle, as part of a portfolio of assets, and this depends on the interest rate and the level of activity. The supply, of course, is determined by the banking system, given Central Bank policy, and taking the price level as given. Money is treated as being ‘at rest’; the issue is why individual agents choose to hold it, when it yields no interest, or less interest than other assets. Yet an account of *why* money is *held* does not explain *how* money is *used* (Hicks, 1967, 1977, 1989).¹ An account of the demand by individual agents for (real) cash balances (the average demand over a period) tells us nothing about the sources and destinations of inflows and outflows or about their regularity. The approach assumes that balances are attributable to individual decisions, based on

¹The supply and demand approach focuses on a certain kind of individual decision; ‘how money is used’ asks about the working of a system. How do we know that enough money will be made available to monetise all transactions? How much money is required for that? Will it matter what serves as the ‘money-article’? Will credit be needed, and if so, how much, and why? To answer such questions, we need to turn to an account of circulation.

preferences, and does not consider the way agents interact with each other as they carry out their duties according to their institutional roles.

The early Classical economists definitely thought in terms of a pattern of circulation (Green, 1992). Money—coins of precious metal—would be advanced, say by merchants, for the purchase of goods produced. Such monies would be received by producers and spent on wages and materials, and perhaps tools and equipment; the wages would be spent by workers, while the producers of materials would spend their receipts, in turn, on wages and materials, etc. At some point in the process, the initial goods purchased by the merchants would be required, and the money would return to them as it was spent on those goods.¹ This would close the circuit—in a manner to be defined below. The value of money—of the article serving as money—depended on the labour required for mining (and also on the seignorage charged for minting); the quantity required depended on the volume of the goods (of equal value in labour to the money) for which the money exchanged. The ‘velocity’, never very precisely defined, referred to the number of times money exchanged for goods during its circuit. However, hoarding—pulling money out of circulation to hold it as an asset—could disrupt the circulation and bring on a crisis (Hicks, 1977, ch. III(i); Nell, 1998, ch. 6). In contrast, hoards could supply coin or bullion to the system in the event of a shortage (e.g., due to purchases abroad or an unexpectedly large harvest).

Copernicus compared the circulation of money to the circulation of the spheres; Petty compared it to the circulation of blood. For the older analysts, money was seen as based on precious metals. Tooke, however, showed how flexible the links were between holdings of metal and circulating paper and bank deposits. Wicksell carried this a step further and modelled the circulation of an economy operating with a currency based on a pure banking system, contrasting such a system with one based on intrinsic value. Keynes in the *Treatise* held money to be a creature of the State, and identified two ‘circulations’—the Industrial and the Financial. In *The General Theory*, however, he ignored the process of circulation, treating the money needed for circulation as a ‘transactions demand’ for money. Later Hicks (1967, 1977) partially restored the older view, by treating money for transactions as determined by the working of the system, in contrast to precautionary and speculative funds, which were matters of choice. Post Keynesians and circuit theorists have returned circulation to centre stage in studies of the ‘endogeneity’ of the monetary system.²

In the *Treatise*, Keynes examined two circuits of bank money, the industrial circuit and the financial circuit. The former concerned ‘the normal process of current output, distribution and exchange and paying the factors of production their incomes . . .’, while the latter referred to ‘the business of holding and exchanging existing titles to wealth, including Stock Exchange and Money Market transactions’³ (p. 243). For Keynes, money was

¹ When firms pay wages, their balances decline and household balances rise; when households buy consumer goods, their balances fall while business receipts rise, and so on. We shall see that constancy of the total circulating fund in this process is more a matter of the way sectors are connected than of individual decisions. See Cencini, 1995.

² By contrast, modern mainstream theorists have largely abandoned the study of circulation and seldom question the nature of money itself, seeing the problem of monetary theory as explaining the supply and demand for an asset without an obvious yield.

³ Money held as an asset may correspond to Keynes’s financial circulation—although there is a *circulation* associated with fixed capital, see below and Nell (1998). But money is held as an asset in portfolios, and it moves in the course of portfolio adjustments, including the retirement of bonds and repayment of debts. Such money is considered *held*; that is to say it does not actively circulate, or does so only slowly. It can be thought of as a *stock*, whereas money circulating in production is a *flow*. The total amount of such ‘idle’ money normally tends to be larger, often substantially larger, than the amount of the actively circulating medium (although recently in advanced countries it has been less. See Nell, 1998.) Funds can be switched back and forth, from asset-holdings to active circulation—with implications for interest rates. So the ‘supply of money’ can refer to money offered to the circulation either through money creation—mining and minting, bank advances—or through portfolio adjustment, changing the composition of asset holdings.

basically money-of account, established by the State, and its principle form was bank-money.¹ His analysis has been developed further in subsequent years,² and we shall make use of the distinction between the circuits. But many economists today do not consider circulation, preferring instead to concentrate on explaining the reasons for holding an asset with no obvious yield. Money certainly serves as an asset, but it is also a medium of circulation, and treating it only as an asset provides no help in determining the amount of money required to monetise all transactions in a stationary economy or how that sum could grow to service an expanding economy.

The problem which needs explanation appears to be straightforward—at first glance: How is it that a sum of money, starting from one or more definite points, can exchange against all the goods and services produced in the economy, and end up where it began, ready for the next round of production? Money circulates—that is, it begins somewhere and ends up back there, on the way having ‘monetised’ all economic activities. And if the economy has expanded, through investment, so that it is ready to produce on a larger scale, the money for the next round will also have expanded.

A second glance reveals a number of complications.³ Where does the money begin—is it issued by the State, or does it have a private origin? How much money is required? How can a given sum of money be put into circulation as working capital—buying inputs—and then a larger sum be received back as revenue including profits? We understand productivity—how given inputs of goods and services are transformed into a larger set of outputs but where does additional *money* come from?⁴ What about speculation and financial transactions—how do they fit in? Finally, given answers to such questions, is there only one pattern of circulation, or might there be many routes which money could take?⁵ Might different routes require different sums of money? And what if some activities are left ‘unmonetised’?

A capitalist, industrial economy is one of ‘monetary production’, in Keynes’ memorable phrase. That is, production is carried out by purchasing inputs of goods and services for money, and obtaining money from the sale of the produced output, where the motivation is

¹ ‘Today all civilised money is, beyond the possibility of dispute, chartalist’ (p. 5). ‘A Money-of Account comes into existence along with Debts, which are contracts for deferred payments. Money itself, namely that by delivery of which debt-contracts and price contracts are *discharged* . . . derives its character from its relationship to the Money-of-Account . . . it is a peculiar characteristic of money contracts that it is the State or Community not only which enforces delivery, but also decides what it is that must be delivered as a lawful or customary discharge of a contract . . .’ (pp. 3–4).

² In England, for example by Newlyn and Bootle (1978) and by Tew (1969). Newlyn and Phillips built a hydraulic model, in which movements of water illustrated the flows of money in the economy. In the US, Rousseas has developed the argument. But none of these has shown how a given sum of money could in successive steps circulate the entire output, or how that sum could be determined.

³ Wicksell (1898), for example, set up a formal model with workers, entrepreneurs, capitalists and bankers, and established a pattern of circulation that would be in equilibrium if, and only if, the ‘natural’ rate of interest were equal to the money rate, as he defined them. However, the scheme required banks to accept a costly deposit for which they had—and could have—no use. Nor did it explain where the banks obtained the funds with which to pay interest (Nell, 1967).

⁴ Marx described ‘Mr. Moneybags’, the capitalist, advancing wages to his workers, ‘throwing money into circulation’, and somehow receiving back more than ‘he’ (standing for the capitalist class as a whole) initially advanced. To explain how more could be returned than was advanced—given that money is sterile—was, for Marx, the central problem of the theory of the circulation of money. The theory of surplus value explained how more could be *produced* than was used up in production; but this rested on the productive powers of labour. By contrast, money has no productive power, and is given in amount (Luxemburg, 1963). See also Keynes (CW, xxix, pp. 81–2).

⁵ It is often argued that monetisation reduces transactions costs or otherwise improves efficiency. But the opposite is just as likely (Nell, 1968). The number of transactions is increased compared with barter. What is undeniable, however, is that money of account is required to calculate rates of return and capital values, and a medium of circulation is needed if capital is to be able to flow quickly into the most profitable channels.

to make a profit in monetary terms. In such an economy, non-monetised activities will be at a disadvantage, so monetisation must cover all economic activities. The circuit must be *complete*. Also, the cycle of production is repeated indefinitely, each time on an expanded scale as the economy grows through investment; hence (abstracting from rollovers and perpetual securities) money must return to its starting point in order to repeat the circuit; so the circuit must be *closed*. Likewise, the path followed cannot be open to question, or unknown; the system must be *determinate in path*. And, it must be *stable*, or capable of being made stable, for otherwise it could fall into crisis—as indeed it often has.

Finally, a definite sum of money must perform the circulation. This must be clearly determined, and minimally costly, and this sum must grow regularly as the economy grows. So the circuit must be *determinate in value*. There are two cases here. First, where the article serving as money has a production cost, and so a value, as with precious metals, or bank notes convertible to metal, the quantity of money will be determined in a manner reminiscent of the ‘Quantity Theory’. Second, in a system of modern banks, the size of the banking sector and a minimum rate of interest will be determined.

1.2 *The circuit and the two-sector model*

We shall follow the flow of money in three cases, differing in the kinds of money and agents, but all tracing out the same circuit. The three cases range from relatively primitive to relatively modern, and a transformational growth approach suggests that the later developed out of the earlier under market pressures. (This will not be pursued here, but see Nell (1998, chs 5, 6).) First, we have merchants with coins, who buy the net output of the consumer goods sector. This money is accepted because it has ‘intrinsic value’, determined and supported by the market—but managed by the Mint. Second—and this will be the case we follow in detail—banks advance working capital to the firms of the capital goods sector, in the form of paper money or bank deposits backed by reserves of precious metal. The advances are made in the form of paper notes and deposits in accounting units. They are accepted because they are ‘convertible’, that is, backed by bullion and coin reserves of ‘real’ value. Third, the most appropriate case for a modern economy, banks advance funds for the purchase of capital equipment to the firms of the consumer sector. These funds are ‘chartal’ accounting money, created and extinguished by the banking system in response to economic incentives. Such money creation is constrained by bank capital, while the resulting medium is accepted by the public because the state will take it in payment of taxes. It is ‘State’ money.¹

The approach can be illustrated on the two-sector diagram in Figure 1 (Nell, 1988, 1992, chs 20, 21, 23).² The underlying production model is presented in the Appendix. One sector produces equipment, using equipment and labour; the other uses equipment and labour to produce the goods that support labour. (Management and services are included in labour.) Equipment will initially be assumed to last one period only, and to be replaced

¹ Arguably, all forms of stable, universally accepted money have a ‘chartal’ aspect. (Wray, 2000). Coinage is managed by the Mint; fiduciary money is ultimately backed by State guarantees, and is managed. The acceptability of any form of money is greatly enhanced when the State is willing to take it in payment of taxes. A banking system must provide clearing balances; such balances rest on State backing.

² This diagram summarises a two-sector model of production, which can be used to explore the theory of Effective Demand, as in Nell (1975/1992). The multiplier, for instance, can be examined by increasing Investment on the LHS and observing the resulting impact on employment and output in both sectors. Changes in Productivity and the Real Wage can likewise be studied. A more general n -equation model of production can be reduced to two sectors and represented in this way. The significant fact for its use here is that the condition that $W_k = P_c$ is independent of distribution, i.e., of the rate of profit. This is proved in Nell (1998, ch. 7) and in the Appendix.

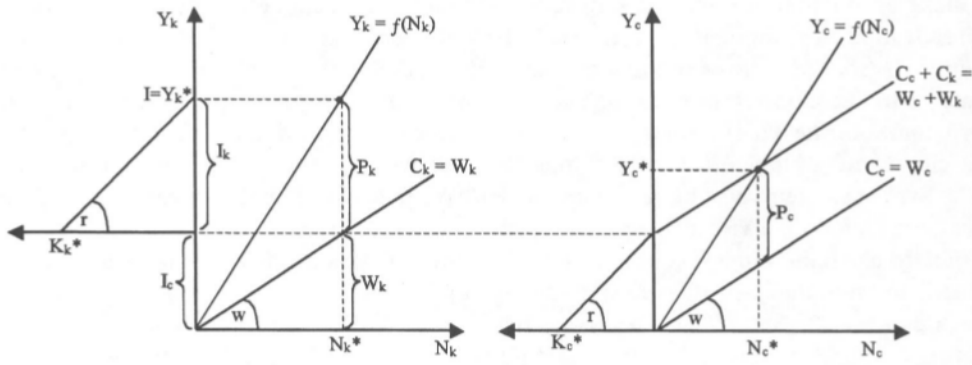


Fig. 1. Two-sector model.

every period. (Alternatively, as in Hicks (1965), equipment could be assumed to last forever.¹) Fixed capital will be examined later. (In terms of the two-sector model in the Appendix, sectoral outputs Y_k and Y_c respectively equal to P_1X_1 and P_2X_2 , equal, in turn, to I and C . See the discussion for interpretation of the diagram in regard to both primal and dual.)

All aggregates are expressed in a monetary unit of account. (The validity of this unit is established in the Appendix.) The rate of profit is the same in both sectors, as is the wage rate: $P_k/K_k = P_c/C_c$, and $W_k/N_k = W_c/N_c$. (The net rate of profit is r , the gross rate $1 + r$.) The diagram on the left shows output of capital goods on the vertical axis, and measuring to the right, capital goods employment, N_k , on the horizontal. The steeper line is the output function, the shallower the wage bill, assumed equal to consumption by capital goods workers, C_k . Investment demand $I = I_k + I_c$ is marked off on the vertical axis; this determines output and employment in the sector, and so its wage bill. This wage bill represents demand for consumer goods, and so is mapped onto the diagram on the right, showing output of consumer goods as a function of employment in the consumer goods sector. It is clear that the wage bill of the capital goods sector equals the gross profit of the consumer goods sector, $W_k = P_c$.

The level of investment spending here is assumed to fall in the 'normal' range of capacity utilisation; that is, in the range in which constant returns to utilisation prevail. When investment demand is higher or lower, the level of employment in capital goods will be higher or lower in proportion. If I changes from I_0 to I_1 , W_k will go from W_{k0} to W_{k1} , and $I_0/I_1 = W_{k0}/W_{k1}$. Therefore profits in the consumer goods sector will vary in proportion to the overall change in investment; profits in the two sectors will be affected in the same proportion. So the uniformity (or divergence) of the rate of profit will be unaffected, whatever the level of investment, so long as it falls in this range. (Changes in investment, of course, have a multiplier effect on overall output and employment.)

Total output during a period of production is

$$Y = Y_k + Y_c = W_k + W_c + P_c + P_k$$

and total expenditure

$$E = C_k + C_c + I_c + I_k$$

¹ As we shall see, fixed equipment, lasting for several periods, implies both users who must save up for large purchases at widely spaced intervals, and suppliers who must produce for big-ticket sales at the same intervals. This sets up the possibility, outlined below, that the former may lend their accumulating savings to support the operating expenses of the latter during the intervals (Nell, 1998).

where all profits are saved and all wages consumed, so that $P = I$. (These will be modified in a moment, since *next* period's higher wage bill will be spent on *this* period's output of consumer goods; but the equations as written will then hold for the *period of circulation*.)

The expressions for Y , output, and E , expenditure, therefore present the variables that have to be monetised; the Classical Savings function greatly simplifies the argument, since if a component of either Y or E is monetised, the corresponding component of the other is also.

Profits are retained and used to finance investment purchases. Managerial consumption is subsumed in household consumption, underwritten by wages and salaries. Initially, the Golden Rule is assumed; later, saving out of wages and consumption out of profits will be considered.¹

Production times in the two sectors are coordinated, and all firms act in concert. If production times were not coordinated, some firms would have a competitive advantage. Here, it will be assumed that production, rather than running continuously, can be broken into distinct periods, each with a definite beginning and ending. Within each period, inputs are used up, incomes are paid and outputs produced. Growth takes place from period to period, as new capital goods are put in place, enabling production to take place on a larger scale.²

Circulation is a repeated process. We can choose to break into the cycle at any point, but it will be convenient to start with inventories of consumer goods assumed to be on hand, the result of production in the previous period. New capital goods and replacements, however, have been sold and are in place. This, then, will be the starting point, and also must be the end-point.

Note the implication mentioned earlier: *last* period's output of consumer goods must support *this* period's household consumption. This requires an adjustment to the production equations. Since there is growth from period to period, last period's output must exceed last period's total wage bill by gC . This inventory must be carried over from period to period, as part of the capital of the consumer goods sector. But in the current period of circulation, total wages paid and spent will equal total consumption.

For purposes of exposition, circulation here will be assumed to start with banks advancing funds to the capital goods sector as needed;³ these funds will eventually total W_k . (This is not the only possibility, as we shall see. For example, merchants could buy the net output of consumer goods, equal to the profit of the consumer sector. Alternatively, banks could advance funds to firms in the consumer sector equal to the cost of their proposed purchases of capital equipment for replacement and new investment. Both of these will be explored later.) By an 'advance' or by 'financing', we mean a loan *for the duration of the period of circulation*. The loan is advanced at the beginning of the period and is repaid (or rolled over, but it *could* be repaid) at the end. Shorter loans, bridging loans of various kinds in anticipation of immediate revenue, are not at issue.

Using the advances, capital goods firms begin production, paying wages to households, who begin purchasing consumer goods, running down inventory, but building up the bank

¹The Golden Rule is important because it justifies the basic relationship between the wage bill of the capital goods sector and the gross profits of the consumer sector. However, it can be shown that, under plausible conditions, deviations from the Golden Rule will not have large effects (Nell, 1998, ch. 7).

²These assumptions are convenient but not necessary; even continuous production has a beginning and ending and lasts for a definite time. Continuous production can be modelled as a case of overlapping periods.

³The advances are drawn on open lines of credit negotiated by the firms as part of their operating capital, and backed by collateral as required by banks, based on the firms' balance sheets.

balances of consumer sector firms, so that their capital remains intact. As soon as these balances begin to accumulate¹ consumer goods firms can draw on them to pay wages, which promptly return to them as consumer goods workers buy consumer goods. (The funds not circulating as consumer sector wages will be temporarily idle, and can be borrowed by banks that need them.) When production is complete in both sectors, the consumer sector will have the entire wage fund, and its inventories will be wholly depleted. It will then spend its profits, P_c , on its desired gross investment, I_c , returning the wage fund to the capital goods sector.

At this stage W_k has been spent on C_k , and has provided the funds for W_c , which has been spent on C_c , while W_k has re-emerged as P_c , which has been spent on I_c . All have been monetised by the circulation of bank advances equal to W_k . To complete the circulation, I_k must be sold for money, realising profits P_k , and enabling the advances to be repaid, closing the circuit.

The capital goods sector may be divided into two subsectors, one of which sells to the consumer goods sector, the other being the rest of the firms. Then the first subsector receives the whole revenue, I_c ; it repays its wage advances and spends its gross profit purchasing its desired investment goods. The second subsector receives the gross profits of the first as its revenue; it repays its wage advances, and spends its gross profits on its desired investment goods. This process will be repeated until a subsector is reached—the ‘machine tool’ sector—which makes its own capital goods. More precisely, the consumer goods sector spends $P_c = I_c$, which will be the initial revenue of the capital goods sector, received by the first subsector. Call this I_{k0} , and from it will be subtracted the wage bill of the first subsector, W_{k1} , which will be repaid to the banks, leaving P_{k1} as profits. Hence

$$\begin{array}{ll}
 I_{k0} - W_{k1} = P_{k1} \rightarrow & \\
 I_{k1} - W_{k2} = P_{k2} \rightarrow & \text{where } \sum_{i=1}^n W_{ki} = W_k \text{ and} \\
 I_{k2} - W_{k3} = P_{k3} \rightarrow & \sum_{i=1}^n P_{ki} = P_k \\
 \dots\dots\dots \rightarrow & \sum_{i=1}^n I_{ki} = I_k \text{ and } \sum_{i=0}^n I_{ki} = I \\
 I_{kn-1} - W_{kn} = P_{kn} = I_{kn} &
 \end{array}$$

The n th sector is the ‘machine tool’ sector, which makes its own capital goods. However, the firms in this sector may have preferences for each other’s goods, or may prefer to contract out with each other. Firm A may use funds in a ‘short circuit’ to buy from B which in turn buys from C, who buys from A.² At this point, the sale for money of I_k will be complete, P_k will have been realised, and all advances repaid. Bank advances equal to W_k will have circulated the entire output Y .

Money income can be expressed in terms of circulation, then:

$$Y = W + P = W_k + W_c + P_c + P_k = 2W_k + W_c + P_k$$

Now let us assume that the labour coefficient will be the same for every subsector in consumer goods. Then $W_c = Y_c - P_c = W_k[(1/(1 - wn_c)) - 1]$, a simple multiplier relationship.

¹ Notice that the wage bill of the capital goods sector *gradually* accumulates in the hands of the firms of the consumer goods sector. It is not spent until the production process in capital goods is complete, at which point the entire wage bill will have been paid out and spent. Part of it will be used to circulate the wage bill of the consumer goods sector itself, but a good part will be idle while production is still being carried out in the capital goods sector. These temporarily idle funds can be used for clearing purposes.

² *Payments* in the capital goods sector take place in succession, but production is simultaneous. There is no implication that firms must wait for an earlier stage to be complete. A clearing system would simplify the process.

Assume that the labour coefficient will be the same in all subsectors in capital goods, and further assume that the machine tool subsector is vanishingly small. Then the first subsector receives $P_c (= W_k)$ in revenue from its sales of capital goods to the consumer sector. It withdraws $wn_k P_c$ to repay its loans, and spends $(1-wn_k)P_c$ purchasing its replacements and new capital goods from the second subsector. This second subsector will withdraw $wn_k(1-wn_k)P_c$ and spend $(1-wn_k)(1-wn_k)P_c$. The resulting sequence, taken to infinity, will sum to $(1/wn_k)P_c$. So we have¹

$$Y = W_k[2 + \{1/(1-wn_c)-1\} + \{1/wn_k\}] = W_k[1 + 1/(1-wn_c) + 1/\{wn_k\}]$$

The LHS shows income expressed in money of account; the RHS shows the sum required for circulation, in units of account, multiplied by an expression showing how that sum circulates. This expression is analogous to the ‘velocity of circulation’.

Let us take stock. Income has been paid in money, and the entire output has been sold for money; the circuit is therefore *complete*. Capital goods have been sold and are in place; consumer goods have been produced and are held as inventory—so the circuit ends up at its starting point. It is therefore *closed*. The path is not *unique*, however, until the amount of money required for the ‘short circuit’ in consumer goods is specified, an institutional matter. Given that, it will be unique. But the *value* of the circuit, W_k , is unique. For no sum of money can make a complete round before the initial payout for wages in the capital sector is complete, because of the competitive assumption that all firms in an industry must act in concert. So no less than W_k will be put into circulation. But no more is needed, and since borrowing is costly, no more will be put into circulation.

This account of circulation therefore *monetises* a production model which itself has been shown to provide a basis for the theory of effective demand. This unifies the theories of money, production, and effective demand.

1.3 T-accounts

The process can be illustrated in a succession of T-accounts, giving the balances in a period of circulation for Banks, Worker Households, Consumer Goods Firms and Investment Goods Firms.

	<i>Assets</i>	<i>Liabilities</i>
Banks:	W_k -loans	W_k -deposits
Workers:	W_k -wages	C_k -expenditure
Consumer Firms:	C_k -inventory	P_c -profits
	C_c -inventory	W_c -wages
Investment Firms:	I_c -inventory	W_k -wage loans
	I_k -inventory	P_k -profits

Evidently, at each stage, the T-accounts balance, and the progress of the circulation can be traced simply by reading down the table. Banks advance the wage bill as loans, creating deposits, which are used to pay wages, W_k , which Workers spend on consumer goods, C_k , acquiring the inventory of Consumer Firms, and providing those firms with profits, P_c . Part

¹ If the two sectors were to have the same capital-labour ratios, then this expression would simplify to: $Y = W_k[1 + 1/\{wn(1-wn)\}]$. The multiplier can be expressed in terms of the wage and the productivity of labour because *profits* are saved, and thus are the ‘withdrawal’. This accords with the stylised facts for the US, in which gross business savings have long been far larger than household savings.

of the money so received will be used, repeatedly, in a 'short circuit', to pay wages, W_c , which are spent on the rest of the inventory of consumption goods, C_c , returning the funds to the firms of the consumer goods sector. Profits P_c are then used to purchase inventory I_c from Investment Firms, providing them with funds to realise profit P_k in the successive sale of inventory I_k , which at the same time provides for repaying wage loans W_k , closing the circuit.¹

At every stage, the total of bank deposits (and assets) in the system will be W_k . When wages are paid, business deposits diminish, but household deposits rise by the same amount; when households spend, their deposits decline, but those of business revenue rise. No payments ever leave the system. As cheques are deposited, the receiving bank's liabilities are increased, and those of the drawing bank are marked down as they clear. The assets of the receiving bank increase, and those of the drawing bank are reduced—the latter bank issues the former a claim to title to a portion of its assets matching the liabilities.

Profits appear as a liability; they are owed by firms to owners of the capital. Retained profits which are invested increase the value of the firm; so on the assumption that stock prices accurately reflect the value of capital, the stock of a firm will appreciate in proportion to investment (assuming that the market agrees that the investment is wise). We can relax the assumption that profits are retained: instead, firms in both sectors pay profits as dividends to holders of their securities, who save and re-invest them in the financial market. To finance investment, firms in both sectors issue new securities, the sale of which returns the savings to them. In this scenario, stock prices remain unchanged, but new issues increase relative to stock outstanding in the same proportion that new investment stands to existing capital. (Profits in the capital goods sector will not be realised until the consumer goods sector spends on investment, setting off the secondary circulation in capital goods. This circulation will therefore be revised to show firms paying out profits as dividends, then issuing new stock to recoup the funds.)

Other complications can be treated similarly. Suppose workers save; then wages will not all be returned in spending. The profits of the consumer goods sector will be lower by the amount of saving by workers in the capital goods sector, and will also be lowered by any saving by its own workers. Figure 2 shows the consumer good sector; the demand for consumer goods by capital goods workers will be $(1-s)W_k$. Saving by consumer goods worker households will lower the angle of the consumption line. So aggregate demand in the sector will be $C^c + (1-s)W_k$.

However, this need not affect investment. The firms in the consumer goods sector can borrow the workers' saving in the bond market, thus returning the funds and closing the circuit. If workers keep their savings in the bank, then firms could borrow from the banks.

'Capitalist consumption'—consumption out of profits—is financed by loans taken out by firms with which they pay dividends, the collateral being the inventories. These dividends are spent on consumption and flow to the consumer goods sector as additional profits. In this case, then, total profits equal investment plus capitalist consumption. Profit in the consumer sector equals the wage bill of capital goods plus capitalist consumption. (Note here again that profits are realised from expenditure financed by loans).²

¹Of course, the circuit need not actually be closed; rather than firms repaying and banks reissuing, the debt could be rolled over, with interest accumulated on the balance sheet. The point is that it *could* be closed.

²Taxes and Government spending have not been considered only because they add so many complexities. There are many cases with widely varying implications for the economy. But the role of the State is fundamental; money is accepted because it can be used to pay taxes (Mosler, 1995).

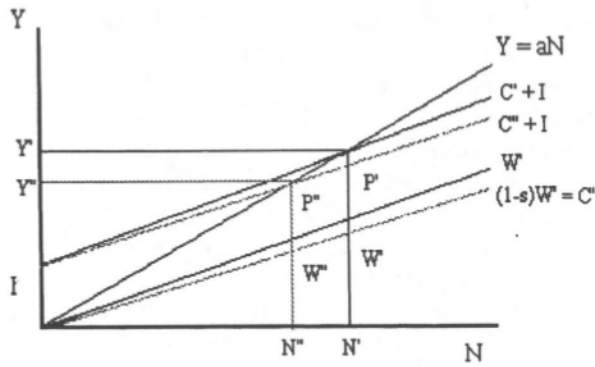


Fig. 2. Consumer-goods sector.

1.4 Interest

The advances that circulate output have to be repaid with interest—where does the money to pay interest come from? The business of banks is to make profits through earning interest on their advances. Interest is their revenue. Anticipating it, banks advance funds to themselves in two categories. On the one hand, they have operating expenses—wages, salaries, supplies—and, on the other, they spend for investment—new offices and equipment. Or they may buy additional reserves. The expenditure of these funds enters them into the circuit, so that they return to the banks as interest on the advances to business (Robinson, 1956). Banks earn what they spend. Total investment thus consists of the investment of the capital goods sector, of the consumer goods sector, plus that of the banks, and the profits of the consumer goods sector equal the wage bill of the capital goods sector, plus the wages of the banking sector. The banks advance the wage bill of the capital goods sector and their own wage bill. In addition, however, they must grow from period to period, so—in a system of real reserves—they will have to purchase additional bullion from the mining sector. They advance these investment funds, too.

The same is true regarding the interest on loans to support capitalist consumption. The funds to pay the interest on such loans are again put in motion by the banks, which anticipate their earnings, and advance investment funds to themselves. They may use these funds to purchase equipment, to acquire additional reserves, or they may buy securities. Any of these will provide the other sectors with the wherewithal to pay the interest, while at the same time increasing the banking system’s own capital.

From period to period, then, the banking sector will hire more employees, and pay out more in wages and salaries, and in investment spending, *pari passu* with making more loans. The funds created and spent on business expenses then return to the banks as interest on their loans. In an extension of Kalecki’s principle, banks, like businesses generally, collectively earn what they spend.

1.5 An earlier form of the circuit: metallic money

An earlier version of the circuit features a different starting point and different agents—merchants instead of banks—but follows the same route. Merchants might purchase the net output of the consumer sector for coins. Having sold its net output, it could now purchase replacements and new capital goods. This would then set off the ‘secondary circulation’ in

the capital goods sector. That sector would then purchase its replacements and net new equipment, and begin production, paying out wages. As its workers spend, they buy consumer goods from the merchants; as the funds return to the merchants, they buy further goods from the consumer sector, which in turn pays its workers, who return the funds to merchants as they buy the goods. When production is complete, the consumer sector has a new set of goods, the funds are all back in the hands of merchants, and the system is ready for another round. The merchants will have made their profit, in the form of an inventory of goods, from the difference between buyers' prices and sellers' prices. Some of this they may consume; another portion can be sold abroad or to mining for the additional coins or bullion needed for the larger output of the next round.

In this case, we can see how the 'value of money' can be established through the interaction of demand and supply, where the former takes a form reminiscent of the Quantity Equation, and the latter is governed by cost of production.

First, money must be expressed in a standard. Assuming that it is supplied appropriately, in order for money to have a stable value in circulation it has traditionally been thought enough to express it in a stable standard.¹ But the study of the circuit shows that expressing money in terms of something stable is not the issue. It must be put into circulation in a manner that will enable it to complete a *stable, closed, complete circuit*. Then, but only then, can we say that it has a stable value in circulation. So it is not a single transaction, e.g., with a 'standard', that defines the value of money, but a stable *sequence* of transactions, determinate in value and following a definite order.

From the account of the circuit, we have the answer: the required quantity of whatever serves as the medium of circulation must be equal in value to $W_k (= P_c)$. So the product of price \times quantity of the medium is equal to a fixed sum; hence the relationship between price and quantity will be a rectangular hyperbola (see Figure 3), the area under the curve being W_k .² This, of course, is the simplest case, abstracting from institutional detail. Allowing for realistic complexities may change the shape somewhat, but the relationship will be inverse and will centre on W_k . This inverse relationship echoes the old Quantity Theory: the lower the value, the more of the money-article will be required to carry out the circulation.³ (The equation does *not* say, the more there is of the money article, the lower will be its value.)

The supply, in the simplest case of constant costs, will be a horizontal straight line, set by the cost of mining and minting, marked up by seignorage. (Seignorage is analogous to profit; suppliers require it, and the market is willing to pay it for the convenience of having reliable coins.) If there are increasing costs in mining (or minting), the supply function will rise. The intersection of the two determines the *quantity* of money that will enter circulation. The *value* of money at every level is determined by the supply function, which shows

¹ Suppose a good is chosen as standard which does not change in value when the economy changes from growing on a Golden Rule path to a condition of extreme divergence from the Golden Rule. As shown in the Appendix, $W_k = P_c$ along the Golden Rule path, and in general not otherwise. So money would be expressed in terms of a stable standard, but would not have a stable value in circulation. (If divergences are small, however, they will not matter for industrial economies, Nell, 1998, ch. 7.)

² As Walras, Marshall, Pigou and Fisher all held, cf., the long discussion in Patinkin (1965, ch VIII, and Supplementary Notes C, D, F, and G). For qualifications, especially when variations are large, cf., Canaan (1926).

³ This does not imply that throwing more coins into circulation will necessarily lead to a rise in prices. In the short run it might, or it might not, depending largely on whether merchants and money-lenders were inclined to hoard. That, in turn, would depend on what they thought the true value of money was, which brings the argument back to the supply function.

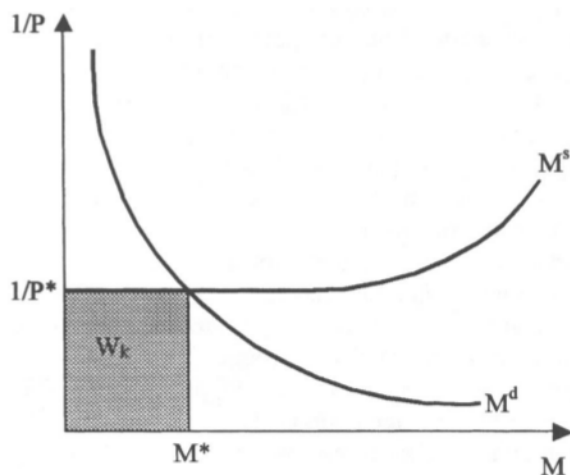


Fig. 3. Determination of the quantity of circulating money.

marginal cost; the value required in circulation, in contrast, is determined by the analysis of the circuit. The intersection determines the value of money at the margin and this is imputed to the money that will circulate. This is in line with Classical Theory generally, in which prices and values are determined chiefly on the supply side, quantities by demand. (Although this is a ‘supply and demand’ theory, it is *not* a typically neoclassical theory: scarcity plays no role.)

On the vertical axis, measure the value of money—reciprocal of the price level—on the horizontal measure the quantity of money. Demand, the money needed for circulation, will be a hyperbola, the area under which will be W_k . Supply, the quantity of the money-article being offered for circulation, will be a horizontal straight line starting from a point on the vertical axis representing the production costs of no-rent mines, marked up by the costs of minting plus seignorage. At a certain point on the horizontal axis, the line will stop and shift up, indicating that at this point higher cost mines will have to be operated.

Maintaining the value of money requires adjustments when, for example, output fluctuates because of an unusual harvest. Mark off ‘gold points’ on the diagram. Destabilising movements are possible if agents, for example, believe that prices will *continue* to fall; they would then withdraw money from circulation. If they believed prices would *continue* to rise, they would convert hoards of money into stocks of goods. The role of the Mint is to buy and sell at the gold points to prevent destabilising movements.

From Figure 3, it is clear that $M^*/P^* = W_k$. Substituting in the equation for circulation derived earlier yields

$$P^*Y = M^*V, \text{ where } V = [1 + (1 + wn_k - wn_c) / \{wn_k(1 - wn_c)\}]$$

This has the form of the traditional quantity equation, but here money is endogenous. Putting more gold into circulation might temporarily raise prices, but then a unit of gold would exchange for goods worth less, so gold mining would be curtailed. Changes in money may affect prices in the short run—the traditional quantity theory—but in the long run money adjusts. (Cf. Moore, 1998.)

Of course, a system of currency based on full value coins is hugely expensive—and does not conveniently generate credit.¹ But coins can be replaced by paper or even by writing cheques on deposits, so long as the paper notes or deposits can be converted at will into coins. The supply and demand functions will determine the total money in circulation. Gold will chiefly serve as reserves, rather than for the currency itself, and reflux—aided by bank policy—will keep the value of paper equal to gold. (This will automatically determine the amount of paper; e.g., if the market judges too much has been issued, it will convert, reducing reserves and constricting the issue.)

Fiat money with no reserve backing might appear to be different. Indeed, both Pigou and Marshall draw the diagram with a unitary elasticity curve and a *vertical* supply function. The stock of money is taken as given, with the current flow small enough to be neglected.² Yet this should be rejected. Even on Marshall's grounds, this can only be a degenerate case. There are no marginal costs; there is no true supply function. The *amount* of money supplied has to be expressed in monetary units. These units, in turn, have to be the same as those in which the demand curve is expressed—and those units have a value at each point. In other words, the issuers of fiat money produce notes denominated in the abstract unit currently in use, therefore having a value given by history—and by the prices the Government will pay. This gives the point on the vertical axis; reading over to the circulation (unitary elasticity) curve tells them how much to issue. If they issue more, the value will fall; if they issue less it will rise. (The great danger, of course, as seen by most writers, is that politics may generate pressures to over-issue.)

1.6 Investment

The circuit in a more advanced economy with a stronger commitment to growth would begin from (and end at) a point at which production in both sectors was complete, but no goods of either kind had been marketed. At the outset, banks lend to the consumer goods sector to finance spending on the acquisition of replacements and new capital goods, i.e., they underwrite purchases of capital goods by the firms of the consumer goods sector. (This finance is one period funding for transactions.) The sum so advanced would equal the profits of the consumer sector. This sum would then flow to the capital goods sector as revenue, and circulate in that sector. This would enable capital goods firms to purchase investment goods and pay wages, whereupon the spending of the wages of the capital goods sector would return the funds as profit to the consumer sector, enabling businesses in that sector to repay their loans. At this point, the previous inventory of both sectors would have been circulated, and a new set of goods of both kinds produced, but not yet sold.

Clearly, any actual economy might operate with a mixture of this circuit and the circuit of advanced working capital.³

¹In a metallic currency system, there is a definite supply function for money. Each unit of the money commodity produced has a definite cost. Supply and demand functions intersect at a margin, and the value determined there is imputed to the intramarginal units. These units (or paper based on these units) circulate. However, such a currency system has no organised supply curve in the market for loans; in the absence of a banking system, there is no uniform or well-defined product. Every loan will be a different product, and it will be difficult to identify a dominating rate of interest, nor will there necessarily be market forces pressing all rates to move together.

²The circulation curve shows current requirements for circulating output; it cannot be combined with a *stock* supply. The latter includes money held for financial and speculative purposes, and issued against different kinds of collateral. The turnover times for production and for financial and speculative balances will differ.

³In practice, banks do both kinds of business; real economies operate both kinds of circuits. For a mixed system to operate smoothly, it must be possible for banks to transfer business back and forth from advancing working capital to supporting investment purchases, for example when the real wage (rate of profit) changes. A rise in the real wage, increasing consumption, implies a lower rate of profit and lower investment (assuming full

Note that banks do *not* finance investment spending as a whole, but only that of the consumer goods sector. ('Finance' means funds are advanced at the beginning of the period of circulation, and not repaid until the end, with interest charged for the entire period.) In particular, there could not be an 'investment-saving circuit', in which the advance of the entire sum to be invested then 'generated' total income through multiplier re-spending.

At first glance, such a circuit might seem attractive and plausible: banks would advance investment funds, which would then be spent and re-spent, according to the multiplier, with savings withdrawn at each stage, and re-channelled into the securities market, enabling firms to issue securities to pay off the bank advances. If investment is I and the multiplier m , then if investment is sufficiently large, $Im = Y_b$ and full employment output would be circulated by money. (With a Classical Savings function, $m = Y/P$, where P is profits.)

But the flaw is easily seen: the capital goods sector will sell investment goods to the consumer goods sector for a sum equal to the profits of that sector. This sum, in turn, will circulate through the investment goods sector, as described above. Why should the capital goods sector borrow to purchase capital goods when its internal exchanges of capital goods will be monetised by the funds it will receive, equal to its wage expenses, from sales to the consumer goods sector?¹ Such borrowing would simply burden the firms of the capital goods sector with unnecessary interest charges² (Nell, 1998, ch. 5, Appendix).

1.7 Fixed capital

To provide an account of circulation in a modern, mass-production economy fixed capital must be brought into the picture. Perhaps surprisingly, this is not so difficult. Fixed capital lasts for a number of periods; hence producers of fixed capital can only sell replacements at distinct intervals. In the meantime, users of fixed capital must save up for the purchase, by setting aside depreciation funds each period, while the producers must borrow funds for their current working capital (to stand in for what they cannot get from sales revenue). Users of fixed capital lend, producers borrow. But this takes place in the context of a growing economy.

Initially, consider steady growth, and divide the economy into two sectors, one sector produces fixed capital goods, using fixed and circulating capital, the other produces circulating capital (chiefly wage goods) using fixed and circulating capital. Both sectors grow at the same rate, which we can initially assume equal to the rate of profits. Social accounting requires that the circulating capital of the fixed equipment producers must equal the annual charges on the fixed capital of the circulating goods producers (Nell, 1998, ch. 7). (This can be considered a generalisation of the condition that the wage bill of the capital goods sector

employment). Banks will reduce their lending for investment purchases, and increase their advances of working capital. For banks to stay 'fully loaned up', the increase in one kind of business must just offset the reduction of the other. That this will generally be the case is proved in Nell (1998, chs 5, 7). See Appendix.

¹ Parguez and Seccarecia (2000, p. 109) do not seem to understand the relationships between the investment and consumer goods sectors. Indeed, in earlier work (in Deleplace and Nell, 1996) they seemed to think that *all* production, consumer goods and capital goods, must be financed by loans. This suggests that loans equal to the total wage bill, W , would be advanced, on the one hand, and in addition, loans equal to total investment, I , would also be advanced. But they assume that $W = C$; so the loans would add up to total output, $C + I$! So loans plus interest would be greater than output! In any case, according to their present account, 'all profits are now used to repay the loans which had been the prior source of investment finance . . .'. If profits equal investment, where does interest fit in? Moreover, as argued above, it implies that some businesses borrow money that they do not need.

² Once again, this has no bearing on very short period borrowing in anticipation of sales. Such 'bridging' loans are not at issue; the question is whether loans equal to the entire volume of Investment must be advanced for the whole period of circulation, or whether only the Investment spending of the consumer goods sector needs to be financed.

equal the gross profits of the consumer goods sector.) Note further that this implies that, when fixed equipment is initially installed, it is capable of being utilised more intensively as time passes.¹ Users earn interest on their accumulated depreciation funds; each period they plough back their earnings, both their operating profits and their interest on their depreciation, so their capital grows. Hence, each period they set aside a larger amount. Producers borrow an expanded amount each period so that they grow from period to period, keeping pace with the rest of the economy. Clearly, for depreciation funds to expand in pace with replacement plus growth requirements, the rate of interest must equal the rate of growth.

Users of fixed equipment set aside depreciation funds, and buy securities dated to fall due at the time of replacement. The receipts from the sale of securities are channelled (possibly through intermediaries) to producers of fixed equipment, to provide their operating funds. Each period the revenue of users grows at rate $1 + g$; it compounds from there to the replacement date. Starting at the installation date, fixed equipment F , in conjunction with labour and materials, produces output Y (at given prices). After n periods, F must be replaced; n periods later output must be $(1 + g)^n Y$. For the *normal* or *planned* capital-output ratio, F/Y , to be unchanged,² F must be replaced by an augmented supply of equipment, $(1 + g)^n F$. Let the initial depreciation allowance be $A = F/n$ in the initial period. The total depreciation fund in period n will be:

$$(1 + i)^n A_1 + (1 + g)(1 + i)^{n-1} A_2 + (1 + g)^2 (1 + i)^{n-2} A_3 + \dots + (1 + g)^n A_n$$

so that if $i = g$, we have

$$(1 + g)^n (A_1 + A_2 + \dots + A_n) = (1 + g)^n nA = (1 + g)^n F$$

Suppliers of fixed equipment must also expand at rate g , if the economy is to grow in balance. Their initial output is F , which must expand after n periods to $(1 + g)^n F$. Each period, they borrow the circulating capital they need to operate. Since it took n periods to stockpile F , A must be the circulating capital needed in the initial period, $(1 + g)A$ will be that of the next and so on. Interest will compound for n periods on the first, $n-1$ on the second, etc. Producers will therefore owe exactly the sum of the depreciation allowances, namely

$$(1 + i)^n A_1 + (1 + g)(1 + i)^{n-1} A_2 + (1 + g)^2 (1 + i)^{n-2} A_3 + \dots + (1 + g)^n A_n$$

Notice that all terms in this expression are equal, provided $i = g$. When users set aside the funds corresponding to the last term, they will not lend them. Instead, they will advance these funds as a payment for the new, expanded batch of equipment to replace the old. At the same time, they will cash in their securities, selling each in reverse order of the date purchased. As the securities are sold, the funds will be returned to them, allowing them to make the next payment. After $n-1$ batches of securities are cashed, all the securities will be cancelled, and the payments made will have exactly covered the full cost of the new equipment. The funds will then be spent by the producers of fixed equipment on the first period's

¹ Each period the economy grows in balance; but by assumption, and realistically, fixed equipment is installed only at distinct intervals. This implies that the output of industries using *given* fixed equipment must expand. Labour and materials inputs can be increased, but the given fixed equipment must be used more intensively—it may become more productive as it is broken in, or in anticipation of growth it may have been deliberately 'overbuilt' when installed.

² Here F/Y designates the ratio designed by engineers and expected by the managers of day-to-day operations. Of course, as time passes, the given fixed equipment will be used more intensively as output and labour and materials inputs grow. This would suggest that the ratio F/Y was falling. While true in *current* terms, such a conclusion would be a mistake—the expected and designed productivity of the equipment is unchanged. The ratio F/Y must be calculated as an average over the *lifetime* of F —the equipment is designed to be used more intensively over time, and this is built into the investment decision.

circulating costs of the *next* set of fixed equipment. Fixed capital, then, fits in as a generalisation of our basic relationships—but it requires a new kind of finance.

1.8 Conclusions

A circulation approach looks at flows of money between interdependent sectors and classes of income recipients, rather than at the decisions of independent individuals. The question to be answered is, what amount of money will be required to ‘monetise’ all the transactions of the economy, and how will it circulate to accomplish this? The amount required is the wage bill of the capital goods sector, and the pattern of circulation can be summed up in an expression reminiscent of the traditional concept of velocity. The result is something akin to the Quantity Equation, but in which money is endogenous.

Transactions can be calculated and expressed in terms of an abstract *numéraire*, but capitalist production requires money—capital ‘turns over’ in a regular process. The turnover can be carried out either with metal-based money advanced by merchants or with bank deposits created through credit lines provided by banks. In the latter case, the realisation of interest must be linked to capital-income realisation generally—banks, like businesses, ‘earn what they spend’. Understanding these processes provides the key to the expenditure multiplier; this links the Classical Equations of production through money to effective demand.

2. The Theory of Interest

Interest must cover the expenses of the banking/financial system. This is a minimum requirement. There are two cases, one where the money and credit issued is backed by real reserves, the other where it is not. When money and credit are backed by real reserves, the supply is constrained. This constraint leads to pressure on interest rates, so that supply and demand tend to stabilise financial markets, and the market rate tends to gravitate around the long-term cost-covering level. By contrast, when money and credit are not constrained by real reserves, interest rates are underdetermined, and supply and demand provide no automatic stabilising forces.

2.1 The long-run interest rate for real reserves

For the case of real reserves, then, a basic explanation of interest can be developed, in terms of costs and the requirements of circulation, allowing questions of stability to be examined. We begin from our earlier observation that the banking system collectively earns what it collectively spends. This has an important implication. The rate of interest clearly depends on the condition that the earnings from interest cover the cost of the banking sector and its growth, given the size of the sector (the number of employees, etc. needed to service the loans, etc.) A higher rate of interest will tend to raise bank profits; bank profits re-invested will permit the banking system to grow at the same rate as the economy, and thus provide a growing supply of bank services and liabilities. In a reasonable sense, then, the rate of interest can be considered a kind of long-run supply price. This provides an answer to an age-old question in economics: why is the risk-free rate of interest positive?¹

¹To determine interest rates, productivity and thrift are traditionally invoked. Borrowers have the use of funds for a time, during which, invested, they yield a return. Savers have foregone the consumption they might have enjoyed during the time of the loan. The thrift of the savers must be compensated; productivity makes this possible. Alternatively, the Keynesian tradition sees interest as compensation for sacrificing liquidity. Heinsohn and Steiger (2000) have recently proposed that interest is compensation for sacrificing a ‘property premium’. These answers all focus on the opportunity costs of individuals. By contrast, the argument above finds an explanation of the level of the interest rate in the real costs of providing a circulating medium for the system. (See also Panico, 1988.)

Let us suppose that service fees and similar charges cover office expenses and rents. Then the costs to be covered are those of wages and salaries. That is, iW'_k has to cover W_b , where W_b is the circulating capital of the banking system, and $W'_k = W_k + W_b$. Bank capital is invested in reserves, so earns nothing. There will be a more or less socially and technically determined number of bank offices and personnel, especially loan officers, needed to process and manage the advances of working capital required to circulate the output. So

$$W_b = \nu_b W'_k$$

where W_b is the capital advanced for providing the loans, that is, of paying the salaries of the bank tellers and loan officers. This has to be covered from the earnings from the loans. But, in addition, the capital of the banking system has to grow with the economy, from period to period, in order to continue to supply the need for advances. It will do so

$$iW'_k = W_b + P_b.$$

But since W_b is the circulating capital, if banking earns the general rate of profit, $P_b = rW_b$. So, $iW'_k = (1 + r)W_b$, and since we assume that $r = g$, this implies

$$i = (1 + g)W_b/W'_k$$

where g is the rate of growth. That is, banks invest their profits in the purchase of additional bullion, so their reserves grow at the same rate as the economy.

But since $W_b = \nu_b W'_k$

$$i = (1 + g)\nu_b$$

This gives the long-run equilibrium level of the rate of interest in terms of the rate of growth and the ratio of labour employed in the banking sector to the labour force.¹

Long-term loans, such as bond sales of household saving to business, will be offered at an interest rate that is a mark-up on the basic rate. Banks must pay a competitive rate on deposits of household savings; to cover their costs and earn a profit they must charge a mark-up on their loans.² (In addition they will want to consider risk and liquidity loss.) So the yield curve will rise in a 'normal' fashion.

2.2 A closer look at interest and bank money

In the above case, it was assumed that W'_k was given. Now consider the banking system facing variable demand for loans. A *service* is provided by the supplying institution, which is fixed in size (W_b is given). This tells us that costs take the shape of a rectangular hyperbola (Figure 4). There is no definable unit cost; marginal costs are virtually zero. The cost of writing up a loan is the same whether the loan is large or small; the costs of supplying the currency—bank liabilities—are those of the banking system as a whole. Revenue is earned by making loans. There is no short-term or medium-term independent supply function for loans, based on costs; bank credit is demand determined but reserve constrained. Banks decide only who is credit-worthy; for any given bank demand is largely outside its control,

¹This is the rate which in the long run the rate of interest must realise on average, not the rate of interest on long-term bonds. This equation must be interpreted carefully. It applies to the banking system as a whole. For any individual bank, efficient operations and reduced labour requirements will be advantageous. So n_b is determined competitively. But the volume of business of the banking system is W_k , which cannot be increased by more efficient banking. Hence, the interest rate will be higher the greater the (competitively determined) level of n_b . (That banks must earn the normal rate of profit was recognised by J. S. Mill, as early as 1844; cf., *Essays*, IV, 1874, p 115.)

²Bank earnings on such loans will be realised as before, through the collective spending of the banking system.

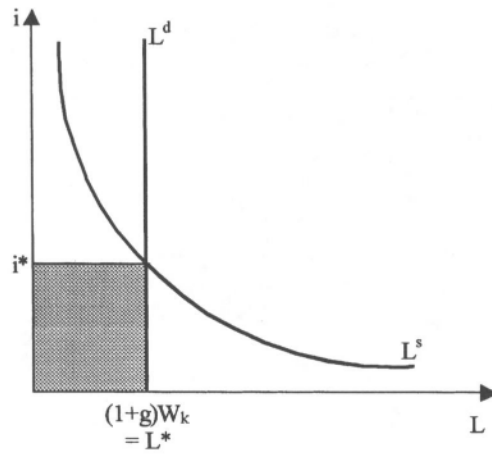


Fig. 4. Loan demand and loan supply in the determination of the rate of interest.

and it may or may not be possible for it to attract reserves. But for the banking system as a whole, both demand and reserves depend significantly on the rate of interest.

On the vertical axis, measure the rate of interest, and on the horizontal axis the volume of loans. The fixed costs of the banking industry will be the area under the rectangular hyperbola representing supply. Included in this may be the fixed sum of profits required for bank capital to grow at the required rate. The constraint will be a vertical line determined by reserves; assuming they were adjusted to the needs of circulation, the line will rise from the point $(1 + g)W_k$ on the horizontal axis. This diagram shows the rate that would cover the banking sector's costs and required investment, for a given level of lending activity. The actual rate of interest will depend on supply and demand for reserves.

2.3 *Stabilising the interest rate with real reserves*

In a boom, then, banks, pursuing profits, will try to lend more and will seek to create new deposits or issue additional notes; to support these activities, they will have to attract additional reserves. This will lead them to bid up interest rates, as they seek to attract idle reserves from each other and from hoards. In a slump, they will issue less and lend less, and will seek to shed reserves, lowering interest rates. In other words, while long-term average rates are determined by costs and competition, current interest rates reflect the balance of supply and demand in the market. They move pro-cyclically. Let us look more closely

A simple model may be suggested. On the one hand, the rate of interest (in relation to the rate of profit) is likely to affect investment inversely, and investment, in turn, will have an impact on prices and employment. Changes in prices and employment will call for changes in reserves. (Even in the conditions of the Craft Economy, the monetary system and the real economy cannot be separated—the Classical Dichotomy does not hold.) On the other hand, shortages or surpluses of reserves will lead to market pressures driving interest rates up or down. Monetary adjustments do complement and reinforce the stabilising effects of the price mechanism.

2.3.1 *Interest rates, investment and the demand for reserves*

To see these relationships clearly, we can plot them on a diagram (Figure 5). The origin will represent the normal position—assumed to be the average over the cycle. Here the rate of

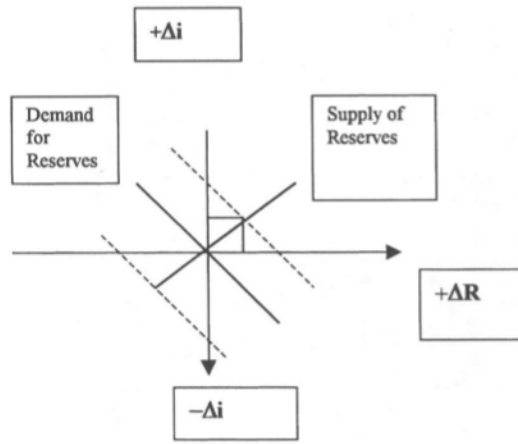


Fig. 5. Reserves and the interest rate.

profit will be at its normal level, and the rate of interest will be just sufficient to cover the expenses of the banking system, allowing it just enough profit to grow at the same rate as the rest of the economy. The vertical axis will represent increases in the interest rate above and below the normal level. The horizontal axis will show the level of reserves, similarly above and below the normal level.

First, we have the relationship between above or below normal interest rates and Investment, which translates into demand for Reserves.

Investment depends inversely on the interest rate—the more expensive it is to borrow in relation to normal profits, the lower will be investment spending, and the less expensive the higher it will be.

Savings depends positively on prices (in relation to wages) and on employment—when prices are bid up, profits increase and profits are saved.

Prices in relation to wages, and employment and output, all depend positively on investment.

For equilibrium, savings has to equal investment.

Taken together, these propositions tell us that, if interest is low, for example, investment will be high, and the resulting demand pressure will drive up prices in relation to money wages, lowering the real wage and increasing employment (and output), raising profits and thus savings; this will go on until savings has risen to match the high level of investment. The process works in reverse; if interest rates are high, investment will be low, prices will fall relative to money wages, so with real wages high employment will fall, and profits and savings will be down. So savings *adapts* to investment, in a Keynesian fashion. But savings depends on profitability and employment, and not on the interest rate, so the latter cannot be determined here.¹

¹ Supporters of the Loanable Funds approach may object, but there is little evidence that saving has ever been very responsive to interest. More importantly, there is a theoretical objection. When interest increases, wealth-holding households find their income higher, but by the same token, business firms find that their profit income net of interest is down. When everyone has adjusted to the change, wealth-holding households may save more, but businesses will be saving less. On the assumption that all capital income is saved, these changes will cancel. If wealth-holding households consume a portion of capital income, but businesses, in the light of competitive pressures, save it all, then a rise in interest would *reduce* saving.

The preceding implies that the level of reserves required depends on the money value of output, so on prices and employment. Thus, the required reserves will be large when interest rates are low, and smaller when they are high. Banks start out with a 'normal' level of reserves; this will have to be adjusted as the interest rate changes. So the curve passes through the origin, falling from the left to the origin and continuing through it. Higher than normal interest rates mean low investment, output and employment, so lower than normal demand for reserves. Lower than normal interest means higher activity, and higher required reserves.

2.3.2 Interest rates and the supply of reserves

To complete the simple diagram, we must explain how reserves will be increased when interest rates are raised; and when interest rates are lowered, reserves will be reduced. That is, the supply of reserves varies positively with interest rates. This needs careful explanation.

Since g has risen, it might seem that gold output would rise *pari passu*. If so, things would be simple. Gold producers would expand mining and processing, and so increase their purchases of capital inputs by an amount necessary to raise gold output to the higher level necessary to ensure that new gold/existing gold stock will equal the new higher growth rate, g_1 . But the increased purchase of capital inputs would mean an increase in the supply of gold; reserves would therefore increase. The increased purchase of capital inputs/normal purchase of capital inputs would equal the new growth rate, assuming equilibrium ratios. Hence, gold reserves would expand in the necessary ratio. But what *incentives* would lead gold producers to this?

Consider: the pressure of the boom will tend to bid up prices (relative to money wages), which would seem to *reduce* the value of money, whereas a *rise* in the value of money is needed to induce mining. That is, there should be a shortage of money relative to goods, as would be expected with a rise in employment and output in a boom; this would call for more output from the mines. But the rise in prices suggests an increase in money relative to goods!

This would mean a lower value of money (monetary gold), which would hardly be an inducement to produce/supply gold.

However, gold producers do not necessarily have to sell gold. They could hold it and deposit it with the banks, in return for interest. In equilibrium, the rate of interest should equal the rate of profit. So gold producers should be indifferent between *investing* new gold, e.g., ploughing it back into the business (selling gold to buy capital inputs), to earn the rate of profit, or *lending* it to earn the rate of interest. They will sell enough to keep the supply of gold in circulation at the level required, and will lend enough to keep the rate of interest at the equilibrium level.

So when a boom develops, the growth rate will rise, and prices will be driven up relative to money wages, raising profits. That is, goods prices measured in the *actual circulating medium*—paper and credit—will rise, relative to wages in that medium. By itself, this would appear to lower profits in gold, since gold's command over goods would be less; but the boom is fuelled by an *excessive* issue of paper and credit. So papering reserves come to be inadequate; so to back their issues, banks will compete to borrow reserves from one another. (There will always, normally, be periods in which a given bank's reserves will temporarily be excessive, making them willing to lend.) This will engender a widespread increase in borrowing pressure, driving up interest rates. Higher rates are likely to increase reserves somewhat, as they will attract 'hoarded' gold, from inventories. But for reserves to increase more substantially, the rise in interest rates must lead directly to a rise in mining. Why?

Mining will increase because the higher rate provides gold producers an opportunity to increase earnings.¹ Since the banks need additional reserves, the mines can increase their output and deposit the additional gold with the banks, and collect interest at the higher rate. As long as the costs of mining and minting are constant, the value of money will be unaffected; the additional output will simply be loaned to the banks, providing them the extra reserves they need. (When the boom deflates, the deposited gold can be withdrawn and put on the market; if necessary, current output can be curtailed. There need be no pressures either way on the price of gold.)

Thus, adjustments in the interest rate will tend to attract or release reserves so that the actual level on hand will be brought into line with that required by the level of activity. Higher prices and output call for higher reserves; to attract more reserves a higher interest rate will be needed, but this will also tend to reduce investment and so bring about lower levels of output and employment. A lower interest rate, in turn, would tend to release reserves, but it would also tend to stimulate investment, and so increase activity and thus raise the demand for reserves. At some intermediate level of the interest rate, the level of reserves attracted should just match the level required. This will be the temporary equilibrium level, as determined by market forces.²

The rising supply curve of reserves has the effect of dampening the impact of the shift in the investment function—the steeper it is, the more it reduces the effects.

Putting the two functions together, with the normal equilibrium at the origin, we see that a shift upward in Investment—due, perhaps, to boom psychology—will raise the demand for reserves (upper dotted line). This will now intersect with the supply at a higher rate of interest, and higher than normal level of reserves. A slump would lead to just the reverse (cf., lower dotted line).

As in the case of supply and demand for circulating funds, this should be a stable market—by the conventional account. If actual reserves are inadequate, those demanding reserves will be willing to pay higher interest charges, and the higher interest will attract more reserves. But note that if there are delays or time lags, the reactions could take the form of a ‘cobweb’. In this case, with a steep Demand function and a moderate Supply function, the cobweb might well be divergent. In that case, Government would be needed to maintain stability by adjusting the interest rate. The policy of the Central Bank would be ‘to lean against the wind’—so the interest rate will move in the right direction—but the Bank can also ensure that the changes will be large enough to have the desired impact.³

The chief reason for shifting from gold and convertible paper to modern bank money was that gold and real reserves could not be quickly adjusted (Day, 1957; Davis, 1996). The quantity in circulation adjusted through the price mechanism; but as the economy shifted to mass production, employment became highly variable; the wage bill would fluctuate with layoffs and re-hiring. The wage bill, however (together with spending on investment), provided the chief market for bank lending, which would now have to adjust quickly to changing demand. But a gold-based system could not expand without first increasing reserves,

¹It might seem that instead of selling gold spot, they could now sell it forward, in the meantime depositing it at interest with the banks. If the spot and forward prices are fixed and equal by law, as under the Gold Standard, this would work. If the prices are set by the market, however, the relation between the spot price of gold and its price for delivery at a forward date must reflect the interest rate. A rise in the interest rate means a rise in the discount applied to the delivery price of gold. The gain in interest would be offset by a lower forward price.

²This provides an explanation for what Keynes called ‘Gibson’s Paradox’, namely the strong and centuries-long correlation between the level of prices and the interest rate (Keynes, *Treatise*).

³Note that fiscal policy cannot be very effective here. A deficit does provide a stimulus, but at the same time it will shift the demand for reserves up, raising interest rates; a surplus provides a drag, but will lower interest rates.

which had to be attracted by higher interest rates. In a slump leading to extensive layoffs and reduced demand for loans, reduced lending would leave banks with excess reserves—an unnecessary expense—leading to reductions in rates and efforts to shed reserves. The countercyclical movement of the interest rate certainly helped to stabilise the economy, but it put a great strain on the banking system (Currie, 1934, 1968; Graham, 1937, 1944).

2.4 *Financial markets and the interest rate without real reserves*

In a system constrained by real reserves, where these have been adjusted to the needs of circulation, banks cannot easily create loans for turning over or speculating in financial assets. As a consequence, financial markets will tend to be ‘thin’, which means that asset-holders cannot easily cash in capital gains. They will therefore press firms to pay out profits, rather than retain them. So companies will distribute earnings, and then issue securities to attract the funds back for investment. Owners and asset-holders will receive the paid-out profits, save them and re-invest them in the securities of companies of their choice. Financial markets channel savings to investment opportunities. Companies will seek to attract savings by offering the most advantageous terms; new issues of shares will compete with bonds. So the ratio of new shares issued to existing ones, which will tend to equal the ratio of investment to existing capital, will tend to equal the rate of interest.

Financial markets work quite differently when banks are no longer constrained by real reserves. Once freed of the reserve constraint, banks can make loans based on securities as collateral. This allows for much greater activity in asset markets, which consequently deepen. Owners can easily realise their capital gains in the market, and choose their portfolios. They do not have to receive profits in the form of dividend pay-outs. As a result, companies can retain their earnings, and plough them back, which allows them greater scope for long-term planning.

In such a system, the rate of interest will not be determined by demand and supply for reserves. For what are called ‘reserves’ have become arbitrary and can be manipulated, reserve ratios can be evaded, and above all reserves are not a special kind of money, like gold. They are just accounts at the Central Bank—which can be created by the Central Bank. As a result, there are no marginal costs, and so no independent supply function.¹ The interest rate cannot be determined by ‘supply and demand’.

So what happens? Banks can now invest in earning assets and no longer have to hold reserves in bullion. At this point, two separate notions of interest can be distinguished—interest as the price for the service of supplying the medium of circulation, and interest as the reward for lender’s risk (and lender’s sacrifice of liquidity). The latter, of course, spills over into being a reward for speculation. So in figuring the rate of profit in banking, we must consider the rate of interest on bank capital separately from the rate of interest on circulating funds. The rate of interest on long-term assets must be set so that bank assets earn a profit, which when combined with earning from circulating loans and invested, will enable them to grow at the same rate as the economy in general.

$$g_b = g = i_L = \Delta K_b / K_b$$

This gives the target rate. But does arbitrage generate pressures that will tend to *establish* $i_L = g$? Competition in the purchase and holding of assets should respond to differences in

¹Wicksell (1898) developed this idea. If the ‘money rate of interest’ lay below the ‘natural rate of return’, then the economy would have a tendency to inflate, if the money rate were above the natural rate, the economy would deflate. If there are no real reserves to constrain the system, the inflation and deflation will tend to continue indefinitely.

rates of return. In particular, there will be arbitrage between long-term bonds and growth stocks. The latter will tend to grow at the rate of growth of the underlying companies, which is to say, at the growth rate of the economy. So the interest rate and the growth rate will be subject to pressures and, certainly, will not for long be able to move in opposing directions. But that does not mean they will be pulled together.

Suppose that $i_L = g$ is the case; if the short rate earned *more* than what was needed to cover wage and salary expenses in banking, plus profits for expansion, then the overall banking rate of profit would exceed the growth rate, $r_b > g$. In this case, bank capital would tend to expand faster than the demand for working capital loans, so that competition would tend to force the short rate down. Conversely, if the short rate failed to just cover the wage and salary expenses of banking, the profit rate in banking would fall short of the growth rate, and bank capital would expand more slowly than the demand for working capital loans, and the short rate would tend to rise. Hence

$$i_S W'_k = W'_b + P_b$$

which implies $i_S = (1 + g)wN_b/wN_k$, so assuming a common level of wages,

$$i_S = (1 + g)N_b/N_k = (1 + g)v_b$$

This, of course, is the same as the basic—expense-and-growth-covering—rate of interest established for the system with real reserves. But arbitrage may not be able to guarantee that growth rates and interest rates will be uniform. So it may no longer be the case that i_L will always be greater than i_S . Yield curve reversals are now possible, and speculative activity may contribute further to ‘disorderly’ financial markets.

2.4.1 Arbitrage

The financing of fixed capital requires the development of markets in securities; this sets the stage for a different circulation, in the financial sphere. The issue of money for this activity cannot be restricted by holdings of real reserves; the supply of funds needs to be flexible. Banks will have to have sufficient capital to cover risks, and then will issue against collateral for the purpose of trading in securities. At any time, a fraction of bonds will be retired, and new bonds come on the market; shares in declining industries will be sold and those of advancing industries bought. Growth of capital will be reflected in new issues of shares and bonds, requiring new money. The process will be regulated by the interest rate in relation to the growth rate, as that is reflected in the appreciation of share prices and/or in new issues.

But just how reliable are the forces of arbitrage? Profit rates tend to uniformity because movements of capital in pursuit of higher profit will tend to adjust prices appropriately to wages, bringing uniformity. The growth rate will tend to uniformity, because, in the long run, all rates are governed by the ‘slowest-expanding’ basic good. Competitive pressures will lead to all after-tax profits (net of interest) being invested, apart from necessary but ‘unproductive’ business expenses, so g will tend to equal r , appropriately adjusted. This brings us to interest, where the situation may be different.¹

¹Would not arbitrage by investors tend to set $r = i = g$, allowing for risk? It might seem that i could be brought into equality very easily, since equilibrium implies profits = interest + investment and, in long-period equilibrium (balanced growth), profits/total capital = interest/borrowed capital = investment/equity capital, $r = i = g$, because the debt/equity ratio must stay constant, so interest/investment = borrowed capital/equity. In long-period equilibrium, then, (interest + investment)/total capital = $r = i = g$. But is there any reason to think that the debt/equity ratio must stay constant? Are there any forces tending to set it at a particular level? Why should it not vary with changing circumstances? But if it can vary, then while r , i , and g may move together, there may not be any market forces that directly bring them into equality (Nell, 1998, chs 6, 13; for a different but related view, see Ciccarone, 1998).

First, no product or service can be priced so the usual movements of capital will not adjust the interest rate. Suppose $g-i > 0$ (understanding both in real terms); then funds would tend to move from bonds to real investment. Bond prices would fall, raising i , but this would not eliminate the difference, since the rise in Investment would raise g . The same would hold in reverse if $g-i < 0$.¹ However, in the absence of reserves, bank lending will be constrained by bank capital, and, surely, arbitrage of bank capital with bonds and corporate equity can be expected? Yet the same problem can be seen. Suppose bank earnings are high because interest rates are high relative to growth, so bank capital is growing faster than corporate capital or output. This could lead corporations to invest their retained earnings in banks, rather than in their own activities. Such expansion of bank capital will lead to further lending and money creation, tending to lower interest rates. But the very move that leads to this—corporate investment in banks—also reduces output growth. The initial divergence between interest and growth rates will not be eliminated by this arbitrage.²

2.4.2 Supply and demand

There is no supply curve in the usual sense, and arbitrage cannot be relied on to bring the rate of interest into equality with growth or profit rates. Nevertheless, there is a demand for money, which banks do supply. We have seen that covering the costs of the banking system sets a minimum level to the interest rate. With fixed capital there must be a securities market, and output growth gives rise to further demands in that market.

To service the demand for securities, money will be issued and held for a number of purposes. With regard to bonds, growth means that the retirement of old bonds will not fully pay for the new ones. The same holds for stocks, when the economy is growing, the sale of declining industries will not suffice to purchase the shares of newly developing ones. With stocks, also, there are demands that arise because of speculation. The sales by bears may not be enough to finance the desired purchases of bulls (or vice versa). And some holders of stock, each period, may wish to cash in their capital gains. There will have to be money for realising such capital gains, and, in addition, there will normally be purely precautionary holdings. (Nell, 1998, ch. 6)

The demand for money for financial markets, then depends on growth in bonds and growth in stocks. A certain percentage of the value of the additional securities will be needed to purchase them; call this α , where $0 < \alpha < 1$. The rest of the value of the securities will be covered by margin credit, or will not be traded. Initially, there will have to be money in the financial markets to cover normal activity—no doubt varying with the ebb and flow of speculative fever. Let us ignore this and concentrate on the change in the amount required due to growth. This will be

$$\Delta \mathcal{F}_D = g\alpha V_{\mathcal{S}}$$

where $V_{\mathcal{S}}$ is the financial value of the capital stock.

Then the supply of money will be forthcoming in response to the demand. But the supply will be able to grow only if bank capital is growing; the growth of bank capital depends on

¹ Another way of putting this: interest rates must be considered ‘low’ or ‘high’ in relation to the rate of growth of output.

² Suppose $i > g$; the profits of banks will enable them to accumulate bank capital faster than the demand for loans is growing, giving them the capacity to fund speculation. But if stock prices are only appreciating at g , why would anyone borrow at i to play the market? The rate g is an average of declining and advancing industries. Speculators will move to the advancing industries, where price appreciation rates will exceed interest, and the influx of funds will drive up prices even faster.

bank profits, which depend on the rate of interest. For simplicity, assume that bank capital grows at the interest rate. Then the ability to supply additional money will depend on the interest rate. But the loans that create the new deposits will require collateral—and that will depend on V_g . A certain fraction will be unburdened and considered creditworthy, call this β . Then the potential money supply will be

$$\Delta S_s = i \beta V_g$$

Clearly, demand and supply for money in financial markets will not determine the rate of interest. Indeed, in a plausible case, where $g = i$ and $\alpha = \beta$, nothing would be determined at all.¹ The two functions coincide.

Issues of money for speculative purchases in financial markets will require collateral, but the securities themselves can be used as collateral. This has the effect of making an asset boom or slump *self-sustaining*. If it is a boom, then the collateral will increase in value, justifying further issues; if a slump, the collateral will decline in value, requiring loans to be called in. When asset prices rise or fall, bank capital is likely to move in tandem, which means that the money supply will accommodate.

Supply and demand provide no stabilisation; interest rates are not determinate. The levels reflecting costs examined earlier become the *minimum* levels of the short and long rates, as averages over the business cycle. Market activity and financial speculation could raise them above, or depress them below such levels. And speculation can change the shape of the yield curve. To keep financial markets orderly, the Central Bank will find that it has to peg the interest rate—but that takes us to a new topic.

2.5 Conclusions

The Theory of Circulation shows us the exact relationships between monetary exchanges and payments, on the one hand, and production and the distribution of income, on the other. The proper measure of the value of money in circulation is not some good that can serve as a 'standard', but is an *exchange relationship*, a sequence of transactions, that serves to define a circuit. The circuit, then, shows how money and money incomes underlie effective demand, linking demand to money income, and money income to the money value of production. And it also shows just how the monetary system is open to the activities of speculators.

The theory tells us exactly how much money is needed in a monetary system based on a real money-article, or on convertibility. This makes it possible to see how a supply and demand system will work to adjust the money in circulation and maintain its value. This can be expressed in the form of the Quantity Equation, but money adjusts endogenously.

In a system of fiat money, however, there will be no independent supply function; supply and demand have no footing. But the theory here enables us to see how large the banking system will have to be and how much interest will have to be paid to provide revenues supporting that system. This makes it possible to relate interest earnings to costs and production. The reason that the risk-free rate of interest must be positive is that interest covers the real cost of providing the growing supply of the medium of circulation. This implies a relationship between money and real rates of return. It also means that the yield curve will be positive.

When banking develops to the point where it is based on yield-bearing assets—bank capital—rather than convertible reserve holdings, speculative activity can develop more freely, and financial markets will deepen. Interest rates can fluctuate more widely and must be related to the rate of growth of output, which in turn will tend to be reflected in the appreci-

¹Knowing they need money, borrowers will arrange to have collateral on hand that is acceptable to lenders.

ation of security prices. Arbitrage will not necessarily bring interest rates and stock prices or growth rates together, and the relationship between the short and the long rate may move away from 'normal'. The supply and demand for money can no longer provide stabilisation.

Bibliography

- Canaan, E. 1929. *Money*, 6th edn, London, Westminster
- Cencini, A. 1995. *Monetary Theory: National and International*, London and New York, Routledge
- Ciccarone, G. 1998. Prices and distribution in a Sraffian credit economy, *Review of Political Economy*, vol. 10, no 4, October
- Currie, L. 1934, 1968. *The Supply and Control of Money in the United States*, Harvard Economic Studies, 47, Harvard University Press. Reprinted New York, Russell & Russell
- Davies, G. 1996. *A History of Money: From Ancient Time to the Present Day*, Cardiff, University of Wales Press
- Day, A. C. L. 1957. *An Outline of Monetary Economics*, Oxford, Clarendon Press
- Deleplace, G. and Nell, E. (eds) 1996. *Money in Motion*, London, Macmillan
- Eichner, A. 1987. *The Macrodynamics of Advanced Market Economies*, Armonk, NY, M. E. Sharpe
- Goodhart, C. 2000. Metallists versus Cartalists, *European Journal of Political Economy*, 2000
- Graham, B. 1937. *Storage and Stability*, New York, McGraw Hill
- Graham, B. 1944. *World Commodities and World Currency*, NY and London, McGraw-Hill
- Green, F. 1992. *Classical Theories of Money, Output and Inflation*, London, Routledge
- Heinsohn, G. and Steiger, O. 2000. in Smithin, J. (ed.) *What is Money?* London, Routledge
- Hicks, Sir John 1989. *A Market Theory of Money*, Oxford, Clarendon Press
- Hicks, Sir John 1977. *Economic Perspectives*, Oxford, Clarendon Press
- Hicks, Sir John 1967. *Critical Essays in Monetary Theory*, Oxford, Clarendon Press
- Johnston, J. 1970. *Bishop Berkeley's QUERIST in Historical Perspective*, Dundalk, Dundalgan Press
- Kurihara, K. 1951. *Monetary Theory and Public Policy*, London, George Allen & Unwin
- Luxemburg, R. 1963 (1951) *The Accumulation of Capital*, trans. Schwarzchild, London, Routledge & Kegan Paul
- Mill, J. S. 1848, 1987. *Principles of Political Economy*, edited by Sir William Ashley, New York, Augustus M. Kelley
- Mill, J. S. (1844) 1874. *Essays on Some Unsettled Questions of Political Economy*, 2nd edn, reprinted 1968, Augustus Kelley, New York.
- Mosler, W. 1995. *Soft Currency Economics*, West Palm Beach, Triple III Finance
- Moore, B. 1988. *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge, Cambridge University Press
- Nell, E. J. 1967. Wicksell's Theory of Circulation, *Journal of Political Economy*, vol. 75, no. 4, pp. 386–94
- Nell, E. J. 1968. The advantages of money over barter, *Australian Economic Papers*, December, pp. 149–166.
- Nell, E. J. 1970. A note on Cambridge controversies on capital theory, *Journal of Economic Literature*, vol. 8, no. 1, March.
- Nell, E. J. 1975, 1992. The simple theory of effective demand, in *Political Economy at the New School*, reprinted in Nell, E. J. 1992. *Transformational Growth and Effective Demand*, London, Macmillan
- Nell, E. J. 1988. On monetary circulation and the rate of exploitation, in Arestis, P. (ed.) *Post Keynesian Monetary Economics*, Aldershot, UK, Edward Elgar. An earlier version appeared in *Thames Papers in Political Economy*, Summer, 1986
- Nell, E. J. 1992. *Transformational Growth and Effective Demand*, London, Macmillan
- Nell, E. J. 1998. *The General Theory of Transformational Growth*, Cambridge, Cambridge University Press
- Newlyn, W. T. and Bootle, R. P. 1978. *Theory of Money*, Oxford, Clarendon Press
- Panico, C. 1988. *Interest and Profit in the Theories of Value and Distribution*, London, Macmillan
- Parguez, A. and Seccarecia, M. p. 109 in Smithin, J. (ed.) *What is Money?* London, Routledge
- Patinkin, D. 1965. *Money, Interest and Prices*, 2nd edn, New York, Harper & Row
- Rosselli, A. 1999. The origin of the political economy of money, *Review of Political Economy*, vol 11, no. 4, October
- Spaventa, L. 1970. Rate of profit, rate of growth and capital intensity in a simple production model, *Oxford Economic Papers*, vol. 22, 129–47.
- Tew, B. 1969. *Monetary Theory*, Beverly Hills, CA, Oxford Sage

Wicksell, K. 1898. *Interest and Prices*, trans. R. F. Kahn, 1936; reprinted 1962, New York, Augustus M. Kelley.
 Wray, L. R. (2000) Understanding modern money, in Smithin, J. (ed.) *What is Money?* London, Routledge

Appendix: The two-sector model and the Golden Rule

The two-sector model, allowing for changes in productivity, while retaining the assumption of circulating capital, must show the characteristic trade-offs in capitalism between wages and profits on the one hand, and consumption and growth on the other. We shall start from a no-surplus system and develop the relationships. The full model has a degree of freedom; it may be presumed that the rate of growth must correspond to the expected rate of growth of demand (Nell, 1998, ch. 10).

We write the capital goods coefficients as a, A , and labour as b, B , measuring labour however, in the equivalent consumption goods. P_1 and P_2 are the money prices of capital goods and consumer goods, respectively. Then, starting from a given a ‘no-surplus’ system:

$$P_1 = aP_1 + bP_2 \tag{1}$$

$$P_2 = AP_1 + BP_2 \tag{2}$$

let there be a speed-up in production, leading to a more rapid turnover of capital goods, where t represents the number of turnovers in a unit period made possible by the speed-up. This will produce a surplus of consumer goods. However, if the composition of this surplus is to include any of the capital good, the sizes of the sectors will have to be changed; this is done by multiplying by X_1 and X_2 , keeping the total quantity of labour the same.

$$[tP_1 > taP_1 + bP_2]X_1$$

$$[tP_2 > tAP_1 + BP_2]X_2$$

Value relations do not depend on quantities, however. Hence, where $p = P_1/P_2$ and r the rate of profit

$$tp = (1 + r)tap + wb \tag{3}$$

$$t = (1 + r)tAp + wB \tag{4}$$

Here w is the wage rate, meaning the competitively determined ratio of payments to labour to the cost of living. When this rate is unity, or 100%, labour’s earnings just cover the cost of living; when it is, say, 110% labour earns 10% above the cost of living.¹ Then, solving for w in terms of $(1 + r)$, we have

$$w = t(1 - (1 + r)a) / [B + (1 + r)(Ab - aB)] = t[1 - (1 + r)a] / D \tag{5}$$

where $D = B + (1 + r)(Ab - aB) = B[1 - (1 + r)a] + (1 + r)Ab > 0$, since the system’s viability requires that $1 - (1 + r)a > 0$. From (5)

$$dw/dr = -tAb / [B + (1 + r)(Ab - aB)]^2 = -tAb / D^2 < 0 \tag{6}$$

which shows that profit and wage rates are inversely related, and

$$p = b / [B + (1 + r)(Ab - aB)] = b / D \tag{7}$$

so that, $dp/dr = b(aB - Ab) / D^2 > 0$ or $r < 0$ according to the sign of $aB - Ab$, that is, according to the relative capital/labour ratios of the sectors. (We shall take it that the maximum rate of profit comes when $w = 1$, the basic cost of living. If $w < 1$, labour will not be able to function.)

The quantity relations start from the outputs in each sector in relation to their use as input. Initially, we have,

$$X_1 = aX_1 + AX_2 \tag{8}$$

$$tX_2 > bX_1 + BX_2 \tag{9}$$

¹To be analogous to other goods, labour must be represented by the means of subsistence that support it, and these, in turn, are produced by the consumer goods sector. Labour is shown as a column in the price equations, and the consumer goods sector will be the corresponding row. Since labour is represented by the goods that support it, the basic real wage will be unity. The net wage will then be a percentage of this, as market conditions drive the wage above or below unity.

By varying X_1 and X_2 , a positive surplus of the capital good can be created. These two multiply the price equations, establishing the sector sizes. But capital goods will be invested in each sector, and consumer goods will be consumed in each sector; so we must introduce two new variables, g for proposed growth, and c for anticipated consumption per capita. Observe that these variables concern the *current use* of the period's output. In particular, ' g ' does *not* refer to growth of capacity or growth of output. Rather ' g ' is the ratio of the allocated output of capital goods, in each sector, to the amount used as input (Spaventa, 1970; Nell, 1970).

In other words, the relative size of output must reflect the proposed uses in the sectors. Accordingly, the output of the capital good will be allocated to replacement of its use in its own production, expansion in proportion to proposed growth, and use in the production of consumption goods in proportion to consumer demand. Consumer goods output will be allocated to replacement of basic wages in capital goods, expansion in proportion to proposed growth, and providing wage good input in proportion to consumer demand in consumption goods.

$$X_1 = (1 + g)aX_1 + cAX_2 \tag{10}$$

$$tX_2 = (1 + g)bX_1 + cBX_2 \tag{11}$$

so that letting $q = X_1/X_2$, and dividing by X_2 ,

$$q = (1 + g)aq + cA \tag{12}$$

$$t = (1 + g)bq + cB \tag{13}$$

from which we derive

$$c = t(1 - (1 + g)a) / [B + (1 + g)(Ab - aB)] = t(1 - (1 + g)a) / D' \tag{14}$$

where D' differs from D by having g in place of r .

Clearly g and c are inversely related, in the same relationship that holds between r and w . And we also see that

$$q = tA / [B + (1 + g)(Ab - aB)] = tA / D' \tag{15}$$

so that

$$dq/dt = A/D' > 0 \text{ and } dq/dg = tA(Ab - aB)D'^{-2} > \text{ or } < 0$$

according to the sign of $Ab - aB$, the ratio of the capital-labour ratios of the sectors.

When $g = 0$, c will be at its maximum level, and the whole net product will consist of the consumer good; that is the point from which we start when there is a speed-up. To form a net product which includes capital goods, there will have to be an adjustment of the relative sizes of the sectors. The equations for the use of the outputs show a relationship between growth and consumption that is dual to that between profits and wages. The relative quantities of the sectoral outputs are related to the growth rate in the same way prices are related to the profit rate.

The two-sector diagram

In Figure 1, the vertical axis on the left, Y_k , measures tP_1X_1 , and the vertical axis on the right, Y_c , measures tP_2X_2 . We assume constant returns in the short run, so that output is shown as a function of employment. But output must equal income, so the vertical axis must also measure both wages plus profits, and consumption plus investment. That is, each of these axes has two interpretations in terms of the equations. Starting on the LHS, we know from equations (1a) and (2a) and (10a) and (11a), that

$$tP_1X_1 = \text{both } (1 + r)taP_1X_1 + wbP_2X_1 \text{ and } (1 + g)taP_1X_1 + ctAP_1X_2$$

So

$$P_k = (1 + r)taP_1X_1 = I_k = (1 + g)taP_1X_1$$

and

$$W_k = wbP_2X_1 = I_c = ctAP_1X_2$$

When the Golden Rule holds, and $r = g$, the first of these is a tautology, but the second becomes

$$bX_1/X_2 = tAP_1/P_2.$$

Turning to the RHS,

$$tP_2X_2 = (1 + r)tAP_1X_2 + wbP_2X_2 = (1 + g)bP_2X_1 + cB P_2X_2$$

So

$$W_c = wBP_2X_2 = C_c = cBP_2X_2$$

and

$$P_c = (1 + r)tAP_1X_2 = C_k = (1 + g)bP_2X_1$$

Again, when the Golden Rule holds, and $r = g$, the first of these is a tautology, but the second becomes

$$bX_1/X_2 = tAP_1/P_2$$

or

$$bq = tAp$$

The Golden Rule

Even in a simple model, why should there not be saving out of wages and/or consumption out of profits? There is a clear answer: markets and competition. If workers are saving, their wages are above the cost of their socially defined lifestyle, and can be forced down (by encouraging immigration, to build up a 'reserve army', or by labour-saving innovation). If capitalists are consuming out of profits, there is an opportunity for some among them to invest instead, grow faster, reap economies of scale and take over the markets of the spendthrifts. This will keep consumption out of profits down. Many expenditures which appear to be such consumption, however, should more properly be considered competitively necessary expenses of maintaining 'intangible' capital, such as information networks, goodwill, brand loyalty and the like.

With this in mind, note that since

$$p = b/D \text{ and } q = tA/D'$$

if $D = D'$, that is, if $r = g$,

$$pq = \text{value of capital goods per worker} = tAb/D^2$$

But $dw/dr = -tAb/D^2$. Hence,

$$dw/dr = -pq \tag{16}$$

where the RHS could also be written $-K/N$. That is, the slope of the wage-profit trade-off equals the value of capital per worker when the Golden Rule holds.

This means that Ndw , the increase in payments to labour, exactly equals $-Kdr$, the reduction in payments to capital, or vice versa. The implication for circulation is that a change in distribution causes no net change in financing requirements; any change can be managed by shifting funds from financing capital payments to financing payments to labour or the reverse. Banks can shift the composition of their loan portfolios, between financing investment and financing wage advances; merchants can switch between purchases of capital goods and purchases of consumer goods.

It also follows that the elasticity of a point on the wage-profit trade-off equals the ratio of the share of capital to the share of labour (Nell, 1970).

Look again at the equations for prices and quantities, above. When $r = g$, $D = D'$; hence, dividing prices by quantities

$$p/q = b/tA$$

the *ratio* of prices to quantities is invariant to the distribution. To put it another way, the value of the relative output of capital goods, evaluated at the relative price of consumer goods, is a constant. Re-writing

$$p = (b/tA)q$$

a straight line rising from the origin, with slope b/tA . This implies that $dp/p = dq/q$, or that $(p/q)dq/dp = 1$. The relative quantity of capital goods to consumer goods moves uniformly with the relative price. Cross-multiplying, we obtain

$$tAp = bq$$

the cross-sectoral balancing condition! This is the 'standard' for money. A unit of account based on this relationship will always be valid for the calculations needed to carry out a closed, complete

circulation. It holds for all distributions between wages and profits, so long as the Golden Rule is satisfied.¹ (In the diagram, $W_k = P_c$.)

Given the speed-up, then, as long as all and only profits are invested, whatever the distribution, prices and quantities will stand in a fixed, i.e., invariant, relationship. When the real wage (or rate of profit) changes, the central relationship in monetary circulation is not affected.

The 'real' money wage and the nominal money wage

When money is itself produced—a metallic currency, gold—or convertible into such a currency, then the money wage in the mining and minting industry becomes a *real wage*. For the (fixed) normal output will be gold, input prices will be expressed in gold, and wages will be gold. The output price will be unity. The rate of profits in the gold industry will therefore depend on the prices of its *inputs*, and competition requires that these prices adjust so that all industries have the same rate of profit. But when the money wage in gold rises, the output of gold net of wages—the surplus plus the amount to be traded for inputs—necessarily falls. Hence a rise in money wages in gold requires a general fall in the rate of profits.²

When money is not produced, or is no longer tied to a produced metal, the money wage must be set in terms of *fiat* money, that is, inconvertible paper, or bank deposits. In modern economies, no industry or activity produces a measured output identical to the money-article; so there is no industry in which money prices do not have to be set. As a consequence, the money wage is not automatically translated by competitive pressures into the real wage. In the case of a 'real' money wage, setting the money wage in the labour market determines the real wage; but when the wage is paid in fiat money, the real wage cannot be determined without considering the markets for goods as well.³

It should be clear that in neither case will the Homogeneity Postulate hold. If money in circulation increases or decreases, so that its value in circulation is no longer equal to its production price, the resulting adjustments will have real effects, for resources will move into or out of the mining and minting sector. When more or less bank money is required, on a long-term basis, the size of the banking sector will have to be adjusted, requiring an appropriate rate of interest. In each case monetary requirements affect real relationships.

¹This conclusion can be reached by another route. Multiply the price equations by the quantities and the quantity equation by the prices, and cancel. The result is $(1 + r)tAp = (1 + g)bq$, or since we assume $r = g$, $tAp = bq$. These results also hold in a two-sector model in which the profit rate and growth rate are applied to both sectors (Nell, 1992, ch. 15, pp. 322–3.) The two sectors can be extended to many, as in Nell (1998, ch. 7). This result may have been missed because seekers after an 'invariable' measure of value looked for a 'price × quantity' expression, a commodity whose value was invariant, rather than a ratio.

²In terms of the two-sector model in the Appendix:

$$\begin{aligned} P_1 &= (1 + r)taP_1 + wbP_2 \\ P_2 &= (1 + r)tAP_1 + wBP_2 \\ \$ &= (1 + r)tmP_1 + wnP_2 \end{aligned}$$

Here \$ is the unit output of gold. The prices of the capital and consumer goods are money prices, expressed in gold. There are three equations and four unknowns: P_1, P_2, r and w . The wage is now fixed in terms of gold: $w = \mu \$, 0 < \mu < 1$. The wage is a certain fraction of unit gold output. This leaves three equations and three unknowns. From the equations for the prices we find

$$P_1/P_2 = \mu \$b/[1 - (1 + r)ta] = (1 - \mu \$B)/[1 - (1 + r)tA]$$

Solving for r , we have:

$$r = [\mu \$ (b + B) - 1] / \{ \mu \$ t (aB + Ab) - a \}$$

From the equation for \$

$$P_2 = [\$ - (1 + r)tmP_1] / \mu \$n$$

and substituting into the expressions for P_1/P_2 .

$$\begin{aligned} P_1 &= \$[1 + (1 + r)t(m - a)] \\ P_2 &= [\$ (1 - (1 + r)tA) / \{ (1 - (1 + r)t(m - a))(1 - \mu \$B) \}] \end{aligned}$$

³Financial services have to be priced like any other services. They do not produce a 'money-article'; they handle money and the amount is determined by the market. The money wage in gold became the real wage because the output of mining and minting was measured in gold, and was fixed. When gold is replaced by bank deposits, the equation for \$ must be replaced by an interest rate equation. There is no longer a supply function for \$. The price level is indeterminate. Starting from a point in history, it will be driven by dynamic processes—such as the wage–price spiral.