12 The Evolution of the Banking System and the Theory of Saving, Investment and Interest*

Author's Note

Alain Parguez has upheld the cause of post-Keynesian economics at the Institute of Statistics, Mathematics and Applied Economics (ISMEA) in Paris for many years. In 1985 he invited me to speak at a conference there. Through overwork and general muddle, I only had time to prepare a very unsatisfactory paper which, when I arrived, I tried to beg off delivering. Alain, full of charm and persuasiveness, refused to allow me this easy way out. I managed to get my time of presentation postponed to the following day and sat down, feeling somewhat doom-laden, to listen to the other papers. During a presentation in French, which I barely understand, my mind wandered to liability management, a phenomenon which had been much discussed in the UK but only in a money-and-banking frame of reference. What difference did it make, I wondered, in the larger framework of macroeconomics? Suddenly I thought I had the answer. Building on work I had done for *Macroeconomics After Keynes* (1983) and putting my answer together with the evolution scheme (which I had invented during a graduate student's supervision) I gave, next day, something like the chapter which follows. I could never have adopted such a high-risk strategy without a very supportive host and audience. The chapter is therefore dedicated to Alain Parguez, who, I think, does not know this story.

The paper has two direct descendants (Chick, forthcoming, and Chick and Dow, 1988) and one close relative (Chick, 1988).

INTRODUCTION

The theory of saving and the rate of interest can – or at any rate should – never be independent of the state of development of financial institutions.[1] In Chick (1983, ch. 9) it was argued that the reversal of causality in the saving–investment nexus proposed by Keynes (1936) should not be seen as

*The author wishes to thank Sheila Dow, Basil Moore, David Llewellyn, Otto Steiger, Brian Tew, and members of the seminar at the University of Buckingham for their helpful comments and suggestions. The final product is the author's responsibility.

correct theory in triumph over error but as a change in what constituted correct theory due to the development of the banking system.

The purpose of this chapter is to carry this perspective forward to the present day.[2] The first part of the chapter traces the evolution of the English banking system[3] in stylised form. The parallel developments in non-bank financial intermediaries and markets, often a necessary adjunct to banking development, will on the whole be taken as known. In subsequent sections it will be argued that bank behaviour has evolved to a stage sufficiently different from the stage to which Keynes's theory pertained to require alterations to the received theory of investment, saving and interest. Specific changes are proposed and their implications for macroeconomics are drawn.

STAGES OF BANKING DEVELOPMENT

The banking system has developed continuously; there are no sharp breaks between the 'stages' outlined below. Thus the difficulty for economic theory is always one of judgement: to decide what characterisation captures the salient features of a complex system in continuous change, and when a change in characterisation has significance for theory.

More often than not, perhaps, developments in institutions and their practices are not made explicit in theoretical work. The theorist may be inhibited by the general presumption that economic theory ought to be independent of its time. Those who accept this presumption and yet construct 'relevant' theories must process institutional data intuitively.[4]

Even when the processing is conscious, the ambiguities involved make the task difficult and contribute to the considerable lag of theory behind events.[5] The qualitative nature of the evaluation and the time lapse make it unwise to attempt to date the stages of banking development with any precision.

Having issued those caveats, let us begin with the two stages already referred to in the Introduction: those summarising the 'Keynesian transition' to investment-led growth.

Stage 1 Banks are numerous and small, and geographically semi-isolated. Bank liabilities are not widely used as means of payment. Banks were thus chiefly repositories for savings; transactions balances did not circulate through them to any great extent. These features implied that an expansion of lending would entail a substantial loss of reserves even to the system as a whole. Banks are thus dependent on deposits for reserves and on reserves for lending capacity.

Letting R = reserves, L = loans (or advances), D = deposits, we have the following causal chain:

$$\Delta D \rightarrow \Delta R \rightarrow \Delta L.$$

Deposits in this stage are savings, on the whole. Their supply to the banks may be regarded as exogenous to the banks, stemming from an increased supply of high-powered money not wanted to be held in cash or from a change in the public's cash/deposit preferences as confidence in the banks grew. If the former were the only important source the existence of banks would not much disturb the quantity theory.

While banks remained at this stage of their development they were exactly as the bankers themselves claimed, 'conduits' between saving and the employment of those savings for investment; they functioned in much the same way as direct-lending institutions, such as the new-issue market, where one has to have saved money before one lends it. In such a circumstance, saving determines the volume of investment.

Stage 2 The banking system has demonstrated its viability and won the public's confidence. The number of banks is fewer and the average size of banks is larger. Branching is developed, contributing to risk-spreading, reducing the loss of deposits after loan expansion, and contributing to the convenience of deposits as means of payment. The consolidation of clearing arrangements further encourages the shift to deposits as means of payment.

The means-of-payment function is crucial. It means that deposits now represent not just saving but also transactions balances financing the consumption circuit, moving from one bank to another but not leaving the system as a whole to any significant extent. It is this that ensures that deposits arising from loans are not much lost to the banking system. Individual banks with excess reserves are now emboldened to lend 'money they haven't got', setting off the multiple expansion for the system as a whole which we now take for granted.

The causal chain pertaining to this stage is

$$\Delta R \rightarrow \Delta L \rightarrow \Delta D.$$

The 'bank deposit multiplier' is the relevant theory: the banking system can now lend to a multiple of reserves, subject to a conventional or imposed reserve requirement; deposits are a consequence:

$$\Delta D = (1/r)\,\Delta R.$$

The banks, whether by habit or some other force, are still passive with respect to seeking additional reserve-creating deposits; reserves come to the banking system as before through changes in the public's preferences away from cash towards deposits or, increasingly importantly as time progresses, from the central bank via open-market operations.

It must be emphasised that in characterising reserves as the autonomous variable we do not deny that in Stage 2 banks also make use of well-developed financial markets and lender-of-last-resort facilities to make good reserve needs in the event of miscalculation. This action does not undermine the above characterisation of their behaviour under normal circumstances. The experience paves the way for later stages.

The significance of deposits becoming means of payment goes beyond reducing the loss of deposits created by increased lending. Because deposits are now means of payment they represent all income whether destined for consumption or saving. It becomes appropriate to argue that investment can precede saving, because bank loans, based as much on consumption flows as on savings lodged with them, play a significant role in the finance of investment.

The deposits which result from the use of those loans to construct plant and equipment are in the first case held willy-nilly as receipts of new income (*ex post* saving equals *ex post* investment) and later, when the rise in income is perceived, to a great extent they remain willingly held, in order to circulate the new larger volume and value of goods and services.

Stage 3 Interbank lending mechanisms develop, supplementing the call-money mechanism and contributing to the efficient use of such reserves as are available. The possibilities open to individual banks even more closely approach those of the system as a whole. The causal mechanism of Stage 2 still applies; the 'bank deposit multiplier' acts even more rapidly.

Stage 4 This stage may have been approached simultaneously with Stage 3. The lender-of-last-resort principle is now firmly established in situations far short of the crises in which Bagehot (1873) recommended its use. To put it another way, the central bank has fully accepted responsibility for the stability of the financial system.

With the expectation that this policy will continue, banks are further emboldened: lending may now expand beyond the reserve capacity of the system.[6] The central bank can be virtually counted upon to make good the shortfall in reserves. When reserves are provided 'at a price', that is, at a penalty rate and/or via open-market operations resulting in higher market interest rates, banks' profits will be affected. The threat of this may constrain their lending.

If a policy of stable interest rates is in place, however, reserves virtually disappear as a constraint on bank behaviour. Banks are now able to meet any reasonable rise in the demand for loans. Deposits will rise as a result and the shortfall of reserves is met by the system. This mechanism has been used often enough in Britain for the Bank of England to be referred to as 'lender of first resort'.

Even if reserves are being supplied only at higher interest rates, banks may still expand beyond the reserve capacity of the system, but they will only do it deliberately if they expect the expansion to be profitable. This depends on the relative elasticities of the demand for credit and the supply of reserves. Most commentators have agreed that in the 1970s, when Stage 4 was becoming fully-developed, the demand for loans was quite inelastic and the supply of reserves was very elastic, which approximates the fixed-interest case.

The causal sequence appropriate for this stage is therefore

$$\Delta L^d = \Delta L \rightarrow \Delta D \rightarrow \Delta R,$$

where L^d refers to the loan demand and L is the actual volume of new loans. The theory of deposit creation appropriate to this stage, with exogenous loans driving the system, has been proposed by Coghlan (1978).

In Stage 4 the banking system is working towards the point where reserves are irrelevant except for over-the-counter transactions. As Tobin (1963) pointed out many years ago, it is when the banks are reserve-constrained that they exhibit the properties described by the deposit multiplier. The arrival at Stage 4 required the adoption of new theory of banking based on the marginal profitability of loan expansion given the marginal costs of obtaining additional reserves. The literature of the 1970s reflects this change at least in microeconomic theories of bank behaviour. (For an exposition see Sinclair, 1983, ch. 6.)

On the macroeconomic side, although there have been allusions to the move to a 'pure credit economy' (e.g., Hicks, 1965, ch. 23), the debate centres around two polarised views: that the authorities control the reserve base and hence the money supply – a view last appropriate in Stage 3 at the latest – and that the money supply is completely elastic, e.g., Moore (1979). If one has to choose between extremes the latter is preferable. However, if the argument is correct that the behaviour of an individual bank has converged on the description appropriate to the system as a whole (a proposition more tenable in the UK and Europe than in the USA perhaps), the results of microeconomic banking theory can be generalised for macroeconomic purposes.

Stage 5 Banks develop 'liability management'. Until this stage all adjustment, regardless of causal structure, took place on the asset side of banks' balance sheets. Deposits, after Stage 1, were largely the passive consequence of bank lending policy subject to the availability of reserves. In Stage 5, banks have actively sought to attract or retain as deposits savings which might have been held in other financial institutions or as government or corporate financial instruments.

Liability management developed as part of the general increase in financial competition officially encouraged under the 1971 policy of 'Competition and Credit Control', which extended reserve requirements to a wider range of bank-type institutions. The banks were no longer penalised by controls levied only on them, but in return they agreed to abandon their cartelised lending and deposit rates. Banks were thrown into open competition amongst themselves and with other financial institutions. They responded with a policy of aggressive expansion, arguably (though this cannot be proved) expanding beyond the point of profit maximisation, as in a 'sales-maximisation' model. This entailed at least from time to time actively seeking lending outlets rather than merely filling all reasonable loan requests, and 'funding' the asset side by offering higher deposit rates to attract deposits. (See Llewellyn, 1985.)

Insofar as bank deposits came to be held in preference to cash or gilts, the shift to deposits will provide reserves as well as 'funding' the asset expansion. The cost is a higher average interest on deposits.

In 1981 the formal reserve requirement was abolished, so one might be tempted to think that cash and bankers' deposits were no longer important. Banks must still be ready to honour claims on cash, however, and the Bank of England has made it clear that it expects banks to consult about 'appropriate' levels of bankers' deposits. This understanding is presumably designed to avoid a sudden expansion on the basis of existing reserves. Over the longer term reserves may be reduced, relatively to deposits.

The causal nexus for this stage of aggressive expansion is:

$$\Delta L^s = \Delta L \rightarrow \Delta D \rightarrow \Delta R.$$

Stage 5 differs from Stage 4 by the absence of passivity regarding any part of the banks' balance sheets.

One interpretation (Pigato, 1985) is that this stage is merely a transition to some bounded competitive equilibrium. Alternatively, this stage represents a sales-maximisation strategy in a struggle for a dominant market share. This expansion may also be bounded but if the banks are 'winners' in the struggle, the boundary is 'further out' – a larger equilibrium balance

sheet is implied. Still a third interpretation, the one I incline to, accepts the possibility that a well-defined and stable competitive solution may not exist. Financial markets are different from the markets for goods; the demand for loans is finite only in the absence of 'moral hazard', and there is no guarantee of the absence of hazards, especially when dealing with incorporated businesses and governments, with no known finite life. The supply of loans is also a problematic concept, as it depends on the lender's assessment of bad debt risk and this estimate, no matter how carefully made, is subject to many different sources of error (including the estimate of moral hazard). The banks' excursion into sovereign lending in the third world illustrates pretty clearly that the level of bad debt risk acceptable to the banks rose considerably in the 1970s.[7] The rise and fall of the secondary banks on the back of a speculative bubble in property is another case in point.

What these episodes point to is the impossibility of defining the concepts of supply and demand for loans in the absence of a firmly-held exogenous constraint such as rationed reserves. If supply and demand are not well-defined a competitive solution is problematic.

A further implication of the ready availability of the lender-of-last-resort facility (even for some Euromarket loans) is that liquidity is taken for granted. Note that the discussion of bank regulation has shifted from questions of liquidity (and control via liquidity constraints or pressures) to the question of capital adequacy. As the sovereign debt crisis illustrated so well, solvency, not liquidity, is now the issue.

IMPLICATIONS FOR SAVING AND INVESTMENT

Saving and investment represent potential disturbances to steady-state equilibrium: that is the source of their interest to macroeconomists. In largely non-financial economies or in financial economies dominated by *de facto* direct borrowing and lending (including Stage 1 of banking development), saving had to occur prior to investment; investment would be thwarted by a lack of saving. With the arrival of Stage 2 banking, investment could precede saving; the matching saving in the first instance is the new bank deposits resulting from loan expansion. Subsequent banking developments have not changed that process; they have intensified it.

There is a problem of how to describe that initial saving, for though it is not deliberate, it is quite voluntary. A traditional Keynesian would say that consumption had not adjusted to the new level of income and the discrepancy was held as bank deposits. Milton Friedman would probably say there

was an excess supply of money; James Tobin would appeal to portfolio disequilibrium. The new deposits also correspond to Hicks's (1939) 'net acquisition of cash by trading', since the deposits represent payment for whatever it was that the loans were used to buy. There are sympathies, too, with the 'buffer stock' approach.

However one describes these first-round holdings, two features are clear. First, though the deposits are willingly held there is no actual decision to save. The deposits represent a passive (and grateful) acceptance of means of payment by workers and traders. Some of it will doubtless be used for consumption, some of it saved, whether as deposits or in some other form. If used to purchase securities, say, the deposits still remain with the banks, title to them being transferred from the buyers of securities to the sellers. If used for consumption, it still remains with the banks.

Second, while individuals quite happily accepted claims on deposits – acceptability after all is the hallmark of the means of payment – the point on which I wish to insist is that no one actually asks those who subsequently have larger deposits whether the expansion of bank balance sheets, including deposit liabilities, was alright with them. In this macroeconomic sense, though not to the individual, the new saving is involuntary (Keynes, 1936, pp. 81–5). From Stage 2 onwards, 'savers' have no influence over the volume of banking business or the volume of deposits. *Borrowers* may repay bank loans out the newly-generated income, but that is the only offset.[8] For the most part, by the time Stage 2 was reached, most income flows circulated through the banks and did not leave.

Nothing in subsequent stages of banking development challenges the causal priority of investment over saving, The difference made by liability management is largely a matter of timing: the 'accidental' deposit may be persuaded back to the banks with greater speed, or even persuaded to stay. The result is an enhanced lending capacity for the banks.

The absence of any meaningful sanction on bank expansion through the cash–deposit ratio or the reserve ratio means, as remarked before, that we have approached the 'pure credit economy', where money generated by the banking system never leaves the banks thereafter. ('Approached', not 'reached', because portfolio shifts away from deposits to cash or government securities may still affect banks' lending capacity.) Yet received macroeconomic theory works in terms of an exogeneous money supply on the assumption that the authorities can and do exercise sufficient control to dominate banking policy. That assumption may have been relevant in Stage 1 and (more doubtfully) Stage 2. It clearly will not do now: the theory of money supply must be a theory of banking policy with the authorities acting, if they act at all, through interest rates.[9]

While banking developments leave the theory of saving and the priority of investment unaffected, a difference with received theory does arise on the investment side. In Keynesian theory, investment evoked saving through income-creation: the initiative came from entrepreneurs and depended on their expectation of long-run profit potential. If their expectations were correct, the investments undertaken would be sufficiently profitable eventually to pay back any lending which supported the undertakings: that is the essence of investing up to the point where the (expected) marginal efficiency of investment equalled the rate of interest.

Despite Keynes's emphasis on inaccuracies and instabilities of expectations, this story has a comforting aspect: the potential to repay depended on real production and the long-term benefit was increased productive capacity. But not all loans had this effect even in Keynes's time: there have always been consumer loans, which do not add directly to productive potential (though they may stimulate investment subsequently) and loans for financial and property speculation.[10]

In Stage 5 the balance of the initiative to increase lending swung to the banks, at a time (the 1970s) when investment was not buoyant. No one ever knows exactly what bank loans are spent on, but we do know of lending which, if it has a base in tangible production at all, has it at several removes. The ideology (and practice?) of self-liquidating loans has been left far behind.

It is currently fashionable to find good in any removal of restrictions and expansion of competition, but there is no assurance of the social value, or even longer-term economic value, of some of the activities supported by Stage 4 and Stage 5 banking. The proposition that investment evokes the necessary saving feels hopeful and progressive. It is far less attractive to say that speculation in City property (as in 1972) evokes the necessary saving to finance it – but if one counts loan-generated deposits as saving, it did, until the bubble burst.

LIQUIDITY PREFERENCE

The discussion has begun to impinge on the speculative demand for money, in received theory a separate province from the saving–investment nexus. It is perhaps in the area of liquidity preference that the effects of the move to Stage 5 can be seen most clearly.

The received theory of liquidity preference is confused by the transmutation of Keynes's speculative demand theory into the almost antithetical asset demand theory based on static portfolio choice.[11] While initially the

asset demand theory arose out of a misunderstanding of Keynes's theory, ironically the competitive posture of the banks may have given it a relevance it did not have at the time of its development.

Keynes's speculative demand theory explains the holding of 'money' (short assets, including bank deposits) on speculative account as a flight from long-term assets when rates on those assets fall so low that the holder believes the risk of capital loss outweighs the superior returns on long assets. In aggregate, at very low interest rates even those who do not normally speculate would join the 'bears'. This theory presupposes the normality of a rising yield curve based on 'normal backwardation'; indeed in the 1930s the curve rose quite steeply, as the banks offered very low rates indeed on deposits. As 'Stage 2' banks they accepted deposits passively and did not compete for them. Thus deposits would be acceptable to security-holders only when they were seeking a capital-safe refuge from potential capital losses on their securities and only for such time as elapsed between forming the expectation of loss and the actual fall in security prices, after which event the funds would be reinvested.

The influx of funds to the banks would occur when the average level of rates was low – a period of pessimistic expectations with little intended investment. So the inflow of funds to the banks did not provide a basis for credit expansion and did not contribute to the financing of investment directly or indirectly. The loan rate was kept high by the flight from long securities and discouraged the desire to invest. There is a bias toward deflation in this theory.

The breaking of the bank cartel in Stage 3 and *a fortiori* the development of liability management have changed all that. Banks have shown their ability and willingness to compete for deposits across a wide range of interest rates. Thus at least from time to time their interest-bearing liabilities are a serious alternative to securities for long-term investment. Deposits have become attractive not just as a vehicle for avoiding capital losses, but for their yield. There now *is* a legitimate reason to hold money as an asset along portfolio-theoretic lines of risk and return rather than merely as a short-term haven from capital losses.

When deposits are competitive alternatives to securities, at the very least they lower the 'loss coefficient' attached to lending by encouraging a return or preventing the departure of loan-induced deposits. When they supplant government securities, they produce reserves as well as settling the balance sheet. The result is greatly enhanced lending power for the banks.

Keynes's liquidity preference theory was asymmetrical in its effect: holding securities as long-term investments, and even on speculative account, was the norm, for rates on deposit accounts and other short-term

assets were simply too low for those assets to be attractive except when the threat of capital losses was paramount. Then, bear speculation could keep interest rates from falling just when more investment was needed to restore employment. If the present argument is correct, this asymmetry should be replaced by a symmetrical concern for the inflationary potential of a banking system on which there appear to be few constraints.

There are symmetries on both the asset and liability sides of banks' balance sheets: liability management can be used to draw funds in boom or slump, and the new aggressive attitude to lending may also manifest itself at any phase of the cycle. In Keynes's analysis the demand for loans dries up in a slump. The banks have shown that if they want to lend, they will find outlets even in a slump. The question, mooted earlier, is what they will lend *for*. Too much new credit in too short a time is bound to be inflationary. So is lending for 'unproductive' expenditure.

The banks' aggressive lending activity may contribute to inflation. It should also be noted that inflation may contribute to banks' lending capacity. Backwardation ceases to be normal in inflation, the more so as nominal rates rise to reflect inflation.[12] In the late 1970s borrowers fled the long-term debt market, lowering rates of return there and encouraging bank lending. Banks responded by competing for deposits against long-term securities in lenders' portfolios – and once bank balance sheets have expanded the banks must continue to compete, whether the general level of rates is 'high' or 'low' (however one assesses that in inflation).

Finally, it should be remarked that the main area of applicability of the speculative-demand model in these days of floating exchange rates is undoubtedly foreign exchange. The Keynesian conclusion that the root of unemployment was a rate of interest which would not fall sufficiently could be replaced by the assertion, easy enough to back up theoretically and well-known from recent experience, that speculation may cause both overvaluation of the exchange rate with a resultant loss of competitiveness and an undervaluation which exacerbates inflation. And the remarkable freedom of the Stage 5 banking system to expand in almost any circumstance is an important element in the explanation of why inflation may then persist even in periods of high unemployment.

CONCLUSION

This chapter argues for a change in theoretical outlook on two points following from the development of 'sales-maximisation' strategies and liability management on the part of the banks. First, while the changes in

bank behaviour outlined above have not altered, but rather intensified, the subordinate position of saving, the autonomous expenditure which dominates saving may comprise less investment than formerly. This increases the probability that an expansion of finance and money may not be supported by any 'real' power to pay these loans back. This is a source of inflationary pressure not included in standard macrotheory.

Second, the deflationary bias predicted in 'standard' macrotheory, arising from a preference for liquidity at low rates of interest, has been altered by liability management. Banks have learned to make their liabilities attractive not only in the slump but also in better times, at any general level of interest rates. This has greatly enhanced their lending power in all phases of economic activity, giving a consistent inflationary bias to the economic system. There are, undoubtedly, 'real' inflationary forces too, but it would seem misguided to ignore the existence of financial causes as well.

Notes

1. See Hicks (1967, ch. 9) on the significance for theory of the shift from metallic to credit money.
2. The need to do this was stated in Chick (1984). At the time I could not see a way forward, and this paper is itself only preliminary.
3. Such a narrow perspective deserves apology. (I am cautioned not even to say 'British banking system', as the Scottish banks followed a slightly different path.) My reason is insufficient knowledge of other systems and my excuse is a hunch that the history of banking follows broadly universal patterns, with albeit important variations.
4. The only thing 'wrong' with intuition is the difficulty of sharing it. On the role of intuition see Dow (1984).
5. Witness Irving Fisher's (1930) and Keynes's (1930) reluctance to include bank deposits in the definition of money; and Schumpeter (1959) remarked the tardiness of the first appearance of the bank credit multiplier theory of deposits.
6. This potentiality is not new. It was pointed out by Lord Overstone in 1840 (Gregory 1929; 1964, p. 50). In Stage 4 the potentiality became significant in actual behaviour.
7. The counterargument that the banks thought that these debts were safe because they expected lender-of-last-resort facilities to underwrite them makes the point stronger rather than destroying it; it is not the loan itself which the banks thought was safe, or any 'real' collateral, but the indirect financial backing.
8. Kaldor and Trevithick (1981) argue that a credit-based money supply can never be larger than people want to hold because an excess supply will result in repayment. However, the money might not fall into the hands of those with overdrafts in the first instance. The money may be spent and push up prices,

after which the money may be willingly held; or the expenditure could increase profits and the profits used either to finance expectations or to repay bank debt. The latter is only one possibility.

9. In Wicksell's (1936) theory, the pure credit economy exhibits a cumulative process of expansion (contraction) when the 'money rate of interest' (banks' lending rate) falls below (rises above) the 'natural rate of interest' – the real rate of return on productive capital. It has long been acknowledged that this 'natural rate' is, to put it generously, a difficult concept to pin down; both capital and the output it helps to produce are impossible to measure in aggregate except in money terms, and the meaning of such measures is especially problematic in times of inflation.
10. Stock market loans with very low margins were prevalent in the United States before the Crash of 1929; this lending was subsequently regulated.
11. For the differences between the two theories see Chick, 1983, pp. 202–8 and 213–17.
12. There is often a substantial lag in the adjustment of nominal rates to changes – upwards and downwards – in the rate of inflation.