The credit theory of money: 
the monetary circuit approach

Alain Parguez and Mario Seccareccia

Introduction

The starting proposition of what is commonly described as the theory of the monetary circuit (TMC) is that, in a monetary economy in which buyers and sellers engage in economic transactions, ‘money’ is the by-product of a balance sheet operation of a third agent who, in modern parlance, can be dubbed a ‘bank’. In particular, money always emerges as a debt (or liability) issued by this third agent on itself, which has as counterpart a credit simultaneously granted to buyers of goods and services within an economy. In this three-way balance sheet relation, every transaction entails the simultaneous creation or destruction of debt, and every seller of goods and services accepts payment of this bank liability on the basis of its general purchasing power or value. The value of this bank liability (or money), however, is not the consequence of some intrinsic characteristic, be it utility or liquidity. It stems, rather, from the certainty that accepting bank debt as payment is to acquire a right on the existing as well as future output that will be created by the agents who have been granted bank credit. Furthermore, these debts would not be legal titles to acquiring present and future real wealth were it not for the direct or indirect role played by the state in endorsing them.

In this regard, the TMC sheds light, among other things, on the historical origin of money. In accordance with Innes (1913) and Heinsohn and Steiger (1984), there can be private debt and credit contracts before the existence of money. This is because money emerges both causally and historically as a result of prior debt and credit relations. Money appears when a community (usually through the legal apparatus of the state) bestows the characteristic of being a legal title to a share of present and future wealth on debts issued by a specific agent. Money has, therefore, a pure ‘extrinsic’ value (Wray, 1998) that is generated by money’s role in the industrial circulation.¹ In the ‘efflux’ phase (to use Tooke’s original expression) of monetary circulation, debts are issued to allow private firms (as well as the state) to start the production process via the credits granted to them by the issuing banks. These debts are then extinguished or cancelled when firms (and the state)
reimburse the creditor banks by acquiring enough of the bank debt in circulation. In the case of firms, this occurs through the sale of commodities in the product market and/or securities in the financial market, and, in the case of the state, taxes and/or government securities. This is what could be described as the 'reflux' phase of the monetary circuit. Credit money, as a rule, is thus created only to be destroyed in the circulatory process and not to be held.

From this general conception of money, we can deduce the two major propositions of the TMC. First, money is, and has always been, a debt created ex nihilo by bank credit advances that are granted either to permit the generation of real wealth or to acquire existing physical assets. Second, there is no alternative between debt financing on the one hand, and the tapping of existing liquid resources (or accumulated savings) to 'finance' expenditures, on the other. At the macroeconomic level, spending in a monetary economy is always and everywhere in the nature of debt financing.

This chapter is divided into two major sections. We begin with a detailed exposition of the theory of money underlying the TMC. In this first section, we shall try to explain why many of the debates pertaining to this approach are the consequence of a fundamental misunderstanding of the dynamic process of money creation. A second section discusses why the TMC is a major advance relative to other existing approaches to money.

The nature of money in the monetary circuit

The simple model of the monetary circuit

As has been emphasized in our introduction, the economic process has always been supported by a set of debt contracts in all organized societies, including the despotic ultra-centralized systems of ancient Egypt and the former Soviet-style regimes of Eastern Europe (Wittfogel 1959). Historical records clearly suggest that, regardless of the type of pre-existing property relations in ancient societies, there can be debt contracts while currency does not yet exist. For instance, Babylonian bills of exchange, regulated by the Code of Hammurabi, and book entries, such as the wheat deposits in ancient Egypt, were all in the nature of credit-debt relations even before such societies had developed coins or other well-defined circulating media of exchange (Einzig 1966: 328).

Interestingly, and in conflict with the views of Heinsohn and Steiger (1984, 1994), even in societies in which communal property was the norm, forms of credit/debt relations existed as long as individuals held informal personal possessions rather than titles based on codified private property rights. To substantiate this, we shall take the example of the aboriginal people living in the Hudson's Bay area of present-day Canada during the pre- and early post-contact era. Given the rugged climate of the region, it is well known that the original nomadic people living around the Hudson's
Bay had developed primitive debt/credit relations based on gift exchange prior to European contact (Ray 1996: 88, Davies 1994: 11–12). When the Europeans did attempt to establish commercial fur trading relations with them, these people of the sub-Arctic fur trading area had no difficulty in establishing formal credit/debt relations vis-à-vis the European traders. The Amerindians had no concept of European currency, but they understood credit/debt relations perfectly well. Indeed, a cashless credit economy evolved with respect to the European fur trading enterprises, such as the Hudson’s Bay Company. As in the modern TMC, the fur trading companies were to advance credit in the form of European traded goods and, through formal double-entry book-keeping based on an abstract unit of account called the Made Beaver, aboriginal trappers would then seek to extinguish their debts by harvesting beaver and other highly valued fur bearing animal pelts (Ray and Freeman 1978: ch. 9). These credit/debt relations emerging in semi-nomadic societies, such as those of the Indians of the sub-Arctic region of Canada (where land remained collective property, even after European contact), would suggest that private property rights are not a necessary prerequisite to the appearance of such relations in primitive economies. We agree with Heinsohn and Steiger (1984), however, that private credit/debt relations predate more advanced monetary systems based on state coins or private bank liabilities as circulating medium.

Indeed, in more modern monetary capitalist economies with entrenched private property rights, money appears when there exists a set of agents, which we shall call ‘banks’ (including the central bank), whose debts are accepted by all other agents in an economy as a means of payment to settle their own debt commitments. Banks are deemed to be so creditworthy that no holder of their debts would ever ask for reimbursement either in kind or in the debt of another agent. Banking institutions enjoy, therefore, the capacity of freely issuing debt without it being subject to an exogenous resource constraint. This means that banks can create these debts ex nihilo when they grant credit to non-bank agents who must spend them to acquire real resources. Bank credit entails the advancing of loans of newly-created bank debt to economic agents who cannot depend on some pre-existing stock of bank liability, which itself would be the result of previously-incurred outstanding bank credit in the economy.

Once credit has been advanced, however, non-bank agents become committed to paying back their loans at some future date by collecting the required quantity of bank debt out of their cash receipts, either from the sale of newly-produced commodities, or from the sale of titles to existing wealth. When initial borrowers reimburse their loans, there is an instantaneous cancellation of both the individual agent’s debt towards the banks and the debts that the latter had issued on themselves to finance the loans. In the TMC, these conventional loans are what is regarded as credit money. This money is endogenously created ex nihilo when banks grant credit and it is extinguished or cancelled when the outstanding credit is reimbursed.
This dynamic flux and reflux process is the essence of the monetary circuit. Its different phases are directly mirrored in banks’ balance sheets. In the initial phase when banks grant credit, they issue new debts upon themselves which they lend to non-bank agents. Since these debts are money, the latter appears as an increase in banks’ liabilities which is equivalent to the newly-held deposits of the non-bank public. Logically accounted for as an equivalent increase in bank assets, the counterpart of this newly-created money is the forward debt of non-bank agents in the form of loans to be paid back at some definite date in the future.

The second phase of the monetary circuit is the period during which non-bank agents spend the money that they have borrowed to acquire real resources, which are generally labour and produced commodities. Sellers of labour services or commodities acquire the quantity of money which was created in the first phase. In the balance sheets of banks, there appears a mere transfer of deposits or liabilities from one group of individual holders to another: the sellers of commodities and labour services. These new holders of money owe nothing to banks. On the contrary, one may infer that it is the banks that owe something to these holders of money. Some Post Keynesian writers (e.g. Moore 1988) have interpreted this sellers’ holding of bank debts as ‘convenience lending’ to banks. This interpretation has also been more recently advocated by heterodox French economists such as Gnos (1998). Sellers of labour services and produced commodities are indeed the new holders of bank liabilities but, we believe, they cannot also be considered as bank creditors. This is because money is merely debt that the banks have issued on themselves ex nihilo. The amount of this credit money that was initially advanced was not based on what this subsequent group of convenience ‘lenders’ wished to hold but on the amount that the initial borrowers wished to spend.

In the third and last stage of the monetary circuit, the initial holders of bank debts seek to recover them in the reflux process out of their receipts generated by their initial expenditures. They can now replenish their deposits and pay back their loans. At the same time banks can recoup the debts that they had issued on themselves to finance those loans and thereby extinguish their own implicit debts. The counterpart of this process is the cancellation of the initial borrowers’ forward debt to banks. In the balance-sheet operation of the banks, this third stage of the monetary circuit is accounted for as an equivalent reduction in bank assets and liabilities.

The definition of money

From this simple model of circulation of bank debts, we can draw a first definition of money. Money cannot be, and never has been, a commodity whose value stems from its scarcity. Some circuitist writers have deemed it to be a pure token (Graziani 1998). While not disagreeing with this, one
may more correctly argue that it is an abstract or virtual token, since it is merely a debt that banks have issued on themselves. The value of this money (or its purchasing power) cannot be inferred on the basis of some neoclassical scarcity principle, since the quantity of new debts issued by banks is constrained only by the expressed demand for money of non-bank agents. The latter need this money in order to spend it on the acquisition of real resources. In the productive sphere, money exists therefore as the means of payment that gives rise to a sequential chain of transactions leading to the creation of new wealth. Money cannot be defined, as is traditionally done by economists, in terms of its presumed functions of unit of account or store of value. A monetary economy does require a unit of account but, in such an economy, accounts are settled in the prevailing monetary unit. The debts banks issue on themselves are denominated in the unit of account because it is the means of acquisition that ensures a viable circulation process upon which the creation of real wealth is based.

As previously mentioned, economies have existed without the use of money, such as ancient Egypt, or in which money played an insignificant role, but which had a state-imposed unit of account, such as the former Soviet Union. This would suggest that money's existence cannot be understood on the basis of this function. In much the same way, since we have argued that money is created ex nihilo to finance loans, and is destroyed only when these loans are repaid, its existence cannot depend on its store of value function. Such a function is inconsistent with the notion of money. Money exists and has a value only as long as it is spent by non-bank agents for the purpose of creating future wealth. The hoarding of money in the form of bank deposits merely obstructs the process of wealth creation upon which the value of money depends. Indeed, it is this inconsistency which explains why the traditional demand for money as a component of wealth has a potentially destabilizing role within the monetary circuit.

Money is credit-driven

This definition of money requires that there are agents capable of issuing debts upon themselves that are generally accepted by all other agents as a means of payment. Certain conditions must be fulfilled for the emergence of these banking agents in an economy. First, the state must either implicitly or explicitly come to endorse fully bank debts. This endorsement has two major consequences.

Amid the class of debtors, banks are now deemed as so creditworthy that no one would consider asking them to reimburse their debts either in commodities or in debts of other agents. Banks are therefore able to issue liabilities at will, which are implicit debts on themselves. Because of the state endorsement of these debts as the ultimate guarantor of their liquidity, it would be wrong to conceive holders of bank liability as bank
creditors. In a sense, given its endorsement by the state, one can legitimately argue that money is always fiat money, even when the state is not issuing it and the role of the state is marginal. The state’s power to create money at will is the logical consequence of its role in the endorsement of bank activity. Having the legal authority to bestow on bank debt the characteristic feature of money, the state has also the power to issue debt on itself that will be money, freely convertible into bank liability. Regardless of the material support of these conventional debts in the form of gold, silver, copper coins or paper notes such as bank money, these coins or paper notes have an essentially extrinsic value because each holder is certain to possess a legal claim on real resources, and because it can serve ultimately as a means of settling one’s debts. The very notion of a commodity money is an illusion which confuses the material support of money with money itself.

The historical process has been to shed the commodity guise of money and to integrate further state money within the activities of commercial banking, with the last phase of this integration being the creation of the state’s own banking department: the central bank. In this latter case, the state’s endorsement of the monetary activity of banks is now operated through the central bank which, as the ultimate purveyor of liquidity, further empowers commercial banks to lend their own debt to credit-worthy borrowers without constraint. At the same time, the state is now entitled to finance its desired expenditures by credits granted by the central bank. In the accounts of the central bank, money is now created as debts issued by the latter on itself and appears as an increase in liabilities. This money is advanced to the state and accounted for as an equivalent increase in the central bank’s assets. However, a small share of these central bank debts may still necessitate some material support in the form of legal tender banknotes, and these would also be accounted as central bank liabilities.

It follows, therefore, that money, whatever its material support, is always a debt created ex nihilo to finance loans. We believe that this analysis is consistent with the historical record. As was mentioned at the beginning of this section, credit money and banks pre-date pure state money which was directly issued by the state authority as coins and later notes and central bank liabilities (Innes 1913, Heinsohn and Steiger 1984, 1994). In many societies, such as Mesopotamia as early as the seventh century BC, and Greece after Solon’s Reform in the fifth century BC (Davies 1994), the state’s monetary requirements were financed by private bank credits. However, regardless of whether these requirements are financed by private banks or a central bank, loans can never be financed by some pre-existing deposits. Moreover, since credit money is created to finance loans to private agents and/or the state, it is by its very nature an endogenous variable. The neoclassical scarcity principle can never be applied to the creation of money.
We have argued that credit money could not exist without the state, and that all credit-driven money is by its very nature a fiat money, irrespective of whether it takes the form of a commercial bank or central bank liability. On the other hand, state power alone cannot guarantee the existence and survival of a viable monetary system. History records many examples of the complete collapse of monetary systems, such as in Russia before the First World War, immediately after the Revolution, and in the late 1990s, as well as in Germany after both World Wars. In each of these historical episodes, the state failed in its effort to maintain the stability of the existing monetary systems and was faced with an untenable situation of hyper-inflation. But how can one explain it? Mainstream economists point to the excessive creation of money, but one must also explain it in reference to what it is excessive.

Contrary to the traditional quantity theory based on the excessive creation arising from the careless use of the government's printing press, our analysis points to a slightly different explanation and suggests why another condition must be met for the existence of money. Rules of creditworthiness must ensure that banks issue debts on themselves to finance loans that ultimately lead to a creation of new real wealth. If banks were to issue liabilities to finance loans that could never support the generation of future real wealth, these debts would be deprived of value and no one would accept them. We may, therefore, put forth the following more general definition of money:

Money is at all times the liabilities issued by banking institutions which have been endorsed by the state primarily for the purpose of financing the formation of future real wealth. This money has a real extrinsic value because every holder of these liabilities has acquired a claim on the future physical wealth that results from the initial bank credit advances.

The role of firms

Such a definition explains why business enterprises and the state have special access to bank credit. Firms must borrow credit money to finance their desired acquisitions of real resources needed to carry out their production plans. They have first to spend money to acquire the labour services required for the production of the planned output of consumption and capital goods out of their existing capital stock. It is firms' spending (including that of public sector enterprises) on labour services which therefore determines labour income. According to our general principle of circulation, firms (and banks) no longer owe anything to wage earners as soon as income is paid.

Firms have also to spend in order to maintain or increase their stock of capital goods by acquiring the newly-available equipment goods produced by the capital goods sector. Abstracting from any specific assumptions
regarding the degree of integration of business enterprises, we see no logical reason why investment, the acquisition of the newly-produced capital goods, should not be financed by bank loans. Some heterodox economists (see Bailly 1992) have argued that, if that were so, one would be espousing some new version of the old loanable funds theory. They assume that if firms also borrow money to finance investment, it would mean that they are borrowing pre-existing deposits generated by the payment of wages. However, this latter interpretation succumbs to the general criticism that we made previously against the mainstream theory of money. Regardless of whether it goes towards the financing of the wage bill or the purchase of capital goods, the TMC suggests that when banks grant loans they issue new debts that generate new deposits ex nihilo. They cannot, as a group, lend pre-existing deposits. Moreover, to the extent that the purpose of these loans is the creation of real wealth, money emerges to generate new production either for direct consumption (consumer goods proper) or for indirect consumption, by replenishing or increasing the stock of real capital, including inventories.

Consumption goods are usually acquired when income earners spend their income which is initially financed by money creation. This credit money created to finance income should permit the generation of the new output and the realization of the value of a share pertaining to articles of direct consumption. Conversely, the value of the share devoted to future consumption or investment must also be realized by loans entailing creation of money that would finance the acquisition of newly-produced equipment goods (Seccareccia 1996, 1998).

As has been shown in previous articles on the TMC (Parguez 1996, Seccareccia 1996), such an extensive conception of the role of credit advances is the necessary requirement for the validity of money. All debts issued by banks are claims on real wealth. If credit were restricted to the financing of wages, the value of equipment goods would not be realized in accordance with the initial expectations of firms (and banks). The newly-created money would lead to the creation and monetary validation only of articles of current consumption, and not of the capital goods that would also be needed to sustain production and consumption in the long term.

As argued above, banks must apply a set of financial criteria to their borrowers that allow the former to assess the creditworthiness of the borrowing firms so as to measure the latter's ability to generate real wealth, as well as to establish norms for banks themselves, in their capacity to lend their debts to the requisite borrowers. The creditworthiness of firms is usually measured by the profits that they individually target. On the other hand, the creditworthiness of the banks themselves (which sustains the conventional nature of the debts that they issue on themselves) is positively related to their own bank-specific capital values as determined by the monetary value of their equity that forms the collateral of their loans (Parguez 1998). In the case of firms, this ability to generate profits out of
current expenditure is measured by the targeted ratio of profits to income expenditure, which has been described elsewhere as the monetary mark-up (Parguez 1998). The more banks want to ensure that the commitments of their borrowers will generate wealth, the greater would be the monetary mark-up that they impose on business enterprises. Firms’ capacity to earn the targeted profit (via their mark-up policy) is thus the prerequisite for continued access to bank credit, and is a critical factor in establishing the viability of the overall credit money system.

Some circuitist writers have raised doubts about the ability of firms to earn profits within the framework of the monetary circuit (Renaud 1998). Others have postulated that only producers in the consumer goods sector could extract monetary profits (Gnos 1998, Vallaegas 1998). In our opinion, both interpretations seem to thrive on a misunderstanding of the circulation process. As shown above, profits are not the pecuniary difference between the reflux and the initial injection of money. Rather, they are generated by the excess of receipts from the sales of commodities over the initial income expenditure paid to the workers. While the nature of circulation establishes an ultimate equality between the efflux and the reflux of money, this is not contradicted by the fact that the initial flux of money is always superior to the income expenditure by firms.

Firms in the capital goods sector must also be able to earn profits from the sale of their output. These money profits are equal to the difference between the selling price of the new equipment goods and their income costs, with the market value of these new capital goods reflecting long-term expectations of profits by firms wishing to acquire these commodities.

This is merely a restatement of the theory of profits in the investment goods sector put forth by Keynes in the Treatise on Money. Firms wishing to acquire more capital goods in order to meet a forecasted future increase in consumption borrow an amount of money equivalent to the value of the additional capital goods by placing their orders for new equipment from the capital goods sector. In this process, the sellers of both sectors realize their money profits and the monetary circuit is now completed with the flux matching the reflux of credit money. All profits are now used to repay the loans which had been the prior source of investment finance and, at the same time, the credit money originally created is extinguished. It is, in part, for this reason that we reject the view put forth by Nell (1998) who has argued that, in order to minimize transaction and borrowing costs, and thereby the number of transactions in a production economy, the optimal quantity of credit money ought to be just equal to the wages paid in the equipment goods sector. Nell’s view, which is based on the principle of cost minimization as well as on a peculiar notion of money, would perhaps also succumb to further criticism (which we discuss in a separate section) of the neoclassical conception of money.

Since loans are granted to firms by commercial banks chartered by the state, their own creditworthiness will ultimately depend on their ability to
earn profits out of their credit-creating activity. In the long run, the existence of money requires that bank profits generate enough equity value to match the value of assets sustained by banks' own desired monetary mark-up. Consequently, we can spell out this stability condition for the monetary system:

The rate of interest charged on credit advances, which is the primary source of bank profits, must be high enough to support an amount of bank equity equal to that generated by bank assets resulting from the monetary mark-up imposed on firms.

From this, there ensues the concept of exogeneity of the rate of interest. By its nature, the rate is imposed by banks on firms and it can be included as part of the creditworthiness rules to which firms have to comply. There has been some discussion in the heterodox literature on the capacity of firms to pay interest to banks (Léonard 1987). With the existence of interest, firms would not be able to meet their obligations to the banks since the reflux would now exceed the initial efflux. This, however, arises because of a misunderstanding of the requirements of monetary circulation.

The problem is not the incapacity of firms to pay interest but the ability of banks to realize profits consistent with an equality between the flux and the reflux of credit money. We have argued that banks issue debts on themselves, and these ought to allow firms to pay all their production costs: wages, the share of profits advanced to rentiers and interest payments. When banks credit firms with an amount equal to the interest requirements, they increase both their liabilities and their assets. Firms have to pay back the debts issued to finance interest expenditure by running down their deposits to pay interest to banks. In the process, bank liabilities decrease while bank assets increase by an equivalent amount, thereby generating an increase in banks' net worth. Banks recycle a portion of the increased assets to pay wages to their employees and dividends to their shareholders. These expenditures further build up demand for the commodities and services of the non-bank sector. However, while banks may even acquire assets of business enterprises, usually by default, their ability to create money for their own purposes is strictly bounded by the expenditure which is required to sustain their own credit-creating activity.

The role of the state

The nature of monetary circulation also affects the role of the state as producer of public goods in modern economies. To provide society with its desired amount of public goods, the state has to spend money to acquire both a share of the labour force and a share of the available supply of capital goods. However, the state cannot spend the proceeds of taxes to obtain such goods, because taxes cannot be raised unless there is already a
pre-existing money in circulation. Logically speaking, taxes could only be levied in the future as the final outcome of the initial expenditures of firms and the state, when all private agents will have received their gross income and/or spent it. Hence, taxes could be levied on gross employment income (primarily wages and salaries) paid by the state and by business enterprises (income taxes), on expenditures arising from household income (value added and sales taxes), or on firms’ gross profits generated by their sales to households, banks and the state (corporate profit taxes).

Taxes must therefore be conceived as a component of the reflux phase of the monetary circuit, while state expenditures are a necessary component of the flux phase. The latter expenditures must normally be financed by credits granted to the state by the central bank, unless such a bank does not exist (as in nineteenth century USA) or, if it exists, it is prohibited from engaging in such financing (as in the European Monetary Union). In these latter cases, state expenditures would be financed via the holding of government securities by commercial banks. However, in the more common case in which a central bank does exist, the bank would behave much as private commercial banks vis-à-vis other non-bank agents. By crediting the spending branch of the state (the treasury), the central bank grants loans by issuing debts on itself. The counterpart of these loans is a forward debt specifying the amount of these debts that the state has to collect through taxes to extinguish its debt obligations to the central bank. The collection of taxes simultaneously cancels a share of central bank debt issued in the first phase of the monetary circuit as well as an equivalent debt of the treasury.

The remaining quantity of central bank liability arising from government spending is normally referred to as the government deficit: a purely ex post notion which can be accounted for only in the reflux phase of the monetary circuit. Since it is an ex post value, the deficit is by its very nature already financed. When mainstream economists speak of ‘deficit financing’ or the ‘monetization of the deficit’, they display a profound misunderstanding of the nature of monetary circulation. Contrary to firms’ indebtedness to commercial banks, the state has the power to determine the amount of its debt to the central bank, since it is a pure conventional debt, a debt of the state to itself. The state can plan the amount of its future deficit, which is the ultimate consequence of its power of endorsing bank debts as money. Since commercial bank debts are entirely convertible into state money, a significant portion of the state money that is injected into the economy will be transformed into commercial bank debts. Such a conversion allows commercial banks to acquire central bank liabilities which appear in their assets as reserves. Conversely, a share of commercial bank money will, at the same time, also be converted into state money by private agents wishing to hold cash or legal tender. In this process, commercial banks would be losing
reserves, and drawing down their deposits at the central bank. In a modern economy, this commercial bank absorption of reserves from state expenditures will normally exceed their losses of reserves determined by the public’s demand for cash. As a consequence of their legal convertibility, the private sector can pay taxes in the form of commercial bank debts in the final stage of the monetary circuit. Since the state will require the conversion of bank deposits into its own money through its banking branch, commercial banks will lose reserves in the reflux phase of the monetary circuit. Assuming a given and low preference for cash on the part of the public, we therefore reach the conclusion that the higher the state deficit, the greater is the net increase in commercial bank reserves.

From this analysis of the role of the state within the monetary circuit, we may draw some crucial propositions regarding the endogenous nature of money:

- First, the quantity of newly-created money is always determined by the effective or expressed demand for credit from firms or the state.
- Commercial banks only meet creditworthy demand for loans which comply with the required monetary mark-up or profitability criteria that is reflected in the rate of interest that they impose on firms. Effective demand for loans is the result of firms having already adjusted their planned expenditures to the commercial banks' creditworthiness norms. Consequently, the effective demand for loans is always equal to the quantity of newly-created bank money. In part, however, these criteria of creditworthiness also depend on commercial banks' own credit worthiness as reflected in their capacity to convert freely their own liabilities to those of the central bank.
- The central bank must ultimately play an accommodating role. Even if it wishes to act otherwise, it will fail because of the existence of state expenditures. For instance, let us assume that the central bank tries to impose some required ratio of cash reserves to bank liabilities. As has been shown, banks' net increase in reserves merely mirrors the state deficit. On the other hand, the net increase in commercial bank liabilities reflects the amount of the initial debts incurred by firms vis-à-vis the banks which the former cannot reimburse after their sales of commodities and financial securities. In an economy in which the state deficit is significant, the net increase in reserves would usually be greater than the increase in bank liabilities. The ex post ratio of reserves to bank liabilities would be higher than the required ratio, thereby preventing any coercive action by the central bank. If anything, the latter would be engaged in a futile action of trying to mop up excess reserves within the banking system to prevent the interbank funds rate from falling below some targeted level (Mosler 1997–8).
Competing approaches to the theory of money and the TMC

We wish now to compare the TMC, which we believe to be a dynamic theory of money, with three other major approaches that have dominated the literature in monetary economics: the neoclassical theory of money, the Post Keynesian theory and the neo-chartalist approach.

An assessment of neoclassical theory

To discuss the neoclassical approach to the theory of money, we have to try first to encapsulate this approach into a set of major propositions. In drawing up this list, we seek to emphasize only the propositions shared by most neoclassical economists. In each case, we shall inquire about the relevance of the proposition in relation to the theory of the monetary circuit.

Barter and monetary exchange

The starting proposition of the neoclassical approach going back to Jevons (1875) and Menger (1892) is that money is to be introduced within an otherwise pre-existing pure barter economy. All neoclassical economists accept the barter economy analogy as the reference system within which money appears (Clower 1969). The introduction of money as a simple intermediary or means of exchange to facilitate the two-way barter system, however, merely increases the number of transactions in such an economy. For instance, before the introduction of money, barter entailed a two-way direct exchange between firms and sellers of productive services, as in Walras's original production model. The introduction of money merely breaks up such barter transactions into three indirect exchanges. Assuming that there exists a financial capitalist who possesses money as a third agent, firms must now exchange their output for money held by this third agent so as to be able to purchase productive services. Sellers of productive services must then exchange that money for the goods previously acquired by the financial capitalist. This same type of multiplicative effect on the number of transactions is to be found, for instance, in Hayek's analysis of a production equilibrium (Hayek 1941). All monetary transactions are de facto of the same nature as barter exchange with a numéraire money having been introduced to it. They are simply exchanges of commodities of equal pre-existing value with numéraire money having a specific intrinsic value like any other commodity. This value is not created by the exchange but rather exists prior to it. In effect, it is an existence condition for the exchange itself.

Since monetary transactions pertain to two commodities of equal pre-existing value, they cannot be debt relations. Because of their theory of exchange, neoclassical economists have never been able to deal properly with debt-credit relations. Historically two solutions have been offered to this question.
The first explanation implied that banks are pure intermediaries allocating a pre-existing stock of commodity money among agents needing capital. Within the context of this, deposits make loans or, perhaps more correctly, savings make loans. This tradition was followed by Hayek (1931, 1941), Gurley and Shaw (1960), Koopmans (1933) and many others, including those referred to by Realfonzo (1998). If, however, banks were to lend more money than the amount that they had obtained as original deposit liabilities, the law of exchange of equivalent values would have been broken and, as a consequence, a cumulative neo-Wicksellian process of inflation would be the outcome of this excess of credit (Seccareccia 1990).

A second solution offered by neoclassical theorists interprets the role of banks as the supplier of pure accounting services. When all trading is closed, there may still remain some transactors who, because of imperfect information or indivisibility in the barter process, find themselves indebted to other transactors, and therefore require money to settle their debts. Banks emerge in the clearing process with holdings of an exogenous stock of money which is then provided to all those transactors facing a financing constraint. This view is found in some of the early works of Clower (1967), Starr (1972) and Ostroy (1973) and is reviewed in Parguez (1975).

For either of these two solutions, the TMC invalidates their starting proposition. As we have argued previously, banks issuing debts on themselves for the financing of loans can never be intermediaries between savers and borrowers. Even when banks are forced into a position of advancing new loans in order to compensate for the high liquidity preference of the public, they do not in fact lend savings (Graziani 1996) since savers do not ‘lend’ their deposits to banks. Banks have to issue new loans in order to refinance any outstanding debts by firms. Consequently, they can never be pure intermediaries nor suppliers of accounting services that do not create money.

Money and utility

A second major proposition of the neoclassical conception of money as a pure medium of exchange is based on the presupposition that money exists because it has a specific intrinsic value. As put forth by Koopmans (1933) and later by Patinkin (1965), the value of money must be explained in the same way as the value of any other commodity, by the principle of marginal utility. This proposition has two major consequences.

First, money has a specific utility if, by using it as an intermediary in exchange, it increases the real wealth of transactors. Unfortunately, neoclassical economists have never provided a substantive proof for this proposition. As implied by the first proposition, money is introduced in an already perfect system of exchanges leading to a general equilibrium consistent with utility maximization for every transactor (Hahn 1982).

Second, even if a proof for the utility of money were to be found, it would merely explain the existence of a specific demand for money.
However, this Patinkinesque demand for money is not an indirect demand for commodities and services, but a reflection of the share of the aggregate stock of wealth which is held in money form instead of being exchanged for physical commodities and services. Herein lies the well-known paradox that money would derive its specific value from its role as intermediary in transactions, but its utility can only be derived if it is held by economic agents. In equilibrium, therefore, the whole stock of commodity money would be hoarded, and money would not play any role in circulation! For such reasons, this second proposition of neoclassical monetary theory contradicts what we have described as the dynamic nature of money analysed within the TMC, in which money is never a commodity possessing any intrinsic value. It is a pure conventional debt issued to generate a sequence of transactions leading to the formation of real wealth.

Money exogeneity

In neoclassical economics, money must also be scarce to have a specific value. The scarcity principle implies that the available quantity of money must be independent of the set of demand schedules for both money and other commodities. As emphasized by the neoclassical school from Walras onwards, the requirement of exogeneity became a logical necessity for the application of the scarcity principle. In this way, as an exogenous variable, money can never be credit-driven and hinge on the loan-granting actions of commercial banks. To achieve this result, banks are assumed not to create money because they are considered as pure financial intermediaries, with the supply of money being determined by the exogenous actions of the central bank, as in the standard monetarist model. However, the error of the monetarist school was to confuse logical necessity with empirical reality which invalidates this third proposition of neoclassical economics. Within the TMC, the quantity of money is strictly endogenous since it is loan-determined and cannot be set by the exogenous actions of the central bank.

The Post Keynesians and the TMC

Unlike the situation in neoclassical economics, it is more difficult to compare the Post Keynesian theory of money vis-à-vis the TMC because there exists no unified Post Keynesian conception of money. We shall try, therefore, first to address what the broad Post Keynesian approach and the TMC have in common, then to discuss what is deemed to be unclear, or perhaps missing, within the Post Keynesian tradition.

Points of convergence

All Post Keynesian economists agree with the theory of the monetary circuit on the rejection of the first proposition of the neoclassical theory of
money. They have accepted the credit/debt theory of money whereby money is always a debt issued by banks on themselves to finance loans. For this reason, they also reject the third proposition of neoclassical economics. Money is always endogenous because its quantity is demand-determined. Numerous Post Keynesians have thus broken away from Keynes's conception of money as it was spelled out originally in Chapter 17 of the *General Theory*. Money cannot be conceived as a commodity having a scarcity value amid other existing commodities in the system to which the equimarginal condition is applied.

However, within the Post Keynesian literature, the so-called 'horizontalist' view is perhaps the closest to the circuit theory of money (Rochon 1999). Within the Kaldor-Moore horizontalist model, banks always issue debts on themselves to finance loans granted to firms that are seeking to acquire real resources. In our opinion, the famous interest/money space horizontal diagram of the money supply rightly spells out that, for a given level of interest rates, banks accommodate all creditworthy demand for loans. Furthermore, horizontalists endorse the circuitist view that the central bank cannot directly control the quantity of credit money. What the central bank can control is the level of interest rates charged by the banks on their loans, with the latter setting their loan rates by merely fixing a mark-up on the base rate established by the central bank. Interest rates are, therefore, exogenous since they are not determined by a market mechanism that adjusts the demand for money to a fixed supply. Given bank mark-ups, the level of interest rates is determined by the behaviour of the central bank in targeting a specific inflation rate.

Indeed, horizontalists reject two crucial aspects of the *General Theory* conception of money. First, like the theory of the monetary circuit, they question the link between money and uncertainty. Keynesian uncertainty cannot *in esse* explain the existence of money, since firms would need bank loans to finance expenditures even within a stationary or ergodic environment. Money is a right to acquire real wealth generated by previous loans, and if sellers of real resources were absolutely uncertain about the outcome of the loan process, they would never accept payments in bank liabilities. Uncertainty can only have an impact on the creditworthiness rules imposed by banks, thereby altering the effective demand for loans. Second, horizontalists have emphasized the distinction between the demand for loans and the Keynesian demand for money, the latter of which is merely the demand for liquid balances emerging *ex post* out of the credit money previously issued to finance loans. Much like the TMC, therefore, they conclude that there can never be an excess supply of money (Lavoie 1992).

*Diverging opinions*

Some Post Keynesian economists, who have remained too strongly connected with the monetary economics of the *General Theory*, have been
unable to grasp fully the consequences of the TMC and do not dissociate themselves completely from the second proposition of neoclassical economics. The latter proposition pertaining to the stock demand for money was of crucial importance for the General Theory explanation of the non-neutrality of money. As is well known, the Keynes of the General Theory emphasized the demand for money as a component of aggregate wealth along Marshallian lines. Portfolio holders choose to hold money instead of other commodities because of its intrinsic liquidity characteristic. On the other hand, he assumed an exogenous quantity of money in accordance with the neoclassical scarcity principle. From his hybrid analytical framework, in Chapter 17 of the General Theory Keynes could then explain why the endogenous rate of interest, determined on the basis of the scarcity principle, could be the source of the average expected rate of profit in the economy. A strong rejection by Keynes of the second proposition of neoclassical monetary economics would certainly have jeopardized and, perhaps, could have led to an unravelling of the whole General Theory analytics.

These difficulties notwithstanding, many Post Keynesians, for instance, Dalziel (1995) and Howells (1995, 1997), have explicitly maintained the Keynesian demand for money as a crucial analytical device in explaining the dynamics of a monetary economy even in an endogenous money context. Moreover, some such as Chick (1992) and Howells (1995, 1997) do not accept the principle of an absolute credit-driven money. There would always be the possibility that money could be created through pure portfolio operations by the central bank. Others refuse to recognize that the central bank cannot have power over bank money creation via quantity constraints. That would supposedly depend upon the willingness of the central bank to be accommodating in satisfying the liquidity needs of commercial banks. Even verticalists display some reluctance on this question. Moore (1996), for instance, accepts that there had been times when the central bank was not accommodating. However, much of this debate arises because Post Keynesians postulate a central bank without the state and ignore the impact of state expenditure on bank reserves. In contrast with what we have tried to do, they do not apply the endogeneity principle to the state itself (Moore 1988).

Even when they address the role of credit, many Post Keynesians, including some horizontalists, impose constraints on what can be financed by credit money. For instance, in Moore (1988) bank loans only go towards the financing of firms’ circulating capital requirements. Like Wray (1990), Moore himself conceives of loan financing as a means of bridging the gap between firms’ expenditure and revenue. Hence, banks issue debts for the sole purpose of financing deficits. This view is completely at odds with the TMC which shows that firms’ or governments’ deficits can only be accounted in the reflux phase of monetary circulation. Such a misunderstanding of the financing process is perhaps best revealed by the way in which profits are conceived in investment financing. Post Keynesians,
especially of Kaleckian pedigree, make the distinction between internal financing of investment out of business-retained earnings and external or debt financing. As has been discussed elsewhere (Seccareccia 1998), this distinction is somewhat misleading since in the TMC the internal generation of finance via business profits is merely the macroeconomic outcome of the new indebtedness of other agents in a monetary economy.

This inability on the part of some Post Keynesians to sever links with the second proposition of neoclassical monetary theory may perhaps also explain why some, such as Chick (1992), have endorsed an evolutionist conception of money. Accordingly, there would have been at least three stages of monetary evolution. In the first stage, banks are pure intermediaries which recycle previous savings since they cannot freely issue debts on themselves. In a second stage of this historical evolution, banks can issue debts on themselves because of the accommodating role of a central bank. There is apparently also a third stage, the disintermediation one, in which banks are presumed to go to the financial markets actively to raise funds via liability management. Funding out of these markets to acquire necessary cash reserves would be substituted for pure bank credit, thereby further eroding the reserve constraint on the ability of banks to provide credit. As our earlier discussion of the history of money has shown, however, such an evolutionist view of monetary developments is somewhat equivocal since banks have never been pure intermediaries.

The neo-chartalist approach and the TMC

While neoclassical economists point to the efficiency gains of introducing money in the context of barter exchange, and while Keynes had emphasized its liquidity role in the context of fundamental uncertainty, there is a third view on the existence of money. We may describe this third approach, as it has been put forth by Mosler (1997–8), Bell (1998b) and Wray (1998), as the neo-chartalist approach.2

The starting proposition of the neo-chartalist approach is that money is always a creation of the state and that the state can ultimately impose any token as money simply by requiring that tax liabilities be paid in such a unit. In this regard, taxes appear as the fundamental prerequisite to the existence of money. By varying the public’s tax liability, it is assumed that the state can even determine the value of money, its purchasing power. From these propositions, neo-chartalists derive a conception of monetary circulation that is closer to the TMC than, say, that put forth by Keynes in the General Theory.

Taxes are required to impose state money by inducing private agents to sell real resources to the state in order to obtain sufficient money to settle their accounts with the latter. Taxes cannot, therefore, be used to finance state expenditures. The state must first inject money into the economy by issuing at will the token in which it has decreed taxes must be paid.
According to the neo-chartalist approach, in modern economies the state token is essentially created when the central bank issues debt on itself at the request of the treasury. When taxes are paid *a posteriori* in the reflux process, they extinguish an equivalent quantity of the state-created money. From this circuitist analysis of the financing process of the state, neo-chartalist writers are brought to reject the conventional analysis of budget deficits.

The orthodox notion of government deficit financing and the so-called 'monetization' of deficits is highly misleading and, indeed, a misnomer. Since deficits are merely *ex post* accounting values measuring the net flux/reflux process pertaining to the state sector, they have already been financed. As was shown earlier in our discussion of the TMC, the counterpart of the state budget deficit is both an increase in the private sector holdings of money and an increase in bank reserves. However, as soon as bank reserves become excessive, the banks will stop borrowing reserves in the funds market and thereby deprive the central bank of any power of intervention. Indeed, if reserves are excessive, banks will actually try to purchase new bonds, whose effect would be to push up bond prices and thus bring about a collapse in interest rates. We can also assume that other private sector holders will attempt to do the same, thereby further accelerating the fall in the level of interest rates in the economy. From this, we can infer that the state must issue new bonds just to keep interest rates at some target level defined by the monetary authorities. Neo-chartalists would generally agree with what is essentially a Post Keynesian horizontalist position, that the central bank sets the base level of interest rates (see Wray 1998). However, any uneasiness with the horizontalist position by neo-chartalists may be due to the fact that Post Keynesians tend to have a theory of the central bank without the state. A careful understanding of this link, and the role of the treasury, along the lines that we have spelled out previously, would probably bring these two views closer to the circuitist position.

However, an important difference between the TMC and the neo-chartalist view relates to the emphasis that the latter places on taxes. As we have defended in our historical discussion, viable monetary systems existed during periods of economic history when taxes were quite insignificant. What matters, therefore, was not whether tax liabilities were of any significance but rather whether, largely through the legal system, the state endorsed existing banks by allowing them to issue debts on themselves. For very long historical periods, state money had been quite negligible in relation to the circulation of bank liabilities. By linking the existence of money exclusively to that of taxes, neo-chartalists are led to find money in societies where there are heavy taxes levied *in natura*. Taxes can more easily be levied if there is a state-imposed unit of account which would serve as money (Wray 1998). Ironically, such a definition of money as a pure unit of account is very close to the Walrasian conception of *numéraire* money in which it is the (state) auctioneer that chooses a *numéraire* before efficient trading takes place!

The state can endorse central and/or private bank liabilities, but it cannot
impose the value of money. While the neo-chartalists seem to identify a positive relation between the value of money and the amount of tax liabilities in an economy, what actually matters is rather the nature or composition of state expenditure. In other words, what is consequential to the value of money is primarily the ability of state expenditure to increase the real wealth of society either directly, through the production of public goods, or indirectly, through their capacity to foster private investment expenditures. If state money is issued merely to finance wasteful expenditures that have no serious positive consequences on the private or collective wealth of a community, the effect in the long run would be to depreciate the value of money, regardless of the power of taxation of the state.

Finally, the neo-chartalist theory of money logically implies that state money precedes private bank money. In other words, the creation of state money upon which taxes can be collected is followed by the emergence of commercial bank money when the state accepts private bank liability in payment of taxes. As emphasized by Bell (1998a), bank money is hierarchically inferior or subordinate to state money. Banks' credit activity is a leverage on the existing stock of state money (Wray 1998). This presupposition signifies that state money is endogenous but bank money is not since, in the final analysis, it is constrained by the variation of the former. This conception of the monetary system which, as it stands, hovers closely to a textbook analysis of how bank money is created, needs to be better articulated. As we have sought to show in our discussion of the relations among banks, firms and the state, the TMC can provide neo-chartalists with a more comprehensive framework of analysis.

Conclusion

In this chapter, we have discussed the main outlines of the TMC approach to money and contrasted this with alternative approaches, including neoclassical theories and contemporary heterodox views. As far as the latter are concerned, we have tried to argue from the beginning that both Post Keynesians and neo-chartalists will find a more cogent and coherent integration of money and commercial banking within the general framework of circulation as enunciated within the TMC. With its broad analytical framework, the TMC is able both to accommodate and to offer meaningful solutions to many of the enigmas presently confronting competing heterodox theories of money.

Notes

The authors wish to thank Marc Lavoie for his helpful comments.

1 As defined by Keynes (1930) in the Treatise on Money.
2 The approach is originally anchored in the work of Knapp (1924).
References


