

Money Creation in the Debt Economy

By Dix Sandbeck, Dec 2016.

Web: attac-ontario.org

Contact: dix@dixsandbeck.ca

Trust and Taxing Power

The question about how money is created has been a critical question ever since the creation of money slipped out of the hands of kings and emperors, who once controlled all minting of coins from precious metal. At first, the size of the coinage were dependent upon the availability of the used precious metals, which were fixed in the short term, but the rulers quickly discovered that they could expand the money supply by debasing the coinage, i.e. mixing the gold or silver with baser metals, but keeping the face value by fiat (royal decree). This method of creating money, which gave the rulers extra spending powers, was called 'seigniorage', since it was the prerogative of the 'seigniors', or the rulers. The term, seigniorage, is still used for the creation of central bank money that pay for government expenditures in excess of revenue.

When proto-banks (goldsmiths, money-changers, etc.) started to expand the money supply by issuing redeemable promissory notes in excess of the gold in their vaults, money creation was partially wrestled away from the feudal rulers and into private hands. This process began to proliferate among the goldsmiths of 17th century Britain, but it should be noted that as long as the promissory notes were directly linked to deposited gold, even when they were written as bearers' notes and therefore could circulate, this was not money creation in a quantity sense. As such it was not materially different from the activity of the moneychangers for instance in medieval Bruges where it has been shown that money was transferred among accounts by oral pledge.¹

Only when promissory notes were issued as loans, which meant that they were issued to people either without gold deposits or perhaps with gold deposits but in excess of the deposits, was money created. This is the earliest example in a European context of proto-bank money creation through the credit-debt channel, and incidentally also the beginning of sustained use of paper money.

The money question grew in complexity when the rulers learnt from the goldsmiths and began to issue convertible paper money on a fractional basis. Bank of England was founded in 1694, and began shortly thereafter to issue paper money backed by a fractional holding of gold. State issued paper money fractionally backed by gold reserves (interrupted by wars at which times convertibility often was suspended) continued as a main money form until the gold standard finally was abandoned in the 1930s.

In the 17th century, the early private banks that had emerged from the goldsmiths'

¹ Raymond de Roover: "Money, Banking, and Credit in Medieval Bruges." The Journal of Economic History, Vol. 2 (1942). Bruges was at the time the leading commercial centre in North-west Europe.

activities began to allow their customers to write their own promissory notes against deposits in the banks, a practice that was the origin of the cheque.

Currency based on fractional gold (and sometimes also silver) reserves, was at first issued both by private and central banks but since the end of the nineteenth century almost exclusively by CBs. After the 1930s, when essentially all economically developed nations abandoned holding gold reserves for the currency they issued, all state money has been fiat money, i.e. currency only backed by trust in the issuing state and its taxing powers.

The tendency to hold money in bank accounts, which the goldsmiths had pioneered, proliferated with the introduction of the private cheque. When credit was added to the mix, for instance by permitting account overdraft or by issuing loans, an element of the money supply was created that only had indirect links to the quantity of monetary gold and silver.

Components of the Quantity of Money

Modern economies have three interrelated forms of money. Currency (banknotes and coins), central bank (CB) reserves, and commercial bank deposits.

Currency is issued by central banks, albeit this form of money plays a falling role in the economy. The main avenue through which currency enters the economy is through commercial banks. When needed they will exchange some of the electronic reserves held at their CB reserve accounts into physical vault cash and then dispense it to their customers when these demand withdrawals as currency (for instance, from ABMs²). Conversely, if a commercial bank holds too much currency (vault cash) they can return it to the CB, which then credits the given bank's reserve account with the money. This means that currency and CB reserves are fungible; it doesn't matter in which form a bank holds its reserves (although if CBs pay interest on reserves, which has become increasingly widespread in recent years, there is an opportunity cost difference).

Modern money circulation is overwhelmingly a function of the system of accounts held by the commercial banking sector. Most incomes occurs as electronic transfers to bank accounts, and the currency used for the fraction of retail payments that still occurs as currency transactions is acquired by withdrawing it from these bank accounts.

Aggregate money quantity and its expansion are often interpreted as being closely dependent upon the credit function of banks. Be as it may, all economies also contain substantial amounts of money that are inserted into the economy through non-credit channels, e.g. seigniorage that accumulates over time, reserves arising from central banks net purchases of financial asset from the private sector, and the conversions of earlier periods commodity monies into fiat currencies including the part deposited into transactional bank account monies. While the latter—the part of the money supply that originate in heritage money—today is a static pool (commodity money no longer being issued), seigniorage and asset purchases are in varying degrees active forms of controlling the amount of circulating money in most countries.

² UK and US English = ATM (Automated Teller Machine).

The conversion of earlier periods monies is a function of the fact that money is never destroyed by use, but only by accidents,³ the latter being very peripheral phenomena with little influence on the aggregate quantity of money. Thus, earlier periods' commodity money supply continued to form the base of the money supply in the next historical period, albeit sometimes in a changed denominative system.⁴ This is particularly true in a commodity money system, but also in fiat money systems as long as the political economy continues in an evolutionary process from period to period.

Sources of Money Creation

With respect to the fundamental question of money creation, there are in modern economies several main sources for creating the money needed to sustain the economy and its growth (as measured by nominal GDP):

A. State seigniorage, which is a state issuing more money (generally through its central bank) to pay for expenditures in excess of what it receives in taxes and other forms of government incomes. In most countries taxes are not paid to the central bank (CB) as such, but to an account which the government maintains at the CB, which then in turn is the account to which government cheques—the cheques paying for government expenditures—are debited. If the government wishes to spend more than its account balance would allow for, it can, technically quite easily, have it replenished by money that the CB creates and credits the account. In most countries this is not done whimsically, but money created through this channel occurs within a strict accounting frameworks where there on the liability side of the government's CB account balance would appear some kind of debt instrument owed to the CB. On the CB's balance sheet this debt would be an asset balancing out the extra money liability. Generally speaking, this kind of money creation is called fiscal expansion, since the spending has to appear as postings in the government's budget and the money is inserted through government spending.

However, in recent decades there has been a strong tendency among the Western developed economies to finance deficits by selling debt claims to private markets (government bonds which has an interest cost). The purpose of financing government deficits by selling debt instead of issuing seigniorage is to absorb the extra, presumably inflationary, liquidity that traditional monetary theories believe expansionary government expenditures will incur. This view is founded on the monetarist belief that all inflation is monetary; a view, which however ignores the cost-push element of prices, linked to scarcity/excess balances of inputs and outputs. For instance, if there are unused inputs, in particular labour inputs (unemployment), targeted seigniorage-financed spending by a government can activate such resources without having any serious influence on price levels. It should be emphasized that seigniorage influence and expand the economy through real sector channels.

³ or in earlier cultures sometimes by sacrifices.

⁴ It should be noted that in medieval European and in the early modern period economies, for instance colonial Canada, there were often competing units of accounts. Currencies issued by weak (and often quickly changing) feudal rulers would in transactions of some size be deselected in favour of the currencies of chief trading economies, or perhaps the silver weight of a main trading centre.

B. Banks can expand the volume of transactional money when they create new credit in ways that expands their balance sheets.

C. Central banks buying financial assets with newly created CB reserves, for instance through open market operations, which involve both buying and selling government debt in attempts to adjust the level of liquidity, both up and down. In general, open market operations are therefore seen as being rather neutral with regards to their overall impact on monetary aggregates.

In case of more dedicated programs for expanding the money supply by CB purchases of assets, as for instance quantitative easing (QE) represents, such programs will, in contrast to seigniorage, impact the real economy indirectly through the economy's financial channels. The recent experiences with QEs appear to have shown that they had subdued effects on real variables and might in fact have had their main effects in driving asset inflation into higher gears.

Different Views on Money Creation

As we saw, modern governments, certainly in the economically advanced western countries, are generally unwilling to finance expansion directly via the seigniorage channel. If we for the moment ignore the experimental QE programs, the bulk of modern money is therefore created by commercial banks. That means that understanding the processes through which this kind of money creation materializes is of utmost importance for a general understanding of how modern economies function.

But views on these matters are not uniform and have undergone substantial changes in step with the rise and fall of different economic and monetary theories, a matter which has led to considerable confusion in debates and analyses of the current political economy.

In the traditional view, banks were seen as intermediaries between household savers with extra cash on their hands and firms in need in of money to facilitate investments in a growing output. Under the assumption that deposits create loanable reserves, a bank's ability to make loans was seen as directly linked to its portfolio of deposits, more or less on a one-on-one basis. Essentially, it was a view born in a period when money was still mainly envisaged as physical cash, and the implication of money also existing as postings in banks ledgers was not fully digested.

Later on, standard economics began to realize the implication of the existence of 'account money', which had opened up for banks' ability to create money as credit and quashed the assumption of a direct link between deposits and loans. After having acknowledged that commercial banks can in fact create credit money, the view emerged that central banks could control this money-creating ability via the so-called money multiplier. This was understood to work as a function of the requirements for banks to hold 'required reserves', which was a certain percentage (typically around 10%) of their assets that had to be held as vault currency or on account with the central bank (CB reserves). However, today many central banks have abandoned reserve requirements, in the process removing the idea of a money multiplier—i.e. the CBs' presumed ability to control money levels via

the level of required reserves—from its former governing place, although this explanation of money creation is still common in many textbooks. Moreover, where reserve requirements remain in place, central banks will normally be accommodative to banks needs for reserves and therefore not try to control the money supply by strict control of reserves. The rationale is that reserve control in a strict, almost mechanical fashion contains the danger that it can disrupt the general payments systems.

The Double Entry Balance Sheet

The logic of banks holding reserves points to the fact that many money and banking phenomena can best be understood if analyzed within a balance sheet accounting framework. Balance sheet accounting stems from the invention of double entry bookkeeping, which appeared in North Italy in the 15th century. The main principle is that if there is a change on one side of the balance sheet, there must be a balancing change on the other side. To fully appreciate the system, one must be able to see that all monetary changes have physical components, even if it might only be electronic redistribution over the Internet of magnetized computer memory bits. Therefore, fungibility⁵ between physical currency and electronic account cash rules in the transactions occurring between banks and their customers.

Let's have a look at an example of a simple bank's balance sheet:

Simple Bank's Balance Sheet

Assets		Liabilities	
CB reserves	60	Deposits	700
Vault cash	60	Loans from money markets	300
Government bonds	180	Equity	100
Loans to customers	900	Retained earnings	100
Sum	1200	Sum	1200

In the above example—with balance item weights arbitrary and not attempting to resemble real conditions—the total reserves that Simple Bank have are 120 units or 10% of assets. On the liabilities side the risk is that a large number of depositors withdraw their money more or less at the same time (a bank run). If they all want cash and Simple Bank only have 60 units it might run out of cash when trying to comply with the withdrawal request. Of course, it can, as just mentioned, with immediate effect exchange its CB reserves for currency but if deposit withdrawals continue a bank might not be able to raise enough cash from its asset side to cover its main commitment on the liability side, which is to let deposit customers at will withdraw the money sitting in their transaction accounts. Consequently, it would become insolvent.

The other general risk that a bank faces is that loan customers' default on their loans. When that happens, the bank has to, first, on its balance sheet write the asset side down with the amount of the defaulted loan, and secondly (due to the principles of double entry bookkeeping), write a corresponding amount down on the liability side, which has to

⁵ The idea that two things are perfect substitutes for each other.

come from the capital portion. If default write-downs continue beyond the capital (equity) portion, and assets thereby become less than liabilities, a bank is technically bankrupt.

Reserves and Deposits

Traditionally, a problem for banks was that currency and CB reserves didn't earn incomes. Therefore they wanted to keep them at a minimum. Currency of course still don't earn interests, but its role in the economy, including in bank operations, has diminished; mainly being used to feed ABMs and supply the retail sector with its needs of cash change. However, a recent innovation in monetary policies has been that many CBs have started to pay interest on CB reserves, which also is an electronic money form in modern systems. Ostensibly, this new development is implemented in order to enhance CBs ability to better control short-term interest rates.

If we return to the commercial banks' abilities to create credit money, a point of contention has been to what extent they are constrained in that endeavour by deposits and reserve positions.

Newer economic theories realized that when banks create loans, the process in fact materializes through crediting the loan receiver's bank account with the amount, or said differently, by *depositing* the amount of the loan into his or hers account (in some cases by first opening a new account). This led to the view that "loans create deposits", and the amount of pre-credit deposits was therefore in this story of money creation not seen as important, a view that still dominates parts of modern monetary debates.

After deposits had been dismissed as important for banks' ability to create loans, the turn came to reserves. The story became that banks when they extend loans don't need to worry about reserves, because the central bank will always extend extra reserves to them in a pinch. Bank first makes loans and then looks for reserves, being assured that if they don't already have enough reserves and neither can find them in the interbank market, then the central bank surely will step in and extend it to them. This assumed willingness to always accommodate banks' needs for reserves was seen as caused by a key concern of central banks, which is at all times to protect the soundness and uninterrupted function of the economy's interbank clearing and electronic payment systems.

Since neither having pre-credit deposits or reserves was seen as a hindrance for the process of creating money through the credit channel, it has led to the widespread assumption that banks can at will "create money out of thin air" whenever they find profitable loan opportunities.

In consequence of the above, a common modern view has become that well-run banks seldom have problems with accessing extra reserves. Either they can access them in the short-term interbank market for reserves (those banks holding excess reserves lending them short-term to those in deficit). Or they can sell government bonds to the central bank, and in exchange the CB inserts new reserves on the given bank's reserve account at the CB, the account from which interbank settlements are drawn. Accessing extra reserves from the CB often occurs as repos (repurchase agreements) that are short-term transactions where the CB after the prescribed period sells the bonds back to the bank,

with the repo rate added as the cost of the credit transaction for the bank.⁶

Reserves Still Matter

Banking systems of modern developed economies are systems that generally obey the probability laws of large numbers in nice Gaussian curves. Black swans fly this world by and most of the time experienced bankers have a good feel for how these probabilities presently are aligned and how they will align in the near future with regard to their banks' operations.

In a well-functioning banking system, the notion that reserves don't matter might therefore be true in practical terms. Under such conditions, a bank's reserve position is in effect a cost problem, balancing different levels of costs of acquiring extra reserves in the different reserve markets (new deposits, interbank markets, CB discount windows, etc.) set against of the opportunity cost of holding excess reserves (thus having a lesser need for outside reserves).

However, that reserve positions don't matter is not true in a stringent theoretical sense and it's particularly not true when the general economic system finds itself in a crisis and the interbank channels are under stress. Nor is it true when an individual bank experience large loan losses or sudden large deposits withdrawals.

During crises, or if for instance large-scale fraudulent activities are inserted into the parts of the financial system, informed expectations to the soundness of counterparties become difficult to gauge. Under such crises situations, a bank's reserves can quickly disappear and access to new reserves through markets freeze up. This was one of the things that happened in the wake of Lehman Brothers collapse in September 2008 when interbank channels froze up. Even though it only lasted for a few days, during that time it threatened a total monetary collapse.

Furthermore, the view considering deposits and reserves unimportant for the creation of credit money also overlooks that deposits created by loans are different from deposit originating from, say, customers depositing their earnings. The fact is that deposits created by a bank's own credit operations are endogenous deposits and create no new reserves, while deposits originating from customers earnings are exogenous deposits and therefore create new reserves for the bank.

This leads to another essential aspect of balance sheet understanding. A key condition influencing the abilities of modern banking systems to expand aggregate money quantities is the processes that net out interbank claims.

With regards to the loans made to customers, when they spend them by issuing cheques (or other payment forms) it requires reserves from other sources to be available for the loan-issuing bank when the payments are being deposited in other banks and returned as interbank claims. But when one bank have to transfer reserves to other banks on behalf of

⁶ The bank rates that CBs publish are the interest rates that—somewhat dependent upon the specific national institutions—govern interbank clearing transactions.

customers' spending of credit money, customer of other banks will also obtain loans and issue business and consumption payments debited to their accounts filled up with newly acquired credit money. Some of this spending will end up deposited in the first bank and thereby replenish its reserves.

In this way, reserves also work on a fractional basis and small holdings of reserves can underpin large flows of transactions due to constant netting-out through the bank clearing channels. During normal times daily discrepancies in mutual claims during the netting-out processes will not be big. Consequently, they will be manageable along the avenues for accessing extra reserves outlined above, but also including acquiring reserves in money markets, etc. Only when the payment flows deviate from their normal patterns and substantial imbalances occur will reserve positions return to focus.

Is Money Really Destroyed Again?

After verifying that deposits and reserves cannot be ignored when explaining the money creating process of commercial banks, it becomes clear that money is not created out of thin air, but out of *institutional* air; that is to say, it is only possible for a bank to create loans without being assured of having directly matching reserves because there exists a set of institutional networks, such as the clearing institutions, the interbank markets for reserves, central banks with discounts windows, etc., all of which allows banks to make loans without first checking for reserves on a one-on-one basis. This of course means that the required volume and consequent cost of holding liquidity when engaging in credit expansion is sharply reduced.

After clearing up the difference between *thin* air and *institutional* air, we come to a related view, which is that when loans are repaid the money is destroyed again. In other words, the money that was created from thin air reverts to thin air again. This is of course a pure metaphysical view that cuts lose the chain of institutional processes that makes the money system go around.

When contemplating loan processes, things can appear muddy if one thinks in the terms of elusive electronic money. But since physical cash and electronic current account money⁷ are fungible, one can instead conceptualize the loan process by thinking of it in terms of currency, i.e. transacted by notes and coins.

If money is withdrawn as currency (physical cash) from a transaction account into which a loan has been credited, and later repaid with interest, again by notes and coins, the bank issuing the loan must be ready to pay it out as currency, i.e. have the currency ready as reserves in its vault or be assured of immediate access to extra currency.

When later the bank receives the repayment of the loan as cash banknotes, it doesn't indulge the bank manager to order these notes burnt in an oil drum in the alley behind the bank. What is important to visualise is that the money, when they originally was paid out from the debtor's bank account to be used for whatever use he or she had for it, it is through the netting out process—in other words through the banks participation in the

⁷ Before computer accounting, this kind of money might be called 'ledger account money'.

aggregate banking sectors simultaneous expansion of their balance sheets—that the loan represents new money, created fractionally on a given reserve base.

Thus, what really happens is that the cash money is not destroyed when repaid, but put into the vault where they change status from M1 money to M0 money⁸. In other words, on the bank's balance sheet a loan asset converts to a vault cash asset. The interest part, the profit from issuing the loan that is added to the repayment, becomes earnings on the liability side, where it becomes an added liability to the owners. Insofar the interest part, being currency, is retained and not paid out as dividends, it will further add to cash reserves on the asset side in accordance with double entry balance sheet principles. It should be noted that the interest part is gross earnings and before it is added to the bank's balance sheet, costs of the underlying credit banking operations will be netted out.

If we now instead conceive the whole transaction conducted as electronic money, exactly the same mechanisms take place, except for that the returning money now is not physical vault cash, but electronic 'vault' cash, which in reality will appear as an entry to the bank's reserve account held with the central bank.

Balance Sheet Money Creation

Thus, in isolation the single bank cannot create new money above its direct access to reserves, but only transform whatever idle hoards it holds (CB reserves, vault cash) into transactional money (M1)⁹ by the credit process. Although it is indeed a fact that central banks in order to preserve the confidence in the payment systems will stretch itself to considerable length in order to accommodate banks' needs for extra reserves, this willingness is not unconditional and will normally require collateral and an interest charge in exchange for extended extra reserves.

To recapitulate, seen in isolation when a bank issues loans fully backed and paid out by reserves it directly holds, it only exchange items on its balance sheet, it doesn't expand it. This in itself can have expansionary effects on the economy through velocity enhancement but will have no quantitative effects.

We can thus see that the principal mechanism through which banks create money in the aggregate is by the loan 'netting out' process that swaps interbank liabilities. The netting out arises from spending of credit money on other banks' transaction accounts. In that case, a bank expands its balance sheet: the loan expand the assets side and a returning deposit from another bank's customer spending replenishes the liability reduction that occurred when the bank's loan customer withdrew the proceeds of the loan.

In the quantitative perspective, money is created when balance sheets in the aggregate expand, and likewise is only destroyed when aggregate balance sheets contract, such as in the case of creditor defaults. The changes in balance sheets link to another important aspect of the supply of money, which is velocity. Balance sheets of banks in modern economies are, it goes without saying, very busy affairs with large numbers of deposits and

⁸ According to this interpretation, currency when it circulates outside of banks are M1 but only M0, or base money, when it sits as vault cash in banks or in CB reserve accounts.

⁹ In order to simplify the argument; I consider all transactional money M1.

withdrawals made every day, and likewise, loans (incl. mortgages) being made and repayment flows returning to the bank in large numbers on a daily basis. The velocity of these flows can go up and down, which in the aggregate is an important macroeconomic function that aligns with the general economic cyclic changes. Therefore, one of the hallmarks of a crisis will typically be that the velocity of the money system slows down, which effectively reduces the nominal money supply, expressed as $P \times V$.

This makes it appear that we have returned to the neoclassical credit multiplier, but through a quite different path. The neoclassical credit multiplier was thought to be an exogenous function in the shape of the central banks control over reserve positions. What we have instead is a credit velocity multiplier that is an endogenous function of the netting out and credit recontracting that constantly occurs in an economy. This multiplier is mainly a function of default rates and the expectations this creates to the profitability of further expanding balance sheets. It is not particularly related to reserve positions, due to the almost automatic response from central banks when banks need extra reserves above what interbank markets can provide. This is to say that central bank's ability to control the credit function is rather weak and mainly arises from its ability to set the interest rates that influence bank's marginal cost of reserves.

The above underscores the problem with the adage that loans create deposits: while it theoretically is so, the view overlooks the mentioned crucial aspect of deposits: deposits that originate from a bank's own lending require fractional reserves, while other deposits add reserves. This distinction will be important when banks come under cyclical stress, etc.

It must therefore be concluded that banks' abilities to create money in the form of new credit that are added to transaction accounts cannot be separated from the institutional networks that govern access to reserves. The fact is that understanding the interplay between reserves and balance sheet expansions is central to understanding modern money.

We can also see that although a single bank on its own is limited in its ability to create money as credit, the banking system as a whole is not. But there is a contingent requirement: that balance sheets are expanded in some measure of symmetry. If not, crashes and deleverage, with contraction of money ripping through the whole system, is a risk when some overextended positions topple.

By that token, if we abstract from the possibility of seigniorage and velocity changes [and flows in foreign currency markets], it becomes clear that the aggregate quantity of money is not expanded by spending that fraction of bank account money that originate in deposits from non-credit channels. Only the mutually netted out part of new credit money has the capacity to expand balance sheets beyond already existing reserves and thus the aggregate quantity of money.

Finally, the deleveraging that occurs during crises involves two processes. First, a rise in defaults that directly destroy capital will impair balance sheets. Secondly, a slow-down of credit recontracting reduces the velocity of the credit processes and can lead to build-up of excess reserves in some quarters, of course dependent upon the extent to which returning loan payments first must replenish capital lost to defaults.