ON THE QUANTITY THEORY OF MONEY, CREDIT, 
AND SEIGNIORAGE
Gerasimos T. Soldatos, American University of Athens, Athens, Greece
Erotokritos Varelas, University of Macedonia, Thessaloniki, Greece

ABSTRACT
This paper argues the predictive power of the sectoral approach towards a quantity theory of credit is weak. A quantity theory of commercial-bank-seigniorage approach is proposed in its place. It suggests that the financial system may be held responsible for price and output fluctuations to the extent commercial bank seigniorage alters the stock of money in circulation. If not, the financial sector can become the source of instability by influencing profitability in the real sector through a Goodwin-type interaction. These trends could be countered by an interest rate rule based on deposit habits and on the deposit rate, and supplemented perhaps by a policy of influencing these habits and manipulating the deposit rate.

JEL: E3, E4, E5

KEYWORDS: Quantity Theory, Credit Theory, Commercial Bank Seigniorage, Instability

INTRODUCTION
Commercial banks enjoy the privilege of using their liabilities as a form of money; their lending decisions based on private interest affect in the aggregate the stock of money and thus social welfare. To maximize profit, banks tend to minimize their level of deposit reserve holdings, loaning out as much of their funds as possible at the market rate of interest. This has provoked repeatedly market instability and fluctuations in the money supply as documented, for example, by Friedman and Schwartz (1963) for the United States. Banks create money, “out of thin air” as the saying goes; money, chasing a given volume goods, decreasing afterwards individuals’ buying power, no differently than what would have been done with forged money.

The Austrian School of Economics claims that this is a win-win state of affairs for the banks and the government: The government tolerates the lending practices of banks in exchange for banks’ commitments to buy debt from the government to fund government’s ambitions (Myers, 2012). But, according to anthropologist Graeber (2011), for most of human history, money has been widely understood to represent debt. So, this is the environment in which the institution of commercial banking had to develop. The point is that nowadays most of the money in circulation is credit money. It is income earned by banks by expanding lending under fractional reserve banking, and labelled through the term “commercial bank seigniorage”; £ 50 billion in the UK, and about € 240 billion in the Euro area in 2013: “At 82–97% (depending on country and monetary aggregate) bank money today represents the lion's share of the money supply in public circulation.” (Huber, 2014, 1).

And, as already noted, credit is powerful in disturbing economic activity at least in the short- and medium-run (see also Iacoviello, 2015). Theoretically, the credit or debt approaches to money that have been advanced fall into a broader category of work which postulates that monetary creation is endogenous (Wray, 2004). Their origin is traced by Schumpeter (1954) back to Plato, and their modern formulation owes to Macleod (1889) and Mitchell-Innes (1914). Recently, Professor Werner (1992, 1997, 2009, 2012) tried to
make the credit viewpoint of money compatible with the quantity theory of money. The purpose of this paper is to look at this attempt closer and take it a step further if possible.

LITERATURE REVIEW

Professor Werner has been recommending a quantity theory of credit as a means of answering questions about *inter alia* defining money, the declining velocity, and financial crises. Fisher’s (1911) equation of exchange is broken down into one part referring to the real sector of the economy, and into a second part concerning the financial sector, as proposed originally by Keynes (1930). Werner is doing so focusing in essence on the propagation of the business cycle by the financial system, which was a matter of similar concern by the early Chicago version of the quantity theory of money, (see e.g. Simons, 1936). And, his approach has been found to have enough explanatory power.

Nevertheless, although breaking down the equation of exchange as proposed not only by Keynes but by Fisher himself too, can be empirically important, such a break-down is shown herein to be of limited predictability. The correctness, for instance, of the quantitative easing prescribed by Werner against the current recession, does not suggest predictive power since, this prescription follows directly from the current conditions of insufficient demand (see e.g. Soldatos, 2015): Give people money to spend and spur recovery.

From the viewpoint of the quantity theory of money, the real issue is the decomposition between central and commercial bank money rather than between a real and a monetary sector. The real issue is what happens within the monetary sector given the “Classical Dichotomy” but also given the correct version of this dichotomy; the version which distinguishes between the invariance of real variables to money changes and the invariance in the demand of goods: To hold the “demand-invariance” view is shown to be simply wrong (Hahn, 2002).

What happens within the monetary sector does affect the demand for goods. And, from the viewpoint of the credit theory of money, what happens depends mostly on commercial bank money. So, a successful marriage between the quantity theory of money and the debt theory of it requires the development of a quantity theory of commercial bank seigniorage. This task is undertaken in the next Section. Section 3 concludes this work by putting the topic under investigation into further theoretical perspective.

A QUANTITY THEORY OF SEIGNIORAGE

In this section, we first show the limited predictability entailed by Werner’s Sectoral Quantity Theory, we then proceed with a hypothesis based on central and commercial bank seigniorage decomposition, and we finish with some policy considerations.

The Sectoral Quantity Theory

Let total nominal transactions, $Y$, be the sum of the transactions $Y_1$ and $Y_2$ in sectors 1 and 2 of the economy, respectively:

$$Y \equiv Y_1 + Y_2 \quad (1)$$

And, let the total stock of money, $M$, consist of the stocks employed in these two sectors:

$$M \equiv M_1 + M_2 \quad (2)$$
Following Keynes (1930), the overall velocity of the economy, \( V = Y/M \), is the sum of the true, sectoral velocities, \( V_1 = Y_1/M_1 \) and \( V_2 = Y_2/M_2 \), with the former velocity being weighted by \( m = M_1/M \) and the latter by \( 1 - m = M_2/M \):

\[
V = mV_1 + (1 - m)V_2
\]  
   (3)

The total differential of (3) is: \( dV = mdV_1 + (1 - m)dV_2 \). Setting \( dV = 0 \) and solving for \( dV_2/dV_1 \), one obtains that:

\[
dV_2/dV_1 = -m/(1 - m) = -M_1/M_2
\]  
   (4)

which is depicted as the “iso-velocity” line \( AB \) in the space \( V_1 - V_2 \) of Figure 1; the slope of this line is:

\[
\varphi \equiv -m/(1 - m) = -M_1/M_2.
\]

Figure 1: The Sectoral Decomposition Diagrammatically

Next, substituting the sectoral velocities, \( V_1 = Y_1/M_1 \) and \( V_2 = Y_2/M_2 \), in (2) yields:

\[
M = (Y_1/V_1) + (Y_2/V_2)
\]

whose total differential gives when \( dM = 0 \) that:

\[
dV_2/dV_1 = -(1 - y)/y = -Y_2/Y_1
\]  
   (5)

where \( y = Y_1/Y \) and \( 1 - y = Y_2/Y \). The “iso-money” line \( \Gamma\Delta \), having slope \( \psi \equiv -(1 - y)/y = -Y_2/Y_1 \), is thus obtained in Figure 1. It may coincide with \( AB \), or cut it from above, as \( \Gamma\Delta' \) does, or below, as \( \Gamma''\Delta'' \) does. Two points need to be made now: Firstly, given \( \Gamma\Delta \), its intersection point with \( AB \), point \( E \), is the point at which both \( dV = 0 \) and \( dM = 0 \). Suppose, for instance, that \( AB, \Gamma\Delta' \), and \( Z \) is the case. \( Z \) is on a \( \Gamma\Delta \) below \( \Gamma\Delta' \); that is, given the \( V \) connected with \( AB \), the stock of \( M \) required to sustain the volume of transactions associated with \( \Gamma\Delta' \) is too small. Increasing this stock, the iso-money line passing through \( Z \) will start shifting upward, sliding from \( Z \) to \( E \). The second point needed to be made is that \( \psi \) gives the sectoral composition of \( Y \), and that a single \( \varphi \) can be consistent with many \( \psi \)'s as Figure 1 illustrates by having \( \Gamma''\Delta' \) and \( \Gamma''\Delta'' \) passing through the same point \( E \) on \( AB \). Figure 2 illustrates that a single \( \psi \) can be consistent with various \( \varphi \)'s, too.
Figure 2: The Vagueness of the Sectoral Decomposition

Therefore, a change in the sectoral velocities with unchanged the overall one, $V$, i.e. a pivot of $AB$ centered at $E$ in Figure 1, provides no information about a would-be change in the sectoral composition and volume of total transactions. Similarly, a pivot of $\Gamma \Delta$ centered at $\Theta$ in Figure 2, offers no insight as to would-be changes in $V$ and in sectoral velocities. Also, a change, an increase, say of $V$ given its sectoral composition, i.e. a parallel shift of $AB$ upwards, cutting $\Gamma'\Delta'$ and $\Gamma''\Delta''$ at $E'$ and $E''$, respectively, can be taken to mean one out of four things. Given the composition of transactions, it might mean either that $\Gamma'\Delta'$ has shifted to the right passing now through $E''$, or that $\Gamma''\Delta''$ has shifted upwards passing now through $E'$. And, given the overall volume of transactions, either that its composition has changed from that described by $\Gamma\Delta$ to the one captured by $\Gamma''\Delta''$, or the opposite. These are the four eventualities in case of a shift in $AB$ including a change in slope as well. Analogous remarks can be made about the shifts of $\Gamma\Delta$ given $AB$ in Figure 2, all pointing to the conclusion that without guidance from empirical observation towards the identification of that “eventuality” which is empirically relevant,… anything goes.

The Seigniorage Decomposition Quantity Theory

There do exist a few empirical investigations of sectoral velocities, notably by Selden (1961), McGouldrick (1962), Garvey and Blyn (1969), and Ireland (1991). To our knowledge, they are the only ones, and they too reflect implicitly or explicitly that the ultimate concern is the financial instability coming out of commercial bank seigniorage, in the spirit always of the early Chicago tradition. Indeed, the nexus between such seigniorage and real economic activity is what prompts in the first place the interest in disaggregation, in a quantity theory of credit. But, disaggregation, differentiation of the source of instability, is one thing, and decomposition, the explicit sectoral modeling, another. The focus is on the interaction between these sources and not on the sources per se. Towards this direction, note that the distinction between central and commercial bank seigniorage does reflect such a source differentiation. Consequently, once a quantity theory of seigniorage is advanced, once central and commercial bank seigniorage are incorporated directly in the equation of exchange as a sum, disaggregation becomes built in this sum. It is an approach to instability, which will not have to cope eventually with the vagueness characterizing sectoral modeling results as follows:

In the absence of commercial banks or the same, under a 100% reserve system, and central bank only seigniorage, $S$, we have that, $S = i(H/P)$, where $H$ is the monetary base, $P$ is the price level, and $i$ is the nominal interest rate. From the real-sector quantity-theory equation, $H = kPQ$ and hence, $S = i(kPQ/P) \Rightarrow S = ikQ$, where $Q$ is real income and $1/k = V$ is the velocity of circulation. In the presence of a commercial banking system benefiting from commercial bank seigniorage, $S = [i(1 - \rho) - r_n]D/P$, under a required reserve ratio $\rho \leq 1$ on deposited money $D/P = \lambda Q$ and under a deposit rate $r_n$, total seigniorage is the sum $S + S = (ik - \lambda r_n)Q$, given that $H = F + D$, where $F$ is cash (see e.g. Baltensperger and Jordan, 1997). Now, note that what banks do mostly is producing bank money out of their own bank
money given a token of $D$. That is, $PS$ is the output and money of the banking system, or in terms of the sectoral notation above, $Y_2$ and $M_2$, implying that $V_2 = 1$, $dV_2 = 0$, and $dV = dV_1$: An additional unit of bank money can be produced instead of rotating an already existing one, and any observed change in velocity comes out not from change in the financial sector per se, but from the impact of financial change on the real sector.

So, letting $M \equiv S + S = (ik - \lambda r_n)Q$ the overall quantity equation, $MV = PQ$ becomes:

$(ik - \lambda r_n)QV = PQ \Rightarrow [i(1/V) - \lambda r_n]V = P \Rightarrow (i - V\lambda r_n) = P \Rightarrow i = P + V\lambda r_n$

which in conjunction with the definition that $i = r + \pi$, implies that: $P + V\lambda r_n = r + \pi \Rightarrow$

$$V = (r + \pi - P)/\lambda r_n$$

(6)

where $r$ is the real rate of interest and $\pi$ is the inflation rate. Setting $dV = 0$ in the total differential of (6), yields that: $dr + d\pi - dP = (r + \pi - P)(1/\lambda)d\lambda + (1 + r_n)dr_n$. In view of this expression and given that from (6), $r + \pi - P = \lambda r_n$, the velocity, $V^*$, which is consistent with $dV = 0$ is:

$$V^* = (dr + d\pi - dP)/(r_n\lambda + \lambda dr_n)$$

(7)

Noting that $\pi = (P - P_{-1})/P_{-1} and \pi_{-1} = (P_{-1} - P_{-2})/P_{-2}, (7)$ may be rewritten as follows:

$$PP_{-2} - P^2_{-1} + P_{-1}P_{-2}[dr - V^*(r_n d\lambda + \lambda dr_n)] = 0$$

(8)

where subscripts “−1” and “−2” denote time lags. We have clearly a complicated cubic equation in prices, capturing price instability with constant $V$ at $V^*$. Setting $P = P_{-1} = P_{-2}$ in (8), which would be the case of price stability, the condition that: $dr - V^*(r_n d\lambda + \lambda dr_n) = 0$ has to be satisfied. It is a necessary but not sufficient condition, because zeroing the bracketed term in (8) gives that $PP_{-2} - P^2_{-1} = 0$, which again is a necessary but not sufficient condition by itself for price stability. Stability presupposes the satisfaction of both conditions. Setting for simplicity but plausibly $P_{-2} = 1$ in (8), the latter condition becomes: $P = P^2_{-1}$; it is a condition for geometric price reduction.

Next, setting $dM = 0$ and $dV = 0$ in the total differential of $M = [i(1/V) - \lambda r_n]Q$, one obtains that $Qdi + idQ = V^*(r_n d\lambda + \lambda dr_n)Q + V^*\lambda r_n dQ$, or letting $dQ = Q - Q_{-1}$ and $Q/Q_{-1} = 1 + g$:

$$Q[(1 + g)[i + di - V^*(r_n d\lambda + \lambda dr_n) - V^*\lambda r_n] - (i - V^*\lambda r_n)] = 0$$

(9)

According to (9), given $M$ and $V$, the composition of $M$ alone does provoke output instability too, as expected, because keeping $M$ and $V$ constant, price changes should be offset by output changes. This can be seen by solving (8) for $V^*(r_n d\lambda + \lambda dr_n)$ and inserting the resulting expression of $P$’s in (9). Note that under price stability, that is, under $dr - V^*(r_n d\lambda + \lambda dr_n) = 0$ and $d\pi = 0$, and given that $di = dr + d\pi$, (9) becomes: $gQ(i - V^*\lambda r_n) = 0$, which implies that steady growth under price stability presupposes an $i$ equaling to $V^*\lambda r_n$. This interest rate rule that $i = V^*\lambda r_n$, takes the place of the condition of geometric price decline, needed for price stability beyond the condition that $dr - V^*(r_n d\lambda + \lambda dr_n) = 0$. $V$, $\lambda$, and $r_n$ are not quantities that change every day. $V$ and $\lambda$ reflect inter alia consumer habits while $r_n$ changes sporadically. Consequently, in practice, the interest rate rule is really one about interest rate stability; price stability would be indeed corroborated by interest rate stability. This is what the condition $dr - V^*(r_n d\lambda + \lambda dr_n) = 0$ is about too, because if in practice $d\lambda = 0$ and $dr_n = 0$, this condition amounts to $dr = 0$, which when coupled with $d\pi = 0$, implies that $di = dr + d\pi = 0 + 0 = 0$. It appears that in practice, one should be the policy rule, namely that $di = 0$. 
It also appears that the stricter, Friedman’s rule that $i = 0$ would apply only if $r_n = 0$; both $S$ and $S$ would be zero in this case. This of course would be in the spirit under which this rule was advanced initially assuming away commercial banking: namely, elimination of central bank seigniorage. Incorporating commercial banks into the discussion, the elimination of commercial bank seigniorage too, would come as the natural extension of Friedman’s rule. But, would such an extended Friedman rule be sensible policy-wise beyond the weaker requirement that $di = 0$? According to (8) and (9), failure to abide by this weaker rule does produce price and output instability even if $M$ and $V$ are held constant; but much more so when $M$ changes since $M = (ik - \lambda r_n)Q$. And, responsible for this change should be held the financial system given that the bulk of $M$ is easily expandable commercial bank seigniorage. It seems to the authors that under these circumstances the extended Friedman rule should be a must if of course it was decided to be followed. Nevertheless, the $k$-percent rule advanced by Friedman (1960) too, is more practical and a policy of $di = 0$ should be seen as its natural companion.

Instability and Monetary Policy

Bank money has own life for which discretionary monetary policy cannot do much. For example, a policy keeping $M$ constant under a Goodwin-type interaction between bank and firm profitability, would be ineffective as follows: Total firm revenues are $PQ \equiv Y \equiv MV$, and are made possible through lending $\Pi_b$ to pay capital and labor expenses in such a manner that:

$$d\Pi_f/dt = \Pi_f(Y - \beta \Pi_b)$$  \hspace{1cm} (10)

where $f$ and $b$ designate the firms and the bank sector, respectively, $t$ is time, $\beta$ is some positive constant reflecting the availability of lending, and $\Pi$ captures profits. That is, the rate of change of the profit of firms depends on whether lending is enough to keep sustaining a given volume of transactions. Variations in lending that leave $Y$ constant by manipulating prices and quantities on the part of firms, become fully mapped into profit variations.

The profit of the banking sector is the bank money loaned to the firms, and based on $D = \xi d\Pi_f$, i.e. $(i - r_n)\xi d\Pi_f$, in such a manner that: $(i - r_n)\xi d\Pi_f = \beta \Pi_b + \xi$, and that:

$$d\Pi_b/dt = \Pi_b[(i - r_n)\xi d\Pi_f - \xi]$$  \hspace{1cm} (11)

The term $\xi = (1 - \beta)\Pi_b$ is a constant, reflecting normal profit by the banks, covering their opportunity cost; it is the profit that would keep them in operation under the worst of circumstances. In effect, bank money, bank lending, and commercial bank seigniorage become synonymous to supernormal bank profit. Equations (10) and (11) are Lotka-Volterra ones, having firms being the prey of predating banks for bank profit beyond the normal one. The two critical points for stationariness are: $\Pi_f^* = \xi/(i - r_n)\xi$ and $\Pi_b^* = Y/\beta$ in connection with a cycle of period equal to $2\pi i/\sqrt{\xi}$ and with the cycle of the profit of firms leading by (1/4)th of this period, where $\pi = 3.14159...$ Now, having $S$ changing over the cycle to be keeping $M$ constant would not alter this course of things given that $(i - r_n)\xi d\Pi_f = S$. Such a policy would be useful only to the extent that $V$ does not respond adequately to the cycle as it did happen with the Great Crash in 1929.

CONCLUDING COMMENTS

To sum up, Werner’s sectoral approach towards a quantity theory of credit is too vague in its predictions. A quantity theory of commercial bank seigniorage approach is proposed in its place, arriving at the conclusion that the financial system may be held responsible for price and output fluctuations to the extent
commercial bank seigniorage alters the stock of money in circulation considerably. If not, the financial sector can become the source of instability by affecting profitability in the real sector through a Goodwin-type interaction. These trends could be countered by an interest rate rule based on deposit habits and on the deposit rate, and supplemented perhaps by a policy of influencing these habits and manipulating the deposit rate.

Many will have noticed a contradiction in the term “quantity theory of (commercial bank) seigniorage”. The quantity theory of money wants money to be exogenous, with independent causal role, whereas commercial bank seigniorage refers to money endogeneity. Such has been the standard doctrinal standpoint. Thirlwall (1999) distinguishes three versions of the quantity theory: the equation of exchange, the income quantity theory, and the cash balance equation. But, none of these is actually a theory. They are all equations, alternative expressions of a single analytical tool adapted to the needs of the particular hypothesis explored each time. A single tool, when money is seen from the viewpoint of financial products within the accounting structure of a banking firm; money defined according to flows of income and expense and stocks of assets and liabilities as, for instance, Fixler and Zieschang (1998) do.

And, a tool does not form a hypothesis, a theory by itself; it helps instead establish one. Keynes and monetarism, for example, wanted it to be a hypothesis per se, but Old Chicago used it to rationalize “A Program for Monetary Reform” (see e.g. Douglas et al., 1939). The source of the confusion has been Hume’s (1987 [1742], II.IV.2) view that: “Were all the gold in England annihilated at once, and one and twenty shillings substituted in the place of every guinea, would money be more plentiful or interest lower? No surely: We should only use silver instead of gold.” As Diaz-Gimenez and Kirkby (2013, 2) note, the keyword in this expression is “at once”, which only in the long-run is expected to hold as a reflection of an average trend: “When central banks conduct monetary policy, changes in the quantity of money are not introduced evenly and ‘at once’… [since] money is injected into… typically the banking system, and it spreads out gradually from there.” And, as a matter of fact, there can be no short- and medium-run theory of money at all, since: “Money is means to effect transactions and savings; what will be used as money depends upon such a miscellany of factors that no ‘theory’ can be expected to emerge” (Rashid, 2001, 1).

Finally, note that the view of commercial bank seigniorage employed herein is in line with the Fixler-Zieschang accounting approach to the definition of credit money: Money stemming from the microeconomic theory of financial firms and of household consumption of financial asset services, by taking “the total sales of financial institutions to be the net interest income on “produced” asset and liability products” (Fixler and Zieschang, 1998, 1): This consistent with any of the three quantity equations.

REFERENCES


ACKNOWLEDGEMENTS

We are grateful to two anonymous reviewers and the editor for useful comments and suggestions. Any remaining errors are our own.

BIOGRAPHY
