Journal of Prices & Markets (2013) 1.1: 17-30

The Quantity Theory of Money

David Howden[1]

Abstract: For an innocuous statement based on a trivial tautology, the quantity theory of money is sorely battered. This paper has three goals. First, it exposes the various flavours of the quantity theory as special cases of a simple application of the law of diminishing marginal utility. Second, it provides an overview of some typically controversial aspects of the quantity theory. Finally, it reformulates the quantity theory in light of these now resolved controversies. Although I use the term "quantity theory of money", by the end of this article I reformulate the concept as an "exchange theory of velocity".

Key Words: Quantity theory of money; velocity; bank-created credit; credit; deleveraging

1 St. Louis University – Madrid Campus. Email: dhowden@slu.edu. A version of this paper was presented at the Austrian Economics Research Conference, March 22, 2013, at the Ludwig von Mises Institute in Auburn Alabama. I wish to thank participants of the conference, Joe Salerno, John Cochran and an anonymous referee for improving the arguments herein. Since I have listened to only some of these comments, it is with some trepidation that I take sole responsibility for all resultant errors.

Although I have chosen "The Quantity Theory of Money" as the title for this article, I do not particularly like it. The name and the theory, perhaps the most famous theory in all economic science and definitely the most famous to be formalized in the 20th century, carries with it much baggage. This article takes its title to keep some semblance of consistency in terminology, but as should be clear the theory developed by the end will bear only superficial resemblance to the more accepted doctrine of the quantity theory. More correctly, by the end of this paper we shall see that the traditional formulation of the quantity theory of money, presented in its various guises, is but a special case of a broad theory of prices, unduly restricted by some unnecessary and detrimental assumptions.

All debates and controversies surrounding the quantity theory of money (QTM) distil to ill-defined terms and concepts. The equation of exchange, the logical statement through which the QTM emerges, is tautologically true – both by way of its interlocking definitions and the way that its terms are defined (Yeager 1994: 159-60). As a simple accounting identity, the nominal value of spending over a period of time must equal the volume of money spent to settle these transactions. Problems with the application of this simple insight have traditionally come from poorly explained causal relationships joining the terms in question.[2]

The present paper starts from the ground up. It first defines the terms in question and which heretofore have received relatively scant treatment compared to the theory's conclusions. In defining terms this reformulation, for lack of a more original verb, of the QTM shares much in common with existing presentations.

One area of departure in the present paper is the focus on the "velocity of money". As the lone unobserved variable in the equation of exchange, velocity has been typically treated as a balancing item – the necessary product when one divides nominal spending by the money supply. Though still treating velocity as an unobserved variable, this paper redefines it in such a way that it is not subject to relegation as a place holder in the general theory. We will also see that changes to money´s velocity have a greater degree of bearing on other variables – both independent (e.g., certain components of the money supply) and dependent (e.g., credit expansion and the level of nominal spending).

The QTM is sorely battered, especially so as this recession wears on. Its detractors have no lack of fodder for their attacks. The rapid expansions of the money supplies of various nations over the past few years have resulted in a steadiness of inflation and inflationary expectations and have had little affect on nominal spending. Just as John Maynard Keynes developed the marginal propensity to consume as a backlash against the QTM to explain the dramatic drop in incomes and prices during the Great Depression, so too does the current malaise provide an opportunity to provide an alternative to a damaged piece.

The Quantity Theories of Money

The four famous letters in the equation MV = PY, are among the first that the budding economist learns. No sooner than he learns the identity, however, is it likely that he sheds the term "equation of exchange" from his memory to replace it with the "quantity theory of money". N. Gregory Mankiw's widely popular intermediate macroeconomics text, for example, introduces the equation of exchange to many young economists (Mankiw 2009: 86-89). After devoting three pages to explaining the variables, Mankiw makes the jump to assuming velocity is constant and thus providing the foundation for the more common quantity theory of money. This subsequent theory, although sharing the same foundation as the equation of exchange, is a causal statement explaining inflation by changes to the supply of money. After a brief formulation of the aggregate demand function in terms of the equation of exchange (Mankiw 2009: 269-71) the remainder of the book couches all discussions of the equation's relevance in terms of the quantity theory of money.

Broadly speaking there are two ways to express the equation of exchange. Both make similar statements, though in different ways. Both rely on a vacuous conceptualization of velocity to act as a placeholder variable to make the relationship between money flows and income balance.

Irving Fisher's version of the QTM started from the formulization of the truth that over any period of time, the volume of money expenditures must equal the sum of cash payments received (Fisher 1911). The former is the

2 Laidler (1991: 302-04) argues that there are also ideological controversies in the development of the QTM, as authors used it as a platform for policy prescriptions. Notable among these was the Monetarist ideal in need of a theory linking money supply growth to inflation, or Joan Robinson's (1970) argument that inflation is everywhere and always a political phenomenon.

product of the quantity of money *M*, and how quickly it circulates to settle transactions *V*. The latter is determined by the gross number of transactions occurring *T*, at the average price of each transaction *P*. Fisher's income approach to the equation of exchange written as $MV = PT$, is not the QTM, though it is the accounting identity that forms the basis for the theory.

The QTM emerges from this foundation once one makes some basic assumptions about the nature of the variables and their interactions with one another. Thus, if one assumes velocity to be constant than inflation becomes always and everywhere a monetary phenomenon.

The "Cambridge" or transactions approach to the QTM argues that if any economy has a given stock of money, the purchasing power of this stock is determined by the demand to hold it. The first and perhaps most precise formulation of this version claimed that the demand to hold money would vary proportionately with nominal income (Keynes 1923). Altering some variables to change nominal spending into nominal income as the product of real national income *Y* and some appropriate price level *P*, the product must equilibrate with the stock of money *M* as,

$$M = kPY$$

which can be rewritten as

$$M(1/k) = PY.$$

The left-hand side expresses a money supply function which must by necessity result in the money demand expressed on the right-hand side.

The similarities between the income and transactions versions are more than superficial. Provided there is a stable relationship between the volume of transactions and real national income, there will also be a stable relationship between Fisher's transactions velocity of circulation *V*, and the Cambridge income velocity $1/k$.[3]

Indeed, both formulations say the same truth – the only distinction is in defining the terms. Although both denoted as *M*, the money supplies in question are distinct (Friedman 1970: 200). Fisher's transactions approach makes use of an *M* primarily concerned with money for transactions purposes, and the most important quality of money is that it is transferred. The income version places emphasis on money held. Fisher is concerned with all transactions in the economy, while the income approach concerns itself more narrowly with only those generating final income. Likewise, the price levels suggested by each *P* differ in that the former version relies on an abstract price level for all goods transacted for, while the Cambridge version looks at prices for only finished goods, the sales of which generate income.

If three of the variables change, by definition each of the velocities will also differ. Fisher's *V* is a residual that equilibrates the volume of money circulating to settle transactions with that stock of money broadly defined as being used in payment – it is a transactions velocity. The income approach shares the similarity that *V* is a residual, though it serves to equilibrate the amount of money directed at generating only income-related output, and thus it represents an income velocity.

It is not that either approach is any more correct than the other: they are both simple tautologies. The vacuous nature of each approach should be apparent. Defining the terms without regard to some basic fundamentals of what the essence of each term results in an empty conclusion. Consider that

> [w]e can readily imagine a "chairs" version of the equation of exchange. In $CV_c=PQ$, P and Q would be the same as before, C would be the number of chairs in existence in the country on average during a year, and V_c would be the "velocity" of chairs, meaning the ratio of nominal income to the number of chairs. Thanks to interlocking definitions, $CV_c=PQ$ is just as formally valid as MV=PQ; but because of facts about how money functions that are not also true of chairs, the money version of the equation has a usefulness that the chairs version lacks. (Yeager1994: 160)

Yeager's illustration demonstrates the point, yet also suffers the same deficiency as the traditional renditions of the QTM. As simple tautologies they are unassailable. However, it is not that money is special that makes the traditional QTMs more appealing than a chairs version. The QTM has always been developed without much mind for what money actually is, and instead focuses after the fact on what money must necessarily be in order to satisfy the equation. For example, in both versions above

3 Indeed, in an early formulation of the Cambridge version, Pigou (1917: 174) noted as much, remarking that "It is thus evident that there is no conflict between my [Cambridge] formula and that embodied in the quantity theory."

	Included in TMS	Excluded from TMS
M1	(1) Currency	(1) Traveller's checks
	(2) Demand deposits	
	(3) Other checkable deposits, e.g., NOW accounts	
M2	(1) Savings deposits (including MMDAs)	(1) Money market mutual funds
	(2) Small-denomination time deposits	
Memorandum Items	(1) Demand Deposits Due to Foreign Commercial Banks	
	(2) Demand Deposits Due to Foreign Official Institutions	
	(3) U.S. Government Demand Deposits and Note Balances	

Table 1: Components of the true money supply
Source: Salerno (1987)

the initial emphasis is on defining the nominal sum of expenditure on the right-hand side of the equation, whether nominal income or gross transactions. It is from that point that money is appropriately defined and then velocity is introduced as a placeholder.[4]

It may strike the reader as strange that the quantity theory of *money* leaves the definition of money until the end. If the QTM is in need of reformulation, it must be pursued so as to make clear at the outset why each item of the reformulation is necessary and how it interacts with the other variables in the theory.

Agreeing on Terms

Present depictions of the QTM suffer a similar infliction – by defining their terms only loosely, they result in a theory which, although logically quite valid and unassailable, is of such a special case as to have almost no bearing on the monetary world. I shall start by defining what money is used for, and why it is held. In this way, the subsequent QTM I will develop will abide by "Wallace´s dictum": namely, that money should not be a primitive in a monetary model (Wallace 1998). By first defining money and its uses, I will then define more narrowly the other three terms that must interlock to form the theory.

Money

Ludwig von Mises (1949: 14, 249) argued that money is held only to satisfy felt uncertainty.[5] Thus, if an individual was certain of all future expenditures – both in terms of time and amount – he would have no need to hold money and incur its opportunity cost. In making this argument, Mises took the opportunity cost for granted without demonstrating what that cost would be.

4 Of course, Yeager must realize this as he notes that "[o]ne might quibble over exactly what counts as a chair, just as over what counts as money, but such quibbles would be relatively peripheral to the logic and usefulness of either equation (1994: 160). Bagus and Howden (2012a) bring up a similar point, noting that a chairs version of the QTM is not strictly comparable to the money version, despite being logically consistent – the number of chairs circulating confers a direct utility from their use value while money, and particularly fiat money, offers no such affect on utility.

5 Of course, Mises did not also focus narrowly on money in only this role, but also more broadly as a unit of account, especially in his equilibrium construct, the evenly rotating economy (Mises 1949: 244-51). On this point see Howden (2009: 8fn8).

Money is a unique financial asset. It is the only asset which redeems on-demand and at par value. Money is the only asset which serves as the final means of settlement for contractual obligations and is generally accepted as such by the economic community. By these standards, we can define the money supply that is available to provide final payment for all purchases.

The common M measures of the money supply include some assets which qualify as money as per our criteria above, and some which do not qualify. Following Rothbard (1963: 87-91) and Salerno (1987), the following table outlines those types of "money" which are usually grouped into each M category, and whether they will be included in our definition of the "true money supply" (TMS). Notably, traveler's checks are excluded as money because they are not the final means of settlement, and money market mutual funds do not count as they are not necessarily instantly redeemable, fixed statutorily at par value, or the final means of settlement. (The reader interested in further details surrounding the inclusion or exclusion of an item may consult Salerno 1987.) It is commonly said that "money is as money does", and serving as a means of payment is not the only role money serves. Per Mises (1949: 249) money is also that asset which is held as an uncertainty hedge. This form of money has an appearance of idleness as it is typically represented as a deposited sum which is only used by an individual upon an unforeseen event.

Thus money serves two distinct roles and it is useful to distinguish between them when defining the money supply. One common division is made between holding money for reservation purposes and exchange purposes, as in Rothbard (1962: 756-62). The former is held as a hedge against perceived future uncertainty; the latter to facilitate payments. Although it is difficult to disentangle the two motives for holding money, there are some cash balances that are clearly held to serve one role rather than the other.

Vault cash held by banks, for example, is a sum of money which is only held for precautionary motives. Banks do not use this sum to facilitate payments per se, and as such we can treat it as a quantity of money not used for settlement of exchange obligations.

We can separate the total money supply into two categories, 1) that sum which functions as the means of settlement and 2) that sum held to ease felt uncertainty.[6] The former is not held but circulates continually through the economy as it serves to settle transactions. The former has no circulation, although its level will be adjusted as felt uncertainty changes or as an unforeseen expense occurs that requires financing.

In sum, the total amount of money available at any period to facilitate transactions is the true money supply less the reservation demand for money: $TMS - M_R$

Quantities

That sum of money which circulates to settle transactions has a partial counterpart in the quantity of goods produced in the area over which this money circulates. Final goods consist of final consumers' goods *C* and final capital goods *K*. The common gross domestic product figures are summaries of these final output levels, whereby consumers´ and capital goods are also included with the level of government spending and net exports.

In addition to expenditures on final goods in an economy there is also much expenditure on goods in process. Menger (1976: chap. 1) distinguished between different "orders" of capital as a way to differentiate final output from those goods produced but still some distance away from final consumption. In his terminology, higher order goods are those furthest from completion for final use while lower orders are those closest to final use. (Use in this case can be for either consumption or investment use, depending on whether the good is a consumers´ or a capital good.) Goods of the lowest order, the zeroth order, are those available for final use (i.e., $C + K$).

For our purposes we will consider that money settles transactions for all: $(C + K)$ plus all goods of a higher order, or intermediate goods, *N*.[7]

Besides monetary expenditures on goods, we can also consider that money is used to settle debt transactions.

6 This separating of the demand for money into the demand to *hold* money as a reservation fund, and the demand to *trade* money to facilitate payments does much to rectify the misgiving of the quantity theory noted by Wallace (1998: 21fn3), "[W]ho is holding and trading the money in the quantity equation?"

7 Traditionally, intermediate goods were called "circulating capital", a term I reject here because of the confusion that may arise by calling an unfinished consumers´ good a type of capital. In a similar vein, I reject the term accepted by the United Nations System of National Accounts, US National Income and Product Accounts, and the European System of Accounts, of "intermediate consumption" for the similar confusion created by referring to unfinished capital goods as a form of consumption.

Debt payments have typically been excluded from the equation of exchange on the basis that they represent a wealth transfer from one party to another. We include them here as they are just one means through which one can spend his money income on. Likewise income can be used to facilitate new equity purchases E_t.

Thus the total of transactions that money can be used to facilitate the payment of includes five categories – consumers and capital goods produced in each period, intermediate goods still in progress, any net debt repayment, and any net purchases of equities.

The common gross domestic product figure captures the first two of these components. GDP limits itself, however, by not including the intermediary goods produced and as such represents income earned in a period but not the total of all transactions. Broader based figured such as gross output or gross domestic expenditure, both of which include all intermediary transactions as well as final ones, are a much more accurate representation of total money expenditure in an economy during any period of time (Skousen 2012). These two figures too are deficient for our purposes, however, as they lack the inclusion of money expenditure on net debt repayment and equity purchases.

Thus the sum of all monetary transactions in the economy is given as:

$$C_t + I_t + N_t + (\text{net debt repayment})_t + (\text{net equity purchases})_t$$

Prices

Of the variables discussed so far, prices are the easiest to conceptually define yet the most difficult to integrate into the analysis. Each transaction has a price. In general these prices are determined in one of two ways. They may transpire at par value, that is, some pre-defined value not subject to change. Alternatively prices can be established at market value, that is, as per the whims of supply and demand at any given time and very much subject to change.

Since every quantity transacted for must have an associated price, we see that debt transactions are settled at par value while the sum of GDE components and equity transactions is determined at market.

Par value is conceptually easy to analyze, and as it is not subject to change by market forces there is no change in these prices from period to period. To speak of price inflation, for example, is of no meaning with debt-based transactions.

Market prices must be summated in some way to obtain an average price at which all market-value transactions take place at. This exercise is fraught with peril, as numerous critiques concerning the relevance of price level computations makes clear (Anderson 2001). Still, the concept of the general price level is not offensive and indeed it can be concretely defined within the context of the total of nominal spending which has occurred over a time period.

When combined with the transactions occurring in the economy above we find that total expenditures equals,

$$p(C_t + I_t + N_t) + (\text{net debt repayment})_t + (\text{net equity purchases})_t$$

where p is some sufficiently designed and weighted average price level for all goods and services transactions.

Since p itself is a contentious issue, it may prove instructive to just reckon all transactions not in specific quantity and price terms, but as the resultant product of money expenditures by way of some aggregate spending figure. Thus, as GDE is just the current money value of all expenditures on consumer, capital and intermediate goods, we can rewrite the above as:

$$GDE_t - \Delta L_t + \Delta E_t,$$

where ΔL_t represents the change in the total level of indebtedness in the economy and ΔE_t represented net new equity purchases, both during some time period t. A positive ΔL_t implies that the total amount of debt is increasing (i.e., the economy in the aggregate is leveraging) while a negative value implies that the total amount of debt is decreasing (i.e., a decrease in the degree to which the aggregate economy is levered), and thus requires some monetary expenditure to cover those loans not re-backed by fresh debt issuances.

Velocity

Finally we reach the velocity of circulation variable. Velocity can be viewed in one of two ways. In typical expositions of the equation of exchange it is an equilibrating variable, the result of nominal spending divided by the money stock. In this way it also contains an error component (Friedman 1987). While there is little wrong with this approach to defining the "velocity" of money, it is not very fecund: it serves only to satisfy the other terms in the equation.

An alternative approach is to recognize that velocity is a real variable. By real I do not mean to imply that its value hinges solely on non-monetary factors. Instead I mean that it is reasonable to refer to the rate at which each unit of money circulates as its velocity. It really does exist outside of the narrow confines of economic theorizing.

In order to make it a meaningful concept, however, there are some barriers to address.

The first is that velocity is an unobserved variable. The fungibility of money implies that no one unit can be tracked easily to see how many times it changes hands. This is not only an applied problem with currency, but it is impossible given the transfer of perfect money substitutes such as money deposits. Any reckoning of money's velocity of circulation must be made in a roundabout way.

Second, the velocity of circulation will critically hinge on what role money is performing. Debates and controversies surrounding the applicability of meaning of velocity in past renditions of the equation of exchange have often centred on this point – what is money and why is it used?

This barrier has already been addressed because we have not used an ad hoc definition of money. Money for our purposes uses both its roles – that of facilitating exchanges and as being held as an uncertainty hedge. As a result, velocity is the ratio of total expenditures to the stock of money available to settle transactions:

$$V = (GDE_t - \Delta L_t + \Delta E_t) / (TMS - MR).$$

Taken in such a way, velocity is a half-way point between its more typical definitions. On the one hand it is a transactions velocity, like in the income tradition, as it looks at the necessary speed at which money must circulate to facilitate all monetary transactions. On the other hand it has an affinity to the Cambridge tradition as it incorporates the demand to hold money as a reservation balance.

Accounting for Unbacked Debt

Debt has typically been excluded from various equations of exchange because it represents a wealth transfer and not an outright use of purchasing power. I have included debt repayment as a use of money above, and as such one might also note that a corresponding change to the money supply should be made to include such a factor – if debt affects the right-hand side of the equation is it not reasonable that it too should affect the left-hand side also?

This is not an unreasonable claim, but has heretofore been addressed unsatisfactorily. Traditional expositions of the QTM exclude debt transactions for one of two reasons. On the one hand they do not represent the final means of payment. Thus, even though a good or service can be "purchased" by incurring a debt, this is just delaying the inevitability of repayment. Accounting for debt-based transactions is unnecessary as the use of debt just shifts the period of payment, but does not significantly alter the fundamental nature of eventual payment.

On the other hand, it is commonly viewed that debt-based transactions do not represent gains in purchasing power. Rather, they are a strict transfer from one spender to another in the economy (as in Salerno 2006: 49). This is true for some though not all debt obligations. In particular, there are two lending operations that do not entail a sacrifice in expenditure by the "lending" side of the exchange.

The first case we shall look at is lending from foreign sources. Financial inflows through the current account are the result of a foreigner lending money or buying a financial asset in order to finance domestic expenditure. The current account represents a funding source financed through debt that does not have an offsetting decrease in expenditure by someone in the domestic economy. (Though there is an expenditure decrease in the foreign economy by the lender.) As a result, negative current account balances act as a "free lunch" of sorts. They are free in the sense that a foreigner has enabled someone in the domestic economy to spend income which has not been lent through some other member of the domestic economy. The nature of this free lunch is, however, fleeting. Positive current account balances will reverse this state of affairs, and imply that

a domestic citizen is using his own savings to finance a foreign expenditure. The renunciation in expenditure in the domestic economy will result in an increase in funding for expenditure purposes in a foreign economy.

As a result of current account flows we can see that the means of payment available to settle a transaction is not limited to the stock of money in the economy at any given point in time. It is also comprised of expenditure "gifts" provided by foreigners on the current account. Domestic individuals will have to repay these "gifts" at some point in time, which will result as the current account turns positive and the flow of funding turns outward.

The second "lending" operation that we must account for is the maturity mismatch that results from bank-created credit. The fractional-reserve banking system makes use of deposited funds to finance its lending operations. These deposits, however, are not the bank's to use. They are the result of a conscious decision on the part of depositors to hold a sum of money as an uncertainty hedge (Huerta de Soto 2006; Bagus and Howden 2009; 2012b; 2013; forthcoming).

I have used quotations above when referring to the nature of fractional-reserve bank lending practices because it is distinct from usual lending activities. All other loans in the economy are enacted through a temporary renunciation on the claim to an asset. When an investor purchases a $1,000 bond, for example, he gives up the use and availability of those thousand dollars for the maturity of the bond, and the company borrowing the sum gains the use of the same.

Note that this renunciation of the use of the lent sum is not apparent if the loan is financed through a deposit. The depositor may not be actively using his deposited funds at any given time, though he is still using them in the sense that he is awaiting an uncertain event to make their use necessary (Bagus and Howden 2013: 239-41). This original step in the fractional-reserve money creation cycle may not seem insurmountable to the traditional variants of the QTM because there is no spending taking place with the original deposit at the time in question (i.e., a deposit only represents money that *might* be spent in the future contingent on a now uncertain event).

Subsequent iterations of the fractional-reserve cycle are of greater consequence. The original loan financed with a deposit is itself ultimately deposited in an account. From there a fraction of it will fund a subsequent loan, and the usual fractional-reserve credit creation process proceeds. Each of these iterations represents an expenditure financed with a loan which did not entail a renunciation of expenditure on the part of the "lender" (who was, after all, the original depositor).

As a result, during any given period an expansion in the amount of bank-created credit will represent a "free lunch" in much the same way as funds entering the country on current account. As such, during any given time period the current account balance CA_t and the amount of new bank-created credit B_t must be included in the means of payment use to settle all transactions. Also note, however, that there is no concept of "circulation" with either of these funding sources unlike is the case with money.

Putting it all Together

We are now in a position where we can put the terms together to construct a new equation of exchange.

On the payments side of the equation we find that

$$(TMS - MR)V - CAt + \Delta Bt.$$

A negative current account balance represents a positive financial account inflow, implying a "payment" for goods and services not stemming from a domestic source or representing a domestic transfer of purchasing power. ΔB_t represents the change in bank credit over some time period t, while the term $TMS - M_R$ represents the amount of the total money supply available for transactions motives less the amount held to satisfy the reservation demands.

Since the expenditure side is just the sum of debt repayment, new net equity purchases and gross expenditures (represented by GDE), the complete equation of exchange becomes:

$$(TMS - M_R)V - CA_t + \Delta B_t \equiv GDE_t - \Delta L_t + \Delta E_t .$$

The right-hand side of the equation includes all transactions that require money to settle. The left-hand side implies that payment for such services comes not just from the amount of available money set aside to satisfy peoples' transactions demand circulating at its own velocity V, but also the amount of unbacked funding in the form of the current account and new bank-created credit.

At this point the equation is still stated as an

equality. What is lacking is a dose of causality to point to how the equation should be rewritten, and which variables are dependent on or independent of each other.

Traditionally, economists have treated the money supply as being a given in the QTM, mostly owing to the fact that it is exogenously fixed in a commodity standard or central bank controlled fiat regime. In our rendition, the relevant money supply for settlement purposes is chosen by individuals. This is a result of the choice to hold money to satisfy the reservation demand, which thereby reduces the portion of the TMS which can circulate for transactions purposes.

In fact, there are four avenues through which an individual can direct his money income: 1) money can be held to satisfy the desire for an uncertainty hedge, 2a) money can be used to facilitate the purchase of consumers and capital goods in the present (both of which are a form of "consumption" expenditure to the extent that they confer a benefit in the present), or 2b) money can facilitate the movement of intermediate goods in a production process, which will confer a benefit in the future, 3) money can be used to settle the payment of services rendered in the past and financed through a debt, and finally, 4) money can be injected into equity markets by stock purchases, in effect purchasing a claim on future profits. Thus money is a separate class of goods used to facilitate the payment of past, present and future services through its role as medium of exchange, in contrast to the ability of consumers goods to confer benefits in the present or equity investments, capital and intermediate goods to confer their benefits in the future.[8]

The ability to pay for services with credit reduces the need for an individual to hold a sum of money to satisfy his reservation demand (Salerno 2006: 48). As credit, especially short-term credit, enables a funding source in not just routine but also emergency situations, an individual is able to direct a greater portion of his money supply to facilitating transactions and dedicate a smaller amount to fulfilling his need for a security hedge. Indeed, Rothbard (1962: 826-27) refers to very short-term credit as a form of "quasi money" because of its ability to substitute for an individual's cash balance held for the reservation demand.

Taking this two-pronged approach to defining the money supply by dividing cash balances into transactions and reservation demands does much to rectify the immediate problem in the equation of exchange that Laidler (1991: 296) identifies, namely, how best to define money. Previous attempts to define the relevant money supply relied either on an overly narrow or too broad definition of money. One way to solve the apparent problem of the indeterminate nature of the monetary stock was to define it as the aggregate whose demand function is mostly stable (Laidler 1969). The relevant monetary stock for transactions purposes is very malleable and unstable, and is determined not only or even principally by the supply of assets serving as money but rather by the amount that people desire to spend after satisfying their reservation demand. Indeed, due to its role in eradicating the continual threat of felt uncertainty, satisfying the reservation demand for money may be the first decision an individual makes with his income prior to assessing the additional expenditure avenues he can explore (Bagus and Howden 2013: 236).

Finally, I wish to comment on price formation. It is not the flow of spending that determines the price level, and neither is it some exogenously determined level of output which is available to direct this spending stream to (Salerno 2006: 51). Rather, it is money prices and the four spending options available to individuals – 1) "purchasing" a reservation stock of money, 2) purchasing consumers, capital or intermediate goods, 3) debt repayment, or 4) new equity purchases – that determine the stream of money spending. This causality is perhaps the starkest difference between the QTM presented here and the more typical versions of it (our version finds affinity in this sense with the "theory of money prices" found in Salerno 2006).

It is not the total value of money spent that determines the aggregate level of expenditure in an economy, but the other way 'round. The level of expenditures that all participants incur will determine to what extent money must circulate to satisfy these transactions. Recognizing this point eliminates the uncertainty and circularity of the reasoning in Friedman and Schwartz (1963: 695) whereby the bulk of the causality in the QTM runs from the money supply to economic

8 This fact gives rise to the trichotomy of goods in existence – medium of exchange, consumers and capital goods (Mises 1971:79). Claims that money is a form of capital good because it is not directly consumed are misplaced (for example, in Barnett and Block 2005; 2007) as they fail to recognize that money's role is not in directly satisfying future wants, but in facilitating our wants in both the present and future. To this more typical characterization of money's role we can also add that money serves to settle the payment for our past wants, as is the case when it is used to settle a debt.

activity in the long run, while in the short run there is also a case for the corollary. Actually, the truth lies somewhere in the middle. The money supply for transactions purposes and the amount of desired expenditure are co-determined in the sense that once one decides how much money to hold in his reservation balance, the expenditure decision is one of allocating the remaining income among the four expenditure options.

Finally, the average price level for all transactions comprising *GDE* is useful to include as a practical matter. Instead of dividing gross expenditures into their component parts, it is easier to recognize that GDE_t is the sum of all money transactions for goods and services at their respective prices. Substituting we get:

$$V = (PQ - \Delta L_t + \Delta E_t + CA_t - \Delta Bt) / (TMS \neg - M_R),$$
$$\text{where } PQ = Cp_c + Kp_k + Np_n.$$

Instead of being a vacuous concept devoid of any real importance except for its role in equilibrating the equation, velocity here becomes the necessary result of people's conscious expenditure decisions. Evidence pointing to the determinants of the behaviour of velocity gain a theoretical underpinning. Is the velocity of circulation determined or at least influenced by the nominal interest rate (Laidler 1989), real interest rate (Friedman 1956), the expected inflation rate (Laidler 1991), or is it a passively determined variable (Keynes 1923)? Maybe the velocity of money is systematically related though mostly insensitive to interest rates, as evidence contained in Friedman (1987) suggests. In my rendition of the QTM there is no need for discussion as to the degree of influence of one determinant on velocity, just as in standard price theory there is no need to discuss the degree to which certain factors determine prices, e.g., preferences as opposed to incomes.

What is clear is that velocity is necessarily determined by all variables on the right hand-side of the equation. Ceteris paribus, velocity will increase if: 1) the general price *P* level rises, 2) the quantity *Q* of goods and services transacted for increases, 3) total indebtedness or issues of bank-created credit decrease (in which case both ΔL_t and ΔB_t are negative), 4) net equity purchases, 5) positive current account balances, 6) the true money supply *TMS* declines, or 7) the stock of money held in reservation balances M_R increases.

Conclusion

I will end by listing the advantages of using the quantity theory of money developed herein over other approaches. Before doing so, however, I wish to reiterate my hesitation in using the chosen title of this article. The quantity theory of money, loosely stated in all of its variants, is just a statement about how changes to the money supply affect the general price level. Stated in such a way it is really just a formalization of the law of diminishing marginal utility. As the units of a good increase (in this case money) the usefulness of each unit decreases (the value of each subsequent unit decreases). The law of diminishing marginal utility can be formalized for money in a way that it cannot be for other goods owing to the fungible nature of the money supply. All units are valued equally, thus instead of each subsequent unit being valued less than its predecessors, all units will see their value diminished equally. If this is the contribution of the quantity theory of money I would hazard to say that the pages of spilled ink over its validity and importance are much ado about not much.

I would have preferred to call this article "The Monetary Exchange Theory of Velocity", but I doubt many would understand it in the way I intend. "Exchange" in this title refers not narrowly to those that create income but more broadly to those that settle expenditures that will satisfy the purchaser in the past, present and future. It concerns velocity as this is the explained variable. Thought of this way, the traditional QTM would be better stated as the "Quantity Theory of Prices".

This title too is deficient in its use of the "theory". There is nothing conjectural about any of the variants of the QTM, the present case included. It is a tautology not in need of empirical testing. As such, from here forward I prefer to call the statement created here as the "Exchange Theorem of Velocity" (ETV). Admittedly this is not as catchy as the Quantity Theory of Money, but it is more honest.

$$V = PQ - \Delta L_t + \Delta E_t + CA_t - \Delta B_t / (TMS - MR)$$

The reasons for favouring ETV over the QTM are as follows:

1. The price level *P* removes distortions that may result from relying on debt-based financing. Prices are composed of those goods that trade at par value, and those

that trade at market. The debt-based portion of exchanges housed in the numerator of the ETV ($CA_t - (\Delta B_t + \Delta L_t)$) trades at a price, but that price is par and set in advance. As such this variable is not subject to change by alterations to aggregate goods' expenditure PQ or the amount of money directed to the transactions portion of the money supply. On the other hand, the price level P is useful in a new way that is not immediately apparent in traditional QTMs owing to its relevance to the prices of goods and services, accounting for expenditures made to settle prior debts.

2. The money supply in the ETV is better defined than in other variants. Transactions versions of the QTM focus narrowly on money's role as a medium of exchange. Income versions require money to be held as a cash balance. The ETV approach makes use of both roles and as such produces a velocity that incorporates money's ability to facilitate transactions as well as serve as an uncertainty hedge through a cash balance.

3. The ETV shares with the transactions versions of the QTM a broad focus on all monetary exchanges.

4. That old couplet that economists over a certain age will remember becomes useful again: "Money's a matter of functions four, a Medium, a Measure, a Standard, a Store" (Milnes 1919: 55). In the ETV, money functions as a medium of exchange as the residual $TMS - M_R$; money serves as a store of value in the reservation demand M_R; money is a measure of value as the nominal amount of current goods and services' expenditure PQ; finally, money in the ETV acts as the standard of deferred payments for all those loans incurred in the past which are being settled in the present. These debts include foreign loans on current account, banking sector debt from demand deposit deleveraging and broader financial sector debt through changes to ΔL_t.

5. Importantly, loans as a funding source are accounted for due the recognition that not all loans represent a renunciation of purchasing power by someone in the economy.

a) Loans on current account are a transfer of purchasing power. Since they are cross-border there is a "free lunch" created for recipients of such loans, at least until they are repaid. Upon repayment (i.e., when the current account turns positive) domestic individuals trade away their purchasing power to a foreign individual.

b) Through its ability to create unbacked credit, the fractional-reserve banking system allows for expenditures to occur which also do not represent a strict transfer of purchasing power. As a result, any change in bank-created credit over the time period in question will also represent a "free lunch" – money exchanges will be facilitated that did not require a reduction in spending by some other member of the economy. Since the credit facilities of the fractional-reserve banking system are well-known for their ability to instigate crises, an understanding of deleveraging and leveraging through bank lending enables us to better understand the effects on prices in general and changes to the velocity of circulation that must result.

6. Prices are not sticky by assumption, as in much Keynesian literature.[9] At the same time, prices are not the variable necessarily enticing changes to the level of expenditure. Instead the price level P is the result of the conscious choice among individuals to divide their incomes between repayment for past expenditures ΔL_t, payment for present consumption, capital and intermediate goods expenditures PQ, payment for new equity purchases ΔE_t, and repayment of foreign debt incurred in the past CA_t. In this way the ETV is influenced by Hülsmann (1997) and Bagus and Howden (2011) who argue that prices are not the variable guiding purchase decisions but are rather the result of the demands to change the quantities of goods consumed and produced. It differs from this conclusion in the sense that prices in the ETV also serve as a constraint on how many goods can be purchased relative to the reservation demand for money M_R.

7. The ETV rectifies the failure of the QTM during the recent spate of unorthodox monetary policies to explain the lack of price inflation in the face of large expansions of money supply. One explanation that follows from the ETV is that new money creation was absorbed by the stock market as equities were the recipients of much of this fresh money creation (as in Machlup 1940: chap. 4 esp. 47-48).]

9 Keynesians do not have a monopoly on this claim. A recent attempt to formulate an equation of exchange more amenable to the Austrian economist includes some degree of short-term price stickiness (Evans and Thorpe forthcoming).

8. Finally and perhaps most importantly is the emphasis placed on velocity as an explained variable. No longer is it merely included as an error variable that must necessarily balance the relationship between the money supply and the amount of aggregate expenditure it produces. Although conceptually similar to its QTM variants, velocity in the ETV is the necessary outcome that individuals consciously create through their demand to expend income. Furthermore, it is negatively related to debt creation – including international, bank-created and more conventional – and as a result is determined by the propensity to borrow, which ultimately relies on interest and expected inflation rates. As a result velocity is a proxy for the propensity to spend – both in terms of consumption and investment expenditures. One implication of this final point is a new method to identify periods of recession that do not rely the interaction between prices and quantities of goods produced (as is the case with GDP), but rather on the desire (or ability) of individuals to make expenditures. This final possibility is elaborated on in Howden (2013).

References

Anderson, William L.2001. What's Wrong with the CPI? *The Free Market* 19(8).

Bagus, Philipp, and David Howden.2009. The Legitimacy of Loan Maturity Mismatching: A Risky, But Not Fraudulent, Undertaking. *The Journal of Business Ethics* 90(3): 399-406.

Bagus, Philipp, and David Howden. 2011. Monetary Equilibrium and Price Stickiness: Causes, Consequences, and Remedies. *Review of Austrian Economics* 24(4): 383-402.

Bagus, Philipp, and David Howden. 2012a. Monetary Equilibrium and Price Stickiness: A Rejoinder. *Review of Austrian Economics* 25(3): 271-77.

Bagus, Philipp, and David Howden. 2012b. The Continuing Continuum Problem and Future Goods. *The Journal of Business Ethics* 106(3): 295-300.

Bagus, Philipp, and David Howden. 2013. Some Ethical Dilemmas of Modern Banking. *Business Ethics: A European Review* 22(3): 235-45.

Bagus, Philipp, and David Howden. forthcoming. The Economic and Legal Significance of 'Full' Deposit Availability. *European Journal of Law and Economics*.

Barnett, William, and Walter Block. 2005. Money: Capital Good, Consumers' Good, or (Media of) Exchange Good? *Review of Austrian Economics* 18(2): 179-94.

Barnett, William, and Walter Block. 2007. Saving and Investment: A Praxeological Approach. *New Perspectives on Political Economy* 3(2): 129-38.

Evans, Anthony J., and Robert Thorpe. forthcoming. The (Quantity) Theory of Money and Credit. *Review of Austrian Economics*.

Friedman, Milton. 1956. "The Quantity Theory of Money: A Restatement", reprinted in *Studies in the Quantity Theory of Money*, pp. 94-111. Chicago: University of Chicago Press.

Friedman, Milton. 1970. A Theoretical Framework for Monetary Analysis. *Journal of Political Economy* 78(2): 193-238.

Friedman, Milton. 1987. "The Quantity Theory of Money", in *The New Palgrave – Money*. London: Macmillan.

Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States*. Princeton, N.J.: Princeton University Press.

Howden, David. 2009. Fama's Efficient Market Hypothesis and Mises's Evenly Rotating Economy: Comparative Constructs. *Quarterly Journal of Austrian Economics* 12(2): 3-12.

Howden, David. 2013. A Velocity History of the United States. Working paper.

Huerta de Soto, Jesús. 2006. *Money, Bank Credit and Economic Cycles*. Auburn, AL: Ludwig von Mises Institute.

Hülsmann, Jorg Guido. 1997. Knowledge, Judgment, and the Use of Property. *Review of Austrian Economics* 10(1): 23–48.

Keynes, John Maynard. 1923. *A Tract on Monetary Reform*. London: Macmillan.

Laidler, David. 1969. The Definition of Money: Theoretical and Empirical Problems. *Journal of Money, Credit and Banking* 1(3): 508-25.

Laidler, David. 1989. "Radcliffe, the Quantity Theory and Monetarism", in D. R. Cobham, R. Harrington, and G. Zis (*eds*.) *Money, Trade and Payments: Essays in Honour of Dennis Coppock*. Manchester: University of Manchester Press.

Laidler, David. 1991. The Quantity Theory of Money is Always and Everywhere Controversial – Why? *Economic Record* 67(4): 289-306.

Machlup, Fritz. 1940. *The Stock Market, Credit and Capital Formation*. London: William Hodge and Co.

Mankiw, N. Gregory. 2009. *Macroeconomics*, 7th ed. New York: Worth Publishers.

Menger, Carl. [1976] 2007. *Principles of Economics*. Auburn, AL: Ludwig von Mises Institute.

Milnes, Alfred. 1919. *The Economic Foundations of Reconstruction*. London: MacDonald and Evans.

Mises, Ludwig von. [1912] 1971. *The Theory of Money and Credit*. New York, The Foundation for Economic Freedom, Inc.

Pigou, Arthur C. [1917] 1952. "The Value of Money", reprinted in *Readings in Monetary Theory*. London: Allen & Unwin.

Robinson, Joan. 1970. Quantity Theories Old and New. *Journal of Money, Credit and Banking* 2(4): 504-12.

Rothbard, Murray N. [1962] 2009. *Man, Economy, and State: A Treatise on Economic Principles*, 2nd edition. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. [1963] 2000. *America´s Great Depression*, 5th edition. Auburn, AL: Ludwig von Mises Institute.

Salerno, Joseph T. 1987. The "True" Money Supply: A Measure of the Supply of the Medium of Exchange in the U.S. Economy. *Austrian Economics Newsletter* (Spring): 2-6.

Salerno, Joseph T. 2006. A Simple Model of the Theory of Money Prices. *Quarterly Journal of Austrian Economics* 9(4): 39-55.

Skousen, Mark. 2012. Gross Domestic Expenditures (GDE): The Case for New National Aggregate Statistic. Working paper.

Wallace, Neil. 1998. A Dictum for Monetary Theory. *Federal Reserve Bank of Minneapolis Quarterly Review* 22(1): 20-26.

Yeager, Leland B. 1994. Tautologies in Economics and Natural Sciences. *Eastern Economic Journal* 20(2): 157-169.