Eastern Econ J (2019) 45:250–268
https://doi.org/10.1057/s41302-018-00127-y

ORIGINAL ARTICLE

# Bubbles and Broad Monetary Aggregates: Toward a Consensus Approach to Business Cycles

**Cameron Harwick**[1]

Published online: 16 November 2018


**Abstract** A challenge for quantity-theoretic explanations of business cycles is that recessions manifest despite central banks' scrupulousness to avoid falls in monetary aggregates, a fact which would seem to indicate a structural explanation. This paper argues that a broader and theoretically richer Divisia aggregate—which reflects changes in financial market liquidity even without changes in the quantity of any particular asset—can reconcile these two approaches. Liquidity shocks such as the rise and collapse of asset bubbles can drive excess supply of and demand for money, respectively, that quantity theorists point to as determinative of short-run economic fluctuations.

**Keywords** Business cycles · Divisia · Monetary policy · Quantity theory · Asset bubbles

**JEL Classification** E32 · E44 · E51

The literature on business cycles is characterized by a number of broad approaches, each of which emphasizes its own set of stylized facts to explain. Hetzel (2009) distinguishes between "quantity theory" approaches, which focus on changes in the money supply and/or the volume of spending over the course of the business cycle (e.g., Friedman 1968; Yeager 1956), and "credit cycle" approaches, which—while not necessarily ruling out ultimate monetary causes—emphasize changes in the structure of credit markets as the proximate cause (e.g., Hayek 1933; Keynes 1936; Minsky 1982).

✉ Cameron Harwick
charwick@brockport.edu

[1] The College at Brockport, 350 New Campus Drive, Brockport, NY 14420, USA

The two are not obviously compatible, and both can claim equivocal empirical success.[1] Quantity-theoretic explanations tend to see boom and bust as causally unconnected, except to the extent that a boom represents a recovery from a previous bust (Friedman 1993). Credit cycle explanations tend to see recessions as following causally from previous booms (Boissay et al. 2013; Garrison 2001). Credit cycle explanations tend, however, to be difficult to operationalize, often relying on ill-defined notions like "roundaboutness" (Lewin and Cachanosky 2018), ad hoc expectational phenomena like "animal spirits," or fine-grained microtheoretic concepts for which limited data exist. Quantity theory explanations, by contrast, seem at first blush much more conceptually straightforward: all their central concepts correspond to readily available and well-documented statistical aggregates, and the basic logic goes through under either adaptive or rational expectations. Nevertheless, since about 1980, quantity theory explanations have lost a great deal of explanatory power across the world (Friedman and Kuttner 1992) and have since then been largely lain aside in favor of moneyless models of economic activity (e.g., Woodford and McCallum 2008).

Recent work on monetary aggregation, however, suggests that the straightforward operationalization of quantity theory explanations may have masked deeper conceptual problems. In particular, the monetary aggregates typically used in empirical work either fail to account for the vast majority of what is actually used as money in a modern economy (e.g., M0), or add together imperfectly substituting monetary assets in a way that vitiates its meaningfulness (e.g., M2).

This paper argues that this more recent work has the potential not only to rehabilitate quantity theory explanations of the business cycle, but also to connect them to long-standing credit cycle explanations in a deeper way than the few reconciliations that have to this point been offered. Specifically, we argue that changes in financial market liquidity driven by shifts in expectations will change the quantity of a broader weighted monetary aggregate, not by changing the quantity of any particular component asset, but by causing additional assets to serve as money, and existing assets to substitute more closely for it. Such a story has the potential to subsume the explananda of the two approaches and to serve as the foundation for a consensus approach to business cycles, a consensus which has not existed in the discipline at least since Keynes.

The following section covers the theoretical background, and a sketch of an account of business cycle dynamics centering in asset bubbles and liquidity dynamics follows in the subsequent sections. We conclude with a reinterpretation of several credit cycle theories in light of a focus on broad monetary aggregates.

[1] The quantity theory evidence is discussed below. Various credit cycle theories include the Austrian theory (on which see Luther & Cohen [2014; 2016]) and the credit channel theory (on which see Boldin [1994]; Bernanke and Gertler [1995]).

## Theoretical Background

### The Quantity-Theoretic Account of Business Cycles

The basic quantity theory logic starts from Walras' law, which states that the excess demands in all markets must identically sum to zero. In other words, unsold inventories of some good imply unsatisfied demand for some other. It is an extension of Say's law, which has sometimes been taken to deny the possibility of a general glut or recession. But, as Yeager (1956) notes, "while an excess supply of some things does necessarily mean an excess demand for others, those things may, unhappily, be money." Because money constitutes one side of nearly every exchange in a modern economy, "an economy wide excess demand for money shows up not as specific frustration in buying money, but as dispersed, generalized frustration in selling things and earning incomes" (Yeager 1956). An excess *supply* of money, by the same token, does not show up as a specific frustration in selling money—for it would not be money if one could be frustrated in selling it—but as a general increase in spending: a "euphoria," or boom.

This much is tautological. An excess demand for money is *necessarily equivalent* with an excess supply of goods, and vice versa. The empirical content comes in with the additional assumption of stable velocity and the long-run neutrality of money. Consider the equation of exchange,

$$MV = Py \tag{1}$$

where $M$ is a monetary aggregate, $V$ is the velocity of that aggregate, $P$ is a price index, and $y$ is real income. If $V$ is stable, and if $y$ is long-run invariant to changes in $M$, then changes in the rate of money growth will pass through one-for-one to changes in the inflation rate.[2]

The long-run neutrality of money, such that $y$ returns to a "natural" level in the long run regardless of changes in $M$, remains well accepted both theoretically and empirically. It is the other assumption, stable velocity, that has in recent decades failed the quantity theory.

Without stable velocity, Walras' law and the equation of exchange are still both tautologically true, but their usefulness is rather circumscribed. A long-standing Keynesian critique of the quantity theory is the notion of the "liquidity trap," which can be expressed as saying that, on certain margins, $V$ is a function of $M$. If the demand for money increases one-for-one with the supply of money, then no amount of monetary expansion can satisfy an excess demand, and the equation of exchange—though still "true"—sheds little light on business cycle dynamics.

In fact, the velocity of M2 was quite stable over at least the first 6 decades of the twentieth century in the USA. (Friedman and Schwartz 1963; Hendrickson 2017), a

[2] The logic of Walras' law can also be restated in these terms: to the extent that $P$ is sticky, an increase (decrease) in $M$ will produce an excess supply of (demand for) money at the given price level, which will result in an excess demand for (supply of) goods. In other words, if $P$ cannot move, equality is maintained in the short run by movements in real income $y$.

fact which made a monetary targeting regime feasible. Nevertheless, as noted above, financial advances and the rise of near-monies around 1980 destabilized velocity to the point that the quantity theory is now generally regarded as nonoperational (Friedman and Kuttner 1992). The demand for the components of M2 became more volatile but also more interest-rate sensitive, driven by the availability of near-monies with nonzero rates of return. Since that time, therefore, the conduct of monetary policy has largely focused on interest rates rather than monetary aggregates as a policy target.

Even so, the quantity theory retains some appealing explanatory features. In particular, the fact that recessions and depressions are economy-wide phenomena, rather than being concentrated in specific sectors or industries (Kydland and Prescott 1990), suggests their origin in an excess demand for money rather than the malinvestment or credit market frictions pointed to by structural credit cycle explanations. The volume of spending, in some form or another, appears to be the relevant transmission mechanism for at least an important subset of economic fluctuations (Beckworth and Hendrickson 2012). The challenge for an operational quantity theory, then, will be to account for the procyclical movement of velocity. Two additional notions will be helpful in doing so.

## Liquidity and the Pyramid of Credit

The background of the quantity theory suggests an explanation of boom and bust in terms of an excess supply of and demand for money, respectively. But what is the appropriate quantity of money? The traditional answer has been an empirical punt: whichever aggregate tracks economic activity the most closely. In Friedman's day, this was M2. Even at that point some argued that monetary aggregation was sufficiently complicated to render the monetarist project moot (e.g., Tobin 1963). Since 1980, the financial developments that have decoupled M2 from economic activity have validated Tobin's concerns, a situation which forces us to step back and define "the quantity of money" in a more constructive manner.

The first formulations of monetary theory considered "money" as a single asset that could be exchanged for any other good in the economy. Menger's (1892) conjectural account of the origin of money out of barter, for example, focused on the process by which a single good becomes far-and-away more saleable—or *liquid*—than other goods. On Menger's telling, the primary inconvenience of barter exchange is the necessity of a "double coincidence of wants": in order to purchase something from my neighbor, I must have something he wants as well. If I want to purchase cabbage from him, and he is in need of a goat, but I raise cows, then we must find another party to the exchange who is willing to exchange cows or cabbage for goats. Needless to say, finding such a willing party for every transaction can become quite costly. In order to mitigate these costs, agents in the economy can hold inventories of some relatively more saleable good, even if they have no use for it themselves, and thereby increase the chance of being able to achieve a double coincidence of wants and to acquire goods they do have a use for. The key to the story is that holding such an inventory constitutes a demand for that good, which *further increases its*

*saleability*. Because the desire of each individual to most effectively satisfy his own wants on the market leads him to hold an inventory of the most saleable good, one good will eventually be converged upon with universal saleability—money.[3]

Menger's conception of money as involving exceptionally high liquidity gives us an easily operationalizable notion to apply to more modern financialized economies. We can define an asset's liquidity as the immediacy with which one can find a willing buyer at the appraised market price—or, alternatively, *illiquidity* as the discount one has to accept on an asset's appraised value if one needs to sell it in a given amount of time, or the time necessary to sell it at a given price. In this sense, we can think of *saleability, liquidity*, and *moneyness* as synonymous properties of a good or asset.[4]

It would be a mistake, however, to imagine that Menger's process implies a singular money asset. Anderson (1917) argues,

> the development of money, while it adds to the saleability of the money-commodity, *also adds to the saleability of other goods*…. The fact that goods have money-prices, which can be compared with one another quite easily, in objective terms, makes barter, and barter equivalents, a highly convenient and very important feature of the most developed commercial system. [emphasis in original]

Kroszner (1990) sees this as a continuation of the process Menger described:

> The process whereby money's saleability brings about the increased marketability of other goods may be able to bring about the end of money itself as *the* medium of exchange…. Financial instruments besides money are accepted as the final means of payment, and a growing range of transactions take place without money…. More and more acceptable substitutes arise, and an ever growing number of transactions in the economy occur without money as we know it. Various assets—perhaps interest-paying bearer bonds and equity instruments—are used to settle accounts.

More precisely, these various assets are employed to satisfy the demand for the various services of money—as a medium of exchange, and a store of value. An asset's liquidity, therefore, will be a joint function of its suitability in these two roles.

Mehrling (2012) arranges these assets into an asset-liability structure, with the most liquid money on the bottom ("base money"), deposits redeemable in base money placed atop it, other assets redeemable in bank money placed atop that, and so on. The quantity of each in terms of the unit of account is indicated by the width at that level, giving the figure the shape of an inverted pyramid (Fig. 1). Much monetary economics takes account of only two levels, namely base money and deposit

[3] On the applicability (or lack thereof) of Menger's conjectural history to the actual history of monetary exchange, see Harwick [2018].

[4] Liquidity as a property of assets is distinct from the popular usage of the term as a property of organizations, for example the liquidity of a bank. Liquidity in this latter sense is more aptly "the bank's liquidity *position*."

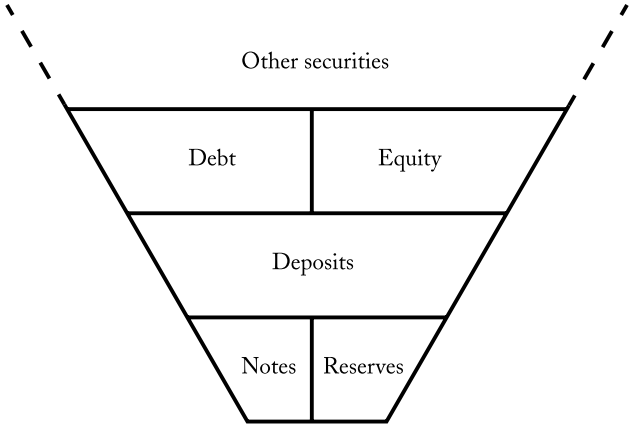


**Fig. 1** The "pyramid of credit." Each layer is redeemable for the asset type in the layer beneath, and the slope indicates aggregate leverage ratios. Similar diagrams appear in Mehrling (2012) and Gabor and Vestergaard (2016)

balances, the ratio between which being called (somewhat misleadingly) the "money multiplier".

This hierarchical structure is of critical importance over the business cycle, as the quantity of each layer constrains the issue of assets in the layer above. Assuming banks target a particular leverage ratio, a decline in the quantity of reserves will lead to a contraction in the quantity of deposits multiplied by the existing ratio.[5] This much is appreciated by orthodox monetary theory. The usefulness of the pyramid is to show that the contraction does not stop there: the contraction in deposits will lead to even larger contractions in the markets for bonds, equities, and securities, all of which serve monetary functions to a greater or lesser extent in a modern financial economy. More broadly, an autonomous deleveraging at any one layer will have an identical contractionary effect on the layers above it.

## The Divisia Method of Monetary Aggregation

The upper levels of the pyramid do not substitute one-for-one with the lower layers more typically regarded as money, even if they do substitute on some margin. The effect on the volume of spending of a bond issue, for example, is less than that of the issue of an equivalent nominal value of deposits due to its lesser liquidity. An economically meaningful monetary aggregate, therefore, cannot simply add together the nominal values of bonds, equities, securities, and so on, to the quantities of base and bank money.

[5] In fact, the situation may be more severe even than this: Adrian & Shin [2010] show that if banks target constant value-at-risk rather than constant leverage, a decline in the quantity of reserves may lead to a contraction of deposits *greater* than the initial decline times the leverage ratio—an exponential contraction rather than a multiple contraction.

For this reason, the pyramid logic is often coupled with a criticism of quantity-theoretic explanations of business cycles. Its asset-liability structure lends itself to more structural credit cycle explanations as shocks are transmitted up the pyramid from institution to institution, and the complex substitutability structure would seem to be a major impediment to operationalizing a quantity theory using it. Nevertheless, recent work on monetary aggregation offers the possibility of retaining the quantity-theoretic concern over the volume of spending while maintaining a conception of money as rich and microfounded as that suggested by the logic of the pyramid. Such is the purpose of the Divisia aggregates, described in Barnett (1980).

A Divisia aggregate is computed by weighting various money-like assets according to their user costs, defined as the spread between the interest rate on an illiquid benchmark asset and the lower interest rates offered by the monetary assets themselves. The Divisia weighting scheme can be used to construct increasingly broad aggregates, M1 through M4, that approximate the supply of the liquidity services of money, regardless of what sort of asset happens to serve those purposes.

One major advantage of a Divisia aggregate is to internalize shifts in the *composition* of money demand. Suppose a fall in the interest rate on Treasury bonds renders bank deposits a closer substitute for them as a store of value. On the margin, investors shift into deposits, and vice versa for a rise in interest rates, hence the negative relationship usually posited between interest rates and the demand for money. But, as Barnett (1980, p. 12) argues,

> the value of an *economic* aggregate (by its definition) *cannot* change as a result of internal substitution effects. Hence the money market substitution effects destabilizing velocity should be completely internalized by proper aggregation over the money market.

To the extent that treasuries serve a monetary role in this example, changes in the *composition* of the demand for money, even if the total demand for the services of money is unchanged, will appear as spurious changes in the demand for money (and a spurious interest elasticity) if we use an inapt definition of the money stock.

That the Divisia aggregates signify something close to what we have in mind by a broad money stock is indicated by the fact that "the velocity of money is increasingly stabilized as the level of aggregation is increased" (*ibid*, p. 12), effectively addressing the Keynesian critique of the quantity theory. In other words, by including less liquid monies than are usually accounted for in monetary aggregates, we have a quantity with a direct causal impact on the volume of spending, without registering internal shifts in the composition of demand as velocity shocks. And indeed, there is much evidence to suggest that movements in Divisia aggregates predict changes in output and interest rates according to the quantity-theoretic story better than simple-sum M1 or M2 aggregates (Hendrickson 2014; Serletis and Chwee 1997; Barnett 2016). This should not be surprising if the logic of the pyramid of credit is correct: it is not the dollar value of some monolithic "quantity of money" which matters for the monetarist story, but the total volume of spending, which is determined by the supply of and demand for a large variety of imperfectly substituting financial assets.

The conjunction of the pyramid concept and a Divisia index of monetary services allows us to connect changes in the leverage structure of financial markets to

changes in the economically relevant quantity of money that registers as a broader excess supply or demand. An illustrative sketch of these dynamics over the course of the business cycle follows.

## Liquidity and the Bubble

An asset must trade in a thick market against a variety of other assets in order to substitute for money in its medium of exchange function. For this reason, moneyness has traditionally been thought to end for all intents and purposes with the deposits layer of the pyramid, perhaps with a limited monetary role for Treasuries. Otherwise, debt and equities—stocks, commercial paper, mortgages, and so on—are far too heterogeneous to be accepted without question against a sufficiently wide variety of other assets to be considered media of exchange. The fact that reducing information costs is an important *sine qua non* of monetary exchange (Alchian 1977; Banerjee and Maskin 1996) suggests that an effective monetary asset must be homogenous in all important respects.

It would be a mistake, however, to imagine that liquidity decreases monotonically as one moves up the pyramid. Indeed, one of the major financial innovations of the past several decades has been securitization, a process of rebundling heterogeneous assets and redistributing them in a more homogenous—and therefore more liquid—form. A mortgage-backed security, for example—the epicenter of the 2008 crisis—represents a claim to a portion of the revenue stream from a large number of mortgages. These claims are then *tranched*: only when the first tranche has been entirely paid from the proceeds of the underlying mortgages is the second tranche paid, and so on for the third and fourth tranches. In this way, heterogeneous mortgages can be repackaged and sold as assets totally homogenous with respect to risk (Coval et al. 2009).

Assets like bank deposits substitute for base money in their capacity as media of exchange primarily on the basis of their stability in purchasing power. For practical reasons of cognitive convenience, interest-bearing assets are not typically accepted in payment for consumer goods. In financial markets, however, trading is specialized and occurs on a sufficiently large scale that aspects of an asset's usefulness *as a store of value* become more important determinants of liquidity. On the margins relevant in financial markets, and holding risk constant, expected return will be more important than stability of purchasing power in determining saleability, *provided* the asset is sufficiently homogenous. In other words, the portfolio demand for money can be satisfied by a different set of assets than the transactions demand for money, but changes in the composition of either can have identical monetary effects over the business cycle.

This logic is consistent with, if only implicit in, a great deal of work in monetary theory. The usual downward slope of the demand for money in the interest rate refers to changes in the portfolio demand for (narrow) money as wealth holders substitute into less liquid monies. Likewise, Williamson (2012) accounts for liquidity traps by arguing that Treasuries and reserves are practically identical assets at the zero lower bound—i.e., that reserves have little transactions role, and ordinary

"open market operations are just swaps of identical assets." Even so, because both substitute on certain margins for media of exchange, a shortage of safe interest-bearing assets—even those with no transactions role—can withdraw media of exchange from circulation.

We will call these assets substituting for money as stores of value rather than as media of exchange *financial quasi-monies*. Quasi-monies have no necessarily fixed price in terms of basic money, so changes in the asset's price are decoupled from identically inverse changes in the price level, as would be the case for a change in the price of bank deposits, for example. Without such a promise anchoring price expectations, speculative shifts become possible. Frankel and Rose (1995) explain:

> Expectations can be described as stabilizing when the effect of an appreciation today—relative to some long-run path or mean—is to induce market participants to forecast depreciation in the future…. Expectations can be described as destabilizing, on the other hand, when the effect of an appreciation is to induce market participants to forecast more appreciation in the future.

More aptly, we should call *speculation* stabilizing or destabilizing, with expectations determining which prevails for a certain asset. Because the asset constituting an economy's unit of account requires some stability of purchasing power vis-a-vis goods prices over time, it must be subject to stabilizing speculation (Burns and Harwick 2017; Harwick 2016). Where this is not the case—for example in hyperinflations—the currency quickly loses its usefulness as a medium of exchange, and thus also as a unit of account, even if legal restrictions prevent any other media from being used. For such monies, the switch from stabilizing to destabilizing speculation, for example a bank run or a convertibility crisis, is a catastrophic event.

No such requirement binds financial quasi-monies. The switch from stabilizing to destabilizing speculation is not necessarily catastrophic and may even augur financial "euphoria." In the case of a self-feeding appreciation, this is what we will call a bubble. *Pace* some asset pricing models that define a bubble as any deviation from a calculation of fundamentals (e.g., Blanchard and Watson 1982), our own definition will be narrower and closer to common usage: a prolonged instance of upward destabilizing speculation in a particular asset.[6]

Bubbles in particular assets, of course, most often coincide with general booms. This is to be expected: per the above quantity-theoretic account of boom

[6] Precisely what causes expectations to shift from stabilizing to destabilizing is difficult to pin down with much exactness. Both the "rational bubbles" literature, where price increases of a sufficient magnitude can "launch" an asset into a self-fulfilling speculative bubble, and "animal spirits" or "sunspots" explanations give us an external description of a pattern, but no internal insight by which the formation of expectations can be made intelligible at a sufficiently abstract level. For reasons reminiscent of the Lucas Critique, however, we may satisfy ourselves that this must necessarily be the case: if the particular justification for the expectation were known in advance, or were similar to a recent occurrence, recognition of that fact would render the expectation self-limiting. We see, therefore, a variety of justifications for particular bubbles, each with a strong sense of "it's different this time." It is enough to observe that a sufficiently abrupt and unexpected rise in an asset's price can, under necessarily unpredictable circumstances, induce the expectation of further price rises [Krugman 1991; Lachmann 1956: 30ff] and launch the asset into a bubble.

and bust, an increase in total spending not only increases real income until prices rise to compensate, but also raises some prices more quickly than others. The most violent price increases we will expect to occur in financial markets, for three reasons:

1. Modern monetary systems mediate increases in the supply of the most liquid base money through financial markets. Price signals therefore reach financial markets first and most strongly.
2. Financial markets are thick, and assets are heavily traded. Their prices are, for this reason, highly sensitive to changes in supply or demand—they are the least "sticky."
3. Dornbusch and Fischer (1980) show that a flexibly priced asset among a constellation of sticky prices can result in the former "overshooting" its equilibrium value in response to a general demand shock.

Stabilizing speculation will take over for many goods at this point. Asset prices increase before goods prices and then fall back somewhat as the new money permeates the wider economy. But—recalling the necessary inexactness of analyzing expectations here—investors often converge on one focal asset with expectations of continuing price increases.

This being the case, an initial spending shock is sufficient both for the initial launch and the continuing rise of the bubble. The foundation of the bubble asset's liquidity is its suitability as an investment vehicle. Rising demand propels its price higher, and as a store of value, it becomes more saleable precisely *because* its price is rising, quite unlike monies further down the pyramid, whose liquidity is founded in part in a more or less stable purchasing power vis-a-vis other goods. The bubble finds ready acceptance *so long as* its price continues to rise at a rate sufficient to avoid disappointing its speculative investors and to compensate for any increased perception of risk.

The usefulness of the pyramid construct at this point is in highlighting the equivalence of a positive velocity shock at one level of the pyramid with a positive supply shock at higher levels. Privately issued monetary assets tend to be demand elastic: an increase in the demand for a particular issuer's liabilities allows the issuer to take on more leverage and issue more liabilities. Suppose, then, that consumers substitute out of cash and into bank deposits. The velocity of the monetary base rises, by definition. This substitution is internalized by M2, which includes bank deposits; however, the *quantity* of M2 rises as banks lever up. From a quantity-theoretic perspective, it makes no difference whether we regard the velocity of M0 or the quantity of M2 as having risen.

A similar process obtains with monies further up the pyramid. A substitution out of bank deposits and into, say, commercial paper due to a rise in interest rates, constitutes a rise in the velocity of M2, hence again the negative interest elasticity supposed to characterize the demand for money. To the extent that this commercial paper is sufficiently liquid to serve as money, however—even if not a perfect substitute—the rise in the velocity of M2 will be internalized in a

Divisia aggregate, whose quantity will rise, not due (in the short run, at least) to an increase in *issues* of commercial paper, but to increases in the *aggregate nominal value* of commercial paper. An expansion of the upper layers of the pyramid in response to a general substitution away from money and into quasi-monies, therefore, can be expressed equivalently as a rise in the total money stock, or a rise in the velocity of the narrower money stock. Importantly, both expressions are consistent with the quantity-theoretic account.

A bubble therefore represents a significant expansion of the broad money supply, over and above the initial increase which launched it, reinforcing its own price rise with further increases in the volume of spending. In this way, optimistic expectations of future growth can result in an excess supply of money without any expansion on the part of the central bank (cf. Hendrickson and Beckworth 2015), an excess supply which can be understood equivalently as a velocity shock to a narrower monetary aggregate, or a supply shock to a broader monetary aggregate. Mehrling (2012) argues that during the boom, the pyramid widens (more assets exist at each level), and the layers flatten (assets on different levels become closer substitutes). I would suggest a more apt metaphor for the latter phenomenon is not the *flattening* of the pyramid, but *vertical additions*. Closer substitutability between two levels reflects the fact that each level has become more liquid. This implies as well the addition of further layers at the top, as assets which were previously relatively illiquid become sufficiently liquid to substitute for proper monies, not least among which is the bubble asset itself.

## Burst and Recession

The bubble asset's liquidity inheres in its status as a focal store of value, driven by a rapidly increasing price. This means that not only must investors eventually be disappointed, but also that the burst must come as a sudden crash. Once the expectations of investors are disappointed by a slowdown in the asset's rising price, the foundation of its liquidity vanishes at once: it loses its status as a focal store of value, and for that reason its substitutability for money. This can lead to three mutually reinforcing effects on the broader economy: (1) a fall in the *real value* of financial quasi-monies as asset prices fall, (2) a fall in the *liquidity* of quasi-monies as uncertainty makes them more difficult to sell, and (3) a fall in the *quantity* of quasi-monies as issuers default or fail. All three have the effect of diminishing the broader and economically relevant money stock.

First, the collapse of the bubble is *ipso facto* a nontrivial collapse in the broader money stock. As with the classical quantity-theoretic account, what matters is not—despite its name—the quantity of monetary assets, but their real value. Unlike changes in the real value of the narrower money stock, which entail either a change in the quantity of nominal assets or in the price level, the real value of the aggregate stock of quasi-monies can move independently of the price level and without a change in quantity because their prices are free to vary against the numeraire. Firms forced to mark their balance sheets to market—for example, when rolling over loans

(cf. Gertler and Kiyotaki 2015)—may find themselves suddenly lacking in means of payment.

Second, in addition to the fall in aggregate value, uncertainty makes it more difficult to sell assets. Finding their portfolios on the whole suddenly less liquid, investors in the bubble asset scramble to move their asset holdings further down the pyramid. In light of the connection between price changes and liquidity for financial quasi-monies, this scramble depresses both the price and the liquidity of the assets they sell, causing investors in diverse classes of assets to increase their demand for liquidity as well. The loss of liquidity can be represented by a removal of some of the top layers of the pyramid, as they are no longer viable substitutes for narrower monies.[7] On top of this, the shift in asset demand down the pyramid constitutes a deleveraging, a narrowing of the entire pyramid. With the money stock thus diminished, the volume of spending falls, sending the economy into a recession.

Finally, if issuers of assets should find themselves insolvent as these first two effects play out, these assets may—deprived of their underlying revenue stream—become valueless, effectively diminishing the quantity of quasi-monies, independently of the aforementioned changes in their price and liquidity. Further assets and securities pyramided atop these may also become valueless. This was the case most spectacularly during the Great Depression, when a string of bank failures cut the effective US money supply by a third, with most of the difference consisting in the deposits of failed banks rather than currency. More recent crises of the "shadow banking" sector have proceeded along similar lines, with less liquid monies.

This last effect is the most straightforward to forestall. The central bank can, in its lender of last resort capacity, prevent illiquidity from bankrupting any particular issuer, or buy the failed issuer's assets in order to prevent knock-on effects. In other words, by counteracting the scarcity either of media of exchange or of interest-bearing quasi-monies,[8] it can in principle prevent panic deleveraging (it is an open question whether it can in fact act quickly enough and in the appropriate location). It cannot, however, generally prevent formerly money-like assets from losing their liquidity, a fact which impinges on the broader money stock even if the central bank successfully prevents price-depressing fire sales.

This is the answer to our original puzzle: a bursting bubble in financial markets can send an entire economy into recession, despite the best expansionary efforts of the central bank, because an immense quantity of money has either been destroyed through deleveraging or lost its moneyness through diminished liquidity—a shift which, reckoning further down the pyramid in terms of M1 or M2, will look like a velocity shock. As Hummel (2011) notes,

[7] Divisia aggregates are computed for a given set of money assets, so this effect will not be reflected in any Divisia aggregate. The construction of weighted monetary aggregates with endogenous selection of component assets would, however, be theoretically valuable and awaits future research.

[8] Williamson [2012] discusses the different policy responses called for in each situation and argues the Great Recession—unlike previous panics in the USA—was characterized by a scarcity of liquid quasi-monies rather than of media of exchange.

> A fall in the broader money stock and a fall in the velocity of the monetary base are exactly the same thing… Thus, whether we label a particular decline in aggregate demand a monetary shock or a velocity shock can depend on how broadly or narrowly we define the money stock.

This is certainly the case during a scramble for liquidity. It is this fact that allows us to regard the velocity shocks of quantity-theoretic explanations not as exogenous preference shifts, but as manifestations of structural shifts in credit markets over the course of the business cycle as pointed to by credit cycle explanations. And it is the effect of those structural shifts on the quantity of broad money, on the other hand, that transmits the spending shock from financial markets to the wider economy.

Shifts in liquidity do not appear in official monetary aggregates, a fact which has led to a premature rejection of monetary explanations of the business cycle. Sight unseen, taking the monetary base as the relevant monetary aggregate would lead one to expect heavy inflation since 2008 (Fig. 2, left panel)—or, also taking into account the equal and opposite shifts in the velocity of the base, one might point to a "liquidity trap." M2, on the other hand, gives the impression that nothing at all has been amiss since the turn of the century (Fig. 2, right panel). It is only when one includes a sufficiently broad array of substitutes for the services of money that the supposed velocity shock of the Great Recession reveals itself as—after all—a shock to the quantity of money (Fig. 3) (Beckworth and Hendrickson 2015).

## Credit Cycle Theories in Light of Broad Money Aggregates

Quantity-theoretic explanations of business cycles tend to be relatively minor variations on the same causal mechanisms. The connection of the previous sections should therefore be apparent. Credit cycle theories, however, are more heterogeneous and detailed, and differ on more margins. Some require rather little emendation—Bernanke and Gertler (1995), for example, state that "we don't think of the credit channel as a distinct, free-standing alternative to the traditional monetary transmission mechanism, but rather as a set of factors that amplify and propagate conventional interest rate effects." Others, however, have traditionally been thought of as mutually exclusive alternatives. It will be worth a brief sketch of some important credit cycle theories in light of the argument here, each of which can point to particular conditions in credit markets with implications for a broader money stock. More detailed evaluations await future research.

### The Austrian Cycle Theory

The Austrian cycle theory focuses on artificially cheap credit initiating malinvestment over the course of the boom, which must then be liquidated during the bust as the "natural" rate of interest reasserts itself. Using a structure-of-production framework (Hayek 1933; Garrison 2001), the argument is that the profitability of investment in "higher-order" industries—that is, those further away from final consumption goods in a supply chain—will be more sensitive to changes in interest rates than

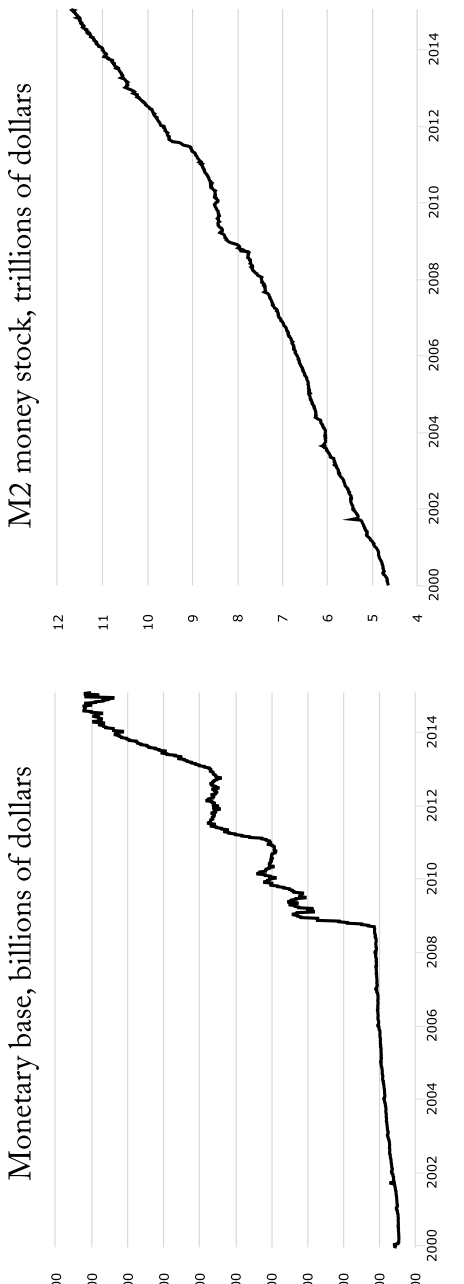


**Fig. 2** The dramatic expansion of base money (left) and the less dramatic expansion of simple sum M2 (right) in the USA from 2000 to 2015

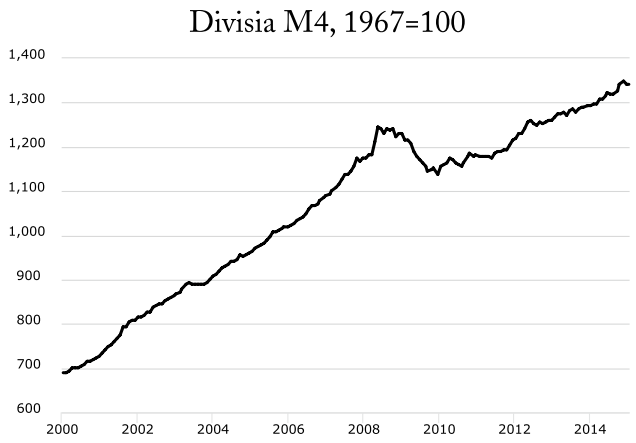


**Fig. 3** The divisia M4 index over the same period shows a decline from 2008 to 2010 corresponding to the Great Recession. Barnett and Gaekwad-Babulal (2017) derive a similar result for the Eurozone

that of "lower-order" industries specializing in consumption goods. Though some do acknowledge the importance of a quantity-theoretic "secondary deflation" (Hayek 1933; Horwitz 2000), recession is always referred primarily to the necessary liquidation of malinvestments brought about by previous overexpansion of credit.

The canonical story is, unfortunately, beset in its details by a number of conceptual problems, most importantly an ambiguity between funds and resources underlying the structure of production construct (Lewin and Cachanosky 2018). For this reason, explanations pointing to malinvestment in higher-order industries have generally not had much empirical success in the postwar era (Luther and Cohen 2014; though see Luther and Cohen 2016).

In broad strokes, however, the story has some merit. Artificially cheap credit may lead to the mispricing of systemic risk, as was revealed in the mortgage market in 2008, and the shock that launches a bubble may indeed stem from an initial credit expansion. In addition, rather than higher-order goods, we may locate the bulk of malinvestment in assets that exhibit a more violent price response to shifts in demand—not merely those with a high price elasticity of demand, but more importantly, those with a *more immediate* response. These are most plausibly liquid and long-term (and hence more interest-elastic) investments, which as an empirical matter overlap substantially with goods that have been traditionally considered higher order. A bubble in mortgage-backed securities, for example, drove relative overinvestment in housing. A bubble in dot-com stocks funded overinvestment in Internet start-ups in the late 1990s, and a bubble in stocks funded overinvestment in mechanization during the 1920s. The subsequent liquidation of all of these investments, quite apart from the central bank's control over narrower monetary aggregates, implies a fall in a broader money stock ceteris paribus.

It follows, then, that—per the Austrian story—the liquidation of malinvestments accounts for the fact that recessions cannot be avoided simply by countercyclical credit policy. Contra the Austrian story, however, countercyclical credit policy does

*not* necessarily forestall the liquidation of malinvestments. Accelerated credit expansion during recession may serve not to "prop up" unprofitable investments (though it may also do this), but to satisfy demand for the services of money that was formerly being satisfied by now-unprofitable investment vehicles. To the extent that an asset serves general liquidity purposes rather than indicating a specific investment—and this will be more nearly the case if the asset experienced a bubble—the satisfaction of that demand for liquidity will be consistent with the pattern of consumer demand in equilibrium in a way that the demand for specific investment in (for example) housing was not. To this extent, countercyclical credit policy can productively ameliorate a recession following on the heels of the collapse of a bubble, even if it cannot avert it entirely.

### The Keynesian "Liquidity Trap"

That the endogeneity of velocity to the supply of money disappears when one uses a broader monetary aggregate has already been established. There remains, however, the practical problem of targeting a broad monetary aggregate. It may still be the case that central bank control over the monetary base means less and less as one increases the level of aggregation. If this is the case, targeting short-term interest rates, as the Federal Reserve does currently, may still turn out to be the most effective way to control the quantity of a broader money stock. The quantity theory—though analytically operational and empirically meaningful—would not necessarily have straightforward implications for monetary policy, especially in the choice of instruments.

This problem is beyond the scope of this paper. But given that velocity increasingly stabilizes as one increases the level of monetary aggregation, it follows from the equation of exchange that targeting a Divisia aggregate would be approximately identical to targeting NGDP ($=Py$) (Belongia and Ireland 2015). The recent literature on the practicability of NGDP targeting would therefore be straightforwardly applicable (e.g., McCallum 2000; Woolsley 2015; Sumner 1997).

## Conclusion

Quantity-theoretic and credit cycle explanations for business cycles have developed in divergent directions since the distinction first became salient. Recent work in monetary aggregation has the potential to pull them back together. This paper has argued that a richer conception of the quantity of money points toward a fundamental connection between liquidity changes in financial markets and changes in a quantity of money with a direct causal connection to aggregate demand. In particular, exceptional liquidity events—the rise and collapse of asset bubbles—may be central to the business cycle, rather than symptomatic of it.

This analysis suggests many lines of further research, first of all in economic history, linking the development of liquid financial instruments with the emergence of a business cycle and identifying the particular shocks that effect the switch in

particular bubbles from stabilizing to destabilizing speculation. Second, as argued in the previous section, some credit cycle theories have perhaps been too hasty in declaring themselves incompatible with quantity theories and will require some degree of emendation. Third, quantity-theoretic and credit cycle explanations are not exhaustive of the broad approaches. A consensus approach to business cycles must also address the Real Business Cycle literature, a task which has not been attempted here. Despite the difficulty of the task, however, such a consensus may be more within reach than the divergent literatures would suggest.

**Acknowledgements** I am grateful to Lawrence White, Scott Burns, Peter Ireland, and several anonymous reviewers for helpful comments on earlier drafts. Errors and ill-conceived ideas are my own.

## References

Adrian, Tobias, and Hyun Song Shin. 2010. Liquidity and leverage. *Journal of Financial Intermediation* 19 (3): 418–437.

Alchian, Armen Albert. 1977. Why money? *Journal of Money, Credit, and Banking* 9 (1): 133–140.

Anderson, Benjamin. 1917. *The value of money*. New York: The MacMillan Company.

Banerjee, Abhijit, and Eric S. Maskin. 1996. A Walrasian theory of money and Barter. *Quarterly Journal of Economics* 111 (4): 955–1005.

Barnett, William A. 1980. Economic monetary aggregates: An application of index number and aggregation theory. *Journal of Econometrics* 14: 11–48.

Barnett, William A. 2016. Friedman and Divisia monetary measures. In *Milton Friedman: Contributions to economics and public policy*, ed. Robert A. Cord and Daniel Hammond. Oxford: Oxford University Press.

Barnett, William, and Neepa, Gaekwad-Babulal. 2017. *Multilateral divisia aggregates for the EMU and EU*. University of Kansas.

Beckworth, David, and Joshua R Hendrickson. 2015. Money still matters. Working paper.

Beckworth, David, and Joshua R. Hendrickson. 2012. Great spending crashes. *The B.E. Journal of Macroeconomics* 12 (1): 1–28.

Belongia, Michael, and Peter Ireland. 2015. A 'working' solution to the question of nominal GDP targeting. *Macroeconomic Dynamics* 19: 508–534.

Bernanke, Ben, and Mark Gertler. 1995. Inside the Black Box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives* 9 (4): 27–48.

Blanchard, Oliver, & Mark Watson. 1982. Bubbles, rational expectations, and financial markets. NBER working paper no. 945.

Boissay, Frederic, Fabrice Collard, & Frank Smets. 2013. Booms and systemic banking crises. European Central Bank working paper 1514.

Boldin, Michael. 1994. Econometric analysis of the recent downturn in housing: Was it a credit crunch? Federal Reserve Bank of New York research paper no. 9332.

Burns, Scott & Cameron Harwick 2017. Giving credit where credit is due: The benefits of bank-issued money. Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2688617.

Coval, Joshua, Jakub Jurek, and Erik Stafford. 2009. The economics of structured finance. *Journal of Economic Perspectives* 23 (1): 3–25.

Dornbusch, Rudiger, and Stanley Fischer. 1980. Exchange rates and the current account. *American Economic Review* 70 (5): 960–971.

Frankel, Jeffrey, and Andrew Rose. 1995. Empirical research on nominal exchange rates. In *Handbook of international economics*, vol. Vol, ed. Gene Grossman and Kenneth Rogoff, 3. Amsterdam: Elsevier.

Friedman, Milton. 1968. The role of monetary policy. *American Economic Review* 58 (1): 1–17.

Friedman, Milton. 1993. The 'plucking model' of business fluctuations revisited. *Economic Inquiry* 31: 171–177.

Friedman, Benjamin M., and Kenneth Kuttner. 1992. Money, income, prices, and interest rates. *American Economic Review* 82 (3): 472–492.
Friedman, Milton, and Anna Jacobson Schwartz. 1963. *A monetary history of the United States, 1867–1960*. Princeton: Princeton University Press.
Gabor, Daniela & Jakob Vestergaard. 2016. Towards a theory of shadow money. INET working paper. https://www.ineteconomics.org/research/research-papers/towards-a-theory-of-shadow-money.
Garrison, Roger W. 2001. *Time and money: The macroeconomics of capital structure*. New York: Routledge.
Gertler, Mark, and Nobuhiro Kiyotaki. 2015. Banking, liquidity, and bank runs in an infinite horizon economy. *American Economic Review* 105 (7): 2011–2043.
Harwick, Cameron. 2016. Cryptocurrency and the problem of intermediation. *Independent Review* 20 (4): 569–588.
Harwick, Cameron. 2018. Money and its institutional substitutes: The role of exchange institutions in human cooperation. *Journal of Institutional Economics* 14 (4): 689–714.
Hayek, Friedrich A. 1931. Reflections on the pure theory of money of Mr. J.M. Keynes. *Economica* 11 (33): 270–295.
Hayek, Friedrich A. 1933. *Monetary theory and the trade cycle*. London: Jonathan Cape.
Hendrickson, Joshua. 2014. Redundancy or mismeasurement? A reappraisal of money. *Macroeconomic Dynamics* 18: 1437–1465.
Hendrickson, Joshua. 2017. An evaluation of Friedman's monetary instability hypothesis. *Southern Economic Journal* 83 (3): 744–755.
Hendrickson, Joshua & David Beckworth. 2015. The supply of transaction assets, nominal income, and monetary policy transmission. Working paper.
Hetzel, Robert. 2009. Monetary policy in the 2008–2009 recession. *Economic Quarterly* 95 (2): 201–233.
Horwitz, Steven. 2000. *Microfoundations and macroeconomics: An Austrian perspective*. New York: Routledge.
Hummel, Jeffrey R. 2011. Ben Bernanke versus Milton Friedman: The Federal Reserve's emergence as the U.S. Economy's central planner. *Independent Review* 15 (4): 485–518.
Keynes, J.M. 1936. *The general theory of employment, interest, and money*. London: Palgrave MacMillan.
Kroszner, Randall. 1990. On the microfoundations of money: Walrasian and Mengerian approaches reconsidered in the light of Richard Rorty's critique of foundationalism. In *Economics and hermeneutics*, ed. Don Lavoie. New York: Routledge.
Krugman, Paul. 1991. Target zones and exchange rate dynamics. *Quarterly Journal of Economics* 106 (3): 669–682.
Kydland, Finn E., and Edward C. Prescott. 1990. Business cycles: Real facts and a monetary myth. *Federal Reserve Bank of Minneapolis Quarterly Review* 14 (2): 3–18.
Lachmann, Ludwig. 1956. *Capital and its structure*. London: Bell and Sons Ltd.
Lewin, Peter, and Nicolás Cachanosky. 2018. The average period of production: The history and rehabilitation of an idea. *Journal of the History of Economic Thought* 40: (forthcoming).
Luther, William, and Mark Cohen. 2014. An empirical analysis of the Austrian business cycle theory. *Atlantic Economic Journal* 42 (2): 153–169.
Luther, William, Mark Cohen, and Mark Cohen. 2016. On the empirical relevance of the Mises–Hayek theory of the trade cycle. *Advances in Austrian Economics* 20: 79–103.
McCallum, Bennett T. 2000. Alternative monetary policy rules: A comparison with historical settings for the United States, the United Kingdom, and Japan. *Economic Quarterly* 86 (1): 49–79.
Mehrling, Perry. 2012. The inherent hierarchy of money. In *Social fairness and economics: Economic essays in the spirit of Duncan Foley*, ed. Lance Taylor, Armon Rezai, and Thomas Michl. New York: Routledge.
Menger, Carl. 1892. On the origin of money. *The Economic Journal,* 2(6). Trans. by C.A. Foley.
Minsky, Hyman P. 1982. *Can "it" happen again?: Essays on instability and Finance*. Armonk, NY: M. E. Sharpe.
Serletis, Apostolos, and Victor Chwee. 1997. Resolving the liquidity puzzle. *Macroeconomic Dynamics* 1: 720–739.
Sumner, Scott. 1997. Reply to Garrison and White. *Journal of Money, Credit, and Banking* 29 (4): 542–545.

Tobin, James. 1963. Commercial banks as creators of 'money'. In *Banking and monetary studies*, ed. Deane Carson, 408–419. Honewood, IL: Irwin.

Williamson, Stephen D. 2012. Liquidity, monetary policy, and the financial crisis. *American Economic Review* 102 (6): 2570–2605.

Woodford, Michael, and Bennett T. McCallum. 2008. How important is money in the conduct of monetary policy? *Journal of Money, Credit, and Banking* 40 (8): 1561–1598.

Woolsley, W.William. 2015. Index futures targeting and monetary disequilibrium. In *Renewing the search for a monetary constitution*, ed. L.H. White, V.J. Vanberg, and E.A. Köhler. Washington, DC: Cato Institute.

Yeager, Leland. 1956. A cash-balance interpretation of depression. *Southern Economic Journal* 22 (4): 438–447.