Backed Money, Fiat Money and the Real Bills Doctrine

by

Michael Sproul

Working Paper 789
Department of Economics
University of California, Los Angeles
Bunche 2263
Los Angeles, CA 90095-1477
March 14, 1999
msproul@csun.edu

JEL Code: E51, E52, E58
Backed Money, Fiat Money, 
and the Real Bills Doctrine

Michael F. Sproul

ABSTRACT

In this paper I argue that the real bills doctrine has been wrongly discredited, and that it ought to displace the quantity theory as the dominant theory of money. The discussion begins with the observation that the issue of backed money will not be inflationary as long as central banks follow the real-bills rule of only issuing money to those customers who offer good security in exchange. I then contend that modern paper currencies, which we normally think of as unbacked fiat money, may in fact be (and probably are) backed. If correct, this would imply that the real bills doctrine, and not the quantity theory, is a correct model of the value of modern money. The paper concludes by discussing some key controversies in the history of the real bills doctrine, and shows that the major arguments responsible for the defeat of the real bills doctrine contain obvious and serious errors.
I. Introduction

When the Directors of the old Bank of England were accused of having allowed the pound to depreciate between 1797 and 1810, their defense was based on the real bills doctrine. They stated that they had only issued money to those customers who offered good security in exchange for the money. Therefore, they claimed, the Bank had only issued as much money as the legitimate needs of business required. The Bullion Committee appointed by the House of Commons in 1810 denounced this defense as "wholly erroneous in principle" (Gilbart, 1882, p. 53). Sixty-three years later, the bankers' answers were still derided as "almost classical by their nonsense." (Bagehot, 1873, p. 86) It would be difficult to count the number of times that similar debates over the real bills doctrine have flared over the centuries. A few episodes are summarized by Mints (1945, p. 9.):

The real-bills doctrine has been a most persistent one. Given its most elegant statement in all its history by Adam Smith in the Wealth of Nations, it has since served as the defense for the directors of the Bank of England during the period of the Restriction. With some changes it reappeared as the banking principle; it was the main reliance of the agitators for banking reform in the United States before 1913; it was as comforting to the Federal Reserve Board following the depression of 1921 as it had been a century earlier to the directors of the Bank of England; more recently it has re-emerged as the doctrine of "qualitative" control of bank credit; and, quite aside from these special uses to which it has been put, it has been consistently defended throughout all these years by a large proportion of bankers and economists.

Since Mints' time, a dissident tradition opposed to the quantity theory (and sometimes favorable to real-bills principles) has been evident in the writings of Tobin (1963), Black (1970), Samuelson (1971), Wallace (1982), and Sargent and Wallace (1982). Still, most
economists' attitudes toward the real bills doctrine have remained far from charitable. G. A. Selgin (1989, p. 489.), for example, comments that

The dead horses of economic theory have a habit of suddenly springing back to life again, which is why it is necessary to beat them even when they appear lifeless.

In what follows I hope to revive this dead horse.

II. Backed Money

Empirical studies by Sargent (1982), Smith (1985a, 1985b, 1988), Calomiris (1988), Siklos (1990), Bomberger and Makinen (1991), and Cunningham (1992) have found that the value of money is more accurately predicted by a real-bills type "backing theory" than by the quantity theory. Opposing views favoring the quantity theory have been presented by Bordo and Marcotte (1987), Laidler (1987), Michener (1987), and McCallum (1992). It is not my purpose to comment directly on these debates. Instead, my goal is to clarify the elements of the real bills doctrine, and to present it as a plausible alternative to the quantity theory. At the same time, I hope to correct certain errors that have crippled past inquiries.

Three main ideas form the basis of my discussion:

1. It is possible that fiat money does not exist. A true fiat money would be unbacked, meaning that the issuing entity would hold no assets against the money, would not recognize the money as its liability, and would be unable (not just temporarily unwilling) to redeem the money for anything of value. None of these conditions is true of the U.S. dollar or any other money, and when we find historical examples of moneys that have lost all backing (such as the Assignats and the Continental dollars) we find that they have simultaneously lost all value.
2. If so-called fiat money is actually backed, then its value is correctly described by the real bills doctrine. That is, banks that follow the real-bills rule of only issuing money in exchange for good security will automatically acquire additional backing every time new money is issued, thus assuring that backing moves in step with the quantity of money.

3. Past debates over the real bills doctrine have assumed the existence of fiat money, and as a result have committed serious errors that have never been exposed. My discussion points out errors in the works of Adam Smith, Henry Thornton, David Ricardo, and Lloyd Mints. The discussion is necessarily brief and selective. My purpose in covering these writers is to point out that certain key refutations of the real bills doctrine are erroneous.

A. The Real Bills View of Backed Money

The real bills doctrine holds that money issued in exchange for sufficient security (usually meaning short-term commercial bills) will not cause inflation. For example, Figure 1 represents a bank which has taken in 100 ounces of gold on deposit and issued 100 'credits' (either bank notes or deposits), each of which is a claim to one ounce of gold.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 oz. of gold</td>
<td>100 credits</td>
</tr>
</tbody>
</table>

**Figure 1**

I assume for simplicity that this bank has zero net worth, and that each credit gives its holder the right to claim one ounce of gold from the bank on demand. Assuming the bank always honors these claims, each credit will be worth one ounce of gold. But should the bank lose some of its gold, the credits would become an equity claim to whatever assets the bank has left. For example, if the bank lost 30 ounces of its gold, the 100 credits
could lay claim to only 70 ounces of gold, so the credits would fall in value to 0.7 ounces each.¹

Granting for the moment that this money is backed, its value would not depend on any of the following factors:
(1) the quantity of money,
(2) the convertibility of the money,
(3) money demand,
(4) the quantity of derivative moneys,
(5) fiscal policy.
I will discuss each of these in turn.

1. The Quantity of Money

Suppose that one of the bank’s customers requests a loan of 100 credits, over and above the 100 credits already issued (Figure 2, line 2), but instead of offering gold in exchange he offers IOU’s (his own or someone else’s) with a current market value of 100 ounces of gold. The banker would have no reason to refuse this offer, and so he would issue 100 more credits, thus doubling the money supply.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) 100 oz. of gold</td>
<td>100 credits</td>
</tr>
<tr>
<td></td>
<td>----------------</td>
</tr>
<tr>
<td>2) IOU’s worth</td>
<td>100 credits</td>
</tr>
<tr>
<td>100 oz. of gold</td>
<td>----------------</td>
</tr>
</tbody>
</table>

Figure 2

¹The value of these credits resembles the payoff to the seller of a put option. If depositors believe that the bank might lose some of its gold, then it seems that they would value these credits at slightly less than one ounce. However, depositors might realize that gold held in their own pockets is also subject to loss, so they might be indifferent between the bank’s uncertain promise of convertibility and their own uncertain ability to hold on to the gold. Alternatively, the argument I make below would not be substantially changed if the bank were assumed to have positive net worth, which would act as a cushion against the loss of gold.
There are now 200 credits laying claim to assets worth 200 ounces of gold, so each credit must still be worth one ounce. The banker can safely issue any amount of money the public desires, provided that he only issues credits to those customers who offer *sufficient security* (i.e., resources worth one ounce of gold). This rule is nothing but the real bills doctrine, except that the security need not be "short-term commercial bills" (i.e., real bills). Anything worth 100 ounces of gold would serve equally well.

A quantity theorist accustomed to using the equation of exchange, \(MV=PT\), would find it strange that the money supply could double without affecting prices. The simplest reply that a real-bills adherent could give is that \(V\) must have fallen or \(T\) must have risen. Newly-issued money could have displaced barter, credit, balance offsets, etc. One should keep in mind, after all, that \(T\) does not represent total output of goods, but only the quantity of goods bought with the particular kind of money represented by \(M\). More importantly, the equation of exchange is tautological. For example, we could let \(M\) represent the number of shares of GM stock in existence, and \(V\) their velocity of circulation; \(P\) would be the number of shares of GM needed to buy a given bundle of goods, and \(T\) would be the number of those bundles bought with GM stock. The equation \(MV=PT\) would be 'correct' for GM stock, but in no sense would it *determine* the value of GM stock. By the same reasoning, the value of backed money is not determined by \(M\), \(V\), \(P\), or \(T\). Money's value is determined by its backing; the equation of exchange merely adds that the total amount of money spent is equal to the amount received. One could just as reasonably state that the amount of rain falling from the sky equals the amount hitting the ground, and call it a theory of rainfall.

While it is true that money-creation will not affect the credits' value *relative to gold*, it is still possible that the issue of credits might reduce the monetary demand for gold and
thus reduce its value. This may seem to support the quantity theory proposition that
money-creation, even on sufficient security, will cause inflation. However, if gold’s value
drops because of competition from the bank’s credits, the drop would reflect increased
economic efficiency, as monetary gold is released for other uses. But this is the effect of
an improvement in monetary technology—not of an increase in the quantity of money.

2. Convertibility

Suppose that the bank in Figure 2 closes over the weekend, thus making its notes
temporarily inconvertible. Then, while the bank is closed, the value of the IOU’s drops to
50 ounces of gold (perhaps because of a default by their issuers). The credits would then
trade for 150/200= .75 ounces for the rest of the weekend. If the bank restored
convertibility at one ounce per credit on Monday morning, it would face a run. The first
150 depositors would get their gold (or something of equivalent value) and the last 50
would get nothing. As the run progressed the expected value of the credits would fall, so
that, for example, after 80 credits had been redeemed at one ounce each the value of each
remaining credit would be 70/120= .58 ounces. If the bank continued to offer one ounce
per note, customers would see it as an empty promise, and they would value the notes at
only .58 ounces. Clearly, it is backing that matters, not convertibility. Put another way,
convertibility requires backing, but backing does not require convertibility.

If banks can suspend convertibility for a weekend, they can suspend it for a hundred
years. For example, a banker might make this offer to his depositors: "Give me resources
worth one ounce of gold today, and in 100 years I will return your deposit plus a
competitive interest yield." Each credit issued on these terms would initially be worth
one ounce of gold, and its value would grow at the rate of interest. If customers preferred
the credits to have roughly constant value, then the banker could make periodic interest
payments, say by adding .05 credits per year to the account of each credit-holder. Note that the banker need not specify the exact date of redemption, or even that he will pay with certainty. All that matters to the customers is that the credits are a claim to something of value.

We are now in a position to make an important observation: It is possible that what we think of as unbacked fiat money is in fact money that is backed but inconvertible. Consider the usual justification for asserting that the dollar is fiat money:

You cannot convert a Federal Reserve Note into gold, silver, or anything else. The truth is that a Federal Reserve Note has no inherent value other than its value as money, as a medium of exchange. (Tresch, 1994, p. 996.)

Observing that the dollar is inconvertible, economists conclude that it is unbacked. The most remarkable thing about this simple non-sequitur is that it has survived virtually unquestioned for centuries. If we want to show that the dollar is not just inconvertible, but *unbacked*, it is not enough to say that the Federal Reserve does not pay out gold on demand. Yet economists' belief in fiat money, and in fact the better part of monetary theory, is founded on nothing but this obviously flawed premise. Add to this the facts that the Federal Reserve (like all central banks) does in fact hold assets against the money it issues, that no dollar is ever issued except in exchange for valuable assets, and that the Federal Reserve's own balance sheet plainly identifies those assets as "Collateral Held Against Federal Reserve Notes", and we have good reason to wonder if fiat money is no more real than the phlogiston, ether, and caloric of early physical sciences.

Consider the Bank of England's suspension of convertibility from 1797-1821. Prior to 1797, the British pound had been convertible into gold at the rate of about 3.9 pounds to the ounce. On February 26, 1797, the Bank suspended convertibility. As Table 1
(Cannan, 1969, p. xliii) shows, the immediate effect of the suspension upon the value of the pound was negligible—a fact which surprised most observers.2

Table 1. British Price Index (1782=100)

<table>
<thead>
<tr>
<th>Year</th>
<th>Prices</th>
<th>Year</th>
<th>Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1792</td>
<td>93</td>
<td>1807</td>
<td>132</td>
</tr>
<tr>
<td>1793</td>
<td>99</td>
<td>1808</td>
<td>149</td>
</tr>
<tr>
<td>1794</td>
<td>98</td>
<td>1809</td>
<td>161</td>
</tr>
<tr>
<td>1795</td>
<td>117</td>
<td>1810</td>
<td>164</td>
</tr>
<tr>
<td>1796</td>
<td>125</td>
<td>1811</td>
<td>147</td>
</tr>
<tr>
<td>1797</td>
<td>110</td>
<td>1812</td>
<td>148</td>
</tr>
<tr>
<td>1798</td>
<td>118</td>
<td>1813</td>
<td>149</td>
</tr>
<tr>
<td>1799</td>
<td>130</td>
<td>1814</td>
<td>153</td>
</tr>
<tr>
<td>1800</td>
<td>141</td>
<td>1815</td>
<td>132</td>
</tr>
<tr>
<td>1801</td>
<td>153</td>
<td>1816</td>
<td>109</td>
</tr>
<tr>
<td>1802</td>
<td>119</td>
<td>1817</td>
<td>120</td>
</tr>
<tr>
<td>1803</td>
<td>128</td>
<td>1818</td>
<td>135</td>
</tr>
<tr>
<td>1804</td>
<td>122</td>
<td>1819</td>
<td>117</td>
</tr>
<tr>
<td>1805</td>
<td>136</td>
<td>1820</td>
<td>106</td>
</tr>
<tr>
<td>1806</td>
<td>133</td>
<td>1821</td>
<td>94</td>
</tr>
</tbody>
</table>

During the suspension, the pound was not a true fiat money. Not only did the Bank maintain assets as collateral against the pound, not only did it continue to recognize the pounds as its liability, but it actually did resume convertibility after a delay of 24 years. There is a difference between a money that is temporarily inconvertible and one that is truly unbacked, but this difference was lost on the major writers of the period (Ricardo, 1811, p. 114; Thornton, 1802, p. 244.), who consistently claimed that the value of the pound was not determined by its backing, but by a limitation of quantity. By way of

2"Looking back from the safety of 1798, 'A Proprietor of Bank Stock' thus summarized the transition: 'In this desponding state, when all men dreaded, with the utmost anxiety, the event that was seen to be inevitable, and not far distant, and which it was supposed would involve the Kingdom in general bankruptcy and intire ruin, the 26th February, 1797, was the crisis that gave the happy turn, and almost instantly dismissed all the horrors and fears that surrounded us, restored complete confidence..." (Ashton and Sayers, 1953, p. 19.)
comparison, the U.S. dollar has been inconvertible to individuals since 1933, and to foreign governments since 1971. Like the old Bank of England, The Federal Reserve has maintained its assets throughout, and has always recognized Federal Reserve dollars as its liability. Unlike the Bank of England, it has not resumed convertibility, and has given no indication that it will. But the simple accounting identity of assets equaling liabilities assures that the Federal Reserve’s assets are sufficient to redeem every dollar it has issued. In spite of this, the proposition that the dollar’s value is determined by a limitation of quantity, and not by its backing, commands a degree of agreement among economists which rivals that of the Law of Demand.

3. Money Demand

If our banker has resources worth 100 ounces of gold, as backing for 100 credits, then those credits will be worth one ounce each regardless of the public's demand for them. If their value exceeded one ounce by (say) 2 percent, then rival bankers could earn easy profits by issuing credits for 1.02 ounces of gold, keeping 1 ounce as backing, and spending the seignorage of .02 ounces on their own consumption. This profit opportunity will exist as long as there is any seignorage, so the only stable solution is for the seignorage to be driven to zero. The same reasoning implies that there can be no such thing as fiat money, since fiat money is money whose whole value is seignorage. That private, competitive banks are unable to issue fiat money does not admit of dispute. That the same thing could also be true of government banks has, to my knowledge, never even been suggested.
Depending on who is talking, we hear that fiat money has value because other people value it (Samuelson, 1980, p. 261), because the government accepts it for taxes\(^3\) (Wicksteed, 1910, p. 619), because it is useful for making exchanges and limited in supply (Marshall, 1922, p. 49), because the government requires banks to hold it (Fama, 1980, p. 56), or because it allows us to transfer wealth to our children (Wallace, 1980, p. 50). The trouble with these theories is that they fail to consider rival moneys. Each theory begins by asserting that there is some force (e.g., liquidity services) that creates a demand for intrinsically worthless pieces of paper. They then assert that it would only be necessary to limit the supply of these pieces of paper in order to give them value. But if this were true, then the issuers of those pieces of paper would be getting a free lunch; rivals would eagerly issue their own pieces of paper, until competition drove their value to zero. Of course, no one believes that private, competitive banks can earn economic profit of this kind; furthermore, if a private bank could issue notes on which it paid no interest, while investing the proceeds at 5%, then competitors would issue rival notes until the interest spread just covered costs of printing, periodic redemption, controlling counterfeiting, etc. Given this, it is strange to see how easily economists accept the proposition that central banks earn seignorage on their note issue, and that note issue therefore gives a free lunch to the Federal Reserve, especially if the dollars go to foreign

---

\(^3\)There have been many cases where acceptability for taxes has conferred value on paper money, but these have been cases where the taxes themselves constituted backing. For example, in the American colonial period, colonies would issue paper money either on loan or in direct payment for government expenses. These notes were declared acceptable for taxes, and they were officially rated in terms of some other money. A New York shilling, for example, was legally rated at 8 shillings to the ounce of silver from 1709-1718 (Brock, pp. 66-67.). Had these shillings been backed in an ordinary sense, the holder of 8 shillings could have claimed one ounce of silver from the colony. But the tax backing used in the colonies gave the holder of 8 shillings the right to discharge a tax (or other debt to the colony) of one ounce of silver. Conventional backing thus depends on the ability of the issuer to pay silver, while tax backing depends on the issuer’s ability to take away silver.
countries. Since most of us are trained to be suspicious of free lunches, this idea deserves some skepticism.

The only reason to believe that the Federal Reserve earns seignorage is that it has the power to suppress rival moneys. But governments cannot suppress commodity money, credit, foreign moneys, or barter. There are also traveler’s checks, gift certificates, and scrip, all of which are bank notes issued by non-bank institutions. (In point of fact the only entities barred from issuing bank notes are banks themselves.) Given this rivalry, it is hard to believe that note issue could yield abnormal profits, even to government banks. Where countries are small, weak, and close together, it seems impossible.

One might argue that the issue of rival money is constrained by reserve requirements, but these only apply to conventional bank accounts— not to credit cards, foreign money, eurodollars, scrip, and so on. In light of this limitless potential for the issue of rival moneys, fiat money seems implausible. In contrast, the view that the dollar is backed but inconvertible only requires us to believe that money is valued for the same reason that any other financial security is valued.

A stock market analogy may help explain the role of reserve requirements. Just as bankers issue checking accounts that are claims to Federal Reserve dollars, stock market traders routinely issue derivative securities which are claims to GM stock. Suppose that those traders were required to hold "reserves" of genuine GM stock against the derivative shares that they issue. Would this requirement increase the value of GM stock? The answer is no, since this requirement would not affect GM's assets or its liabilities. If one accepts the idea that the dollar is backed, then the same reasoning implies that reserve requirements are irrelevant to the value of the dollar.

A reasonably skeptical reader could still argue that a constraint on rival moneys could cause the dollar to sell for a few points above its backing. However, one could also argue
that GM stock could be raised a few points (or lowered!) by a constraint on the issue of rival stocks. But I doubt that this argument would persuade economists to abandon the theory that stocks are valued according to their backing. When applied to money, the same argument is clearly an inadequate reason for believing that the dollar is a pure fiat money.

Why does the Federal Reserve (and every other central bank) bother to hold gold and financial securities if the dollar does not get its value from backing? How could fiat money ever come into circulation in the first place? Why issue dollars through an expensive central bank instead of just printing them and spending them? Why do even the weakest countries seem to be able to maintain "fiat" money in circulation? These questions and many more have inspired a mountain of convoluted monetary theories. But if fiat money is in fact an illusion—if it is actually backed but inconvertible, then these questions do not even arise.

4. The Quantity of Derivative Moneys

Checking accounts issued by private banks entitle depositors to claim Federal Reserve notes on demand. Thus we could call the accounts 'derivative money' (a term I prefer to 'inside money') since they are claims to Federal Reserve dollars. The dollar, in turn, is an inconvertible claim to the assets of the Federal Reserve, and is itself a derivative money, even though we commonly think of it as base money. By analogy, there are derivative financial securities (options, warrants, etc.) that are claims to GM stock. The GM shares, in turn, are a claim (generally inconvertible) against GM's assets. Thus the base stock is itself a derivative security.

The issue of derivative shares of GM stock does not change GM's assets or liabilities, and therefore does not depreciate GM stock. Similarly, if the dollar has value because of
its backing, then the issue of derivative dollars does not reduce the value of the dollar. The quantity theory, however, implies that derivative dollars reduce the demand for base dollars and thus cause inflation. On this view, a legitimate banker is no different from a counterfeiter. Both increase the quantity of money, so both cause inflation! This belief has led to a number of proposals to require all banks, public and private, to maintain 100% reserves against the money they issue (e.g., Friedman, 1948, p. 372.). This idea, besides being out of character for libertarian economists, ignores the fact that banks recognize their money as their liability, while counterfeiters do not.

Derivative moneys raise the question of what is 'money' and what is not. Should we include only gold? Notes issued by the central bank? Private bank notes? What about checking accounts, credit cards, traveler's checks, eurodollars, overdrafts, and gift certificates?⁴

These questions could only matter to someone who believes that the value of money depends on its quantity. No one bothers to wonder whether derivative shares of stock (options, hypothecated shares, etc.) should be counted along with genuine shares. We simply recognize that derivative shares can take many forms, and that each share will be valued in accordance with the resources backing it. If economists understood money as well as they understand stock, they would recognize that derivative moneys also take many forms, but that their quantity is irrelevant to the value of the dollar.

Misunderstandings of this point are widespread. For example, Salin (1984, p. 13.) worries that eurodollars reduce the Federal Reserve's control over the world supply of dollars. On this view, the issuer of a eurodollar is violating the 'brand name' of the dollar. The real

⁴In 1710, people denied that bank notes were money, since they ultimately had to be paid off in coin (e.g., Harley, 1710, p. 42). In 1845, people denied that checking accounts were money, since they ultimately had to be paid in notes or coin (Fullarton, 1845, p. 32). Having learned nothing from this, modern economists deny that credit cards are money, since they must ultimately be paid off by check, notes, or coin.
bills view, however, is that a foreign bank which issues a eurodollar is analogous to a foreign brokerage house which issues a derivative share of GM stock. Viewed in this way, we see that eurodollars are no cause for concern to Americans, and in fact are likely to improve the efficiency of the market for dollars.

5. Fiscal Policy

Sargent and Wallace (1981, p. 176) observe that a government deficit can be financed either by borrowing or by printing money. They conclude that

...once the limit on the federal debt per capita that can be marketed with the public has been reached, the Fed has no choice: It must increase base money. That is, it must "monetize" all of the additional government borrowing by purchasing all real additions to the stock of interest-bearing debt that the treasury issues. More generally, given the time path of fiscal policy and given that government interest-bearing debt can only be sold at a real interest rate exceeding the growth rate n, the tighter is current monetary policy, the higher must the inflation rate be eventually.

That "the Fed has no choice" is incorrect. The Federal Reserve, like most central banks, has the right to buy as much or as little government debt as it chooses. Furthermore, the competitive auction process assures that it pays market value for that debt. So suppose that the Treasury, having spent itself into bankruptcy, tries to raise some cash by offering a bond that promises to pay one million ounces of gold next year. Since nobody trusts that promise, that bond will sell today for (say) 100 ounces. If the Federal Reserve buys that bond with bank notes worth 100 ounces, then its assets will rise as much as its liabilities, and the value of the dollar will not change.

If the central bank is not independent, fiscal policy could affect the value of the dollar, but not for the reasons given by Sargent and Wallace. If the government orders the bank

page 16
in Figure 2 to hand over 100 credits in exchange for bonds that are really worth only 40 ounces, then the value of the credits will drop to $240/300=.8$ ounces. The inflation is not caused by the increase in the supply of money, but by the reduction in the bank's ratio of assets to units of money. This, in turn, can only happen when the central bank is subservient to the Treasury. Sargent (1981) has observed that inflation abates when a subservient central bank becomes independent, but his interpretation is that independence allows the central bank to restrain the growth of the money supply. The real bills interpretation is that independence frees the central bank from government policies that would reduce the bank's ratio of assets to units of money.

A money-issuing bank located in Mexico might own nothing but U.S. government bonds, even though its money circulates in Mexico. Thus, no matter what happens in Mexican fiscal affairs, the bank's money will be stable. Conversely, if U.S. bonds drop in value, the Mexican bank's money will depreciate even when Mexican finances are in order. Note, however, that the U.S. Treasury cannot force a Mexican bank to buy U.S. bonds, and therefore the Mexican bank will never monetize U.S. debt. If the Mexican bank's money became widely-used in the U.S., then the U.S. would have a completely independent central bank which would never monetize U.S. debt, regardless of U.S. fiscal policy.

B. Inflation, Money Shortages, and Inflationary Feedback

Return to the case of a bank that holds miscellaneous goods worth 100 ounces of gold as backing for 100 outstanding credits. Now suppose that the bank follows an 'easy money' policy and begins issuing credits in exchange for securities worth only .95 ounces of gold. If the banker issues 50 new credits on these terms, then the value of each credit will drop to $(100+(50\times .95))/150=.983$ ounces. We get the familiar result that easy money
leads to inflation. Note that the inflation results from a drop in the bank's ratio of assets to money, not from an increase in the quantity of money relative to aggregate output of goods.

When depositors see the bank handing out credits worth one ounce of gold in exchange for securities worth only .95 ounces, they will rush to buy these credits. Thus the quantity of bank money will expand as long as the bank follows its easy money policy. At the same time, the value of the money will fall because of the drop in the bank’s ratio of assets to money. This point has confused economists for centuries. The true cause of inflation is that the bank issues money for insufficient security. The apparent cause of inflation—the increase in the quantity of money—is really just a side effect of the bank's easy money policy.

This analysis provides an interesting view of credit rationing. In discussing an English credit rationing episode, Friedrich Hayek commented that:

   This recourse to a rationing of credit caused renewed stringency in the money market in the spring of 1796 and evoked loud protests from the City (London).
   It is not easy to reconcile these complaints about the continued scarcity of money during this period with the no less insistent complaints about high prices, and with the continued unfavorable course of the exchanges." (Hayek, 1933, p. 40)

From a real bills perspective, it is easy to reconcile the two sets of complaints. If the Bank of England were issuing notes on insufficient security, then each issue would cause the Bank's ratio of assets to money to fall and thereby reduce the value of the pound. Meanwhile, since the Bank was issuing pounds in exchange for security worth (say) .98 pounds, customers would eagerly buy any pounds the Bank offered. Assuming the Bank
dealt with the resulting surge in demand by rationing credit, we would naturally expect people to complain of a shortage of credit.

A problem arises when the bank's assets are denominated in the bank's own credits. (e.g., The Federal Reserve's bondholdings are denominated in dollars.) If the bank's credits were to depreciate for any reason, then the bank's assets would also depreciate. The credits would then fall still further, and so on. Let P represent the price, in gold, of a bank's credits. Assume that the bank initially has 100 outstanding credits, each worth one ounce of gold, and that its assets consist of 30 ounces of gold plus bonds, denominated in the bank's own credits, which have a market value of 70 credits (Figure 3). Now suppose

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
</table>
| 1) 30 ounces of gold
   Bonds worth 70 credits | 100 Credits |
| 2) Bonds worth 6 credits | 10 Credits |

Figure 3

that the bank follows an easy money policy and issues 10 new credits in exchange for bonds with a true value of only 6 credits (line 2). If assets (30 ounces of gold plus bonds worth 76 credits) are to equal liabilities (110 credits), then it must be true that

\[30 + 76P = 110P\]

which implies \(P = 0.888\) ounces of gold. If the bank had instead started with 20 ounces of gold and bonds worth 80 credits, the same easy money policy would have reduced the value of the bank's credits to \(P = 0.833\) ounces. Thus, the bank's holdings of gold or other 'real' assets would limit the force of inflationary feedback. The smaller are real reserves, the more volatile the bank's money.
C. The Faults on Both Sides

Debates between quantity theorists and real bills adherents have flared repeatedly over the last three centuries. As a rule the quantity theory has come out on top, but often the controversies "have slumbered, rather from the exhaustion of the combatants than from the acknowledged defeat of either party." (Farrer, 1898, p. 78.) When a debate becomes as protracted as this, there is good reason to think that both sides are asking the wrong questions. I contend that two key errors lie at the root of the problem. First, both sides accepted the existence of fiat money, without exploring the possibility that paper money could be backed but inconvertible. Second, both sides held that stable prices would be achieved if the money supply rose and fell with the "needs of business" (i.e., real output). They ignored the role of the issuing bank's assets. In what follows I discuss some key episodes in the history of the real bills doctrine in order to show the effects of these and other errors.

1. Adam Smith

What a bank can with propriety advance to a merchant or undertaker of any kind, is not either the whole capital with which he trades, or even any considerable part of that capital; but that part of it only, which he would otherwise be obliged to keep by him unemployed, and in ready money for answering occasional demands. If the paper money which the bank advances never exceeds this value, it can never exceed the value of the gold and silver, which would necessarily circulate in the country if there was no paper money; it can never exceed the quantity which the circulation of the country can easily absorb and employ.

When a bank discounts to a merchant a real bill of exchange drawn by a real creditor on a real debtor, and which, as soon as it becomes due, is really paid by that debtor; it only advances to him a part of the value which he would otherwise be obliged to keep by him unemployed and in ready money for answering occasional demands. (Smith, 1776, p. 322)
The idea that a bank should only lend as much money as its customers would otherwise have kept in their tills is nonsense. We all know that a bank can, with propriety, lend to a merchant any amount reasonably short of the merchant's net worth. Such a bank will always have assets sufficient for its liabilities, and thus its money will maintain its value. Smith's proposition that the amount of bank money should correspond to the amount of gold and silver it replaces implicitly assumes that money's value will be maintained by a limitation of its quantity, and not by a matching of bank assets to the quantity of money. Having no clear idea of this distinction, Smith made a dangerous misinterpretation of central bank policies:

By issuing too great a quantity of paper, of which the excess was continually returning, in order to be exchanged for gold and silver, the bank of England was for many years together obliged to coin gold...at an average (of) about eight hundred and fifty thousand pounds. For this great coinage the bank (in consequence of the worn and degraded state into which the coin had fallen a few years ago) was frequently obliged to purchase gold bullion at the high price of four pounds an ounce, which it soon after issued in coin at 3 l. 17 s. 10 1/2 d. an ounce, losing in this manner between two and a half and three per cent. on the coinage of so very large a sum. (Smith, 1776, p. 286.)

By attributing the bank's loss upon the coinage to excessive note issue, Smith got the chain of causation precisely backwards. Note issue expanded because the bank issued money on insufficient security. But on Smith's interpretation, it was excessive note issue that forced the bank to accept insufficient security for its notes. When the bank pays out four one-pound notes for an ounce of gold that it reissues as coins worth 3 l. 17 s. 10 1/2 d., (about 3.9 pounds) then customers will eagerly bring one ounce of gold to the bank to exchange for 4 pound notes. They will then return 3.9 of those notes to the bank in
exchange for coins containing one full ounce. With the bank losing 2.5% on each exchange, the bank would soon exhaust its treasure.

What would get the bank started on such a ruinous course? Smith himself gave the answer: "the worn and degraded state into which the coin had fallen". When a new one pound coin was issued, it would have contained gold worth one pound. Because of wear and clipping, that same coin might soon contain gold worth 0.975 pounds. The heavy coins would disappear from the circulation and the value of the pound would drop by 2.5%. Thus, as Smith states, it would require four one-pound notes to buy the same amount of gold that it previously bought for 3.9 notes. But the effort to restore the pound to its old value will be hopeless. If the bank recoined the newly-purchased gold and sold it for 3.9 pounds per ounce, it would be buying its own notes on the market for 2.5% more gold than the notes were worth. The pound's backing would fall, and the value of the pound would drop.

Modern central banks make the same mistake when they attempt to support their currency in world markets. Suppose, for example, that the pound trades for $1.60, but that the Bank of England wants the pound to trade for $1.80. The quantity theory prescription would be for the Bank of England to use its dollar reserves to buy pounds in the open market for $1.80. But by paying $1.80 for a British pound that is only worth $1.60, the Bank would lose $.20 on each purchase. Its ratio of assets to currency would drop, and the value of the pound would fall. Small wonder then, that efforts to support various currencies are so often followed by devaluation (Taylor, 1982, pp. 356-68). This empirical result is exactly what the real bills doctrine implies, and exactly opposite to the implications of the quantity theory.
2. Henry Thornton

Henry Thornton (1802) is largely responsible for a popular misconception that bank credit will not be adequately limited by the requirement that loans only be granted on the basis of sufficient security:

"Real notes," it is sometimes said, "represent actual property. There are actual goods in existence, which are the counterpart to every real note. Notes which are not drawn, in consequence of a sale of goods, are a species of false wealth, by which a nation is deceived. These supply only an imaginary capital; the others indicate one that is real."

In answer to this statement it may be observed, first, that the notes given in consequence of a real sale of goods cannot be considered as, on that account, certainly representing any actual property. Suppose that A sells one hundred pounds worth of goods to B at six months credit, and takes a bill at six months for it; and that B, within a month after, sells the same goods, at a like credit, to C, taking a bill; and again, that C, after another month, sells them to D, taking a like bill, and so on. There may then, at the end of six months, be six bills of 100 pounds each existing at the same time; and every one of these may possibly have been discounted. Of all these bills, then, only one represents any actual property. (Thornton, 1802, p. 86.)

Thornton's mistake was in failing to realize that no matter how we look at it, 600 pounds of debt will not be created unless security worth 600 pounds is offered in exchange. The scenario Thornton described is illustrated in Figure 4. Suppose A sells wheat worth 100 pounds to B, and receives B's IOU in exchange. B then sells the wheat to C, in exchange for C's IOU, and the process repeats six times. If B is respected in the community, then his IOU might serve as money, and the exchange would have increased the money supply by 100 pounds. Alternatively, as Thornton states, B's IOU might be discounted by a banker, and the banker's IOU would then serve as money. In either case, each exchange potentially increases the supply of money, and it is possible, as Thornton
states, that six successive sales of the same wheat could increase the money supply by 600 pounds.

Thornton’s error becomes apparent once we realize that A would only accept B’s IOU if it were backed by something worth 100 pounds. For example, B might own property that A could take from him in court. It is as if B’s IOU were actually backed by a lien on B’s property (Figure 4), C’s IOU by a lien on C’s property, etc. Every additional sale of the wheat would create new IOU’s backed by new goods, and no matter how far the process went, the self-interest of the parties involved would assure that every new IOU would be backed by goods of commensurate value.

Figure 4

Thornton’s argument that the six IOU’s are backed only by the single unit of wheat is plainly indefensible, and it is surprising that a banker such as Thornton would have forgotten the importance of collateral to the value of an IOU. The implications for his refutation of the real bills doctrine are straightforward: Thornton thought he had proven that derivative moneys (the IOU’s in Figure 4) can be issued without backing. He also believed that the value of the Bank of England’s pounds (the base money of his day) was unaffected by its backing (Thornton, 1802, p. 244.). Thus Thornton could reach no other conclusion than that the value of money in general derived not from its backing, but from a limitation of its quantity. Once we recognize that derivative money is backed, it
remains only to show that base money is also backed. In the case of the British pound, this is not difficult. As I have mentioned above, the Bank of England maintained its backing throughout the restriction period of 1797-1821. The Bank always recognized its notes as its liability, and actually did resume convertibility in 1821. There seems to be no qualitative difference between the Bank of England during this period and the hypothetical bank in Figure 2, which suspends convertibility of its notes each Friday and resumes it on Monday. The Bank of England suspended convertibility for a longer period, and the public had a more uncertain expectation that convertibility would be resumed, but these are differences of degree only. Economists familiar with Thornton’s work can hardly overstate its importance to the economics profession’s rejection of the real bills doctrine, but his belief in the irrelevance of backing was clearly incorrect.

3. The Bullionist Debates

a. The Quantity Theory Position

The most extensively debated inflationary episode in the annals of monetary theory occurred during the suspension of convertibility by the Bank of England from 1797-1821. (For a history of the period, see Ashton & Sayers (1953)). The controversy centered on the question of why the pound had depreciated during the suspension period. The 'Bullionist' (quantity theory) explanation was championed by Ricardo, who held that money-issuing banks had increased the quantity of money:

Let us suppose all the countries of Europe to carry on their circulation by means of the precious metals, and that each were at the same moment to establish a Bank on the same principles as the Bank of England--Could they, or could they not, each add to the metallic circulation a certain portion of paper? and could they not permanently maintain that paper in circulation? If they could, the question is at an end, an addition might then be made to a circulation already sufficient, without occasioning the notes
to return to the Bank in payment of bills due. If it is said they could not, then I appeal to experience, and ask for some explanation of the manner in which bank notes were originally called into existence, and how they are permanently kept in circulation. (Ricardo, 1811, p. 117.)

In this statement, Ricardo convincingly shows that banks are able to increase the quantity of money. Being imbued with the quantity theory, he considered this as satisfactory proof that banks cause inflation. But the connection between money and inflation should have been the very point under examination. On real bills principles, an increase in the money supply, accompanied by an equal increase in bank assets, will have no effect on prices. But Ricardo, like quantity theorists ever since, ignored bank assets, and did not consider the reasonable proposition that the pound had fallen because the Bank of England's assets (mainly British government bonds) had fallen in value.

b. The Real Bills Position

The Anti-Bullionist position, as stated by Charles Bosanquet, relied on real bills principles:

...(inflation will result whether) the issue be gold from a mine or paper from a government bank. All this I distinctly admit, but in all this statement, there is not a single point of analogy to the issues of the Bank of England.

The principle on which the Bank issues its notes is that of loan. Every note is issued at the requisition of some party, who becomes indebted to the Bank for its amount, and gives security to return this note, or another of equal value... (Bosanquet, 1810, pp. 52-53.)

First, note that Bosanquet admitted the existence of fiat money ("paper from a government bank"), and that an increase in the quantity of fiat money will cause inflation. He then denied that the creation of derivative money will cause inflation, since

\[5\] Bosanquet's ideas, and even his errors, are identical in this respect to those of Fullarton (1845, p. 58), Samuelson (1971, p. 2), and Tobin (1963, p. 415).
every issue of derivative money is matched by an equal increase in bank assets. But once
the existence of fiat money is admitted, it cannot be denied that the issue of derivative
money will reduce the demand for the fiat money and thus reduce its value. This was a
weakness that Ricardo was quick to exploit, though neither man considered that fiat
money might not exist at all.

Ricardo held that during the Restriction period the pound was a true fiat money, whose
value was determined by its quantity. ("...depreciation may arise from the abundance of
the notes alone, however great might be the funds of those who were the issuers of
them." (Ricardo, 1811, p. 114.)) His mistake was in confusing backing with
convertibility. On February 27, the day after the suspension of convertibility, the Bank of
England's ratio of outstanding notes to assets cannot have been much different from the
day before. Thus the real bills doctrine implies that the pound would be stable, as for a
time it was (Table 1). Ricardo, however, asserted that all that was necessary for an
inconvertible currency to have value was a limitation of its quantity. This leads to the
doubtful proposition that the forces determining the value of the pound changed
completely on February 26. Before that date, convertibility would have forced the pound
to be worth its backing. Afterwards, the value of the pound was supposedly determined
by the number in circulation. Ricardo made this assertion in spite of the fact that the
suspension of convertibility was temporary, and in spite of the fact that the Bank of
England continued to hold backing for the pound throughout the Restriction period (just
as the Federal Reserve does with dollars).

Ricardo's argument could still be salvaged by contending that the value of the pound
had been maintained all along by a limitation of the quantity in circulation. We could
suppose, for example, that when the pound was convertible, a certain number were called
into existence by the needs of business. Then, the proper quantity of pounds having been
established by long use, convertibility would gradually become less important, until at last it would become possible to suspend convertibility entirely, and the value of the pound could be maintained strictly by the limitation of quantity.

The trouble with this view is that it still requires us to believe that the forces determining the value of the pound changed over time. No serious economist would deny that a newly created money, issued by a bank that has only just commenced business, must be valued according to its backing and convertibility. And yet quantity theorists insist that the paper pound, and all similar currencies, were and are valued because of a limitation of quantity, and not because of backing. Thus quantity theorists are forced to contend that the factors determining the value of the paper pound must have changed sometime between the Bank of England’s first issue of notes in 1694, and its suspension of convertibility in 1797. Against this unlikely scenario, I propose the more plausible alternative that the pound’s value was determined by backing both before and after the suspension of convertibility; that it was backing that mattered during normal business hours when the notes were convertible, that it was backing that mattered during the nights and weekends when the notes were temporarily inconvertible, and that it was still backing that mattered during the longer suspensions of convertibility such as the period from 1797-1821.

5. Lloyd Mints

In the few textbooks that even mention the real bills doctrine, Lloyd Mints' criticism is still standard:

The fundamental error of all three men (Law, Steuart, and Smith)... lay in the fact that they failed to see that, whereas convertibility into a given physical amount of specie (or any other economic good) will limit the

page 28
amount of notes that can be issued, although not to any precise and foreseeable extent (and therefore not acceptably), the basing of notes on a given money's worth of any form of wealth—be it land or merchants' stocks—presents the possibility of unlimited expansion of loans, provided only that the eligible goods are not unduly limited in aggregate value. (Mints, 1945, p. 30.)

Mints supposed that a bank issued new money based on security that was initially sufficient, but which was denominated in the bank's own money. He then asserted that the increase in the quantity of money would cause inflation, thus reducing the real value of borrowers' debts and allowing them to borrow still more. This in turn would lead to a vicious circle of more inflation and more borrowing. He implicitly assumed, however, that the initial issue of money on sufficient security would cause an initial round of inflation. But on real bills principles this initial inflation would not occur. The value of the bank's money would be determined by its backing, and an issue of new money in exchange for sufficient security would automatically increase backing in step with the new money. Thus the 'unlimited expansion of loans' would be cut off before it started. Mint's refutation of the real bills doctrine implicitly assumed the correctness of the quantity theory—the very point in dispute!

Mints' discussion (1945, p. 31.) was directed mainly at inconvertible paper money, and his presentation makes it clear that he had in mind unbacked fiat money. If we grant the existence of fiat money, then we must implicitly grant that the money's value results from a limitation of its quantity. By extension, we would have to grant Mints' contention that new money—even if issued in exchange for real bills—would cause a self-perpetuating inflation. However, if the money in question is not fiat money, then Mints' scenario does not follow. The question of whether fiat money exists at all is the point on which the entire argument stands or falls, and it is precisely this point upon which Mints and his fellow quantity theorists are completely silent.
It is fitting at this point to remind the reader of the reasons I have given to doubt the existence of fiat money:

1. If any bank, public or private, could issue fiat money, then the issuing bank would earn profits that would entice other banks, both foreign and domestic, to issue their own fiat money, thus driving the value of the original money down to zero, or to a level commensurate with its backing.

2. Fiat money would also lose value in the face of competition from derivative moneys, such as checking accounts, credit cards, eurodollars, etc. Backed money would suffer no such loss of value.

3. All central banks recognize their money as their liability, and in fact maintain assets sufficient to redeem their money.

The reader might raise two objections: (1) Competition from rival moneys might not be strong enough to prevent fiat money from having value. (2) The fact that central banks recognize their money as their liability might just be a relic of the gold standard. However, one would at least have to recognize that economists who have rejected the real bills doctrine have done so without recognizing the possibility that fiat money is an illusion. If this single assertion is true, then the quantity theory must be rejected in favor of the real bills doctrine. It will take a good deal of empirical work before the 300-year dominance of the quantity theory might begin to falter, but until that work is done we have no basis for rejecting the real bills view.

III. Conclusion

The real bills doctrine holds that money issued for good security will not cause inflation. When money is backed, this doctrine is a simple and obvious truth: If every issue of money is matched by an equal increase in its backing, then the issue of new
money will not change the amount of backing per unit of money, and so the value of the money will be stable. On the other hand, the real bills doctrine would not be true of unbacked (fiat) money. Quantity theorists contend that fiat money has value because it is limited in supply. Thus, even if new money is issued for good security, the simple fact that there is more money will cause inflation. The two views are clearly set apart: Any given money is either backed or unbacked. If it is backed, then its value is correctly described by the real bills doctrine. If it is unbacked, then its value is correctly described by the quantity theory.

I take the view that all money is backed, and that fiat money does not exist. I offer three observations in support of this view: (1) Fiat money would create a free lunch for its issuer, thus inducing rivals to issue their own money until competition drove the money's value to zero. (2) The fact that a currency such as the dollar is inconvertible does not imply that it is unbacked. Convertibility can be suspended for a weekend or for a hundred years, but as long as the issuing bank maintains valuable assets against its money, we cannot call that money unbacked. (3) All central banks recognize their money as their liability. They maintain valuable assets as collateral against that money, and they have, on many occasions, used those assets to resume convertibility.

If we accept the view that the dollar is backed, then it follows that the value of the dollar would be unaffected by the quantity of base money, by convertibility, by money demand, by the quantity of derivative money, or by fiscal policy (except where fiscal policy affects the central bank's ratio of assets to units of money).

When the central bank fails to take sufficient security for its money, inflation will follow from the resulting drop in its ratio of assets to units of money. At the same time, the easy money policy will lead to an increase in the quantity of money, and observers will wrongly conclude that the increase in the quantity of money caused the drop in its
value. This faulty perception has allowed the quantity theory to become the dominant
theory of money, while the real bills doctrine has been wrongly discredited.

Adam Smith's view was that as long as banks only issued money in exchange for real
bills, the quantity of money would automatically move in step with the economy's real
output. Ever since, the economics profession has unfortunately viewed the real bills
doctrine as a rule for limiting the quantity of money, which is completely counter to a
proper understanding of the real bills view. The correct view is that the real bills rule
would maintain stable prices by causing the quantity of money to move in step with its
backing.

Henry Thornton paid closer attention to the role of backing, but his "false wealth"
fallacy led him to believe that privately-created money would seldom be backed by any
actual property. His mistake was in failing to realize that the self interest of the parties
involved in money-creation automatically assures that all money is backed by the entity
that issues it.

Charles Bosanquet correctly saw that money which is created by loan differs from
money that is not recognized as anyone's liability. He was unfortunately unable to counter
Ricardo's charge that money created by loan is nevertheless new money, which as such
must be inflationary. Neither man saw that the quantity of money was not what
mattered—that as long as the money was backed, any amount could be created without
affecting the price level.

In the 20th century, the real bills doctrine has been rejected largely on the grounds of
Lloyd Mints' "money's worth" argument, which states that loans secured by a given
money's worth of assets will create a self-perpetuating cycle of more money and more
inflation. But Mints begged the question. He assumed that the initial issue of money on
sufficient security would cause inflation, and thereby assumed the correctness of the
theory he was trying to support. The whole point of the real bills doctrine is that money issued for sufficient security will not cause inflation.

References


Smith, Bruce D. "American Colonial Monetary Regimes: The Failure of the Quantity Theory and Some Evidence in Favour of an Alternative View." Canadian Journal of Economics volume 18, August 1985, pp. 531-65. (a)

Smith, Bruce D., "Some Colonial Evidence on Two Theories of Money: Maryland and the Carolinas", Journal of Political Economy, volume 93, number 6, December 1985, p. 1178. (b)


