

## Cato Institute Policy Analysis No. 17: Gold, Paper, or...Is There a Better Money?

September 23, 1982

David Friedman

David Friedman is a professor in the law school and the economics department at Santa Clara University.

### Executive Summary

Alternative monetary systems are too often discussed in the context of current policy debates: Will gold cure inflation? Is tight money strangling the economy? Should the Fed be independent? Instead, I will ignore the particular problems of 1982 in order to evaluate the monetary alternatives that face any society. I shall try to consider separately two issues that are often confounded: What kind of money we are to have, and who is to produce it. In evaluating the alternatives, my main concern will be how well each monetary system permits people to make plans and contracts with reasonable confidence that the money they use will not change its value over time in unexpected ways. I will also discuss costs associated with different monetary systems.

The three alternatives I will consider are commodity monies, fractional reserve monies, and fiat monies. I will consider each primarily as a mechanism for controlling the quantity of money. This may seem a somewhat odd approach; while it is natural to regard the value of a fiat money as determined by how much of it is in circulation, it seems equally natural to believe that if the money consists of chunks of gold, it is the commodity value of gold which determines their value. It is natural, but I think it is a mistake.

### Commodity Monies

Assume that some convenient commodity -- homogeneous, easily subdivided, and with a high value to weight ratio -- has come into use as money. Individuals find that by holding a stock of that commodity they are able to separate the acts of buying and selling; they can sell goods and hold the money until they find something to buy, or buy goods with some of their stock of money and replenish it by later sales. After a little experimentation, they find that a stock of money equal to, say, a week's income is sufficient to cover their normal requirements. There is now a new demand for the commodity (let us call it gold); in addition to those who wish to purchase it for gilding jewelry or plating electrical contacts, there are others who wish to purchase it to hold as money. The new and higher demand curve will intersect the old supply curve at a higher price; more will be produced, less will be used for the old, non-monetary purposes, and the difference will go to monetary uses.

What does a "higher price" mean? We are accustomed to measuring prices in terms of money; in that sense that money price of money -- the price of a dollar in dollars, for example -- is always one. More generally, one can measure prices in terms of any good, speaking of the apple price of pears (how many apples must you give up for a pear) or the labor cost of weeding your own garden. For measuring the price of money it is possible to use any commodity; if apples sell for 50 cents, the price of a dollar is two apples. In general, the price (or "cost" or "value") of a unit of money is the amount of something else you must give up to get it, hence the inverse of that something else's price measured in money. Since there are many different commodities, and since their money prices may not all change in the same way, it is convenient to use some conventional "bundle" of commodities -- two apples, a pound of steel, and a loaf of bread,

say -- and define the price of money as the number of such bundles a unit of money exchanges for. The price of money then becomes the inverse of a price index, with the weights of the items in the index being the amounts of those items in the bundle.

Although the price of gold measured in gold will not change when it becomes money, the increased demand will result in its exchanging for more of other goods; hence its price will have gone up. Since the price of gold is now being determined by the sum of monetary and non-monetary demands, it seems odd to say that its worth as money is determined by its worth as commodity. I would prefer to say that its price as money is determined by the intersection of the demand curve for money and the supply curve for money, the latter being the supply curve for gold minus the demand curve for gold in non-monetary uses.

Even with commodity money the identity of commodity and money is not complete; the price of the commodity as money is not necessarily the same as the price of the commodity in other uses. The two are equal only if the non-monetary form ("bullion") can be costlessly converted into the monetary form, and vice-versa. In many commodity money systems someone, usually the government, controls coining and charges a fee (called seigniorage) for making bullion into money. It is only the opposite transformation (melting coin into bullion) that is generally free, due to the difficulty of preventing it. Under such a system, the "monetary rule" in effect is that if the money is worth less than its weight in gold plus the seigniorage, its quantity will increase. Such a system is, in a sense, equivalent to a fiat system under which the monetary authority follows an identical rule, printing money when the price index gets below some predetermined level (meaning that the value of the money has gotten "too high") and burning it when the price index gets above some slightly higher level.

What are the disadvantages of the commodity system? One is the commodity itself; since more is being produced and less being used for non-monetary purposes, the resources devoted to additional production and the benefits forgone must be counted as the price of the system. Also, the supply of money, and hence its value, may be affected by anything that affects the supply of gold (new discoveries or new production technologies) or its non-monetary uses. It may also be affected by anything that affects the demand for money -- economic growth, changes in the structure of the economy, or simply the spread of the use of gold to new countries.

Along with these disadvantages comes one very important advantage. Because the supply of money under a commodity system is determined only by the sorts of things I have just listed, its variation is rarely very large. Mining gold takes work, and developing new technologies takes time. Hence, in the worst cases the magnitude of variations in supply is likely to be small, at least in comparison with the alternative of fiat money.

While the cost of producing the money is an essential characteristic of a commodity system, the other disadvantages depend in large part on the particular commodity used; hence, avoiding them may be viewed as one criterion for choosing a commodity. The ideal commodity from this standpoint should be one which is unlikely to have substantial unpredictable changes in value due to changes in its supply curve or its monetary or non-monetary demand. This might be a commodity whose supply curve is stable and highly elastic, allowing it to accommodate shifts in the demand curves without large changes in the value of the commodity; or a commodity whose non-monetary demand curve is stable and highly elastic, allowing it to accommodate shifts in supply or monetary demand; or simply a commodity for which supply and non-monetary demand are both stable and large compared to monetary demand. A fourth possibility would be a commodity for which supply and non-monetary demand, although unstable, are small compared to monetary demand, so that substantial changes in the value of the commodity would at least be limited to those produced by changes in monetary demand.

### **Fractional Reserve Money**

A fractional reserve system is a commodity system without the commodity. The money consists of promises to pay some amount of gold, silver, or whatever the monetary base of the system may be. The organizations that issue these promises (typically banks) arrange to fulfill them by holding an amount of the commodity equal to some fraction of their total obligations, while keeping the rest of the assets that they acquired in exchange for the money in some more productive form. The stock in the commodity that they hold ("reserves" -- hence the term "fractional reserve system," since the reserves are only a fraction of the obligations) permits them to fulfill their promise to redeem their money in

the commodity. If many of those holding their money want the commodity instead, the banks sell some of their other assets and buy more of the commodity to give them. In such a system there are generally at least two monies circulating at once: a commodity money such as coined gold ("currency") and a fractional reserve money, in the form of either banknotes (written promises to pay particular numbers of gold coins or quantities of gold) or deposits (obligations by the bank to give its customer, or anyone he names, up to a certain amount of coins or gold).

The advantage of such a system over a pure commodity system is that it economizes on the use of the commodity. If most people choose to hold their money in notes or deposits rather than in currency, and if the banks find that they can function with reserves that are a small fraction of their outstanding obligations, the total amount of money in the system will be much larger than the amount of gold being used for monetary purposes (currency plus reserves -- sometimes referred to as "high-powered" money). Thus, the society gets its monetary services at a lower cost -- an advantage reflected, at the individual level, in the willingness of banks to pay interest, in money or services, to their depositors.

Such a system has an additional advantage; an increase in the demand for money due to economic growth or the spread of the system to additional countries will have less effect on the value of money than under a pure commodity system using the same commodity, since the additional monetary demand can be satisfied with less gold. It is for the same reason more sensitive to changes in non-monetary demand, since non-monetary demand makes up a larger part of total demand for gold under a fractional reserve system. It has the further important disadvantage of being sensitive to changes in the monetary demand for the commodity caused by changes within the fractional reserve system itself. The amount of monetary gold necessary to support a given quantity of money under such a system depends on both how large a fraction of their liabilities banks hold as reserves and on how large a fraction of their money individuals hold as currency. Both can, under some circumstances, change rapidly. If some depositors lose confidence in the banks, for example, and begin trying to convert their deposits into currency, some banks may be unable to sell other assets fast enough to keep from running out of reserves and being forced to suspend payment. The news that some banks have suspended payment provides an incentive for more depositors to try to get their money out before their bank does the same. Such a "run" on the banks increases the monetary demand for gold in the short run by increasing the fraction of money held as currency, and in the longer run by inducing banks to hold larger reserves. The price of gold goes up (meaning that prices measured in gold go down).

Here again, avoiding these problems may be seen as one criterion for choosing a commodity on which a fractional reserve system is to be based. As before, the ideal commodity would have a stable and elastic "monetary supply" -- its supply curve and non-monetary demand curve would not change very much, or at least change only in predictable ways, and one or the other or both would be sufficiently elastic to absorb changes in monetary demand without large changes in the value of the commodity. The relevant concept of elasticity here involves changes in "monetary supply" considered as percentages of monetary demand; hence it depends both on elasticity in the conventional sense (percent change in supply or demand for a one-percent change in price) and on how large supply and non-monetary demand are compared to monetary demand. A less satisfactory alternative would be a commodity for which supply and non-monetary demand are either stable or small compared to monetary demand, so that they would at least not cause price changes.

Before leaving the subject of fractional reserve systems, I should mention one particularly bizarre variant -- a fractional reserve system based on fiat money. I call it bizarre because the essential function of a fractional reserve system is to reduce the resource cost of producing money, by allowing an ounce of reserves to replace, say, five ounces of currency. The resource cost of producing fiat money is zero; more precisely, it costs no more to print a five-dollar bill than a one-dollar bill, so the cost of having a larger number of dollars in circulation is zero. The cost of having more bills in circulation is not zero but small. A fractional reserve system based on fiat money thus economizes on the cost of producing something that costs nothing to produce; it adds the disadvantages of a fractional reserve system to the disadvantages of a fiat system without adding any corresponding advantages. It makes sense only as a discreet way of transferring some of the income that the government receives from producing money to the banking system, and is worth mentioning at all only because it is the system presently in use in this country.

## **Fiat Money**

In a fiat system, there is no non-monetary demand for the money at all; it typically consists of pieces of printed paper, and the supply is determined by a printing press controlled by whoever issues it. (Its value is maintained entirely by its monetary demand.) This seems paradoxical, since the existence of a monetary demand for it is dependent on its having value. In practice, the problem has usually been solved by gradually creating a fiat system out of an existing fractional reserve system, eventually eliminating entirely the bank's obligation to pay in the commodity. It could also be created out of a commodity system by gradually increasing the seigniorage at a time when monetary demand is increasing, and allowing the monetary demand to raise the price of the coin to a large multiple of the value of the commodity it contains. Once created, fiat systems have proved astonishingly stable; the convenience of using the same money as everyone else is apparently so great that people continue using a fiat money (instead of making their transactions in terms of some convenient commodity, such as gold or cigarettes) even when it is rapidly losing value.

In terms of producing stable and predictable prices, a fiat system is at the same time the best and the worst alternative. It is the best alternative because it is possible, by following some simple monetary rule (such as "keep the amount of money in circulation constant"), to make the supply of money perfectly predictable, or by following some slightly more complicated rule (print money when the price index goes below 1, burn it when the price index goes above 1) to make (average) prices almost perfectly predictable, automatically accommodating the supply of money to the demand. It is the worst of systems because it is possible to expand the money supply virtually without limit (the cost being the cost of adding additional zeros to the newly printed bills to convert tens into hundreds, or hundreds into millions). Whether a fiat system is good or bad, then, depends on your prediction of what the people running it will find it in their interest to do.

### **Private or Public Money**

The question of who -- what individual or institution -- is to control the monetary system is usually considered only within the context of some government system, since it is widely assumed that money must be produced by governments. There are, however, numerous historical examples of private commodity or fractional reserve systems, some of which, such as the Scottish free banking system of the 18th and 19th centuries, functioned successfully for more than 100 years. My discussion will deal not with the question of what government institutions should control the production of money, but with whether government systems are likely to do better or worse than private ones. In government systems I include those, such as the present American system, in which money is produced by private as well as public institutions (checking accounts, which are a form of money, are produced by private banks), but in which the activities of the private institutions are effectively controlled by the government. In our system, for example, the Federal Reserve Board can control the quantity of money produced by the private banking system through its control of the quantity of currency plus deposits in Federal Reserve Banks ("high-powered money") and of reserve requirements.

In analyzing these alternatives, I will use the methodology of public choice theory; I will consider not what governments could do if they were run by benevolent and omniscient despots but what they can be expected to do given the interests they actually have and their observed behavior in the past.

Seen from this perspective, it is clear that the interests of governments that produce money are not always the same as those of the people who use it. Starting with commodity money as the simplest case, the interest of the government has frequently been to debase it, decreasing the weight in order to make a profit on recoinng heavy coins into light ones. This serves the additional function of benefiting creditors at the expense of debtors, by allowing the creditors to pay their debts in the debased money. This is attractive when the debtors happen to be politically powerful or when the government itself is a debtor. The most striking exception to the general tendency of commodity money to be debased is the case of competing international monies. Where one of a country's coins is primarily used outside of the country in foreign trade, the government has an incentive to maintain its value; if the coin is debased it will be replaced in circulation by the more reliable money of some other country. Such international monies have sometimes maintained their weight and fineness for several centuries. Examples would be the Byzantine nomisma and the Arabic dinar during the Middle Ages.

In the case of more government systems using fractional reserve money, additional difficulties may arise. The government frequently has an interest in affecting the domestic price level either as a way of redistributing income

between creditors and debtors or as a way of affecting unemployment and other politically important variables. By expanding or contracting the amount of fractional reserve money, the government can achieve its objectives in the short run. But in doing so, it inevitably affects the domestic price of the commodity in terms of which the money is defined, since the price of that commodity is directly linked to the value of money; anyone who has the money can go to the bank and demand the commodity. If the price of that commodity is high or low relative to other commodities in comparison to its relative price elsewhere, the result is to make it move into or out of the country, canceling the effect of the government's policies. Such flows are normally described as "balance of payments problems"; if a country is exporting the monetary commodity (say, gold) because it is cheap domestically (meaning that domestic prices for goods in terms of money are high) it is, on net, importing other goods. For historical reasons, this is referred to (misleadingly) as an "unfavorable" balance of payments. The consequence is to severely limit the government's ability to control the economy. These limitations may be, and frequently are, avoided by restrictions on trade which prevent goods being imported in exchange for gold, or encourage (by export subsidies) the export of goods whose cost of production is higher than what they can be sold for. Thus, one of the disadvantages of a government-run fractional reserve is that it encourages undesirable restrictions on trade.

The principal difference between a government fractional reserve system and a government fiat system is that under the latter there are no longer any obstacles to prevent the government from doing whatever it likes with the money supply. This has an advantage in that the government no longer needs to restrict trade in order to get "monetary sovereignty." It has a disadvantage in that a government which wishes to print large amounts of money, either as a source of revenue or for other political reasons, is free to do so. Historically, all or almost all rapid inflations have been produced by such systems, usually in situations where the government was using money creation as a major source of revenue.

### **The Private Alternative**

The simplest private monetary system is a commodity money produced by a number of private firms. Each firm mints coins of standard weight and sells them. While such a system has the disadvantage described in my earlier discussion, it should be less vulnerable to debasement than one in which coining of money is a government monopoly. With more than one firm, customers can shift away from one that starts producing underweight coins, hence the opportunities for such fraud would be rare. Such a system is very much like the competing international monies of the Middle Ages. While those monies were produced by governments, they were sold, for the most part, to customers over whom the producing governments had no control. The governments producing them competed like private firms to induce merchants to use their money; the obvious way of doing so was by maintaining its quality.

In a modern society, another sort of commodity money is also possible: warehouse receipts. Instead of carrying around pieces of gold, one carries around receipts for pieces of gold in storage somewhere. In such a system, unlike a fractional reserve system, every piece of paper is backed by a specific piece of gold; hence, it may be described as a 100-percent reserve system. Its advantages over an ordinary commodity system are that it eliminates the wear and tear on the coins and permits the use as monies of commodities which do not have some of the characteristics which would otherwise be necessary. Carrying around enough iron to buy an automobile would be inconvenient, to say the least, but carrying around receipts for enough iron would be no more inconvenient than carrying receipts for enough gold. Since the characteristics of the commodity used for money affect how well a commodity system works, expanding the range of possible commodities may lead to a considerable improvement in the system.

Once a private commodity money is established, there are good reasons why a fractional reserve system is likely to develop. By holding only enough reserves to meet its day-to-day needs a bank can free the rest of its assets to be converted into loans or other interest-bearing assets. The first bank to establish such a system is getting, in effect, an interest-free loan from its customers. Once other banks follow its lead, competition forces them all to pay interest -- in money or services -- to their customers. 100-percent reserve banks -- which must charge their customers for the service of holding their money -- become an unattractive alternative.

All of this assumes that the fractional reserve banks can offer depositors a reasonable certainty of being able to get their money back if they want it. Most of the criticisms of private fractional reserve systems depend on their being either unable or unwilling to do so. It is often argued that the system is inherently unstable; a run due to rumors of

weakness in one bank persuades many depositors to withdraw their money, and since the banking system as a whole has obligations much larger than its reserves the banks are unable to pay, and the system collapses.

But even if a bank's reserves are much less than its obligations, it can still fulfill its obligations in full. All that is necessary is that the bank sell non-reserve assets in order to get back the currency that is being paid out to frightened depositors. Hence, the real problems arise either from having assets that are insufficiently liquid, from having total assets (reserves and other assets) which are less than total liabilities, or from having assets whose market value (measured in money) falls in a panic. This last is likely unless the value of the assets is somehow linked to the value of the money, since in a panic the money supply falls, the value of money rises, and the money prices of commodities (other than the monetary commodity) consequently fall.

There are a number of ways in which banks can, and do, protect themselves. One is to hold assets, such as loans and bonds, whose market value is fixed in money rather than in real terms. Another is to start with total assets larger than total liabilities, so as to guarantee to their depositors that even if the bank loses money it can still fulfill its obligations. A historical example is the Scottish banking system described by Adam Smith. The banks were partnerships, and the partners were generally wealthy men. Since they were not protected by limited liability, the partners were individually liable for the debts of the bank. Hence, the depositors knew that they could lose their money only if the bank's net liabilities exceeded the combined fortunes of the partners. Several of the banks did fail, but in every case the depositors were paid off in full.

It can also be argued that while private fractional reserve banks can be stable, they will choose not to be. Once a bank has built up a reputation for reliability it pays it to convert that reputation into cash by vastly expanding its deposits without any adequate backing, and then convert that cash from an asset of the corporation to a private asset of its owners and officers, leaving the depositors with a worthless shell. While such frauds are certainly possible in private banking (and elsewhere in the economy) there is no obvious reason to expect them to be common, especially in a modern economy with well-developed institutions for generating and transmitting information on the financial condition of firms. If such a problem did develop in a private system, one consequence might be a preference by depositors for banks that were not protected by limited liability.

Two further arguments are sometimes made for why money creation cannot be private; both, I think, are mistaken. The first is that competition is impossible, since it is necessary to have a single uniform money, the alternative being a system where every transaction requires the intervention of a money changer. But this argument confuses standardization with monopoly. It is certainly convenient for the monies of different firms to exchange at a ratio of one to one, just as it is convenient for the nuts made by one firm to fit the bolts made by another. The obvious way to arrange for standardization is for the different banks offering fractional reserve monies to use the same commodity in the same units. If all banks make their money (whether notes or deposits) redeemable in grams of gold, for example, then all monies should exchange at one for one (or five or 10 for one in the case of different denominations). The only exception would be the money of a bank believed to be financially shaky, and so not certain to redeem its obligations. Such money would sell at an inconvenient discount; that would drastically reduce the demand for it and provide an incentive for banks to be careful of their reputations.

The other argument is that money is special because it always pays a bank to produce more of it, and it can be produced costlessly. There are two errors here. The first is that in order to produce money that people will accept, a bank must demonstrate its ability to redeem it. That is not costless, and its costs increase with the amount of money outstanding. The second is the assumption that when a bank gets the use of assets by getting people to hold its money, it does not pay for them; so it always benefits and always wants to print more money (unless the cost of doing business is greater than the interest it can earn). In a competitive market the interest paid for deposits would be bid up until it absorbed any excess, so that banks, like other competitive firms, would receive only enough to cover their costs of operation .

### **What Commodity?**

So far I have ignored the question of what commodity a private system should base its money on. Such systems have used gold or silver. In my opinion, either would be a poor choice in the modern world. Gold and silver are well suited

for a simple commodity money, since they have a high value to weight ratio (making them portable), are easily subdivided and recombined, and relatively easy to measure and evaluate. In a modern society none of these characteristics is important, since the circulating medium is not the commodity itself but claims upon it. On the other hand, silver and gold have very inelastic supplies and relatively inelastic demands. Judging by recent history the value of both (in terms of most other commodities) can and does vary erratically even without the additional instabilities that would be introduced by a fractional reserve system based on them.

The ideal commodity backing for a modern system would not be any single commodity, but a commodity bundle. The bank would guarantee to provide anyone bringing in, say, 100,000 of its dollars with a bundle consisting of a ton of steel of a specified grade, 100 bushels of wheat, an ounce of gold, and a number of other items. The goods making up the bundle would be chosen to make its value correlate as closely as possible with the general price level. While a change in production technology or non-monetary demand might alter the value of one good in the bundle, it would have only a small effect on the value of the bundle as a whole. Since the quantity of such goods being used for monetary purposes would be a tiny fraction of the total quantity of steel, wheat, gold, etc., changes in monetary demand would have negligible influence on the value of the bundle, and hence on the value of money.

It is worth noting that the system I have described would work, in practice, very much like the ideally stable fiat system described earlier. If the money supply increased to the point where the commodity bundle was worth more than 100,000 "dollars," holders of dollars would turn them in for commodities, bringing the money supply and the price level back down. If the money supply fell so that the commodities were worth less than the money, banks would find that they could issue additional money without any of it being turned in for commodities, and the money supply would rise. The system as a whole would therefore stabilize prices in such a way as to make the price of the bundle (a crude price index) stable at its "face value." And it would do this through the private interest of the competing banking firms issuing the money, without relying on the wisdom or benevolence of any public employees appointed to manage the system. Since the nature of the reserves in this system makes it unnecessary for the banks to hold any significant quantity of them, such a system is essentially equivalent to a "tabular standard"--a system under which the issuer of money is not obliged to redeem it, but is committed to maintain that money supply, which keeps a specified price index a particular value. The only difference is that under a fractional reserve system based on a commodity bundle the obligation to redeem the currency in commodities is a mechanism for enforcing that commitment.

### **Preference vs. Prediction**

This is what I would like to see. Is it likely to occur? I think not. To go from one monetary system to another involves a difficult coordination problem. I would rather use a poor money which everyone else uses than an ideal money which nobody else uses. I will therefore continue to use the present system unless I can somehow arrange for everyone else to shift at the same time I do. An inflation rate of 12 percent a year corresponds, for an individual whose money balance is 100 dollars, to an implicit tax of a dollar a month. Even several times that is a small price for the convenience of using the same money as everyone else. This is the reason even quite badly run fiat systems continue to be used. My own opinion is that even if there were no legal bars to the use of private money (and I do not think there are any very important ones) the existing fiat system would remain in use unless it became very much worse than it now is.

For similar reasons, I think it likely that if a private system does come into use it will be based on gold, even though gold is not a very suitable commodity for the purpose. For reasons I have already discussed, it is desirable that banks issuing private money converge on a common commodity standard. It would be very much easier to converge on gold, which has been widely used in the past, than on some complicated commodity bundle, whatever the theoretical advantages of the latter.

Even if gold is not a very suitable commodity, it does not follow that a private system based on gold is worse than what we now have. Historical experience suggests that while a gold standard may produce either inflation or deflation, it is unlikely to produce as serious an inflation as even a relatively successful fiat system (such as our own) and that the inflations which have been produced by particularly unsuccessful fiat systems dwarf anything that might result from new discoveries of gold. The possibilities for contraction under a fractional reserve system based on gold are more serious; since governments profit by printing money, not by burning it, this has rarely been much of a problem

under a pure fiat system.

Finally, it is important to distinguish between proposals for a true private banking system and proposals for some sort of government-run fractional reserve system linked to gold, such as what the United States had in various forms for most of this century. The connection to gold in the latter case may constrain the government's ability to manipulate the money supply and the price level, but it does so at the cost of giving the government an incentive to block the free flow of goods and services in international trade as a way of evading those constraints.