Lessons from the history of money

François R. Velde

Introduction and summary

The use of money began in the sixth century B.C. in what is now western Turkey, when lumps of gold found in rivers were melted and turned into pieces of uniform size imprinted with a stamp. For almost all of the time since then, the common monetary system has been *commodity money*, whereby a valuable commodity (typically a metal) is used as a widely accepted medium of exchange. Furthermore, the quantity of money was not under anyone's control; private agents, following price incentives, took actions that determined the money supply.

Today, the prevalent monetary system is that of *fiat money*, in which the medium of exchange consists of unbacked government liabilities, which are claims to nothing at all. Moreover, governments have usually established a monopoly on the provision of fiat money, and control, or potentially control, its quantity. Fiat money is a very recent development in monetary history; it has only been in use for a few decades at most.

Why did this evolution from commodity money to fiat money take place? Is fiat money better suited to the modern economy or was it desirable but impractical in earlier times? Were there forces that naturally and inevitably led to the present system?

Fiat money did not appear spontaneously, since government plays a central role in the management of fiat currency. How did governments learn about the possibility and desirability of a fiat currency? Did monetary theorizing play any role in this evolution?

In this article, I will argue that the evolution from commodity to fiat money was the result of a long process of evolution and learning. Commodity money systems have certain advantages, in particular in providing a natural anchor for the price level. But they also have certain disadvantages, manifested in particular in the difficulty of providing multiple denominations concurrently. These problems arose early on, in the fourteenth century, in the form of money shortages. Societies tried to overcome these disadvantages, and this led them progressively closer to fiat money, not only in terms of the actual value of the object used as currency, but also in terms of the theoretical understanding of what fiat money is and how to manage it properly.

In the process, societies came to envisage the use of coins that were worth less than their market value to replace the smaller denominations that were often in short supply. These coins are very similar to bank notes; they are printed on base metal, rather than paper, but the economics behind their value is the same. What governments learned over time about the provision of small change is thus directly applicable to our modern system of currency.

In his *A Program for Monetary Stability* (1960), Milton Friedman begins with the question: Why should government intervene in monetary and banking questions? He answers by providing a quick history of money, which he describes as a process inevitably leading to a system of fiat money monopolized by the government (p. 8):

> These, then, are the features of money that justify government intervention: the resource cost of a pure commodity currency and hence its

François R. Velde is an economist at the Federal Reserve Bank of Chicago. *This article draws on joint work with Thomas J. Sargent, University of Chicago and Hoover Institution.*

> tendency to become partly fiduciary; the peculiar difficulty of enforcing contracts involving promises to pay that serve as medium of exchange and of preventing fraud in respect to them; the technical monopoly character of a pure fiduciary currency which makes essential the setting of some external limit on its amount; and finally, the pervasive character of money which means that the issuance of money has important effect on parties other than those directly involved and gives special importance to the preceding features. ... The central tasks for government are also clear: to set an external limit to the amount of money and to prevent counterfeiting, broadly conceived.

This article will find much to validate this view. It turns out that the problem of counterfeiting, identified as central by Friedman, provided obstacles that were overcome only when the appropriate technology became available. As technology changed and offered the possibility of implementing a form of fiduciary currency, various incomplete forms of currency systems were tried, with significant effects on the price level. These experiments led to the recognition that quantity limitation was crucial to maintaining the value of the currency. The need for a government monopoly, however, does not emerge from our reading of the historical record, and we will see that the private sector also came up with its own solutions to the problem of small change, thereby presenting alternatives to the monetary arrangements we have adopted.[1]

Commodity money and price stability

Among the desirable features of a monetary system, *price stability* has long been a priority, as far back as Aristotle's discussion of money in *Ethics*. In the words of the seventeenth century Italian monetary theorist Gasparo Antonio Tesauro (1609), money must be "the measure of all things" (*rerum omnium mensura*) (p. 633). Aristotle also noted that commodity money, specifically money made of precious metals, was well suited to reach that goal: "Money, it is true, is liable to the same fluctuation of demand as other commodities, for its purchasing power varies at different times; but it tends to be comparatively constant" (Aristotle, *Ethics*, 1943 translation).

The commodity money system delivers a nominal anchor for the price level. The mechanism by which this takes place can be described in the context of a profit-maximizing mint, which was how coins were produced in the Middle Ages and later.[2] Suppose there is a way to convert goods into silver and silver into goods at a constant cost (in ounces of silver per unit of goods), which can be thought of as either the extraction cost of silver and the industrial uses of the metal or the "world price" of silver in a small country interpretation. Silver is turned into coins by the mint; the mint (which really represents the private sector) also decides when to melt down existing coins.

The government's role is limited to two actions. It specifies how much silver goes into a coin, and it collects a seigniorage tax[3] on all new minting.

When the mint is minting new coins, its costs are the cost of the silver content, the seigniorage tax, and the production cost;[4] its revenues are the market value of the coins, which is the inverse of the price level. Similarly, when the mint is melting down coins, its costs are the market value of the coins, and its revenues are the value of the silver contained in them.

Whether the mint will produce new coins or melt down existing coins will thus depend on how the price level relates to the parameters: silver content of the coins, production costs, and seigniorage rate. The price level cannot be too low (or the purchasing power of the coins too high) or the mint could make unbounded profits by minting new coins and spending them. Similarly, the price level cannot be too high (or the purchasing power of the coins too low), or the mint would make profits by melting down the coins. The absence of arbitrage for the mint places restrictions on the price level, which is contained in an interval determined by the minting point and the melting point (figure 1).

This system, which prevailed until the late nineteenth century, has some noteworthy features. The quantity of money is not controlled directly by the government; rather, additions to or subtractions from the money stock are made by the private sector, on the basis of incentives given by the price level. The incentives operate so as to make the system self-regulating. If coins become too scarce, their value increases and the price level falls until it reaches the minting point, when more coins are added to the stock. If coins

FIGURE 1

Constraints on the price level caused by arbitrage

become too numerous, on the other hand, their market value reaches their intrinsic value and it becomes worthwhile for the mint to melt them down. The commodity nature of the currency places bounds on the price level, but does not determine the price level within that interval.

Within the interval, the price level depends on how the quantity of money relates to the volume of transactions, according to Irving Fisher's famous quantity theory equation.[5] As long as the price level is inside the interval, the stock of coins, or quantity of money, is fixed. Variations in the volume of transactions or in income would shift the price level up or down, unless such variations were so severe as to push the price level up to the melting point or down to the minting point. In that case, the mint would enter into action and modify the quantity of money in the appropriate way.

Consider now the interval in figure l. Its position on the real line is determined by the world price of silver and the silver content of a dollar coin. Any reduction in the number of ounces of silver per dollar, that is, any debasement of the currency, shifts the interval to the right; the price level is therefore higher. But the width of the interval is determined by production costs and the seigniorage tax. We may take production costs as a technological given, but the seigniorage tax is chosen by the government. In principle, the government could make the tax a subsidy; it could even subsidize the production costs completely. In that case, the interval in figure 1 would be reduced to a point, the minting point and melting point would coincide, and the price level would be completely tied to the world price of silver. This would eliminate any fluctuations in the price level due to the quantity theoretic effects described above. The only variations would be due to fluctuations in the world price of silver. In western European practice, however, the seigniorage rate was positive in almost all countries.

Although governments considered minting a fiscal prerogative, they were constrained in their choice of the seigniorage rate. High rates, a form of monopoly rent, were possible only if the government could effectively prevent competition. But in medieval Europe, all manner of coins circulated in all places and individuals were quite willing to take their metal to the mint of a nearby lord or king, subject to transportation costs, if they found the local seigniorage rate too high. Also, the technology for making coins was rather crude and available to any jeweler or goldsmith, so that counterfeiters would also be tempted by high seigniorage rates. In practice, then, the width of the interval was rather small, and production costs with seigniorage were on the order of 1 percent to 2 percent for gold and 5 percent to 10 percent for silver (the latter being ten times less valuable, transport costs were higher).

Multiple denominations and token coinage

This simple commodity system lacks one feature: multiple denominations. Although it is always possible to express any price in pennies, in practice it is necessary to have a range of coins of various denominations.[6]

In its last incarnation (the so-called classical gold standard), the commodity money system handled multiple denominations in a straightforward way, which is described in textbooks, for example, John Stuart Mill (1857).

The standard formula

The method that Cipolla (1956) calls the standard formula, consists of choosing a principal (large) denomination, which continues to be provided as before at the initiative of the private sector, thus continuing to provide a nominal anchor for the price level. The provision of lower or subsidiary denominations relies on three key elements: 1) monopolization of coinage by the government, 2) issue of *token* coins, and 3) peg of the token coins by having the government convert them on demand into the larger denominations. The intrinsic content of token coins was somewhat or much smaller than the face value at which they circulated. Some authors call such coins partly fiduciary. The opposite of a token coin is a *full-bodied coin*.[7]

In the case of the gold standard, the larger denominations were gold coins, and currencies (the U.S. dollar, the British pound, and the French franc) were defined by the number of ounces of gold per currency unit. The subsidiary coinage consisted of silver and bronze coins, which were token. The government's willingness to peg, say, the silver quarter at 1 / 40 of a gold eagle was implemented by the U.S. Treasury.[8]

Thus, in the standard formula, tokens play the same role as convertible notes issued by the central bank. As with notes, a mechanism serves to regulate the quantity outstanding: Excess quantities of token quarters are turned in at the treasury in exchange for gold eagles, while needed tokens are sold by the mint.[9]

The advantages of a token coinage are the same as the advantages of a representative money system, as pointed out by a long line of writers, including Adam Smith, John M. Keynes, and Milton Friedman. Resources that had been spent forming and maintaining that part of the stock of metallic currency were freed up for other purposes. To quote the French monetary official Henri Poullain, writing in 1612: "In a card game, where various individuals play, one avails oneself of tokens, to which a certain value is assigned, and they are used by the winners to receive, and by the losers to pay what they owe. Whether instead of coins one were to use dried beans and give them the same value, the game would be no less enjoyable or perfect" (Poullain, 1709, p. 68).

Another advantage, from the point of view of the government, is that the issue of tokens is quite profitable. To the extent that tokens circulate for more than their intrinsic value plus the costs of minting, they represent a pure profit, the seigniorage in the medieval and modern sense of the word.

These two advantages (social savings and government revenues) have been understood for centuries, and, as Friedman points out, have provided impetus for the development of money away from a strict, full-bodied commodity version. However, these two motivations do not determine clearly in which direction money will develop; perhaps, in fact, each pushes in a different direction. The tension will be illustrated in the historical process I describe.

Prerequisites of a token coinage

Whatever its advantages, the implementation of the standard formula depended on some prerequisites. With a token coinage, the profits to the issuer are large, and, as Friedman says (1960, p. 6), "In fraud as in other activities, opportunities for profit are not likely to go unexploited." The government's ability to maintain its monopoly on token issue is thus dependent on the prevention of counterfeiting.[10] While nowadays counterfeiting may seem to be a significant but not overwhelming nuisance, which suitable technology can always remedy (such as that embodied in the recently issued $100 and $50 bills), in the past it presented an insuperable obstacle to the development of the standard formula.

One way to prevent counterfeiting is to impose high costs of entry to counterfeiters. Law enforcement provides a second method; as the Italian economist Montanari wrote in 1683, "A die which costs the prince 3 to make, will cost a counterfeiter 8 or 12; because he who works at the mint does not risk his life, and receives only the wage commensurate to his activity; but if a goldsmith has to make a coin at the risk of his whole being, he will not be persuaded if not with a lot of gold." The death penalty[11] for counterfeiters adds a risk premium to the counterfeiters' wage costs, which may or may not be sufficient to wipe out their potential profits. A third method is to make the government currency difficult to imitate, for example, if it is produced with a technology that is not accessible to the private sector in some way; either the government can make better coins or the same coins more cheaply.

If such a cost or technology advantage is not available to the government, then attempts at issuing token coinage will be plagued by counterfeiting or competition from neighboring currencies. Ultimately, the gross seigniorage rate will be driven down to the production costs (common to both government and counterfeiters). Thus, without the appropriate technology, only full-bodied coins can be used for small denominations.

The big problem of small change

This seemingly trifling aspect of the monetary system turns out to have bedeviled Western societies for centuries. Nowadays, the only problem most people see with small change is that we have too many pennies around, but for students of monetary history, the "big problem of small change" (a phrase coined by Carlo Cipolla in 1956) refers to recurrent coin shortages that were prevalent before the adoption of token coinage. The last time the U.S. experienced a shortage of small change was in 1965–66, when quarters and dimes still contained silver; the Coinage Act of 1965 made them completely token (Spengler, 1966).

Full-bodied small change

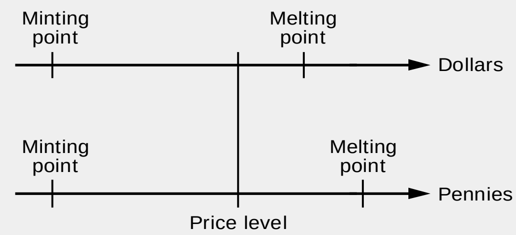
The medieval technology for making coins was very simple. Metal was melted and beaten into sheets, the sheets were cut with shears into blanks, and the blanks were placed between two hand-held dies. The upper die was struck with a hammer and the blank imprinted. Dies were made by goldsmiths using ordinary tools, and the design on coins could easily be copied by any goldsmith. Thus, the government and the private sector had access to the same technology.

Around 800 A.D., Charlemagne unified most of Western Europe and created a uniform currency. Until the twelfth century, Europe only had one coin, the silver penny, initially minted identically across Charlemagne's empire. Thus the commodity money system was in its simple, one-coin form. Around the year 1200, large improvements in the European economy, improved safety, and economic expansion led to greater volumes of trade and the need for larger denominations than the penny. This led to the appearance of silver coins of about five to ten times the content of a penny, called *grossi*. Over time, the denomination structure became richer, with the addition of gold coins in the mid-thirteenth century. Coins throughout the denomination structure remained close to full-bodied.

However, the commodity money system acquires unexpected complications when multiple denominations are introduced. To see this, let us return to the mint's problem, and suppose we have two currencies, dollars and pennies. The same reasoning as before will apply to both coins separately. As a result, the requirement that there be no arbitrage left for the mint will now place two sets of restrictions on the price level, which we can represent by two intervals, as in figure 2.

FIGURE 2

Constraints on the price level of two coins



In order to make the two intervals comparable, the lower one (which corresponds to pennies) is scaled by the market exchange rate between the two coins (expressed in dollars per pennies). This simply means that the mint's calculations about minting or melting pennies are computed in dollars.

The intervals must overlap, of course. Recall that the position of a coin's interval on the real line is linked to the intrinsic content of that coin, so that a smaller intrinsic content of the dollar corresponds to a higher price level. With two coins, the ratio of intrinsic contents must be reasonably close to the intended parity between denominations, although it need not coincide with that parity. But that is not enough: A coin is produced only when the price level reaches the minting point. Therefore, if the lower ends of the intervals do not coincide, one type of coin is never minted. Equating the lower ends of the intervals (by the government's choice of the intrinsic contents and the seigniorage rates) makes the mint stand ready to buy silver for the same price, whether it pays in pennies or dollars.

On the other hand, if the upper ends of the intervals do not coincide, one coin might be melted, but the price level could still rise further and the other coin remain in use. Equating the upper ends of the intervals makes the ratio of metal contents in the two coins equal the exchange rate, in which case pennies are strictly full-bodied. If the melting point for pennies is higher than the melting point for dollars, pennies are relatively light.

Thus, if pennies are not full-bodied, a sufficient rise in the price level will make large coins disappear. If the mint prices differ, a sufficient fall in prices will prompt minting of only one of the two coins. The perpetual coexistence of both coins in the face of price fluctuations requires that pennies be full-bodied and that equal mint prices prevail for both coins; that is, the intervals must coincide and the sum of the seigniorage rates and the production cost must be equal for the two coins.

The state of the technology creates yet another difficulty. We have seen that government had little freedom to choose the seigniorage rate: It had to be positive and could not be large. But making small coins was much more expensive than making large coins, because making a small coin or a large coin involves essentially the same process, independent of the size or content of the coin. In the extreme, if it costs the same to make a penny or a dollar, then the production costs for 100 pennies is 100 times the cost per dollar for the same value of output (the coins). Historical data shows that the cost of making a coin fell with the denomination, but not fast enough. Figure 3 plots the production costs as a function of coin size for various European countries.

This technological constraint presented the mints with a dilemma: provide only full-bodied coins and see pennies never minted, or offer the same price for bullion in pennies or in dollars and face the risk of seeing the price level increase and large coins disappear. Thus, the commodity money system with full-bodied denominations has the potential for either shortages or gluts of small change.

In fact, shortages of small change were a common complaint, running through centuries of monetary history all over Europe and also (in the early nineteenth century) in the U.S. The above argument, although limited to the supply side, shows how vulnerable the commodity money system was to such shortages, given the technology available. An analysis of the demand side reveals even more trouble.

If we think of pennies and dollars as required for consumption purchases (a feature called a cash-in-advance constraint), but we assume that large coins cannot be used in small transactions, whereas small coins can be used in large transactions, it emerges that, within the overlapping intervals of figure 2, there is a certain indeterminacy of the exchange rate between dollars and pennies or the ratio at which pennies enter into the total money stock $M = M_1 + eM_2$ (where M is the total stock in dollars, M_1 is the number of dollar coins, e is the market exchange rate, and M_2 is the number of pennies). As long as there are enough pennies to carry out small transactions (not just in the physical sense M_2 but in terms of their total value eM_2), there can be more or fewer pennies or they can be worth more or less. If, for some reason, the relative share of small transactions changes and more pennies are needed, more pennies will be provided only if the minting points are lined up correctly *and* the price level falls enough. But for the general price level to fall, the shock must affect the volume of *all* transactions, and it is not hard to imagine situations where the existing stock of pennies is insufficient, yet no new pennies are minted.

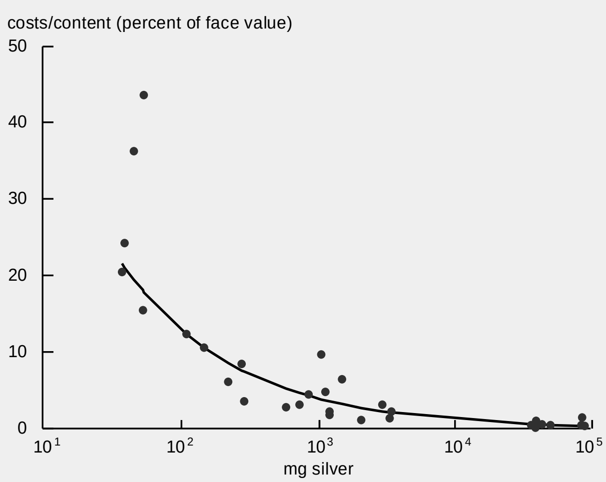
These shortages of small change have a curious feature: In a decentralized economy, agents choose how many pennies to hold. In order for them to hold too few pennies, there needs to be a price incentive for them to economize on pennies. This occurs through a rate of return dominance, that is, the return on holding pennies is lower than the return on holding dollars. In other words, the market exchange rate, e (in dollars per pennies), falls. But this means that the share of pennies in the total stock of coins shrinks further, accentuating the shortage of small change. Furthermore, a fall in the exchange rate shifts the lower interval of figure 2 to the left, making it likelier that the price level will hit the *upper* bound of the interval for pennies, the melting point.

Thus, shortages of small change push the economy in a vicious cycle, by making the shortage even more severe through a depreciation of the smaller denominations, and ultimately bringing about a melting down of pennies, once they have depreciated to the value of their intrinsic content.

Within the confines of the available technology, one partial remedy is for the government to counteract the leftward shift of the interval due to the fall in e by reducing the intrinsic content of pennies, which shifts the interval to the right. Figure 4 plots the evolution of the mint equivalent (the inverse of the intrinsic content) for two medieval Florentine silver coins, the *picciolo* (a penny, or 1d) and the *grosso* (worth 4d), during the

FIGURE 3

Production costs of coins in late medieval Europe, 1350–1500

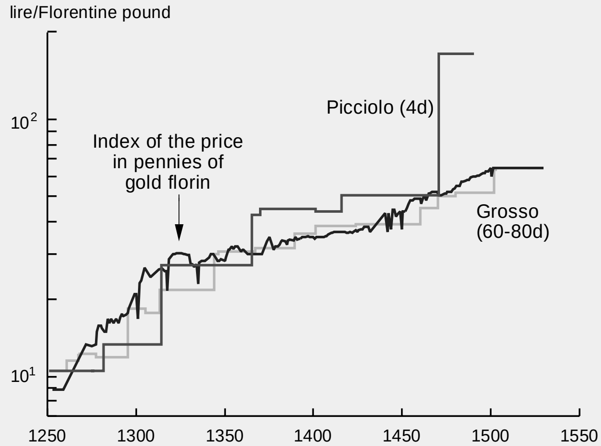


Note: Production costs were calculated by the percent of the face value of the coin as a function of the coin size in milligrams of silver.

Source: T. Sargent and F. Velde, 1997, "The evolution of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-13.

FIGURE 4

The intrinsic content of two Florentine coins, 1250–1530



Notes: The data have been plotted inversely. The coins, picciolo and grosso, are measured in Italian denaro (d), the equivalent of a penny.

Source: T. Sargent and F. Velde, 1997, "The big problem of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-8.

Middle Ages (the gold florin's intrinsic content remained constant). A pattern of recurrent debasements is apparent. The graph also displays the price of the gold florin in terms of silver pennies. This corresponds to the exchange rate of pennies per dollar or $1/e$ (the florin ranged from 240d in 1250 to 1,680d in 1530). One way to interpret this graph is that the periodic debasements, evident as upward steps, occurred to remedy the upward drift in the price of the florin, as our model predicts.

This version of the model takes the price of silver in terms of real resources as constant. In fact, this cost could be taken as variable over time, embodying a variety of shocks (changes in the technology to mine silver, including new discoveries, and changes in the demand for silver in industrial activities). Furthermore, the model assumes that large and small coins are made of the same metal; but small and medium coins being made of silver and large coins being made of gold, the intervals of figure 2 shift around due to changes in the relative price of gold and silver. Depending on the width of the intervals, small shocks might be accommodated, but larger variations lead to the same problems outlined above, unless e is allowed to change. The difficulties in providing multiple denominations render bimetallism (the simultaneous use of two metals in legal tender currencies with a fixed exchange rate) a fragile system.

The evolution of monetary doctrine

These shortcomings of the commodity money system were a result of the state of minting technology until 1550 or so. Moving toward the standard formula, or toward fiduciary coinage, required a better technology. However, the technology would have gone unexploited had monetary doctrine not weakened its attachment to the concept of full-bodied coinage. This evolution of monetary doctrine can be traced in the writings of medieval jurists.[12] This doctrine arose from their efforts to understand observed price patterns and devise ways to deal with the legal consequences for private contracts (the problem of the standard of deferred payments).

Because medieval Europe had begun with the penny and later added larger coins, the tradition was that prices were denominated in pennies, dozens of pennies (shillings), and scores of shillings (pounds).[13] Many nominal debts and contracts were thus expressed in pounds of the small coin, whose constant debasement led to the long-term inflation that is apparent in figure 4.

When the penny was the only coin, monetary doctrine was straightforward. In modern terms, it applied standard price theory to money, treating it as a commodity like any other. When a loan of 100 pennies came due, 100 pennies were owed, irrespective of any fluctuations in the purchasing power of pennies. The Neapolitan jurist Andrea d'Isernia (1220–1316) wrote: "If I lend you a measure of wheat in May when it is expensive and is worth perhaps 3 tarini, and I reclaim it in July after the harvest when it is worth perhaps 1 tarino, it is enough to return the measure of the same wheat in kind, even though it is worth less; likewise if it is worth more, for example if I lent it in July and demanded it in the following May ... the same reasoning applies for money as it does for wheat and wine" (d'Isernia, 1541).

From Charlemagne's reform around 800 A.D. (which restored a uniform currency in Western Europe) to the twelfth century, the penny changed content at various rates, through the action of wear and tear and debasements. Such changes in the intrinsic content of a penny were also treated by jurists in a similar way. The jurist Azo (d. 1220) formulated a simple rule: "The same money or measure is owed that existed at the time of the contract" (in Stampe, 1928, p. 36).

With the appearance of larger denomination coins and the existence of time-varying rates of exchange between denominations, the legal problems grew more challenging, and jurists began to diverge in their answers. A distinction was made between the “intrinsic quality” of a coin (its metal content) and the “extrinsic quality,” taken to mean either its purchasing power (the inverse of the price level) or its rate of exchange with other coins. The general consensus prevailing in the fourteenth and fifteenth centuries called for adjusting debt repayments for variations in the intrinsic quality, but ignoring variations in extrinsic quality; and small coins were considered legal tender to the degree that they were full-bodied and interchangeable with large coins.

However, jurists also observed the existence of positive seigniorage rates (the width of the interval in figure 1), and realized that money’s purchasing power could be greater than its intrinsic value. In other words, they discovered that the price level could move above the minting point. One strand of the legal literature insisted that seigniorage should be set close or equal to 0. Others, who argued that precious metals as bullion and in the form of coins should afford the same utility, recommended that the state subsidize the mint completely (in particular, the jurist Bartolo da Sassoferrato, 1313–1357). As jurists, they tried to define rules for repayment of monetary debts. They correctly perceived that their proposal would eliminate some fluctuations in the standard of value.

In practice, the jurists realized that governments were unwilling to subsidize mints and were tempted to increase seigniorage revenues as much as they could. A small tax was considered acceptable and a larger tax under very specific circumstances, such as a fiscal emergency (paying for a sudden war or the king’s ransom). Some even argued that, in the words of Gabriel Biel (d. 1495) a large seigniorage rate “is the easier way to collect quickly the required funds without fraud and undue exactions from the subjects. It is, moreover, felt less and for this reason more easily borne without protest and without the danger of a rebellion on the part of the people. It is the most general form of taxation embracing all classes, clergy, laity, nobility, plebeians, rich and poor alike” (Biel, 1930 translation, p. 35).

Some jurists like d’Isernia even went further. D’Isernia probably observed episodes such as the siege of Faenza in 1241, when the Emperor Frederic II ran out of money and paid his troops with leather money that he redeemed into gold after the successful conclusion of the siege. D’Isernia argued that, under the specific circumstances already identified by the current doctrine, money could be made of worthless material, like lead or leather, as long as it was redeemed after the end of the emergency into good money. This was the basis for the concept of deficit financing, which would play an important role in the development of fiat money. By the late sixteenth century, these notions were commonly held. The widely cited René Budel (1591) held it “to be indubitable that a Prince in the midst of costly wars, and therefore in great necessity, can order that money be made out of leather, bark, salt, or any material he wants, if he is careful to repair the loss inflicted thereby on the community with good and better money” (Budel, 1591, chapter 1, paragraph 31).

In other words, the intrinsic content could be set to 0, as long as some measure of convertibility, either immediate or in the near future, was implied. In 1481, a small town in Catalonia carried out an experiment to solve its problem of small change: it was authorized by the king of Aragon to issue pure copper coins,[14] whose intrinsic value was about 25 percent of their face value, as long as “the city be known to pledge, and effectively pledge to receive said small money from those who might hold it, and to convert it and return for it good money of gold or silver, whenever and however much they be asked” (in Botet y Sisó, 1911, p. 328). This experiment was imitated by a number of other Catalonian cities, although they were plagued by counterfeiting, which the state of technology made relatively easy.

Technological change and policy experiments

These developments in monetary doctrine, and the early Catalonia experiment, show that technology remained the real barrier to the implementation of a standard formula for small change. The technology did change, in two major waves; and each wave opened up new possibilities that governments exploited.

Recall that the standard formula incorporates several ingredients: monopolization of coinage, issue of tokens, and convertibility of the tokens. The ingredients are logically distinct. The period between the first and the second wave of technological change (1550 to 1800) saw a wide variety of experiments, in which some but not always all ingredients were proposed or implemented. The variety of outcomes offered a rich mine of lessons in monetary doctrine.

Mechanization and the Age of Copper

The first major shift in minting technology took place around 1550. In southern Germany, two processes were independently developed to mechanize the minting process, using machines rather than tools to cut uniform blanks and impress them with a design. One technology (the screw-press) proved to be better than the other (the cylinder-press), but also more expensive, and only prevailed in the late seventeenth century. Until then, the other proved popular in a number of countries, including the various German states and Spain.

The king of Spain heard about the cylinder-press technology from his cousin the count of Tirol, who had been the first to install the new machines in his state mint. The machines were imported and set up in Segovia in 1582, and applied to the silver coinage of pieces of eight. The coins produced in Segovia were much more uniform and round, and more sharply imprinted, than anything done using the old hand tools. The Spanish government soon realized the potential in this technology, and decided in 1596 to produce all small denominations in pure copper with the new machines. King Philip II explained his reasons in an edict:

> We have been advised by people of great experience, that the silver which is put in those billon coins[15] is lost forever and no profit can be drawn from it, except in their use as money, and that the quantity of silver which is put to that use for the necessities of ordinary trade and commerce in this kingdom is large. We have also been advised that, since we have established a new machine in the city of Segovia to mint coins, if we could mint the billon coinage in it, we would have the assurance that it could not be counterfeited, because only a small quantity could be imitated and not without great cost if not by the use of a similar engine, of which there are none other in this kingdom or the neighboring ones. And it would thus be possible to avoid adding the silver (in Rivero, 1919, p. 150).

Until then, copper, silver, and minting costs each represented a third of the face value of billon coinage. With Philip II's decree, the silver was withheld and the copper content reduced.

Philip II had efficiency in mind. He ordered that the new copper coins be issued only to retire existing small denomination coins (M_2) with token coinage and that the mechanism with its melting and minting points be preserved for providing large denomination silver coins (M_1). Retaining the mechanism for supplying M_1 would keep the price level within the appropriate melting and minting points so long as some large denomination coins continued to circulate.

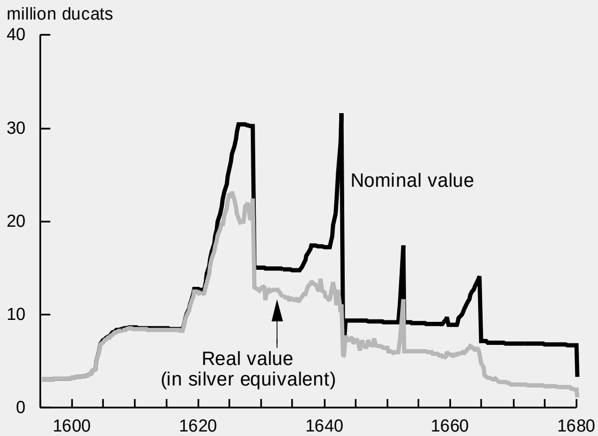
But Philip II's successors, Philip III (1598–1621) and Philip IV (1621–64), saw that the cylinder press offered opportunities to enhance revenues. A first experiment in 1602, whereby the copper content of coins was reduced by 50 percent with no resulting effect on the price level, convinced the government that the intrinsic value of the coins could be made much lower and the seigniorage rate much more lucrative. Another experiment, carried out in 1603, further reinforced the point that individuals did not care about the composition of their money balances. After the 1602 reduction, two kinds of pennies circulated, one twice as heavy as the other; it was decided that all old (heavy) pennies were to be brought to the mint, stamped with a "2" and one two-cent coin returned for every two old pennies presented. The operation was successful and all old pennies were presented, affording the government 50 percent seigniorage on the stock of pennies.

From that point on, the Castilian government knew no restraint, and enormous quantities of *vellón* (as these copper pennies were called) were minted and used to finance government consumption. Figure 5 shows the path of nominal and real balances of vellón in that period; note that the total money stock before 1600 was around 20 to 30 million ducats.

Recall that we express the total quantity of money as $M = M_1 + eM_2$, where M_1 represents the stock of large denomination (silver) coins, and M_2 represents small denomination (copper) coins. The exchange rate between the two types of coins is e, and M_2 is expressed in dollars. The policy followed by the Castilian government consisted in increasing M_2 to the point at which it completely replaced M_1, all the while with no inflation (real and nominal balances coincide). In terms of the total money stock, $M_1 + eM_2$, a

FIGURE 5

Nominal and real values of vellón, 1595–1680



Source: F. Velde and W. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.

progressive displacement of M_1 by vellón is consistent with no change in *e*, and, other things being equal, an unchanged money stock will correspond to a constant price level. However, once M_1 has disappeared, the money stock consists only of copper coins M_2, and all further increases in M_2 result in increases in the price level, as is apparent in figure 5. Once the figure of about 20 million ducats was reached, nominal and real balances diverged, and inflation set in with a vengeance. The disappearance of silver released the price level from the constraints imposed by the melting / minting points for the dollar interval, and unleashed the quantity theory with copper as the determinant of the price level.

The only way to return the price level to its bounds was to engineer a reappearance of the silver coins, either by decreasing M_2 or by decreasing *e*. The Castilian government toyed with the idea of decreasing M_2 by an open-market operation (selling bonds to buy back the copper coinage), but in the end decided to halve *e* overnight, in 1628.

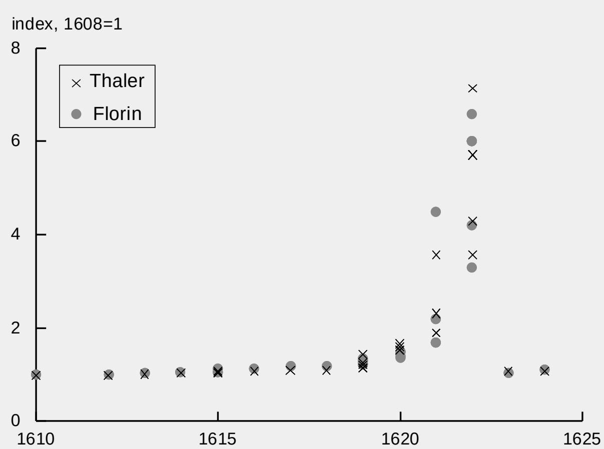
The rest of the movements in vellón balances are due to repetitions of the earlier operations of vellón issue, restamping (multiplying the face value of existing coins by N and extracting a seigniorage of $(N–1) / N)$ and overnight devaluations. As figure 5 shows, Castilians grew weary of the manipulations, which were less successful as balances of vellón fell over time.

The Spanish experience unleashed unprecedented "man-made" inflation, which made the Price Revolution of the sixteenth century (price level increases due to the inflow of American gold and silver) look tame. It was among the first large-scale experiments in inconvertible fiat currency (although the coins were accepted at face value in payment of taxes). It demonstrated the ease with which token coinage could overtake the money stock, the workings of the quantity theory, the need for the issuer of inconvertible token coinage to restrain issues, and the strength of the temptation created by high seigniorage rates for a government unwilling or unable to raise other taxes.

The Spanish experiment was not the only one at the time. During the Thirty Years War, which started in 1618, many German states concurrently debased their small denominations (all the while maintaining silver coinage intact) and issued large amounts of copper coinage to raise revenues through seigniorage. The results are shown in figure 6, which tracks the exchange rate between large denomination coins and small denomination coins and makes it clear why the Germans called this *die große Inflation* (the great inflation), at least until a similar experiment exactly 300 years later (the famous

FIGURE 6

Prices of the gold florin and silver thaler, in Bavarian kreutzers



Source: H. Altmann, 1976, *Die Kipper- und Wipperinflation in Bayern*, München: Neue Schriftenreihe des Stadtarchivs München, pp. 272–273.

German hyperinflation of 1922–23 under the Weimar Republic). Poland and Russia also underwent copper inflations in the 1650s, as did the Ottoman empire in the 1690s. This is why the seventeenth century has earned the name the Age of Copper.

Lessons from the Castilian inflation

The lessons were not lost on contemporary observers. The Spanish episode was discussed not only by writers in Spain, but also in Italy, France, and elsewhere, leading to a consensus on quantity limitations and limited legal tender for small coins.

One of the more famous commentators was the Jesuit Juan de Mariana (1536–1624), who wrote a treatise on the vellón coinage between 1603 and 1606, as the experiment was beginning and inflation had not yet taken off. He lays out arguments pro and con, and thus provides a window on the debates among policymakers around the Spanish king.

The advantages vaunted by proponents of the copper coinage are not limited to the social savings mentioned by Philip II in his edict. Proponents claimed that without a stock of silver coins as a potential reserve to settle trade deficits, Spain would be forced to maintain surpluses and resort to import substitution, thereby stimulating Spanish industry; they also claimed that the copper money was lighter and easier to transport, and that its cheap provision would lower the rate of interest and stimulate agriculture and industry. In other words, arguments were made that, beyond the social savings from forsaking commodity money, increases in the quantity of money could stimulate output.[16]

Mariana was conscious that incentives for counterfeiting created by the overvaluation of copper coins could be resolved by the new machines in the Segovia mills. He was doubtful of the arguments on balance of trade and stimulus of the economy, which could be made to go the other way through an anticipated inflation effect. He predicted that copper coinage would drive out silver, lead to an increase in prices, and induce the government to set price controls that would either be ignored or counterproductive, at which point the government would be forced to reduce the face value of the coins, as indeed happened in 1628. Mariana saw the projected sequence of inflation and deflation as disruptive to trade and contracts and, therefore, to the king's tax revenues. He also viewed the high seigniorage rates of 50 percent in the restamping operations as immoral, because in his view the king has no right to tax his subjects without their explicit consent. Mariana noted that such high tax rates would never be tolerated on any other tax base. The worst consequence he predicted was general hatred of the government. Quoting Tacitus, he recalled that "everyone claims prosperity for himself, but adversity is blamed on the leader" (1994, p. 104).

The Frenchman Poullain, quoted earlier, concluded that token coins could replace other coins for domestic transactions and that this was precisely why their quantity should be limited. Poullain, as a monetary official, successfully fought back various plans to issue copper on a large scale. Only twice, in 1640 and 1653, did France come close to embarking on a Spanish-style inflation, in both cases at times of fiscal emergency.

The Italian Montanari, also quoted above, wrote: "It is clear enough that it is not necessary for a prince to strike petty coins having metallic content equal to their face value, provided he does not strike more of them than is sufficient for the use of his people, sooner striking too few than striking too many. If the prince strikes only as many as the people need, he may strike of whatever metallic content he wishes" (Montanari, 1804, p. 109). Various other writers stressed quantity limitations, as well as limited legal tender for small coins. The latter measure uncouples the two stocks of money in the equation $M_1 + eM_2$, which was critical in the Spanish experience.

Monopoly versus laisser-faire

English coins had always been made of sterling silver, and shortages of small change became particularly acute when pennies and farthings ceased to be minted altogether in the sixteenth century. From that point until 1817, English policy alternated between three regimes for the supply of small change: private monopolies of inconvertible token coinage, government monopoly of full-bodied coinage, and laisser-faire (that is, the absence of government intervention).

Private monopolies (1613–44) were created by royal charter, which granted various individuals in turn (usually well-connected aristocrats) the exclusive right to issue token coinage, although these were never made legal tender and their quantities were limited by the terms of the charter. A government monopoly was asserted in 1672, making private tokens illegal, and the Royal mint issued copper coins, intermittently and insufficiently, until 1754. Although mechanization had been adopted in 1660, England remained committed to full-bodied copper coinage.[17]

The laisser-faire regime (mid-sixteenth century to 1613, 1644 to 1672, and the late eighteenth century) was characterized by the absence of government-issued small denominations and by the issue of tokens by private parties or local governments. In the late sixteenth century, up to 3,000 London merchants issued tokens. In the period from 1644 to 1672, over 12,700 different types of tokens have been catalogued, issued in 1,700 different English towns. From the 1740s on, trade tokens took over when official coinage ceased. Some of these issues were authorized by government. The city of Bristol sought and secured permission to issue farthings in 1652, and went through three different issues over the next 20 years. The Bristol farthings, furthermore, were officially convertible into large denominations. They are also known to have been counterfeited. The government put an end to the laisser-faire regime twice, in 1672 and by the Act of Suppression in 1817; each time, it did so immediately after adopting a new technology.

France's experiences were somewhat parallel. In the early seventeenth century, private monopolies were instituted for brief periods of time. France also had a brief experience with free token issue in 1790–92. The government had decided in September 1790 to issue substantial amounts of large-denomination paper currency backed by a land sales scheme. Soon, thousands of private and municipal banks emerged to intermediate the government's notes with their own small denomination notes and, in some cases, coins. Initially, the government abstained from regulating the industry, which operated on fractional or 100 percent reserves, depending on the institution. But soon the government moved to eliminate its competitors in the business of issuing currency. The government decided to issue medium-sized notes (equivalent to silver coins) in June 1791, followed in December 1791 by small-denomination notes. Technical difficulties postponed the first issue of small notes to August 1792. The government could now impose a monopoly. Within a few weeks, all private banks were forbidden to issue their notes and private coins were outlawed, amid unproven allegations of wildcatting and fraud.

These episodes present parallels with the Free Banking Eras of eighteenth century Scotland and the nineteenth century U.S. One of the ingredients of the standard formula is monopolization of coinage; it is not clear, on theoretical or historical grounds, that this ingredient is needed if the other two (issue and convertibility of tokens) are present. Of the two advantages of the token coinage system, social savings and government revenues, the latter clearly provides an impulse toward monopolization that the former does not.

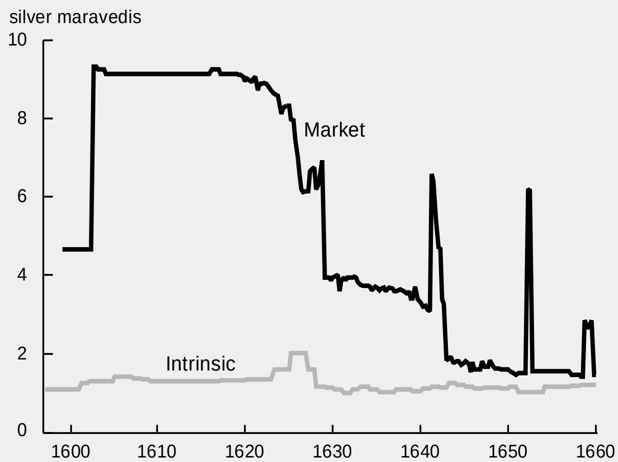
The steam engine and the gold standard

The second major technological innovation following the mechanization of minting around 1550 was the adaptation of the steam engine to minting. In 1787, Matthew Boulton, partner of James Watt, produced trade tokens for the Anglesey Copper Company. A few years later he was producing copper coins for private issuers across England and even in France. In Paris, the most popular token coins in 1790–92, issued by the firm of Monneron, were minted in Birmingham by Boulton's steam presses. The British government contracted with him to produce official copper coinage in 1797, then bought the technology, and in 1817 eliminated its competitors by making private coins illegal. The new steam-driven presses were used to mint the new silver coins, which, under the Coinage Act of 1816, were for the first time issued as partly fiduciary coins, whose intrinsic value was significantly lower than their face value. It took a decade and a half before an implicit agreement was reached between the Bank of England and the Treasury for the convertibility of the silver coinage into gold upon demand. By 1830, the standard formula had been fully implemented.

England's implementation of the standard formula in 1816 applied both to bronze or copper coinage and to silver coins, leaving gold as the single anchor for the price level (the gold standard) and officially abandoning bimetallism. It took other countries some time to follow suit: Germany in 1871, France and the Latin Monetary Union (Belgium, Switzerland, Spain, and Italy) in 1873, the Netherlands in 1875, and the U.S. in 1873–79 (the so-called Crime of 1873). Recently, researchers (Friedman 1990 and Flandreau 1997) have argued that this abandonment was a mistake, and bimetallism was better suited to stabilizing the price level than the gold standard. Nevertheless, there was no substantial difference between applying the standard formula to silver and applying it to copper coinage, and the

FIGURE 7

Market and intrinsic values of a vellón cuarto coin, 1597–1659



Source: F. Velde and W. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.

forces identified by Friedman (1960) and those leading to coin shortages seemed to lead to the outcome effectively adopted by most countries.

Conclusion

The questions raised by Friedman (1960) about the necessary ingredients for an efficient and well-managed currency are old questions indeed. The big problem of small change led monetary thinking on the path to fiduciary currency, at least in the form of intrinsically trifling but convertible tokens; policy followed only after the right technology became available. As technology changed and experience accumulated, various elements of the standard formula were tried separately, including irredeemable copper money. The resulting inflation led to the recognition that a form of quantity theory was at play, and led governments to formulate various ways of limiting the quantity—through convertibility and through monopoly.

Of the main ingredients of the standard formula, the historical trend points clearly to token coinage. Monopolization is less obvious an outcome, especially given the prolonged Free Token Eras of England. Friedman argues that government needs a monopoly on fiduciary currency because free entry into the issue of irredeemable paper would drive down currency to its intrinsic value (namely, 0). As figure 7 shows, this is what happened in seventeenth century Spain, as the market value of copper coinage was driven down to its intrinsic value. Arguably, counterfeiting was widespread, but judging by figure 5, government issues are enough to account for the phenomenon. Surely, experiences with fiat money in the twentieth century (a century replete with hyperinflation) show that governments can drive the value of a paper currency they monopolize to its intrinsic value with great efficacy.

Perhaps it is not surprising that seventeenth century Spain was under an autocratic regime, as was contemporary France (which came close to the same outcome). England, where counterweights to the executive were at least apparent at the time and constitutionally set in 1688, maintained a different policy. Nor perhaps is it a surprise that the standard formula was first implemented in Britain, the most advanced democracy in Europe at the time. The policy was implemented in 1816, just as Britain was emerging from a successful use of inconvertible paper money to finance 20 years' worth of wartime expenditures (in contrast to France's similar attempt in 1790–97, which proved less durable). Irredeemable currency for deficit financing was already a centuries-old idea; the Catalonian town of Gerona used a coin issued as siege money to start a convertible-token system in 1481. Success with deficit financing was probably a good predictor of success with subsidiary coinage; both may have something to do with the degree of accountability of policymakers.

NOTES

[1]Much of the material presented here derives from the work in Sargent and Velde (1997a, 1997b).

[2]The model sketched here is developed fully in Sargent and Velde (1997a).

[3]Seigniorage is literally the lord's right to collect a tax, and is derived from the French term for lord, *seigneur*.

[4]This cost is exclusive of the coin's content. It represents the costs of transforming metal into coins, and is to some degree independent of the content.

[5]The equation is $pY = vM$, where p is the price level, Y is income or the volume of transactions, v is velocity, and M is the quantity of money.

[6]The optimal denomination structure is an unstudied problem; however, see Telser (1995).

[7]In the numismatic sense, token means something that is not officially money, but used as money; numismatists will speak of full-bodied tokens. From an economic viewpoint, the distinction between official and unofficial money is somewhat arbitrary.

[8]Act of June 9, 1879: "Be it enacted ... that the holder of any of the silver coins of the United States of smaller denominations than one dollar, may, on presentation of the same in sums of $20, or any multiple thereof, at the office of the Treasurer or any assistant treasurer of the United States, receive therefor lawful money of the United States" (Statutes at Large 21 [1879]: 7).

[9]The status of silver dollars remained uncertain between the Bland–Allison Act of 1878 and the final defeat of the pro-silver forces after 1896. Only after 1900 did the silver dollar become no different in nature from other subsidiary coins.

[10]Note that the ability to maintain a monopoly on full-bodied coinage is dependent on the same.

[11]The punishment for counterfeiters was particularly severe. In medieval France, they were boiled alive (not poached). A document from 1311 details the costs of executing two counterfeiters, including the price of a large cauldron and the cost of adding iron bars to the cauldron, a detail that suggests a rather long process (Saulcy, 1879–92, Vol. 1, p. 180).

[12]An anthology of their writings is in Velde (1997).

[13]This did not preclude the denomination of many prices in terms of the gold coin and fictitious subdivisions thereof.

[14]Interestingly, the coins were modeled on a currency issued some years earlier as emergency money during a siege and later left in circulation.

[15]A mixture of silver, to give value, and copper, to give bulk, commonly used for small denominations.

[16]Another famous proponent of similar arguments was the Scot John Law (1671–1729), whose experiment in setting up a paper currency in France went spectacularly awry in 1720, during the Mississippi Bubble.

[17]The proclamation of 1672 stated that small coins "cannot well be done in silver, nor safely in any other metal, unless the intrinsic value of the coin be equal, or near to that value for which it is made current." Sir Isaac Newton, master of the mint, wrote in 1720: "Halfpence and farthings (like other money) should be made of a metal whose price among Merchants is known, and should be coined as near as can be to that price, including the charge of coinage. ... All which reasons incline us to prefer a coinage of good copper according to the intrinsic value of the metal" (Shaw, 1896, pp. 164–165).

REFERENCES

Altmann, Hans Christian, 1976, *Die Kipper- und Wipperinflation in Bayern*, München: Neue Schriftenreihe des Stadtarchivs München.

Aristotle, 1943, *Ethics*, H. Racham (trans.), Cambridge, MA: Harvard University Press.

Biel, Gabriel, 1935, *Treatise on the Power and Utility of Money*, Robert B. Burke (trans.), Philadelphia, PA: University of Philadelphia Press.

Botet y Sisó, Joaquim, 1911, *Les Monedes Catalans*, Vol. 2, Barcelona: Institut d'Estudis Catalans.

Budel, René, 1591, *De Monetis et Re Numaria*, Book 1, Cologne: Schann Nackel.

Cipolla, Carlo M., 1956, *Money, Prices and Civilization in the Mediterranean World, Fifth to Seventeenth Century*, New York: Gordian Press.

D'Isernia, Andrea, 1541, *Commentaria in Usus Librum Feudorum*, Lyon, France: J. Giunta.

Flandreau, Marc, 1997, "The gradient of a river: Bimetallism as an implicit fluctuation band," Paris: l'Observatoire Français des Conjonctures Economiques, manuscript.

Friedman, Milton, 1990, "The crime of 1873," *Journal of Political Economy*, Vol. 98, December, pp. 1159-1194.

_________, 1960, *A Program for Monetary Stability*, New York: Fordham University Press.

Mariana, Juan de, 1994, *De Monetae Mutatione*, Josef Falzberger (ed.), Heidelberg: Manutius Verlag.

Mill, John Stuart, 1857, *Principles of Political Economy*, Fourth Edition, London: John W. Parker and Son.

Montanari, Geminiano, 1804, "Della moneta: Trattato mercantile," in *Scrittori Classici Italiani di Economia Politica*, parte antica, Vol. 3, Milan: G.G. Destefanis.

Poullain, Henri, 1709, *Traité des Monnoies*, Paris: Frédéric Léonard.

Rivero, Casto Maria del, 1919, *El Ingenio de la Moneda*, Revista de Archivos, Bibliotecas y Museos, Vol. 40.

Sargent, Thomas J., and François R. Velde, 1997a, "The big problem of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-8.

_________, 1997b, "The evolution of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-13.

Saulcy, Ferdinand de, 1879–92, *Recueil de Documents Relatifs à l'Histoire des Monnaies*, 4 vols., Paris: Imprimerie Nationale.

Shaw, William Arthur, 1896, *The History of Currency, 1252 to 1894*, London: Wilsons & Milne.

Spengler, Joseph J., 1966, "Coin shortage: Modern and premodern," *National Banking Review*, Vol. 3, pp. 201–216.

Stampe, Ernst, 1928, "Das Zahlkraftrecht der Postglossatorenzeit," *Abhandlungen der Preußischen Akademie der Wissenschaften*, philosophisch-historische, Klasse 1, Berlin: Verlag der Akademie.

Telser, Lester G., 1995, "Optimal denominations for coins and currency," *Economic Letters*, Vol. 49, pp. 425–427.

Tesauro, 1609, "Tractatus de monetarum augmento et variatione," in *De Monetarum Augmento Variatione et Diminutione Tractatus Varii*, Turin, Italy.

Velde, François R., 1997, "An anthology of writings on money, 12th–17th century," Federal Reserve Bank of Chicago, manuscript.

Velde, François R., and Warren E. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.