GWARTNEY INSTITUTE

FREEDOM • JUSTICE • FLOURISHING

OTTAWA UNIVERSITY

THE FUTURE OF MONEY

2024 PPE LEAGUE READER

MENARD FAMILY

Student Programs

THE FUTURE OF MONEY

Table of Contents

2024

SESSION 1

What is Money and Bank?

Rothbard "Mystery of Banking" Chapter 1	***Pages 1-13***
Nick Szabo: Shelling Out "The origins of money"	***Pages 14-38***
Rothbard "Mystery of Banking" Chapter 3	***Pages 39-51***

Gwartney Institute

I.

MONEY: ITS IMPORTANCE AND ORIGINS

1. THE IMPORTANCE OF MONEY

Today, money supply figures pervade the financial press. Every Friday, investors breathlessly watch for the latest money figures, and Wall Street often reacts at the opening on the following Monday. If the money supply has gone up sharply, interest rates may or may not move upward. The press is filled with ominous forecasts of Federal Reserve actions, or of regulations of banks and other financial institutions.

This close attention to the money supply is rather new. Until the 1970s, over the many decades of the Keynesian Era, talk of money and bank credit had dropped out of the financial pages. Rather, they emphasized the GNP and government's fiscal policy, expenditures, revenues, and deficits. Banks and the money supply were generally ignored. Yet after decades of chronic and accelerating inflation—which the Keynesians could not begin to cure—and after many bouts of "inflationary recession," it became obvious

to all—even to Keynesians—that something was awry. The money supply therefore became a major object of concern.

But the average person may be confused by so many definitions of the money supply. What are all the Ms about, from M1-A and M1-B up to M-8? Which is the *true* money supply figure, if any single one can be? And perhaps most important of all, why are bank deposits included in all the various Ms as a crucial and dominant part of the money supply? Everyone knows that paper dollars, issued nowadays exclusively by the Federal Reserve Banks and imprinted with the words "this note is legal tender for all debts, public and private" constitute money. But why are checking accounts money, and where do they come from? Don't they have to be redeemed in cash on demand? So why are checking deposits considered money, and not just the paper dollars backing them?

One confusing implication of including checking deposits as a part of the money supply is that banks *create* money, that they are, in a sense, money-creating factories. But don't banks simply channel the savings we lend to them and relend them to productive investors or to borrowing consumers? Yet, if banks take our savings and lend them out, how can they *create* money? How can their liabilities become part of the money supply?

There is no reason for the layman to feel frustrated if he can't find coherence in all this. The best classical economists fought among themselves throughout the nineteenth century over whether or in what sense private bank notes (now illegal) or deposits should or should not be part of the money supply. Most economists, in fact, landed on what we now see to be the wrong side of the question. Economists in Britain, the great center of economic thought during the nineteenth century, were particularly at sea on this issue. The eminent David Ricardo and his successors in the Currency School, lost a great chance to establish truly hard money in England because they never grasped the fact that bank deposits are part of the supply of money. Oddly enough, it was in the United States, then considered a backwater of economic theory, that economists first insisted that bank

deposits, like bank notes, were part of the money supply. Condy Raguet, of Philadelphia, first made this point in 1820. But English economists of the day paid scant attention to their American colleagues.

2. How Money Begins

Before examining what money *is*, we must deal with the importance of money, and, before we can do that, we have to understand how money arose. As Ludwig von Mises conclusively demonstrated in 1912, money does not and cannot originate by order of the State or by some sort of social contract agreed upon by all citizens; it must always originate in the processes of the free market.

Before coinage, there was *barter*. Goods were produced by those who were good at it, and their surpluses were exchanged for the products of others. Every product had its barter price in terms of all other products, and every person gained by exchanging something he needed less for a product he needed more. The voluntary market economy became a latticework of mutually beneficial exchanges.

In barter, there were severe limitations on the scope of exchange and therefore on production. In the first place, in order to buy something he wanted, each person had to find a seller who wanted precisely what he had available in exchange. In short, if an egg dealer wanted to buy a pair of shoes, he had to find a shoemaker who wanted, at that very moment, to buy eggs. Yet suppose that the shoemaker was sated with eggs. How was the egg dealer going to buy a pair of shoes? How could he be sure that he could find a shoemaker who liked eggs?

Or, to put the question in its starkest terms, I make a living as a professor of economics. If I wanted to buy a newspaper in a world of barter, I would have to wander around and find a newsdealer who wanted to hear, say, a 10-minute economics lecture from me in exchange. Knowing economists, how likely would I be to find an interested newsdealer?

This crucial element in barter is what is called the *double coincidence of wants*. A second problem is one of *indivisibilities*. We can see clearly how exchangers could adjust their supplies and sales of butter, or eggs, or fish, fairly precisely. But suppose that Jones owns a house, and would like to sell it and instead, purchase a car, a washing machine, or some horses? How could he do so? He could not chop his house into 20 different segments and exchange each one for other products. Clearly, since houses are *indivisible* and lose all of their value if they get chopped up, we face an insoluble problem. The same would be true of tractors, machines, and other large-sized products. If houses could not easily be bartered, not many would be produced in the first place.

Another problem with the barter system is what would happen to *business calculation*. Business firms must be able to calculate whether they are making or losing income or wealth in each of their transactions. Yet, in the barter system, profit or loss calculation would be a hopeless task.

Barter, therefore, could not possibly manage an advanced or modern industrial economy. Barter could not succeed beyond the needs of a primitive village.

But man is ingenious. He managed to find a way to overcome these obstacles and transcend the limiting system of barter. Trying to overcome the limitations of barter, he arrived, step by step, at one of man's most ingenious, important and productive inventions: *money*.

Take, for example, the egg dealer who is trying desperately to buy a pair of shoes. He thinks to himself: if the shoemaker is allergic to eggs and doesn't want to buy them, what *does* he want to buy? Necessity is the mother of invention, and so the egg man is impelled to try to find out what the shoemaker would like to obtain. Suppose he finds out that it's fish. And so the egg dealer goes out and buys fish, not because he wants to eat the fish himself (*he* might be allergic to fish), but because he wants it in order to *resell* it to the shoemaker. In the world of barter, everyone's purchases were purely for himself or for his family's direct use. But now, for the first time, a new element of demand has entered:

The egg man is buying fish not for its own sake, but instead to use it as an indispensable way of obtaining shoes. Fish is now being used as a *medium of exchange*, as an instrument of *indirect exchange*, as well as being purchased directly for its own sake.

Once a commodity begins to be used as a medium of exchange, when the word gets out it generates even further use of the commodity as a medium. In short, when the word gets around that commodity X is being used as a medium in a certain village, more people living in or trading with that village will purchase that commodity, since they know that it is being used there as a medium of exchange. In this way, a commodity used as a medium feeds upon itself, and its use spirals upward, until before long the commodity is in general use throughout the society or country as a medium of exchange. But when a commodity is used as a medium for most or all exchanges, that commodity is defined as being a *money*.

In this way money enters the free market, as market participants begin to select suitable commodities for use as the medium of exchange, with that use rapidly escalating until a general medium of exchange, or money, becomes established in the market.

Money was a leap forward in the history of civilization and in man's economic progress. Money—as an element in every exchange—permits man to overcome all the immense difficulties of barter. The egg dealer doesn't have to seek a shoemaker who enjoys eggs; and I don't have to find a newsdealer or a grocer who wants to hear some economics lectures. All we need do is exchange our goods or services for money, for the money commodity. We can do so in the confidence that we can take this universally desired commodity and exchange it for any goods that we need. Similarly, indivisibilities are overcome; a homeowner can sell his house for money, and then exchange that money for the various goods and services that he wishes to buy.

Similarly, business firms can now calculate, can figure out when they are making, or losing, money. Their income and their expenditures for all transactions can be expressed in terms of money. The firm took in, say, $10,000 last month, and spent

$9,000; clearly, there was a net profit of $1,000 for the month. No longer does a firm have to try to add or subtract in commensurable objects. A steel manufacturing firm does not have to pay its workers in steel bars useless to them or in myriad other physical commodities; it can pay them in money, and the workers can then use money to buy other desired products.

Furthermore, to know a good's "price," one no longer has to look at a virtually infinite array of relative quantities: the fish price of eggs, the beef price of string, the shoe price of flour, and so forth. Every commodity is priced in only one commodity: money, and so it becomes easy to compare these single money prices of eggs, shoes, beef, or whatever.

3. The Proper Qualities of Money

Which commodities are picked as money on the market? Which commodities will be subject to a spiral of use as a medium? Clearly, it will be those commodities most useful as money in any given society. Through the centuries, many commodities have been selected as money on the market. Fish on the Atlantic seacoast of colonial North America, beaver in the Old Northwest, and tobacco in the Southern colonies were chosen as money. In other cultures, salt, sugar, cattle, iron hoes, tea, cowrie shells, and many other commodities have been chosen on the market. Many banks display money museums which exhibit various forms of money over the centuries.

Amid this variety of moneys, it is possible to analyze the qualities which led the market to choose that particular commodity as money. In the first place, individuals do not pick the medium of exchange out of thin air. They will overcome the double coincidence of wants of barter by picking a commodity which is *already* in widespread use for its own sake. In short, they will pick a commodity *in heavy demand*, which shoemakers and others will be likely to accept in exchange from the very start of the money-choosing process. Second, they will pick a commodity which is *highly divisible*, so that small chunks of other goods can be bought, and size of purchases can be flexible. For this they

need a commodity which technologically does not lose its quotal value when divided into small pieces. For that reason a house or a tractor, being highly indivisible, is not likely to be chosen as money, whereas butter, for example, is highly divisible and at least scores heavily as a money for this particular quality.

Demand and divisibility are not the only criteria. It is also important for people to be able to carry the money commodity around in order to facilitate purchases. To be easily *portable*, then, a commodity must have *high value per unit weight*. To have high value per unit weight, however, requires a good which is not only in great demand but also relatively scarce, since an intense demand combined with a relatively scarce supply will yield a high price, or high value per unit weight.

Finally, the money commodity should be highly durable, so that it can serve as a store of value for a long time. The holder of money should not only be assured of being able to purchase other products right now, but also indefinitely into the future. Therefore, butter, fish, eggs, and so on fail on the question of durability.

A fascinating example of an unexpected development of a money commodity in modern times occurred in German POW camps during World War II. In these camps, supply of various goods was fixed by external conditions: CARE packages, rations, etc. But after receiving the rations, the prisoners began exchanging what they didn't want for what they particularly needed, until soon there was an elaborate price system for every product, each in terms of what had evolved as the money commodity: cigarettes. Prices in terms of cigarettes fluctuated in accordance with changing supply and demand.

Cigarettes were clearly the most "moneylike" products available in the camps. They were in high demand for their own sake, they were divisible, portable, and in high value per unit weight. They were not very durable, since they crumpled easily, but they could make do in the few years of the camps' existence.[1]

---

[1]See the justly famous article by R.A. Radford, "The Economic Organization of a P.O.W. Camp," *Economica* (November 1945): 189–201.

In all countries and all civilizations, two commodities have been dominant whenever they were available to compete as moneys with other commodities: *gold* and *silver*.

At first, gold and silver were highly prized only for their luster and ornamental value. They were always in great demand. Second, they were always relatively scarce, and hence valuable per unit of weight. And for that reason they were portable as well. They were also divisible, and could be sliced into thin segments without losing their pro rata value. Finally, silver or gold were blended with small amounts of alloy to harden them, and since they did not corrode, they would last almost forever.

Thus, because gold and silver are supremely "moneylike" commodities, they are selected by markets as money if they are available. Proponents of the gold standard do not suffer from a mysterious "gold fetish." They simply recognize that gold has always been selected by the market as money throughout history.

Generally, gold and silver have both been moneys, side-by-side. Since gold has always been far scarcer and also in greater demand than silver, it has always commanded a higher price, and tends to be money in larger transactions, while silver has been used in smaller exchanges. Because of its higher price, gold has often been selected as the unit of account, although this has not always been true. The difficulties of mining gold, which makes its production limited, make its long-term value relatively more stable than silver.

4. The Money Unit

We referred to *prices* without explaining what a price really is. A price is simply the ratio of the two quantities exchanged in any transaction. It should be no surprise that every monetary unit we are now familiar with—the dollar, pound, mark, franc, et al.—began on the market simply as names for different units of weight of gold or silver. Thus the "pound sterling" in Britain, was exactly that—one pound of silver.[2]

[2]At current writing, silver is approximately $13 an ounce, and the pound is about $1.50, which means that the British "pound sterling," once

The "dollar" originated as the name generally applied to a one-ounce silver coin minted by a Bohemian count named Schlick, in the sixteenth century. Count Schlick lived in Joachimsthal (Joachim's Valley). His coins, which enjoyed a great reputation for uniformity and fineness, were called *Joachimsthalers* and finally, just *thalers*. The word *dollar* emerged from the pronunciation of *thaler*.

Since gold or silver exchanges by weight, the various national currency units, all defined as particular weights of a precious metal, will be automatically fixed in terms of each other. Thus, suppose that the dollar is defined as 1/20 of a gold ounce (as it was in the nineteenth century in the United States), while the pound sterling is defined as 1/4 of a gold ounce, and the French franc is established at 1/100 of a gold ounce.[3] But in that case, the *exchange rates* between the various currencies are automatically fixed by their respective quantities of gold. If a dollar is 1/20 of a gold ounce, and the pound is 1/4 of a gold ounce, then the pound will automatically exchange for 5 dollars. And, in our example, the pound will exchange for 25 francs and the dollar for 5 francs. The definitions of weight automatically set the exchange rates between them.

Free market gold standard advocates have often been taunted with the charge: "You are against the government fixing the price

---

proudly equal to one pound of silver, now equals only 1/8 of a silver ounce. How this decline and fall happened is explained in the text.

[3]The proportions are changed slightly from their nineteenth century definitions to illustrate the point more clearly. The "dollar" had moved from Bohemia to Spain and from there to North America. After the Revolutionary War, the new United States changed its currency from the British pound sterling to the Spanish-derived dollar. From this point on, we assume gold as the only monetary metal, and omit silver, for purposes of simplification. In fact, silver was a complicating force in all monetary discussions in the nineteenth century. In a free market, gold and silver each would be free to become money and would float freely in relation to each other ("parallel standards"). Unfortunately, governments invariably tried to force a fixed exchange rate between the two metals, a price control that always leads to unwelcome and even disastrous results ("bimetallism").

of goods and services; why then do you make an exception for gold? Why do you call for the government fixing the price of gold and setting the exchange rates between the various currencies?"

The answer to this common complaint is that the question assumes the dollar to be an independent entity, a thing or commodity which should be allowed to fluctuate freely in relation to gold. But the rebuttal of the pro-gold forces points out that the dollar is *not* an independent entity, that it was originally simply a name for a certain weight of gold; the dollar, as well as the other currencies, is a unit of weight. But in that case, the pound, franc, dollar, and so on, are not exchanging as independent entities; they, too, are simply relative weights of gold. If 1/4 ounce of gold exchanges for 1/20 ounce of gold, how *else* would we expect them to trade than at 1:5?[4]

If the monetary unit is simply a unit of weight, then government's role in the area of money could well be confined to a simple Bureau of Weights and Measures, certifying this as well as other units of weight, length, or mass.[5] The problem is that governments have systematically betrayed their trust as guardians of the precisely defined weight of the money commodity.

If government sets itself up as the guardian of the international meter or the standard yard or pound, there is no economic incentive for it to betray its trust and change the definition. For the Bureau of Standards to announce suddenly that 1 pound is

---

[4]In older periods, foreign coins of gold and silver often circulated freely within a country, and there is, indeed, no economic reason why they should not do so. In the United States, as late as 1857, few bothered going to the U.S. Mint to obtain coins; the coins in general use were Spanish, English, and Austrian gold and silver pieces. Finally, Congress, perturbed at this slap to its sovereignty, outlawed the use of foreign coins within the U.S., forcing all foreign coinholders to go to the U.S. Mint and obtain American gold coins.

[5]Thus, Frederick Barnard's late nineteenth-century book on weights and measures has a discussion of coinage and the international monetary system in the appendix. Frederick A.P. Barnard, *The Metric System of Weights and Measures*, rev. ed. (New York: Columbia College, 1872).

now equal to 14 instead of 16 ounces would make no sense whatever. There is, however, all too much of an economic incentive for governments to change, especially to lighten, the definition of the currency unit; say, to change the definition of the pound sterling from 16 to 14 ounces of silver. This profitable process of the government's repeatedly lightening the number of ounces or grams in the same monetary unit is called *debasement*.

How debasement profits the State can be seen from a hypothetical case: Say the *rur*, the currency of the mythical kingdom of Ruritania, is worth 20 grams of gold. A new king now ascends the throne, and, being chronically short of money, decides to take the debasement route to the acquisition of wealth. He announces a mammoth call-in of all the old gold coins of the realm, each now dirty with wear and with the picture of the previous king stamped on its face. In return he will supply brand new coins with his face stamped on them, and will return the same number of *rurs* paid in. Someone presenting 100 *rurs* in old coins will receive 100 *rurs* in the new.

Seemingly a bargain! Except for a slight hitch: During the course of this recoinage, the king changes the definition of the *rur* from 20 to 16 grams. He then pockets the extra 20 percent of gold, minting the gold for his own use and pouring the coins into circulation for his own expenses. In short, the number of grams of gold in the society remains the same, but since people are now accustomed to use the *name* rather than the weight in their money accounts and prices, the number of *rurs* will have increased by 20 percent. The money supply in *rurs*, therefore, has gone up by 20 percent, and, as we shall see later on, this will drive up prices in the economy in terms of *rurs*. *Debasement*, then, is the arbitrary redefining and lightening of the currency so as to add to the coffers of the State.[6]

The pound sterling has diminished from 16 ounces of silver to its present fractional state because of repeated debasements, or

[6]This enormous charge for recoinage is called "seigniorage," payment to the seignieur or sovereign, the monopoly minter of coins.

changes in definition, by the kings of England. Similarly, rapid and extensive debasement was a striking feature of the Middle Ages, in almost every country in Europe. Thus, in 1200, the French *livre tournois* was defined as 98 grams of fine silver; by 1600 it equaled only 11 grams.

A particularly striking case is the *dinar*, the coin of the Saracens in Spain. The *dinar*, when first coined at the end of the seventh century, consisted of 65 gold grains. The Saracens, notably sound in monetary matters, kept the dinar's weight relatively constant, and as late as the middle of the twelfth century, it still equaled 60 grains. At that point, the Christian kings conquered Spain, and by the early thirteenth century, the dinar (now called *maravedi*) had been reduced to 14 grains of gold. Soon the gold coin was too lightweight to circulate, and it was converted into a silver coin weighing 26 grains of silver. But this, too, was debased further, and by the mid-fifteenth century, the maravedi consisted of only 1½ silver grains, and was again too small to circulate.[7]

Where is the total money supply—that crucial concept—in all this? First, before debasement, when the regional or national currency unit simply stands for a certain unit of weight of gold, the total money supply is the aggregate of all the monetary gold in existence in that society, that is, all the gold ready to be used in exchange. In practice, this means the total stock of gold coin and gold bullion available. Since all property and therefore all money is owned by *someone*, this means that the total money stock in the society at any given time is the aggregate, the sum total, of all existing *cash balances*, or money stock, owned by each individual or group. Thus, if there is a village of 10 people, A, B, C, etc., the

---

[7]See Elgin Groseclose, *Money and Man* (New York: Frederick Ungar, 1961), pp. 57–76. Many of the European debasements were made under the guise of adjusting the always-distorted fixed bimetallic ratios between gold and silver. See Luigi Einaudi, "The Theory of Imaginary Money from Charlemagne to the French Revolution," in F.C. Lane and J.C. Riemersma, eds., *Enterprise and Secular Change* (Homewood, Ill.: Irwin, 1953), pp. 229–61.

total money stock in the village will equal the sum of all cash balances held by each of the 10 citizens. If we wish to put this in mathematical terms, we can say that

$$M = \Sigma m$$

where M is the total stock or supply of money in any given area or in society as a whole, m is the individual stock or cash balance owned by each individual, and Σ means the sum or aggregate of each of the ms.

After debasement, since the money unit is the *name* (dinar) rather than the actual weight (specific number of gold grams), the number of dinars or pounds or maravedis will increase, and thus increase the supply of money. M will be the sum of the individual dinars held by each person, and will increase by the extent of the debasement. As we will see later, this increased money supply will tend to raise prices throughout the economy.

Shelling Out: The Origins of Money

Nick Szabo

Originally published in 2002

Abstract

The precursors of money, along with language, enabled early modern humans to solve problems of cooperation that other animals cannot – including problems of reciprocal altruism, kin altruism, and the mitigation of aggression. These precursors shared with non-fiat currencies very specific characteristics – they were not merely symbolic or decorative objects.

Table of Contents

Money

From the very start, England's 17th century colonies in America had a problem – a shortage of coins[D94][T01] The British idea was to grow large amounts of tobacco, cut timber for the ships of their global navy and merchant marine, and so forth, sending in return the supplies they felt were needed to keep the Americans working. In effect, early colonists were supposed to both work for the company and shop at the company store. The investors and the Crown much preferred this to paying in coin what the farmers might ask, letting the farmers themselves buy the supplies – and, heaven forbid, keep some of the profit as well.

The colonists' solution was at hand, but it took a few years for them to recognize it. The natives had money,

but it was very different from the money Europeans were used to. American Indians had been using money for millennia, and quite useful money it turned out to be for the newly arrived Europeans – despite the prejudice among some that only metal with the faces of their political leaders stamped on it constituted real money. Worse, the New England natives used neither silver nor gold. Instead, they used the most appropriate money to be found in their environment – durable skeleton parts of their prey. Specifically, they used wampum, shells of the clam *Venus mercenaria* and its relatives, strung onto pendants.

Necklace of wampum. During trade the beads were counted, removed, and re-assembled on new necklaces. Native American shell beads were also sometimes woven into belts or other mnemonic and ceremonial devices that demonstrated the wealth and commitment of a tribe to a treaty.

Clams were found only at the ocean, but wampum traded far inland. Sea-shell money of a variety of types could be found in tribes across the American continent. The Iriquois managed to collect the largest wampum treasure of any tribe, without venturing anywhere near the clam's habitat.[D94] Only a handful of tribes, such as the Narragansetts, specialized in manufacturing wampum, while hundreds of other tribes, many of them hunter-gatherers, used it. Wampum pendants came in a variety of lengths, with the number of beads proportional to the length. Pendants could be cut or joined to form a pendant of length equal to the price paid.

Once they got over their hangup about what constitutes real money, the colonists went wild trading for and with wampum. Clams entered the American vernacular as another way to say "money". The Dutch governor of New Amsterdam (now New York) took out a large loan from an English-American bank – in wampum. After a while the British authorities were forced to go along. So between 1637 and 1661, wampum became legal tender in New England. Colonists now had a liquid medium of exchange, and trade in the colonies flourished.[D94]

The beginning of the end of wampum came when the British started shipping more coin to the Americas, and Europeans started applying their mass-manufacturing techniques. By 1661, British authorities had thrown in the towel, and decided it would pay in coin of the realm – which being real gold and silver, and its minting audited and branded by the Crown, had even better monetary qualities than shells. In that year wampum ceased to be legal tender in New England. In 1710 briefly became legal tender in North Carolina. It continued

to be used as a medium of exchange, in some cases into the 20th century – but its value had been inflated one hundred fold by Western harvesting and manufacturing techniques, and it gradually went the route that gold and silver jewelry had gone in the West after the invention of coinage – from well-crafted money to decoration. The American language of shell money became a quaint holdover – "a hundred clams" became "a hundred dollars". "Shelling out" came to mean paying in coins or bills, and eventually by check or credit card.[D94] Little did we know that we had touched the very origins of our species.

Collectibles

Native American money took many forms besides shells. Furs, teeth, and a variety of other objects with properties we will discuss below were also commonly used as media of exchange. 12,000 years ago, in what is now Washington state, the Clovis people developed some marvelously long chert blades. The only problem – they break far too easily. They were useless for cutting. The flints were being made "for the sheer enjoyment" – or for some other purpose that had nothing to do with cutting.[G01] As we shall see, this seeming frivolity was, quite likely, actually very important to their survival.

Native Americans had not, however, been the first to make artful but useless blades, nor had they invented shell money. Nor, for that matter, had Europeans, even though they, too, in ages past had widely used shells and teeth for money – not to mention cattle, gold, silver, weapons, and much else. Asians had used all that and faux axes issued by governments to boot, but they as well imported this institution. For archaeologists have found pendants of shells dating to the early Paleolithic that could easily have substituted for Native American money.

Beads made from shells of the pea-sized snail Nassarius kraussianus, that lived in a nearby estuary. Blombos Cave, South Africa, 75,000 B.P.[B04]

In the late 1990s archaeologist Stanley Ambrose discovered, in a rock-shelter in the Rift Valley of Kenya, a cache of beads made of ostrich eggshell, blanks, and shell fragments. They are dated using the argon-argon (40Ar/39Ar) ratio to at least 40,000 years old[A98]. Pierced animal teeth have been found in Spain also dating to this time.[W95] Perforated shells have also been recovered from early Paleolithic sites in Lebanon[G95]. Recently regular shells, prepared as strung beads and dating further back still, to 75,000 BP, have been found in Blombos Cave in South Africa.[B04]

Ostrich-eggshell beads, Kenya Rift Valley, 40,000 B.P. (Courtesy Stanley Ambrose)

Our modern subspecies had migrated to Europe and necklaces of shell and tooth appear there, from 40,000 B.P. onward. Shell and tooth pendants appear in Australia from 30,000 B.P. onward[M93]. In all cases, the work is highly skilled, indicating a practice that probably dates much further back in time. The origin of

collecting and decorating is quite likely Africa, the original homeland of the anatomically modern subspecies. Collecting and making necklaces must have had an important selection benefit, since it was costly – manufacture of these shells took a great deal of both skill and time during an era when humans lived constantly on the brink of starvation[C94].

Practically all human cultures, even those that do not engage in substantial trade or that use more modern forms of money, make and enjoy jewelry, and value certain objects more for their artistic or heirloom qualities than for their utility. We humans collect necklaces of shells and other kinds of jewelry – for the sheer enjoyment of it. For the evolutionary psychologists an explanation that humans do something for "the sheer enjoyment of it" is not an explanation at all – but the posing of a problem. Why do so many people find the collection and wearing of jewelry enjoyable? For the evolutionary psychologist, this question becomes – what caused this pleasure to evolve?

Detail of necklace from a burial at Sungir, Russia, 28,000 BP. Interlocking and interchangeable beads. Each mammoth ivory bead may have required one to two hours of labor to manufacture.[W97]

Evolution, Cooperation, and Collectibles

Evolutionary psychology starts with a key mathematical discovery of John Maynard Smith[D89]. Using models of populations of co-evolving genes, from the well-developed area of population genetics, Smith posited genes that can code for strategies, good or bad, used in simple strategic problems (the "games" of game theory). Smith proved that these genes, competing to be propagated into future generations, will evolve strategies that are Nash equilibria to the strategic problems presented by the competition. These games include the prisoner's dilemma, a prototypical problem of cooperation, and hawk/dove, a prototypical problem of aggression and its mitigation.

Critical to Smith's theory is that these strategic games, while played out between phenotypes proximately, are in fact games between genes at the ultimate level – the level of competition to be propagated. The genes – not necessarily the individuals – influence behavior as if they were boundedly rational (coding for strategies as optimal as possible, within the limits of what phenotypes can express given the biological raw materials and previous evolutionary history) and "selfish" (to use Richard Dawkins' metaphor). Genetic influences on behavior are adaptations to the social problems presented by genes competing through their phenotypes. Smith called these evolved Nash equilibria evolutionary stable strategies.

The "epicycles" built on top of the earlier individual selection theory, such as sexual selection and kin selection, disappear into this more general model which, in a Copernican manner, puts the genes rather than individuals at the center of the theory. Thus Dawkins' metaphorical and often misunderstood phrase, "selfish gene", to describe Smith's theory.

Few other species cooperate on the order of even Paleolithic humans. In some cases – brood care, the

colonies of ants, termites, and bees, and so forth, animals cooperate because they are kin – because they can help copies of their "selfish genes" found in their kin. In some highly constrained cases, there is also ongoing cooperation between non-kin, which evolutionary psychologists call reciprocal altruism. As Dawkins describes it[D89], unless an exchange of favors is simultaneous (and sometimes even then), one party or the other can cheat. And they usually do. This is the typical result of a game theorists call the Prisoner's Dilemma – if both parties cooperated, both would be better off, but if one cheats, he gains at the expense of the sucker. In a population of cheaters and suckers, the cheaters always win. However, sometimes animals come to cooperate through repeated interactions and a strategy called Tit-for-Tat: start cooperating and keep cooperating until the other party cheats – then defect yourself. This threat of retaliation motivates continued cooperation.

The situations where such cooperation in fact occurs in the animal world are highly constrained. The main constraint is that such cooperation is restricted to relationships where at least one of the participants is more or less forced to be in the proximity of the other. The most common case is when parasites, and hosts whose bodies they share, evolve into symbiotes. If the interests of the parasite and the host coincide, so that both working together would be more fit than either on their own, (i.e. the parasite is also providing some benefit to the host), then, if they can play a successful game of Tit-for-Tat, they will evolve into symbiosis – a state where their interests, and especially the exit mechanism of genes from one generation to the next, coincides. They become as a single organism. However, there is much more than cooperation going on here – there is also exploitation. They occur simultaneously. The situation is analogous to an institution humans would develop – tribute – which we will analyze below.

Some very special instances occur that do not involve parasite and host sharing the same body and evolving into symbiotes. Rather, they involve non-kin animals and highly constrained territory. A prominent example Dawkins describes is cleaner fish. These fish swim in and out of the mouths of their hosts, eating the bacteria there, benefiting the host fish. The host fish could cheat – it could wait for the cleaner to finish its job, then eat it. But they don't. Since they are both mobile, they are both potentially free to leave the relationship. However, the cleaner fish have evolved a very strong sense of individual territoriality, and have stripes and dances that are difficult to spoof – much like a difficult to forge brand logo. So the host fish know where to go to get cleaned – and they know that if they cheat, they will have to start over again with a new distrustful cleaner fish. The entrance costs, and thus the exit costs, of the relationship are high, so that it works out without cheating. Besides, the cleaner fish are tiny, so the benefit of eating them is not large compared to the benefit of a small number of, or even one, cleaning.

One of the most pertinent examples.is the vampire bat. As their name suggests, they suck the blood of prey mammals. The interesting thing is that, on a good night, they bring back a surplus; on a bad night, nothing. Their dark business is highly unpredictable. As a result, the lucky (or skilled) bats often share blood with the less lucky (or skilled) bats in their cave. They vomit up the blood and the grateful recipient eats it.

The vast majority of these recipients are kin. Out of 110 such regurgitations witnessed by the strong-stomached biologist G.S. Wilkinson, 77 were cases of mothers feeding their children, and most of the other cases also involved genetic kin. There were, however, a small number that could not be explained by kin altruism. To demonstrate these were cases of reciprocal altruism, Wilkinson combined the populations of bats from two different groups. Bats, with very rare exceptions, only fed old friends from their original group.[D89] Such cooperation requires building a long-term relationship, where partners interact often, recognize each other, and keep track of each other's behavior. The bat cave helps constrain the bats into long-term relationships where such bonds can form.

We will see that some humans, too, chose highly risky and discontinuous prey items, and shared the resulting surpluses with non-kin. Indeed, they accomplished this to a far greater extent than the vampire bat. How they did so is the main subject of our essay. Dawkins suggests, "money is a formal token of delayed reciprocal

altruism", but then pursues this fascinating idea no further. We will.

Among small human groups, public reputation can supersede retaliation by a single individual to motivate cooperation in delayed reciprocation. However, reputational beliefs can suffer from two major kinds of errors – errors of about which person did what, and errors in appraising the value or damages caused by that act.

The need to remember faces and favors is a major cognitive hurdle, but one that most humans find relatively easy to overcome. Recognizing faces is easy, but remembering that a favor took place when such memory needs to be recalled can be harder. Remembering the specifics about a favor that gave it a certain value to the favored is harder still. Avoiding disputes and misunderstandings can be improbable or prohibitively difficult.

The appraisal or value measurement problem is very broad. For humans it comes into play in any system of exchange – reciprocation of favors, barter, money, credit, employment, or purchase in a market. It is important in extortion, taxation, tribute, and the setting of judicial penalties. It is even important in reciprocal altruism in animals. Consider monkeys exchanging favors – say pieces of fruit for back scratches. Mutual grooming can remove ticks and fleas that an individual can't see or reach. But just how much grooming versus how many pieces of fruit constitutes a reciprocation that both sides will consider to be "fair", or in other words not a defection? Is twenty minutes of backscratching worth one piece of fruit or two? And how big a piece?

Even the simple case of trading blood for blood is more complicated than it seems. Just how do the bats estimate the value of blood they have received? Do they estimate the value of a favor by weight, by bulk, by taste, by its ability to satiate hunger, or other variables? Just the same, measurement complications arise even in the simple monkey exchange of "you scratch my back and I'll scratch yours".

For the vast majority of potential exchanges, the measurement problem is intractable for animals. Even more than the easier problem of remembering faces and matching them to favors, the ability of both parties to agree with sufficient accuracy on an estimate of the value of a favor in the first place is probably the main barrier to reciprocal altruism among animals.

Just the stone tool-kit of even early Paleolithic man that has survived for us to find was in some ways too complicated for brains of our size. Keeping track of favors involving them – who manufactured what quality of tool for whom, and therefore who owed whom what, and so on – would have been too difficult outside the boundaries of the clan. Add onto that, quite likely, a large variety of organic objects, ephemeral services (such as grooming), and so on that have not survived. After even a small fraction of these goods had been transferred and services performed our brains, as inflated as they are, could not possibly keep track of who owed what to whom. Today we often write these things down – but Paleolithic man had no writing. If cooperation occurred between clans and even tribes, as the archaeological record indicates in fact occurred, the problem gets far worse still, since hunter-gatherer tribes were usually highly antagonistic and mutually distrustful.

If clams can be money, furs can be money, gold can be money, and so on – if money is not just coins or notes issued by a government under legal tender laws, but rather can be a wide variety of objects – then just what is money anyway? And why did humans, often living on the brink of starvation, spend so much time making and enjoying those necklaces when they could have been doing more hunting and gathering? Nineteenth-century economist Carl Menger[M1892] first described how money evolves naturally and inevitably from a sufficient volume of commodity barter. In modern economic terms the story is similar to Menger's.

Barter requires a coincidence of interests. Alice grows some pecans and wants some apples; Bob grows apples and want some pecans. They just happen to have their orchards near each other, and Alice just happens to trust Bob enough to wait between pecan harvest time and apple harvest time. Assuming all these conditions are met, barter works pretty well. But if Alice was growing oranges, even if Bob wanted oranges

as well as pecans, they'd be out of luck – oranges and apples don't both grow well in the same climate. If Alice and Bob didn't trust each other, and couldn't find a third party to be a middleman[L94] or enforce a contract, they'd also be out of luck.

Further complications could arise. Alice and Bob can't fully articulate a promise to sell pecans or apples in the future, because, among other possibilities, Alice could keep the best pecans to herself (and Bob the best apples), giving the other the dregs. Comparing the qualities as well as the quantities of two different kinds of goods is all the more difficult when the state of one of the goods is only a memory. Furthermore, neither can anticipate events such as a bad harvest. These complications greatly add to the problem of Alice and Bob deciding whether separated reciprocal altruism has truly been reciprocal. These kinds of complications increase the greater the time interval and uncertainty between the original transaction and the reciprocation.

A related problem is that, as engineers would say, barter "doesn't scale". Barter works well at small volumes but becomes increasingly costly at large volumes, until it becomes too costly to be worth the effort. If there are n goods and services to be traded, a barter market requires **n^2** prices. Five products would require twenty-five prices, which is not too bad, but 500 products would require 250,000 prices, which is far beyond what is practical for one person to keep track of. With money, there are only **n** prices – 500 products, 500 prices. Money for this purpose can work either as a medium of exchange or simply as a standard of value – as long as the number of money prices themselves do not grow too large to memorize or change too often. (The latter problem, along with an implicit insurance "contract", along with the lack of a competitive market may explain why prices were often set by long-evolved custom rather than proximate negotiation).

Barter requires, in other words, coincidences of supply or skills, preferences, time, and low transaction costs. Its cost increases far faster than the growth in the number of goods traded. Barter certainly works much better than no trade at all, and has been widely practiced. But it is quite limited compared to trade with money.

Primitive money existed long before large scale trade networks. Money had an even earlier and more important use. Money greatly improved the workings of even small barter networks by greatly reducing the need for credit. Simultaneous coincidence of preference was far rarer than coincidences across long spans of time. With money Alice could gather for Bob during the ripening of the blueberries this month, and Bob hunt for Alice during the migration of the mammoth herds six months later, without either having to keep track of who owed who, or trust the other's memory or honesty. A mother's much greater investment in child-rearing could be secured by gifts of unforgeable valuables. Money converts the division of labor problem from a prisoner's dilemma into a simple swap.

The proto-money used by many hunter-gatherer tribes looks very different from modern money, now serves a different role in our modern culture, and had a function probably limited to small trade networks and other local institutions discussed below. I will thus call such money collectibles instead of money proper. The terms used in the anthropological literature for such objects are usually either "money", defined more broadly than just government printed notes and coins but more narrowly than we will use "collectible" in this essay, or the vague "valuable", which sometimes refers to items that are not collectibles in the sense of this essay. Reasons for choosing the term collectible over other possible names for proto-money will become apparent. Collectibles had very specific attributes. They were not merely symbolic. While the concrete objects and attributes valued as collectible could vary between cultures, they were far from arbitrary. The primary and ultimate evolutionary function of collectibles was as a medium for storing and transferring wealth. Some kinds of collectibles, such as wampum, could be quite functional as money as we moderns know it, where the economic and social conditions encouraged trade. I will occasionally use the terms "proto-money" and "primitive money" interchangeably with "collectible" when discussing pre-coinage media of wealth transfer.

Gains From Wealth Transfers

People, clans or tribes trade voluntarily because both sides believe they gain something. Their beliefs about the value may change after the trade, for example as they gain experience with the good or service. Their beliefs at the time of the trade, although to some degree inaccurate as to the value, are still usually correct as to the existence of gain. Especially in early intertribal trade, restricted to high-value items, there was strong incentive for each party to get their beliefs right. Thus trade almost always did benefit both parties. Trade created value as much as the physical act of making something.

Because individuals, clans, and tribes all vary in their preferences, vary in their ability to satisfy these preferences, and vary in the beliefs they have about these skills and preferences and the objects that are consequent of them, there are always gains to be made from trade. Whether the costs of making these trades – transaction costs – are low enough to make the trades worthwhile is another matter. In our civilization, far more trades are possible than were through most of human history. Nevertheless, as we shall see some kinds of trades were worth more than the transaction costs, for some cultures, probably back to the beginning of *homo sapiens sapiens*.

Voluntary spot trades are not the only kinds of transactions that benefit from lower transaction costs. This is the key to understanding the origin and evolution of money. Family heirlooms could be used as collateral to remove the credit risk from delayed exchanges. The ability of a victorious tribe to extract tribute from the vanquished was of great benefit to the victor. The victor's ability to collect tribute benefited from some of the same kinds of transaction cost techniques as did trade. So did the plaintiff in assessment of damages for offenses against custom or law, and kin groups arranging a marriage. Kin also benefited from timely and peaceful gifts of wealth by inheritance. The major human life events that modern cultures segregate from the world of trade benefited no less than trade, and sometimes more so, from techniques that lowered transaction costs. None of these techniques was more effective, important, or early than primitive money – collectibles.

When *H. sapiens sapiens* displaced *H. sapiens neanderthalensis*, population explosions followed. Evidence from the takeover in Europe, c. 40,000 to 35,000 B.P, indicates that *H. sapiens sapiens* increased the carrying capacity of its environment by a factor of ten over *H. sapiens neanderthalensis* – i.e., the population density increased tenfold[C94]. Not only that, the newcomers had spare time to create the world's first art – such as the wonderful cave paintings, a wide variety of well-crafted figurines – and of course the wonderful pendants and necklaces of seashells, teeth, and eggshell.

These objects were not useless decorations. Newly effective wealth transfers, made possible by collectibles as well as another probable advance of the era, language, created new cultural institutions that quite likely played the leading role in the increase of carrying capacity.

The newcomers, *H. sapiens sapiens*, had the same size brain, weaker bones, and smaller muscles than the Neanderthals. Their hunting tools were more sophisticated, but in 35,000 B.P. they were basically the same tools – they were probably not even twice as effective, much less ten times more effective. The biggest difference may have been wealth transfers made more effective or even possible by collectibles. *H. sapiens sapiens* took pleasure from collecting shells, making jewelry out of them, showing them off, and trading them. *H. sapiens neanderthalensis* did not. The same dynamic would have been at work, tens of thousands of years earlier, on the Serengeti, when *H. sapiens sapiens* first appeared in that dynamic maelstrom of human evolution, Africa.

We shall describe how collectibles lowered transaction costs in each kind of wealth transfer – in the voluntary free gift of inheritance, in voluntary mutual trade or marriage, and in the involuntary transfers of legal judgments and tribute.

All these kinds of value transfer occurred in many cultures of human prehistory, probably from the beginning of *Homo sapiens sapiens*. The gains to be made, by one or both parties, from these major life event transfers of wealth, were so great that they occurred despite high transaction costs. Compared to modern money,

primitive money had a very low velocity – it might be transferred only a handful of times in an average individual's lifetime. Nevertheless, a durable collectible, what today we would call an heirloom, could persist for many generations and added substantial value at each transfer – often making the transfer even possible at all. Tribes therefore often spent large amounts of time on the seemingly frivolous tasks of manufacturing and exploring for the raw materials of jewelry and other collectibles.

The Kula Ring

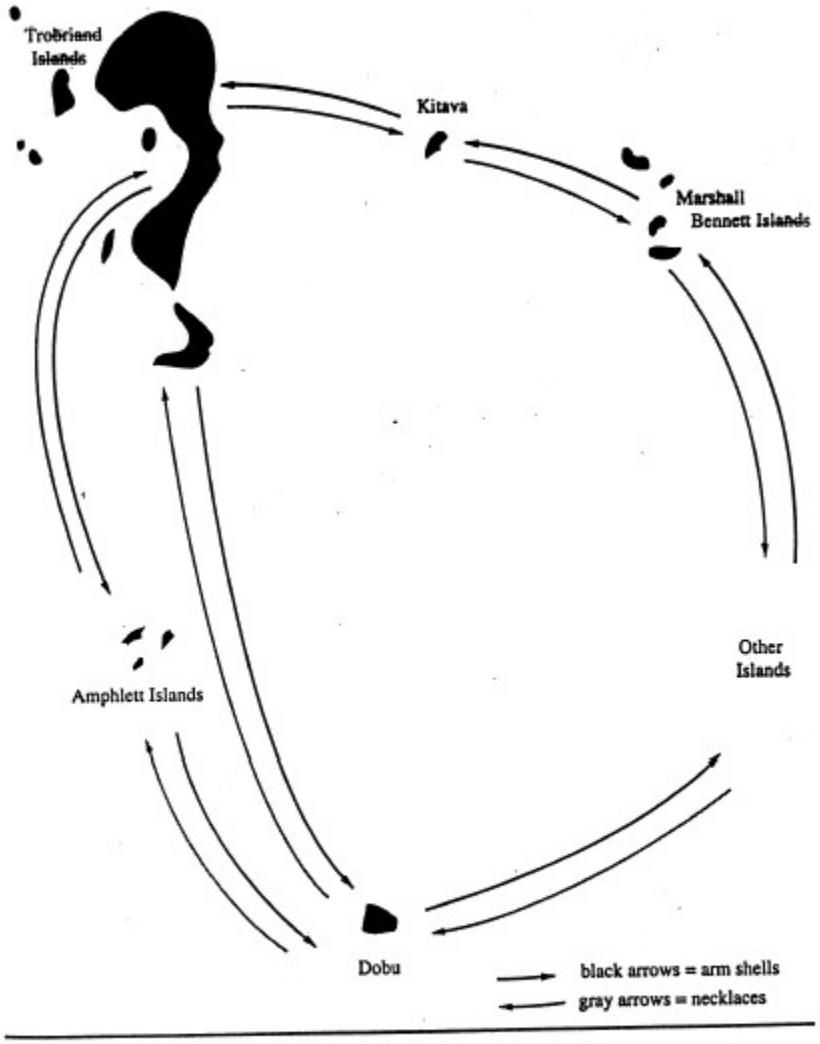
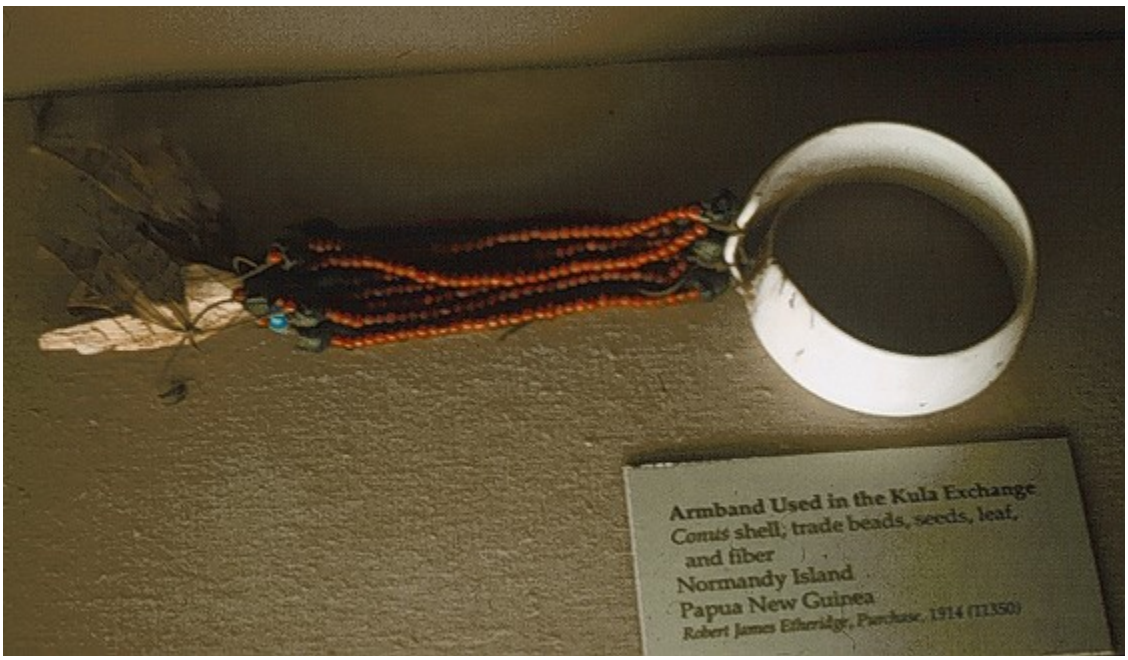


Figure 5. The *kula* ring.

The Kula trading network of pre-colonial Melanesia. The kula valuables doubled as "high power" money and mnemonic for stories and gossip. Many of the goods traded, mostly agricultural products, were available in different seasons, and so could not be traded in kind. Kula collectibles solved this double-coincidence problem as an unforgeably costly, wearable (for security), and circulated (literally!) money. Necklaces circulated clockwise, and armshells counter-clockwise, in a very regular pattern. By solving the double-coincidence problem an armshell or necklace would prove more valuable than its cost after only a few trades,

but could circulate for decades. Gossip and stories about prior owners of the collectibles further provided information about upstream credit and liquidity. In other Neolithic cultures collectibles, usually shells, circulated in a less regular pattern but had similar purposes and attributes.[L94]



Kula armshell (mwali).

Kula necklaces (bagi).

For any institution in which wealth transfer is an important component, we will ask the following questions:

1. What coincidence in time between the event, the supply for the transferred good, and demand for the transferred good was necessary? How unlikely or how high a barrier to the wealth transfer did the improbability of coincidence represent?

2. Would the wealth transfers formed a closed loop of collectibles just based on that institution, or were

> other wealth transfer institutions necessary to complete circulation cycles? Taking the actual flow graph of monetary circulation seriously is critical to understanding the emergence of money. General circulation among a wide variety of trades did not and would not exist for most of human prehistory. Without completed and repeated loops collectibles would not circulate and would become worthless. A collectible, to be worth making, had to add value in enough transactions to amortize its cost.

We shall first examine the kind of transfer most familiar and economically important to us today – trade.

Starvation Insurance

Bruce Winterhalder[W98] surveys models of how and why food is sometimes transferred between animals: tolerated theft, producing/scrounging/opportunism, risk-sensitive subsistence, by-product mutualism, delayed reciprocity, trade/exchange not in kind, and other selection models (including kin altruism). Here we focus on risk-sensitive subsistence, delayed reciprocity, and trade (exchange not in kind). We argue that substituting trade of food for collectibles for delayed reciprocity can increase food sharing. It does so by mitigating the risks of a variable food supply while avoiding the largely insurmountable problems of delayed reciprocity between bands. We will deal with kin altruism and theft (tolerated or not) in broader contexts below.

Food is worth far more to starving people than to well-fed ones. If the starving man can save his life by trading his most precious valuables, it may be worth to him months or even years of the labor it might take to replace that value. He will usually consider his life worth more than the sentimental value of the family heirlooms. Like fat itself, collectibles can provide insurance against food shortages. Starvation from local shortages could be staved off with at least two different kinds of trades – for the food itself, or for foraging or hunting rights.

Nevertheless, the transaction costs were usually too high – bands were far more likely to fight than ever trust each other. The hungry band that couldn't find its own food usually starved. However, if the transaction costs could be lowered, by lowering the need for trust between bands, food that was worth a day's labor to one band might be worth several months' labor to the starving band.

Local but extremely valuable trade was, this essay argues, made possible among many cultures by the advent of collectibles, by the time of the Upper Paleolithic. Collectibles substituted for otherwise necessary but non-existent trusting long term relationships. If there had existed a high degree of sustained interaction and trust between tribes, or individuals of different tribes, so that they gave each other unsecured credit, this would have stimulated time-lagged barter trade. However, such a high degree of trust then is highly implausible – for the reasons stated above regarding reciprocal altruism, confirmed by the empirical evidence that most hunter-gatherer tribal relations have been observed to be quite antagonistic. Hunter-gatherer bands usually broke up into small bands for most of the year and gathered into "aggregates", something like medieval European fairs, for a few weeks out of the year. Despite the lack of trust between bands, an important trade in staples, of the kind illustrated in the accompanying figure, almost surely occurred in Europe and probably elsewhere, such as with the big game hunters of America and Africa.

The scenario illustrated by the accompanying figure is hypothetical, but it would be very surprising if it did not occur. While many Europeans even in the Paleolithic enjoyed wearing shell necklaces, many lived farther inland and made necklaces instead out of the teeth of their prey. Flints, axes, furs, and other collectibles were also quite likely used as media of exchange.

Reindeer, bison, and other human prey migrated at different times of the year. Different tribes specialized in different prey, to the point where over 90%, and sometimes as much as 99%, of the remains from many sites during the Paleolithic in Europe come from a single species[C94]. This indicates at least seasonal specialization and perhaps full-time specialization by a tribe in a single species. To the extent they

specialized, the members of a single tribe would have become experts at the behavior, migration habits, and other patterns surrounding their specific prey species, as well as the specialized tools and techniques for hunting them. Some tribes observed in recent times are known to have specialized. Some North American Indian tribes specialized respectively in hunting bison, antelope, and fishing for salmon. In northern Russia and parts of Finland, many tribes, including the Lapp even today, specialized in herding a single species of reindeer.

Such specialization was probably far higher when more large prey (horse, auroch, giant elk, bison, giant sloth, mastodon, mammoth, zebra, elephant, hippopotamus, giraffe, musk oxen, etc.) roamed North America, Europe, and Africa in large herds during the Paleolithic. Large wild animals unafraid of humans no longer exist. During the Paleolithic they were either driven extinct or adapted to be afraid of humans and our projectiles. However, for most of the time span of *H. sapiens sapiens* these herds were abundant and easy pickings to specialist hunters. According to our theory of trade-based predation, specialization was quite likely far higher when large prey roamed North America, Europe, and Africa in large herds during the Paleolithic. Trade-based division of labor in hunting between tribes is consistent with (although not securely confirmed by) the archaeological evidence from the Paleolithic in Europe.

These migrating bands, following their herds, frequently interacted, creating many opportunities for trade. American Indians preserved food by drying, making pemmican, and so on in ways that lasted for a few months but typically not a full year. Such food was commonly traded, along with skins, weapons, and collectibles. Often these trades occurred during annual trading expeditions[T01].

Large herd animals migrated through a territory only two times a year, with a window most often of one or two months. Without any other source of protein besides their own prey species, these specialist tribes would have starved. The very high degree of specialization demonstrated in the archaeological record could only have occurred if there was trade.

Thus, even if the time-offset barter of meat were the only kind of trade, this is quite sufficient to make the use of collectibles quite worthwhile. The necklaces, flints, and any other objects used as money circulate in a closed loop, back and forth, in roughly equal amounts so long as the value of meat traded remains roughly equal. Note that it is not enough, for the theory of collectibles put forth in this paper to be correct, that single beneficial trades were possible. We must identify closed loops of mutually beneficial trades. With closed loops the collectibles continue to circulate, amortizing their costs.

As mentioned, we know from archaeological remains that many tribes specialized in a single large prey species. This specialization was at least seasonal; if there was extensive trade it could have been full-time. Becoming experts in the habits and migration patterns, and best methods of taking down, a tribe reaped enormous productive benefits. These benefits, however, would normally be unattainable, for specializing in a single species meant going without food most of the year. Division of labor between tribes paid off – and trade made it possible. The supply of food would nearly double from trade just between two complementary tribes. There were, however, rather than two prey species, often up to a dozen that migrated through most hunting territories in areas like the Serengeti and the European steppe. The amount of meat available to a species-specializing tribe would thus likely more than double with such trade among a handful of neighboring tribes. On top of this, the extra meat would be there when needed most – when the meat from a tribe's own species prey would already have been eaten, and without food the hunters would starve.

Thus there were at least four gains, or sources of surplus, from a trade cycle as simple as two prey species and two non-simultaneous but offsetting trades. These gains are distinct but not necessarily independent:

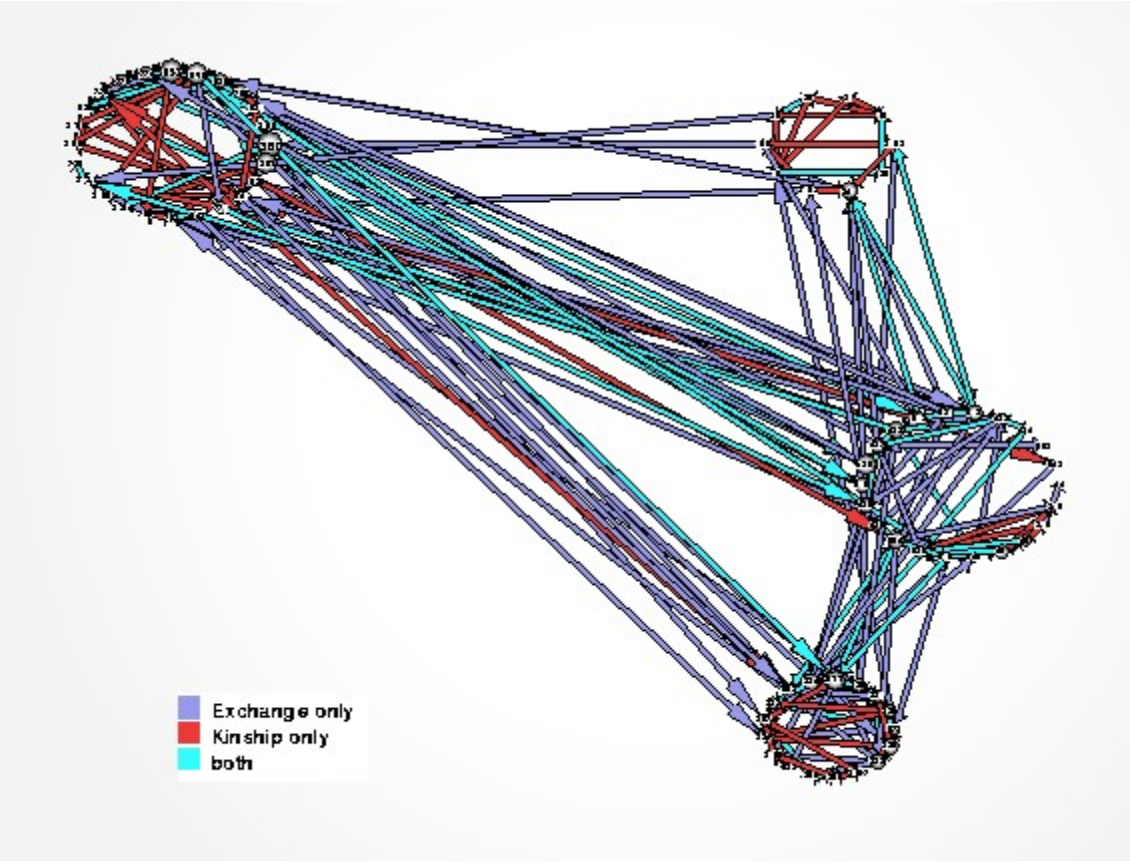
1. An available source of meat at a time of the year when one would otherwise starve.
2. An increase in the total supply of meat – they traded the surplus beyond what they could eat immediately or store; what they didn't trade would have gone to waste.

3. An increase in the variety of nutrition from meat, by eating different kinds of meat.
4. Increased productivity from specialization in a single prey species.

Making or saving collectibles to trade for food itself was not the only way to insure against bad times. Perhaps even more common, especially where large prey items were not available, was territoriality combined with trade in foraging rights. This can be observed even in some of the remnants of hunter-gatherer culture that exist today.

The !Kung San of southern Africa, like all other modern remnants of hunter-gatherer cultures, live on marginal lands. They have no opportunity to be specialists but must take the meager remnants available. They may thus be rather uncharacteristic of many ancient hunter-gatherer cultures, and uncharacteristic of the original *Homo sapien sapiens*, which first seized the lushest lands and best game routes from *Homo sapiens neanderthalensis* and only much later drove the Neanderthals from marginal lands. Yet despite their severe ecological handicap, the !Kung use collectibles as items of trade.

Like most hunter-gatherers, the !Kung spend most of the year in small, dispersed bands and a few weeks of the year in an aggregate with several other bands. Aggregation is like a fair with added features – trade is accomplished, alliances are cemented, partnerships strengthened, and marriages transacted. Preparation for aggregation is filled with the manufacture of tradeable items, partly utilitarian but mostly of a collectible nature. The exchange system, called by the !Kung *hxaro*, involves a substantial trade in beaded jewelry, including ostrich-shell pendants quite similar to those found in Africa 40,000 years ago.



Pattern of hxaro exchanges and kinship relations among neighboring tribes of !Khung San hunter-gatherers.

Necklaces used in the hxaro exchange.

One of the main things the !Kung buy and sell with their collectibles are abstract rights to enter another band's territory and hunt or gather food there. Trade in these rights is especially brisk during local shortages which can be alleviated by foraging in a neighbor's territory[W77][W82] !Kung bands mark their territories with arrows; trespassing without having purchased the right to enter and forage is tantamount to a declaration of war. Like the inter-band food trade discussed above, the use of collectibles to purchase foraging rights constitutes an "insurance policy against starvation", to use the phrase of Stanley Ambrose[A98].

Although anatomically modern humans surely had conscious thought, language, and some ability to plan, it would have required little conscious thought or language, and very little planning, to generate trades. It was not necessary that tribe members reasoned out the benefits of anything but a single trade. To create this institution it would have sufficed that people follow their instincts to make obtain collectibles with the characteristics outlined below. (as indicated by proxy observations that make approximate estimations for these characteristics). This is to various extents true of the other institutions we will study – they evolved, rather than being consciously designed. No one participating in the institution's rituals would have explained their function in terms of ultimate evolutionary function; rather they explained in terms of a wide variety of mythologies that served more as proximate motivators of behavior than as theories of ultimate purpose or origin.

Direct evidence for trade in food has long since decayed. We may, in the future, find more direct evidence than is now available for this article, via comparison of hunting remains in one tribe with the consumption patterns in another tribe – the hardest part of this task likely being to identify the boundaries of different tribes or kin groups. According to our theory, such transfer of meat from one tribe to another was common in many parts of the world during the Paleolithic where large-scale and specialized big game hunting occurred.

For now, we do have extensive indirect evidence of trade, via the movement of the collectibles themselves. Fortunately there is a good correlation between the durability desired for collectibles and the conditions under which an artifact has survived to be found by today's archaeologists. In the early Paleolithic, when all human movement was on foot, we have instances of perforated sea shells found up to 500 kilometers away from the nearest source[C94]. There was a similar long-distance movement of flint.

Unfortunately, trade was severely restricted by high transaction costs in most times and places. The primary barrier was the antagonism between tribes. The predominant relationship between tribes was one of distrust on good days and outright violence on bad days. Only ties of marriage or kinship could bring tribes into a relationship with trust, albeit only occasionally and of limited scope. The poor ability to protect property, even collectibles worn on the person or buried in well-hidden caches, meant that collectibles had to amortize

their costs in a few transactions.

Trade was thus not the only kind of wealth transfer, and probably not the most important kind during the long human prehistory where high transaction costs prevented the development of the kinds of markets, firms, and other economic institutions we now take for granted[L94]. Underneath our great economic institutions are far more ancient institutions that also involved wealth transfer – in prehistoric times, the main kinds of wealth transfer. All of these institutions distinguished *Homo sapiens sapiens* from previous animals. We now turn to one of the most basic kinds of wealth transfer that we humans take for granted but other animals do not have – passing wealth onto the next generation.

Kin Altruism Beyond the Grave

Coincidence in time and locale of supply and demand for trade was rare – so much so, that most kinds of trades and trade-based economic institutions we now take for granted could not exist. Even more unlikely was the triple coincidence of supply with demand with a major event for a kin group – the formation of a new family, death, crime, or victory or defeat in war. As we shall see, clans, and individuals greatly benefited from a timely transfer of wealth during these events. Such wealth transfer in turn was much less wasteful when it was the transfer of a store of wealth more durable and general than consumables or tools designed for other purposes. The demand for a durable and general store of wealth for use in these institutions was thus even more urgent than for trade itself. Furthermore, the institutions of marriage, inheritance, dispute resolution, and tribute may predate intertribal trade, and involved for most tribes a greater transfer of wealth than trade. These institutions thus more than trade served as the motivator and incubator of the earliest primitive money.

In most hunter-gatherer tribes this wealth came in a form that strikes us preposterously wealthy moderns as trivial – a collection of wooden utensils, flint and bone tools and weapons, shells on strings, perhaps a hut and in colder climates some mangy furs. Sometimes it could all be carried on the person. Nevertheless, these motley assortments were wealth for a hunter-gatherer no less than real estate, stocks, and bonds are wealth for us. To the hunter-gatherer tools and sometimes warm clothes were necessary for survival. Many of the items were highly valued collectibles that insured against starvation, purchased mates, and could substitute for massacre or starvation in event of war and defeat. The ability to transfer the capital of survival to one's descendants was another advantage *Homo sapiens sapiens* had over previous animals. Furthermore, the skilled tribesman or clan could accumulate a surplus of wealth from the occasional, but cumulative over a lifetime, trade of surplus consumables for durable wealth, especially collectibles. A temporary fitness advantage could be translated into a more durable fitness advantage for one's descendants.

Another form of wealth, hidden from the archaeologist, were titles to offices. Such social positions were more valuable than the tangible forms of wealth in many hunter-gatherer cultures. Examples of such positions included clan leaders, war party leaders, hunting party leaders, membership in a particular long-term trading partnership (with a particular person in a neighboring clan or tribe), midwives, and religious healers. Often collectibles not only embodied wealth, but also served as a mnemonic, representing the title to a clan position of responsibility and privilege. Upon death, to maintain order, the heirs to such positions had to be quickly and clearly determined. Delays could spawn vicious conflicts. Thus, a common event was the mortuary feast, in which the deceased was feted while both his tangible and intangible forms of wealth were distributed to descendants, as determined by custom, clan decision-makers, or the will of the deceased.

Other kinds of free gifts were quite rare in pre-modern cultures, as Marcel Mauss[M50] and other anthropologists have pointed out. Seemingly free gifts in fact implicitly invoked an obligation in the recipient. Before contract law, this implicit obligation of the "gift", along with community dishonor and punishments ensuing if the implicit obligation was not met, was perhaps the most common motivator of reciprocation in delayed exchange, and is still common in the variety of informal favors we do for each other. Inheritance and other forms of kin altruism were the only widely practiced forms of what we moderns would

call gift proper, namely a gift that imposed no obligation on the recipient.

Early Western traders and missionaries, who often saw natives as childish primitives, sometimes called their tribute payments "gifts" and trades "gift exchanges", as if they bore more resemblance to the Christmas and birthday present exchanges of Western children than to the contractual and tax obligations of adults. Partly this may have reflected prejudice, and partly the fact that in the West by that time obligations were usually formalized in writing which the natives lacked. Westerners thus usually translated the rich variety of words natives had for their exchange institutions, rights, and obligations as "gift". Seventeenth-century French settlers in America were thinly scattered among much larger populations of Indian tribes, and often found themselves paying tribute to these tribes. Calling these payments "gifts" was a way for them to save face with other Europeans who faced no such necessity and found it cowardly.

Mauss and modern anthropologists have unfortunately kept this terminology. The uncivilized human is still like a child, but now innocent like a child, a creature of moral superiority who would not stoop to our kind of base, cold-blooded economic transactions. However in the West, especially in the official terminology used for our laws covering transactions, a "gift" refers to a transfer that imposes no obligation. When coming across anthropological discussions of "gift exchange" these caveats should be kept in mind – modern anthropologists are not at all referring to the free or informal gifts we commonly refer to in our modern use on the term "gift". They are rather referring to any of a wide variety of often quite sophisticated systems of rights and obligations involved in wealth transfers. The only major transactions in prehistoric cultures similar to our modern gift, in that it was neither itself a widely recognized obligation nor imposed any obligation on the recipient, were parents or maternal kin caring for their children and inheritance. (An exception was that inheriting title to a position imposed the responsibilities of the position on the heir as well as its privileges).

Inheritance of some heirlooms might proceed for several generations uninterrupted, but it did not by itself form a closed loop of collectibles transfers. Heirlooms were only valuable if they eventually got used for something else. They often were used in marriage transactions between clans that could form closed loop cycles of collectibles.

The Family Trade

An early and important example of a small closed loop trade network made possible by collectibles involves the much higher investment humans make in raising offspring than our primate relatives, and the related human institution of marriage. Combining arrangements of long-term matches for mating and child-raising, negotiated between clans, with wealth transfer, marriage is a human universal and probably dates back to the first *Homo sapiens sapiens*.

Parental investment is a long-term and almost one-shot affair – there is no time for repeated interactions. Divorce from a negligent father or unfaithful wife usually represented several years of time wasted, in genetic fitness terms, by the jilted party. Fidelity and commitment to the children were primarily enforced by in-laws – the clan. The marriage was the contract between clans that usually included such promises of fidelity and commitment as well as wealth transfer.

The contributions a man and a woman will bring to a marriage are seldom equal. This was even more true in an era when mate choice was largely determined by clans and the population from which clan leaders could choose was quite small. Most commonly, the woman was considered more valuable and the groom's clan paid a bride price to the bride's clan. Quite rare in comparison was dowry, a payment by the bride's clan to the new couple. Mostly this was practiced by upper classes of monogamous but highly unequal societies in medieval Europe and India, and was ultimately motivated by the far greater reproductive potential of upper-class sons than upper-class daughters in those societies. Since literature was mostly written about upper classes, dowry often plays a role in European traditional stories. This does not reflect its actual frequency

across human cultures – it was quite rare.

Marriages between clans could form a closed cycle of collectibles. Indeed, two clans exchanging partners would be sufficient to maintain a closed loop, as long as brides tended to alternate. If one clan was wealthier in collectibles from some other kind of transfer, it could marry more of its sons to better brides (in monogamous societies) or a greater number of brides (in polygamous societies). In a loop involving only marriages, primitive money would simply serve to replace the need for memory and trust between clans over a long period of delay between unbalanced transfers of reproductive resources.

Like inheritance, lawsuit, and tribute, marriage requires a triple coincidence of the event, in this case the marriage, with supply and demand. Without a transferable and durable store of value, the current ability of a groom's clan to supply the current desires of the bride's clan, to a large enough degree to make up the value mismatch between bride and groom, while also satisfying the political and romantic constraints of the match, were quite unlikely to be well satisfied. One solution is imposing an ongoing service obligation from the groom or his clan to the bride's clan. This occurs in about 15% of known cultures[DW88]. In a much larger number, 67%, the groom or groom's clan pays the bride's clan a substantial amount of wealth. Some of this bride price is paid in immediate consumables, in plants to be gathered harvested and animals slaughtered for the marriage feast. In herding or agricultural societies much of the bride price is paid in livestock, a long-lasting form of wealth. The balance, and usually the most valuable portion of the bride-price in cultures without livestock, is paid with what are usually the most valuable family heirlooms – the rarest, costliest, and most durable pendants, rings, and so on. The Western practice of the groom giving the bride a ring – and a suitor giving a maiden other kinds of jewelry – was once a substantial transfer of wealth and was common in many other cultures. In about 23% of cultures, mostly modern ones, there is no substantial wealth exchange. In about 6% of cultures there is mutual exchange of substantial wealth between bride and groom clans. In only about 2% of cultures does the bride's clan pay the new couple a dowry.[DW88]

Unfortunately, some wealth transfers were a far cry from the altruism of the inheritance gift or the joy of marriage. Quite the opposite, in the case of tribute.

The Spoils of War

Death rates from violence in chimp troops and hunter-gatherer human cultures alike are far higher than in modern civilizations. This probably dates at least as far back as our common ancestor with the chimpanzees – chimp troops, as well, are constantly fighting.

Warfare involved, among other things, killing, maiming, torture, kidnapping, rape, and the extortion of tribute in exchange for avoiding such fates. When two neighboring tribes were not at war, one was usually paying tribute to the other. Tribute could also serve to bind alliances, achieving economies of scale in warfare. Mostly, it was a form of exploitation more lucrative to the victor than further violence against the defeated.

Victory in war was sometimes followed by an immediate payment from the losers to the winners. Often this just took the form of looting by the enthusiastic victors, while the losers desperately hid their collectibles. More often, tribute was demanded on a regular basis. In this case, the triple coincidence could and sometimes was avoided by a sophisticated schedule of payments in kind that matched the losing tribe's ability to supply a good or service with the victor's demand for it. However, even with this solution primitive money could provide a better way – a common medium of value that greatly simplified the terms of payment – very important in an era when terms of the treaty could not be recorded but had to be memorized. In some cases, as with the wampum as used in the Iriquois Confederacy, the collectibles doubled as a primitive mnemonic device that, while not verbatim, could be used as an aid to recall the terms of the treaty. For the winners, collectibles provided a way to collect tribute at closer to the Laffer optimum. For the losers, collectibles buried in caches provided a way to "under-report", leading the victors to believe the losers were less wealthy

and thus demand less than they might. Caches of collectibles also provided insurance against overzealous tribute collectors. Much of the wealth in primitive societies escaped the notice of the missionaries and anthropologists due to its highly secretive nature. Only archeology can reveal the existence of this hidden wealth.

Hiding and other strategies presented a problem that tribute collectors share with modern tax collectors – how to estimate the amount of wealth they can extract. Value measurement is a thorny problem in many kinds of transactions, but never more so than in the antagonistic collection of tax or tribute. In making these very difficult and nonintuitive trade-offs, and then executing them in a series of queries, audits, and collection actions, tribute collectors efficiently optimized their revenue, even if the results seemed quite wasteful to the tribute payer.

Imagine a tribe collecting tribute from several neighbor tribes it previously defeated in war. It must estimate how much it can extract from each tribe. Bad estimates leave the wealth of some tribes understated, while forcing others to pay tribute based on estimates of wealth they don't actually have. The result: the tribes that are hurt tend to shrink. The tribes that benefit pay less tribute than could be extracted. In both cases, less revenue is generated for the victors than they might be able to get with better rules. This is an application of the Laffer curve to the fortunes of specific tribes. On this curve, applied to income taxes by the brilliant economist Arthur Laffer, as the tax rate increases, the amount of revenue increases, but at an increasingly slower rate than the tax rate, due to increased avoidance, evasion, and most of all disincentive to engage in the taxed activity. At a certain rate due to these reasons tax revenues are optimized. Hiking the tax rate beyond the Laffer optimum results in lower rather than higher revenues for the government. Ironically, the Laffer curve was used by advocates for lower taxes, even though it is a theory of tax collection optimum to government revenue, not a theory of tax collection optimal to social welfare or individual preference satisfaction.

On a larger scale, the Laffer curve may be the most important economic law of political history. Charles Adams[A90] uses it to explain the rise and fall of empires. The most successful governments have been implicitly guided by their own incentives – both their short-term desire for revenue and their long-term success against other governments – to optimize their revenues according to the Laffer Curve. Governments that overburdened their taxpayers, such as the Soviet Union and later Roman Empire, ended up on the dust-heap of history, while governments that collected below the optimum were often conquered by their better-funded neighbors. Democratic governments may maintain high tax revenues over historical time by more peaceful means than conquering underfunded states. They are the first states in history with tax revenues so high relative to external threats that they have the luxury of spending most of the money in non-military areas. Their tax regimes have operated closer to the Laffer optimum than those of most previous kinds of governments. (Alternatively, this luxury may be made possible by the efficiency of nuclear weapons in deterring attack rather than the increased incentives of democracies to optimize to tax collection). When we apply the Laffer curve to examining the relative impact of treaty tribute terms on various tribes, we conclude that the desire to optimize revenues causes victors to want to accurately measure the income and wealth of the vanquished. Measuring value is crucial to determining the tributaries' incentives to avoid or evade the tribute by hiding wealth, fight, or flight. For their part, tributaries can and do spoof these measurements in various ways, for example by burying collectibles in caches. Tribute collection involves a measurement game with unaligned incentives.

With collectibles, one can demand tribute at strategically optimal times instead of when items can be supplied by the tributary or is in demand by the victor. The victors can then choose when they will in the future consume the wealth, rather than having to consume it at the time the tribute is extracted. Much later, well into the dawn of history, in 700 BC, though trade was widespread, money still took the form of collectibles – made out of more precious metals, but in their basic characteristics, such as lack of uniform value, similar to most of the proto-money used since the dawn of *Homo sapiens sapiens*. This was changed by a Greek-

speaking culture in Anatolia (modern Turkey), the Lydians. Specifically, the kings of Lydia were the first major issuers of coins in the archaeological and historical record.

From that day to this, government mints with self-granted monopolies, rather than private mints, have been the main issuers of coin. Why wasn't minting dominated by private interests, such as private bankers, which did exist at the time in these semi-market economies? The main explanation for government dominance of coin minting has been that only governments could enforce anti-counterfeiting measures. However, they could have enforced such measures in protection of competing private mints, just as they enforce trademarks today and at that time as well.

It was far easier to estimate the value of a coin than that of a collectible – especially at low transaction values. Far more trades could be made with money instead of barter; indeed many kinds of low-value trades became possible for the first time as the small gains from trade for the first time exceeded transaction costs. Collectibles were low-velocity money, involved in a small number of high value transactions. Coins were high-velocity money, facilitating a large number of low value trades.

Given what we have seen about the benefits of proto-money to tribute and tax collectors, as well as the critical nature of the value measurement problem in optimally coercing such payments, it is not surprising that tax collectors, specifically the kings of Lydia, were the first major issuers of coinage. The king, deriving his revenue from tax collection, had a strong incentive to measure to value of wealth held and exchanged by his subjects more accurately. That the exchange also benefited from cheaper measurement by traders of the medium of exchange, creating something closer to efficient markets, and allowing individuals to enter into the marketplace on a larger scale for the first time, was for the king a fortuitous side effect. The greater wealth flowing through markets, now available to be taxed, boosted the king's revenues even beyond the normal Laffer curve effect of reducing mismeasurement between given tax sources.

This combination of more efficient tax collection with more efficient markets meant a vast increase in overall tax revenues. These tax collectors almost literally hit a gold mine, and the wealth of Lydian kings Midas, Croesus, and Giges is famous to this day.

A few centuries later, the Greek king Alexander the Great conquered Egypt, Persia and much of India, funding his spectacular conquest by plundering Egyptian and Persian temples, filled with assemblages of low-velocity collectibles, and melting them down into high-velocity coins. More efficient and encompassing market economies as well as more efficient tax collection sprung up in his wake.

Tribute payments did not form by themselves a closed loop of collectibles. These were only valuable if they ultimately could be used by the victors for something else, such as marriage, trade, or collateral. However, victors could coerce the vanquished into manufacturing for obtaining collectibles, even if it did not serve the vanquished's voluntary interests.

Disputes and Remedies

Ancient hunter-gatherers did not have our modern tort or criminal law, but they did have an analogous means of settling disputes, often judged by clan or tribal leaders or vote, that covered what modern law calls crimes and torts. Settling disputes through punishments or payments sanction by the clans of the disputing parties substituted for cycles of revenge or vendetta wars. Most pre-modern cultures, ranging from the Iriquois in America to the pre-Christian Germanic peoples, decided that payment was better than punishment. Prices (e.g. the Germanic "weregeld" and Iriquois blood money) were assigned to all actionable offenses, ranging from petty theft to rape to murder. Where money was available, the payment took the form of money. Livestock was used in herding cultures. Otherwise, payment of collectibles was the most commonly used remedy.

The payment of remedies for damages in a lawsuit or similar complaint led to the same kind of problem of triple coincidence of event, supply, and demand as occurred in inheritance, marriage, and tribute. The judgment of the case had to coincide with the ability of the plaintiff to pay the damages as well as the opportunity and desire of the defendant to benefit from them. If the remedy was a consumable the plaintiff already had plenty of, the remedy still served as a punishment but would not likely satisfy the defendant – and thus would not curb the cycle of violence. Thus, we here again the value added by collectibles – in this case, in making possible the remedy to resolve a dispute or terminate a cycle of revenge.

Dispute remedies would not form a closed loop if the payments served to entirely eliminate vendettas. However, if the payments did not completely dampen the vendetta, the payments could form a cycle following the cycle of revenge. For this reason, it is possible that the institution reached an equilibrium when it had reduced but not eliminated cycles of revenge until the advent of more densely connected trading networks.

Attributes of Collectibles

Since humans evolved in small, largely self-sufficient, and mutually antagonistic tribes, the use of collectibles to reduce the need for favor-tracking, and to make possible the other human institutions of wealth transfer we have explored, was far more important than the scale problems of barter for most of the timespan of our species. Indeed, collectibles provided a fundamental improvement to the workings of reciprocal altruism, allowing humans to cooperate in ways unavailable to other species. For them, reciprocal altruism is severely limited by unreliable memory. Some other species have large brains, build their own homes, or make and use tools. No other species has produced such an improvement to the workings of reciprocal altruism. The evidence indicates this new development had matured by 40,000 B.P.

Menger called this first money an "intermediate commodity" – what this paper calls collectibles. An artifact useful for other things, such as cutting, could also be used as a collectible. However, once institutions involving wealth transfer became valuable, collectibles would be manufactured just for their collectible properties. What are these properties? For a particular commodity to be chosen as a valuable collectible, it would have had, relative to products less valuable as collectibles, at least the following desirable qualities:

1. More secure from accidental loss and theft. For most of history this meant carriable on the person and easy to hide.
2. Harder to forge its value. An important subset of these are products that are unforgeably costly, and therefore considered valuable, for reasons explained below.
3. This value was more accurately approximated by simple observations or measurements. These observations would have had more reliable integrity yet have been less expensive.

Humans the world over are strongly motivated to collect items that better satisfy these properties. Some of this motivation probably includes genetically evolved instincts. Such objects are collected for the sheer pleasure of collecting them (not for any particularly good explicit and proximate reasons), and such pleasure is nearly universal across human cultures. One of the immediate proximate motivations is decoration. According to Dr. Mary C. Stiner, an archaeologist at the University of Arizona, "Ornamentation is universal among all modern human foragers."[W02] For an evolutionary psychologist, such a behavior that has a good ultimate explanation, in terms of natural selection, but has no proximate rationale other than pleasure, is a prime candidate to be a genetically evolved pleasure that motivates the behavior. Such is, if the reasoning in this essay is correct, the human instinct to collect rare items, art, and especially jewelry.

Point (2) requires some further explanation. At first, the production of a commodity simply because it is costly seems quite wasteful. However, the unforgeably costly commodity repeatedly adds value by enabling beneficial wealth transfers. More of the cost is recouped every time a transaction is made possible or made

less expensive. The cost, initially a complete waste, is amortized over many transactions. The monetary value of precious metals is based on this principle. It also applies to collectibles, which are more prized the rarer they are and the less forgeable this rarity is. It also applies where provably skilled or unique human labor is added to the product, as with art.

We have never discovered or made a product that does really well on all three scores. Art and collectibles (in the sense that word is used in modern culture, rather in the technical sense it is used in this paper) optimize (2), but not (1) or (3). Common beads satisfy (1) but not (2) or (3). Jewelry, made at first out of the most beautiful and less common shells but eventually in many cultures out of precious metals, comes closer to satisfying all three properties. It is no coincidence that precious metal jewelry usually came in thin forms such as chains and rings, allowing for inexpensive assaying at randomly chosen locations. Coins were a further improvement – substituting small standard weights and trademarks for assays greatly reduced the costs of small transactions using precious metals. Money proper was just a further step in the evolution of collectibles.

The kind of mobile art also made by Paleolithic man, (small figurines and the like) also matches these characteristics well. Indeed, Paleolithic man made very few objects that were not either utilitarian, or shared characteristics (1)-(3).

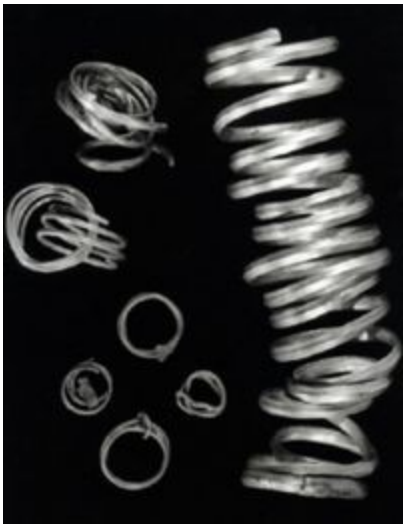
There are many puzzling instances of useless or at least unused flints with *homo sapiens*. We have mentioned the unusable flints of the Clovis people. Culiffe[C94] discusses a European Mesolithic era find of hundreds of flints, carefully crafted, but which micrograph analysis reveals were never used for cutting.

Flints were quite likely the first collectibles, preceding special-purpose collectibles like jewelry. Indeed, the first flint collectibles would have been made for their cutting utility. Their added value as a medium of wealth transfer was a fortuitous side effect that enabled the institutions described in this article to blossom. These institutions, in turn, would have motivated the manufacture of special-purpose collectibles, at first flints that need have no actual use as cutting tools, then the wide variety of other kinds of collectibles that were developed by *Homo sapiens sapiens*.

Shell money from Sumer, c. 3,000 B.C.

During the Neolithic era, in many parts of the Middle East and Europe, some kinds of jewelry became more standardized – to the point where standard sizes and assayability were often valued over beauty. In commercial areas the quantity of this jewelry sometimes greatly exceeded that of traditional jewelry in hoards. This is an intermediate step between jewelry and coins, when some collectibles increasingly took a fungible form. Around 700 B.C., the Lydian kings started issuing coins, as described above. The unforgeable costliness of standard weights of precious metals could now be "assayed" in a marketplace, by wage earners, or by tax collectors via trademark, i.e. trust in the mint's brand, instead of chopping coiled wire at a randomly selected spot.

It is no coincidence that the attributes of collectibles are shared with precious metals, coins, and the reserve commodities that have backed most non-fiat currencies. Money proper implemented these properties a purer form than the collectibles used during almost all of human prehistory.



Silver ring and coil money from Sumer, c. 2,500 B.C. Note the standard size of cross-sections. Many of the pieces had a standard weight, ranging from one-twelfth of a shekel to sixty shekels. To assay a ring or coil, it could be weighed and cut at random locations. (Courtesy Oriental Institute, University of Chicago)

A novelty of the 20th century was the issue of fiat currencies by governments. ("Fiat" means not backed by any reserve commodity, as the gold- and silver-based currencies of previous centuries were). While generally excellent as media of exchange, fiat currencies have proven to be very poor stores of value. Inflation has destroyed many a "nest egg". It is no coincidence that markets in rare objects and unique artwork – usually sharing the attributes of collectibles described above – have enjoyed a renaissance during the last century. One of our most advanced high-tech marketplaces, eBay, is centered around these objects of primordial economic qualities. The collectibles market is larger than ever, even if the fraction of our wealth invested in them is smaller than when they were crucial to evolutionary success. Collectibles both satisfy our instinctive urges and remain useful in their ancient role as a secure store of value.

Conclusion

Many kinds of wealth transfers – one-way and mutual, voluntary and coerced – face transaction costs. In voluntary trades both parties gain; a truly free gift is usually an act of kin altruism. These transactions create value for one or both parties as much as the physical act of making something. Tribute benefits the victor and a judgment of damages can prevent further violence as well as benefiting the victim. Inheritance made humans the first animals to pass wealth to their next-generation kin. These heirlooms could in turn be used as collateral or payment in trade for goods, for food to stave off starvation, or to pay a marriage bride price. Whether the costs of making these transfers – transaction costs – are low enough to make the transfers worthwhile is another matter. Collectibles were crucial in making these kinds of transactions possible for the first time.

Collectibles augmented our large brains and language as solutions to the Prisoner's Dilemma that keeps almost all animals from cooperating via delayed reciprocation with nonkin. Reputational beliefs can suffer from two major kinds of errors – errors of about which person did what, and errors in appraising the value or damages caused by that act. Within clans (the small and immediately local kin group, or extended family, which formed a subset of a tribe), our large brains could minimize these errors, so that public reputation and coercive sanctions superseded the limited motivation provided by the counterparty's ability to cooperate or defect in the future as the main enforcer of delayed reciprocation. In both *Homo sapiens neanderthalensis* and *Homo sapiens sapiens*, with the same large brain size, it is quite likely that every local clan member kept track of everybody other local clan member's favors. The use of collectibles for trade within the small local kin group may have been minimal. Between clans within a tribe both favor tracking and collectibles were

used. Between tribes, collectibles entirely replaced reputation as the enforcer of reciprocation, although violence still played a major role in enforcing rights as well as being a high transaction cost that prevented most kinds of trade.

When costliness becomes forgeable *– Glass trade beads, manufactured in Venice in the 16th or 17th century, excavated from Mali, Africa. Such beads were very popular wherever European colonialists encountered Neolithic or hunter-gatherer cultures.*

To be useful as a general-purpose store of wealth and means of wealth transfer, a collectible had to be embedded in at least one institution with a closed-loop cycle, so that the cost of discovering and/or manufacturing the object was amortized over multiple transactions. Furthermore, a collectible was not just any kind of beautiful decorative object. It had to have certain functional properties, such as the security of being wearable on the person, compactness for hiding or burial, and unforgeable costliness. That costliness must have been verifiable by the recipient of the transfer – using many of the same skills that collectors use to appraise collectibles today.

The theories presented in this paper can be tested by looking for these characteristics (or the lack of them) in the "valuables" often exchanged in these cultures, by examining the economic gains from the cycles through which these valuables move, and by observing preferences for objects with these characteristics in a wide variety of cultures (including modern ones).

With their unprecedented technology of cooperation, humans had become the most fearsome predator ever seen on the planet. They adapted to a shifting climate, while dozens of their large herd prey were driven, by hunting and climate change in America, Europe, and Asia, to extinction. Today, most large animals on the planet are afraid of projectiles – an adaption to only one species of predator [R97]. Cultures based more on gathering than hunting also greatly benefitted. A population explosion followed – *Homo sapiens sapiens* was able to populate more parts of the planet and at a density over ten times that of *Homo sapiens neanderthalensis* [C94], despite weaker bones and no increase in brain size. Much of this increase may be attributed to the social institutions made possible by effective wealth transfer and language – trade, marriage, inheritance, tribute, collateral, and the ability to assess damages to dampen cycles of vengeance.

Primitive money was not modern money as we know it. It took on some of the functions modern money now performs, but its form was that of heirlooms, jewelry, and other collectibles. The use of these is so ancient that the desires to explore, collect, make, display, appraise, carefully store, and trade collectibles are human universals – to some extent instincts. This constellation of human desires might be called the collecting instinct. Searching for the raw materials, such as shells and teeth, and manufacturing of collectibles took up a considerable portion of many ancient humans' time, just as many modern humans expend substantial resources on these activities as hobbies. The results for our ancient forebears were the first secure forms of

embodied value very different from concrete utility – and the forerunner of today's money.

References

1. [A90] Adams, Charles, *For Good and Evil: The Impact of Taxes on Civilization* ↩
2. [A98] Tim Appenzeller, "Art: Evolution or Revolution?", Science 282(Nov 20, 1998), p. 1452. See also the home page of Stanley Ambrose ↩ ↩
3. [B04] The Blombos Cave Project ↩ ↩
4. [C94] Culiffe, Barry, ed., *The Oxford Illustrated History of Prehistoric Europe*, Oxford University Press 1994. ↩ ↩ ↩ ↩ ↩
5. [D89] Dawkins, Richard, *The Selfish Gene*, Oxford University Press 1989. ↩ ↩ ↩
6. [D94] Davies, Glyn, *A History of Money, From Ancient Times to the Present Day*, University of Wales Press 1994. ↩ ↩ ↩ ↩
7. [DW88] Daly, Martin and Wilson, Margo, *Homicide*, New York: Aldine (1998). ↩ ↩
8. [G95] Gilead, I. 1995. "The Foragers of the Upper Paleolithic Period," in Archaeology and Society in the Holy Land. Ed. by T. E. Levy. New York, Facts on File. ↩
9. [G01] [ref: http://www-geology.ucdavis.edu/~GEL115/115CH1.html] ↩
10. [Gr01] Graeber, David, *Towards an Anthropological Theory of Value*, Palgrave 2001.
11. [I98] Ifrah, Georges, *The Universal History of Numbers*, John Wiley & Sons 1998, pg. 73.
12. [K99] Kohn, M. and Mithen, S. "Handaxes: Products of sexual selection?", Antiquity, 73, 518-526.
13. [K99] Kohn, M. and Mithen, S. "Handaxes: Products of sexual selection?", Antiquity, 73, 518-526.
14. [L94] Landa, Janet, *Trust, Ethnicity, and Identity: Beyond the New Institutional Economics of Ethnic Trading Networks, Contract Law, and Gift-Exchange*, The University of Michigan Press, second edition, 1998. ↩ ↩ ↩
15. [M1892] Menger, Carl, "On the Origins of Money" Economic Journal, volume 2,(1892) p. 239-55. translated by C.A. Foley, at http://www.socsci.mcmaster.ca/~econ/ugcm/3ll3/menger/money.txt ↩
16. [M50] Mauss, Marcel, *The Gift*, 1950, English translation by W.D. Halls, W.W. Norton 1990. ↩
17. [M93] (Morse 1993) via http://www.wac.uct.ac.za/wac4/symposia/papers/s095wht1.pdf ↩
18. [R96] Riddley, Matt, *The Origins of Virtue*, Viking 1996.
19. [T01] Taylor, Alan, *American Colonies – The Settling of North America*, Penguin 2001. ↩ ↩
20. [P89] Plattner, Stuart, *Economic Anthropology*, Stanford University Press 1989.
21. [W77] Wiessner, P. 1977. Hxaro: a regional system at reciprocity for reducing risk among the !Kung San. Unpublished PhD thesis: University of Michigan. ↩
22. [W82] Wiessner, P. 1982. Risk, reciprocity and social influences on !Kung San economies. In: Leacock, H. R. & Lee, R.B. (eds) *Politics and history in band societies*: 61-84. London: Cambridge University Press.
23. [W95] White, Randall, "Ivory Personal Ornaments of Aurignacian Age: Technological, Social and Symbolic Perspectives", Institute For Ice Age Studies, http://www.insticeagestudies.com/library/Ivory/Ivorypersonal.html ↩
24. [W97] White, Randall, "From Materials To Meaning", Institute For Ice Age Studies, http://www.insticeagestudies.com/library/materialstomeaning/index.html ↩
25. [W98] Winterhalder, Bruce, "Intra-Group Resource Transfers: Comparative Evidence, Models, and Implications for Human Evolution", http://www.unc.edu/depts/ecology/winterweb/intra_group.html ↩
26. [W02] Wilford, John, "Debate is Fueled on When Humans Became Human", New York Times, February 26th, 2002 ↩

Acknowledgements

My thanks to Jerome Barkow, Andrew Odlyzko, Bruce Smith, K. Eric Drexler, Markus Krummenacker, Mark Wiley, Norm Hardy, and others for their insightful comments.

Please send your comments to nszabo (at) law (dot) gwu (dot) edu

---

Editor's note: Some links may be broken.

Back | Index

- About
- Contact
- Donate BTC
- Atom feed
- GitHub

III.

Money and Overall Prices

1. The Supply and Demand for Money and Overall Prices

When economics students read textbooks, they learn, in the "micro" sections, how prices of specific goods are determined by supply and demand. But when they get to the "macro" chapters, lo and behold! supply and demand built on individual persons and their choices disappear, and they hear instead of such mysterious and ill-defined concepts as *velocity of circulation, total transactions,* and *gross national product*. Where are the supply-and-demand concepts when it comes to overall prices?

In truth, overall prices are determined by similar supply-and-demand forces that determine the prices of individual products. Let us reconsider the concept of *price*. If the price of bread is 70 cents a loaf, this means also that the *purchasing power* of a loaf of bread is 70 cents. A loaf of bread can command 70 cents in exchange on the market. The price and purchasing power of the unit of a product are one and the same. Therefore, we can construct a diagram for the determination of overall prices, with

the price or the purchasing power of the money unit on the Y-axis.

While recognizing the extreme difficulty of arriving at a measure, it should be clear conceptually that the price or the purchasing power of the dollar is the inverse of whatever we can construct as the *price level*, or the level of overall prices. In mathematical terms,

$$\text{PPM} = \frac{1}{\text{P}}$$

where PPM is the purchasing power of the dollar, and P is the price level.

To take a highly simplified example, suppose that there are four commodities in the society and that their prices are as follows:

eggs	$ .50 dozen
butter	$ 1 pound
shoes	$ 20 pair
TV set	$ 200 set

In this society, the PPM, or the purchasing power of the dollar, is an array of alternatives inverse to the above prices. In short, the purchasing power of the dollar is:

either	2 dozen eggs
or	1 pound butter
or	1/20 pair shoes
or	1/200 TV set

Suppose now that the price level doubles, in the easy sense that all prices double. Prices are now:

eggs	$ 1 dozen
butter	$ 2 pound
shoes	$ 40 pair
TV set	$ 400 set

In this case, PPM has been cut in half across the board. The purchasing power of the dollar is now:

either	1 dozen eggs
or	1/2 pound butter
or	1/40 pair shoes
or	1/400 TV set

Purchasing power of the dollar is therefore the inverse of the price level.

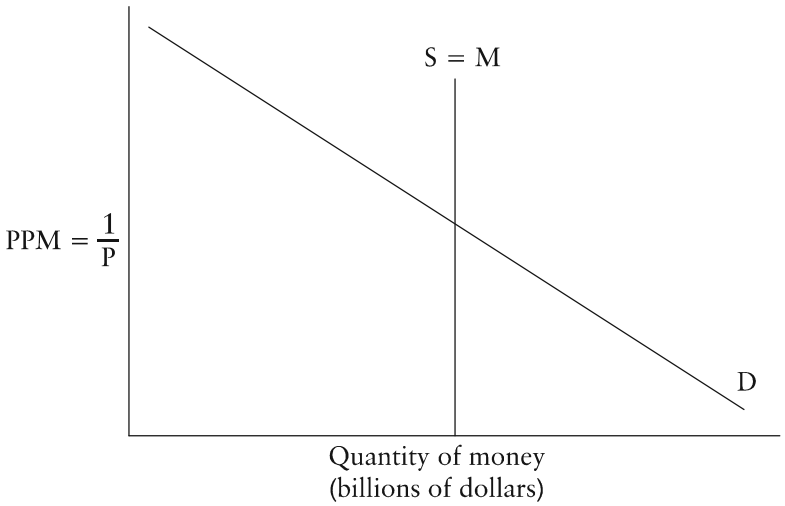


FIGURE 3.1 — SUPPLY OF AND DEMAND FOR MONEY

Let us now put PPM on the Y-axis and quantity of dollars on the X-axis. We contend that, on a complete analogy with supply, demand, and price above, the intersection of the vertical line indicating the supply of money in the country at any given time, with the falling demand curve for money, will yield the market equilibrium PPM and hence the equilibrium height of overall prices, at any given time.

Let us examine the diagram in Figure 3.1. The supply of money, M, is conceptually easy to figure: the total quantity of

dollars at any given time. (What constitutes these dollars will be explained later.)

We contend that there is a falling demand curve for money in relation to hypothetical PPMs, just as there is one in relation to hypothetical individual prices. At first, the idea of a demand curve for money seems odd. Isn't the demand for money unlimited? Won't people take as much money as they can get? But this confuses what people would be willing to accept as a gift (which is indeed unlimited) with their *demand* in the sense of how much they would be willing to give up for the money. Or: how much money they would be willing to keep in their cash balances rather than spend. In this sense their demand for money is scarcely unlimited. If someone acquires money, he can do two things with it: either spend it on consumer goods or investments, or *else* hold on to it, and increase his individual money stock, his total cash balances. How much he wishes to hold on to is his demand for money.

Let us look at people's demand for cash balances. How much money people will keep in their cash balance is a function of the level of prices. Suppose, for example, that prices suddenly dropped to about a third of what they are now. People would need far less in their wallets, purses, and bank accounts to pay for daily transactions or to prepare for emergencies. Everyone need only carry around or have readily available only about a third the money that they keep now. The rest they can spend or invest. Hence, the total amount of money people would hold in their cash balances would be far less if prices were much lower than now. Contrarily, if prices were triple what they are today, people would need about three times as much in their wallets, purses, and bank accounts to handle their daily transactions and their emergency inventory. People would demand far greater cash balances than they do now to do the same "money work" if prices were much higher. The falling demand curve for money is shown in Figure 3.2.

Here we see that when the PPM is very high (i.e., prices overall are very low), the demand for cash balances is low; but when

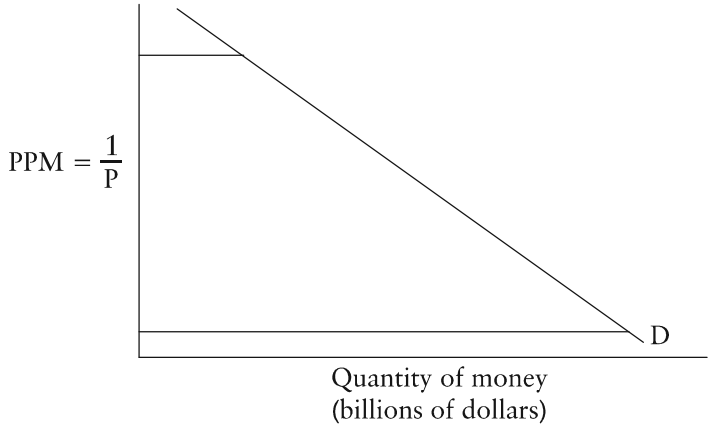


FIGURE 3.2 — DEMAND FOR MONEY

PPM is very low (prices are high), the demand for cash balances is very high.

We will now see how the intersection of the falling demand curve for money or cash balances, and the supply of money, determines the day-to-day equilibrium PPM or price level.

Suppose that PPM is suddenly very high, that is, prices are very low. M, the money stock, is given, at $100 billion. As we see in Figure 3.3, at a high PPM, the supply of total cash balances, M, is greater than the demand for money. The difference is surplus cash balances—money, in the old phrase, that is burning a hole in people's pockets. People find that they are suffering from a monetary imbalance: their cash balances are greater than they need at that price level. And so people start trying to get rid of their cash balances by spending money on various goods and services.

But while people can get rid of money individually, by buying things with it, they can't get rid of money in the aggregate, because the $100 billion still exists, and they can't get rid of it short of burning it up. But as people spend more, this drives up demand curves for most or all goods and services. As the demand curves shift upward and to the right, prices rise. But as prices overall rise further and further, PPM begins to fall, as the downward

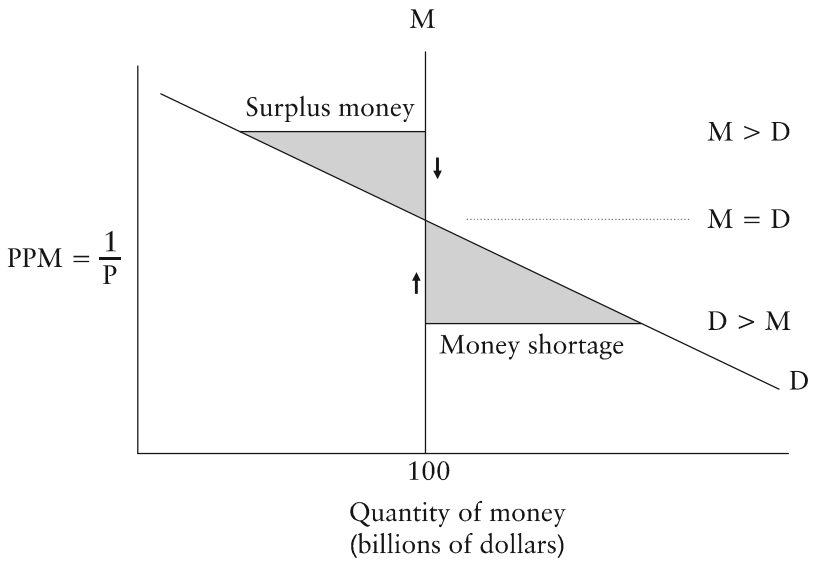


FIGURE 3.3 —DETERMINATION OF THE PURCHASING POWER OF MONEY

arrow indicates. And as the PPM begins to fall, the surplus of cash balances begins to disappear until finally, prices have risen so much that the $100 billion no longer burns a hole in anyone's pocket. At the higher price level, people are now willing to keep the exact amount of $100 billion that is available in the economy. The market is at last cleared, and people now wish to hold no more and no less than the $100 billion available. The demand for money has been brought into equilibrium with the supply of money, and the PPM and price level are in equilibrium. People were not able to get rid of money in the aggregate, but they were able to drive up prices so as to end the surplus of cash balances.

Conversely, suppose that prices were suddenly three times as high and PPM therefore much lower. In that case, people would need far more cash balances to finance their daily lives, and there would be a shortage of cash balances compared to the supply of money available. The demand for cash balances would be greater than the total supply. People would then try to alleviate this

imbalance, this shortage, by adding to their cash balances. They can only do so by spending less of their income and adding the remainder to their cash balance. When they do so, the demand curves for most or all products will shift downward and to the left, and prices will generally fall. As prices fall, PPM *ipso facto* rises, as the upward arrow shows. The process will continue until prices fall enough and PPM rises, so that the $100 billion is no longer less than the total amount of cash balances desired.

Once again, market action works to equilibrate supply and demand for money or cash balances, and demand for money will adjust to the total supply available. Individuals tried to scramble to add to their cash balances by spending less; in the aggregate, they could not add to the money supply, since that is given at $100 billion. But in the process of spending less, prices overall fell until the $100 billion became an adequate total cash balance once again.

The price level, then, and the purchasing power of the dollar, are determined by the same sort of supply-and-demand feedback mechanism that determines individual prices. The price level tends to be at the intersection of the supply of and demand for money, and tends to return to that point when displaced.

As in individual markets, then, the price or purchasing power of the dollar varies directly with the demand for money and inversely with the supply. Or, to turn it around, the price level varies directly with the supply of money and inversely with the demand.

2. Why Overall Prices Change

Why does the price level ever change, if the supply of money and the demand for money determine the height of overall prices? If, and only if, one or both of these basic factors—the supply of or demand for money—changes. Let us see what happens when the supply of money changes, that is, in the modern world, when the supply of nominal units changes rather than the actual weight of gold or silver they used to represent. Let us assume, then, that the supply of dollars, pounds, or francs increases, without yet

examining how the increase occurs or *how* the new money gets injected into the economy.

Figure 3.4 shows what happens when M, the supply of dollars, of total cash balances of dollars in the economy, increases.

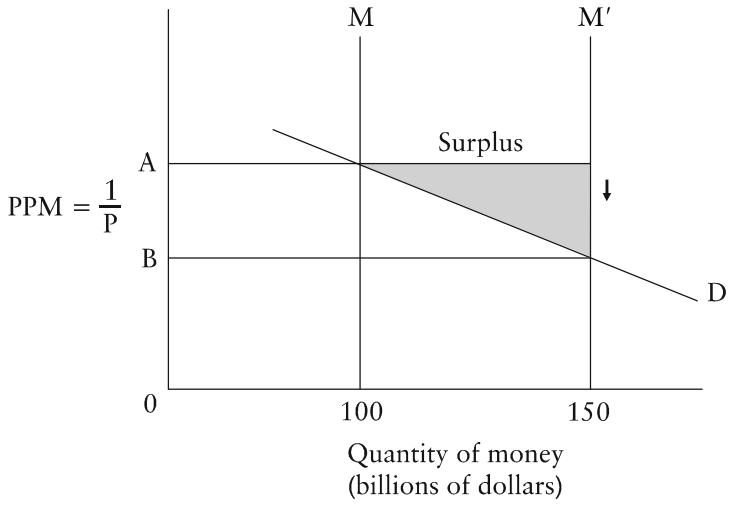


FIGURE 3.4 — INCREASE IN THE SUPPLY OF MONEY

The original supply of money, M, intersects with the demand for money and establishes the PPM (purchasing power of the dollar) and the price level at distance 0A. Now, in whatever way, the supply of money increases to M′. This means that the aggregate total of cash balances in the economy has increased from M, say $100 billion, to M′, $150 billion. But now people have $50 billion surplus in their cash balances, $50 billion of excess money over the amount needed in their cash balances at the previous 0A prices level. Having too much money burning a hole in their pockets, people spend the cash balances, thereby raising individual demand curves and driving up prices. But as prices rise, people find that their increased aggregate of cash balances is getting less and less excessive, since more and more cash is now needed to accommodate the higher price levels. Finally, prices rise until PPM has fallen from 0A to 0B. At these new, higher price levels,

the M′—the new aggregate cash balances—is no longer excessive, and the demand for money has become equilibrated by market forces to the new supply. The *money market*—the intersection of the demand and supply of money—is once again cleared, and a new and higher equilibrium price level has been reached.

Note that when people find their cash balances excessive, they try to get rid of them, but since all the money stock is owned by *someone*, the new M′ cannot be gotten rid of in the aggregate; by driving prices up, however, the demand for money becomes equilibrated to the new supply. Just as an increased supply of pork drives down prices so as to induce people to buy the new pork production, so an increased supply of dollars drives down the purchasing power of the dollar until people are willing to hold the new dollars in their cash balances.

What if the supply of money, M, *de*creases, admittedly an occurrence all too rare in the modern world? The effect can be seen in Figure 3.5.

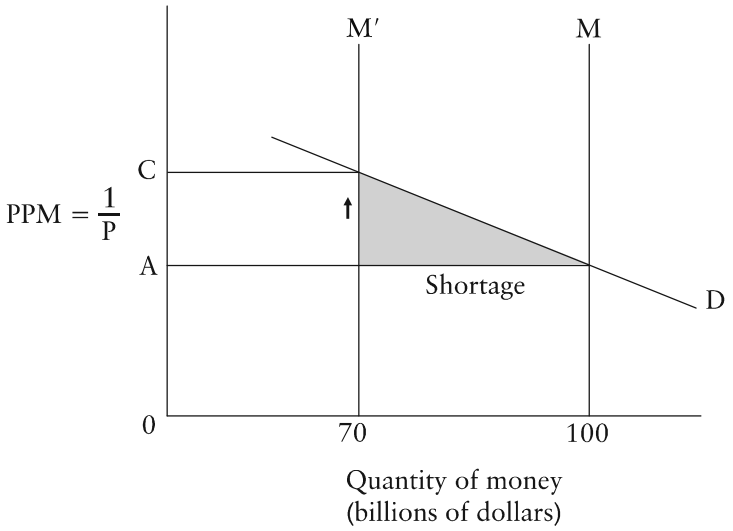


FIGURE 3.5 — A FALL IN THE SUPPLY OF MONEY

In the unusual case of a fall in the supply of money, then, total cash balances fall, say, from $100 billion (M) to $70 billion (M′). When this happens, the people find out that at the old equilibrium price level 0A, aggregate cash balances are not enough to satisfy their cash balance needs. They experience, therefore, a cash balance shortage. Trying to increase his cash balance, then, each individual spends less and saves in order to accumulate a larger balance. As this occurs, demand curves for specific goods fall downward and to the left, and prices therefore fall. As this happens, the cash balance shortage is alleviated, until finally prices fall low enough until a new and lower equilibrium price level (0C) is established. Or, alternatively, the PPM is at a new and higher level. At the new price level of PPM, 0C, the demand for cash balances is equilibrated with the new and decreased supply M′. The demand and supply of money is once again cleared. At the new equilibrium, the decreased money supply is once again just sufficient to perform the cash balance function.

Or, put another way, at the lower money supply people scramble to increase cash balances. But since the money supply is set and outside their control, they cannot increase the supply of cash balances in the aggregate.[1] But by spending less and driving down the price level, they increase the value or purchasing power of each dollar, so that real cash balances (total money supply corrected for changes in purchasing power) have gone up to offset the drop in the total supply of money. M might have fallen by $30 billion, but the $70 billion is now as good as the previous total because each dollar is worth more in *real*, or purchasing power, terms.

An increase in the supply of money, then, will lower the price or purchasing power of the dollar, and thereby increase the level of prices. A fall in the money supply will do the opposite, lowering prices and thereby increasing the purchasing power of each dollar.

---

[1]Why doesn't an excess demand for cash balances increase the money supply, as it would in the case of beef, in the long run? For a discussion of the determinants of the supply of money, see chapter IV.

The other factor of change in the price level is the demand for money. Figures 3.6 and 3.7 depict what happens when the demand for money changes.

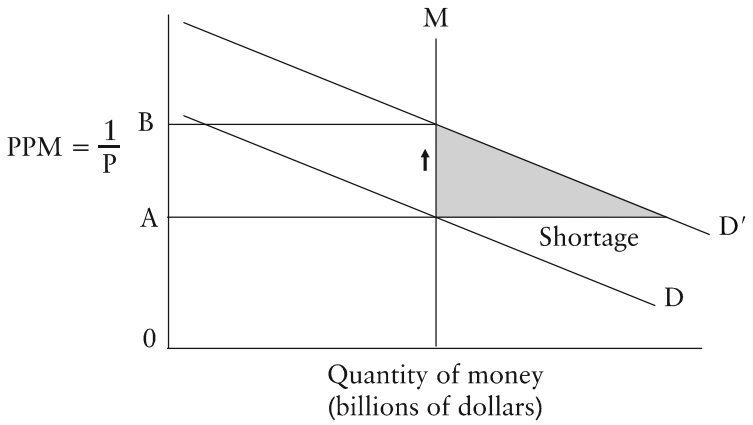


FIGURE 3.6 — AN INCREASE IN THE DEMAND FOR MONEY

The demand for money, for whatever reason, increases from D to D′. This means that, whatever the price level, the amount of money that people in the aggregate wish to keep in their cash balances will increase. At the old equilibrium price level, 0A, a PPM that previously kept the demand and supply of money equal and cleared the market, the demand for money has now increased and become greater than the supply. There is now an excess demand for money, or shortage of cash balances, at the old price level. Since the supply of money is given, the scramble for greater cash balances begins. People will spend less and save more to add to their cash holdings. In the aggregate, M, or the total supply of cash balances, is fixed and cannot increase. But the fall in prices resulting from the decreased spending will alleviate the shortage. Finally, prices fall (or PPM rises) to 0B. At this new equilibrium price, 0B, there is no longer a shortage of cash balances. Because of the increased PPM, the old money supply, M, is now enough to satisfy the increased demand for cash balances. Total cash balances have remained the same in nominal terms, but in *real* terms,

in terms of purchasing power, the $100 billion is now worth more and will perform more of the cash balance function. The market is again cleared, and the money supply and demand brought once more into equilibrium.

Figure 3.7 shows what happens when the demand for money falls.

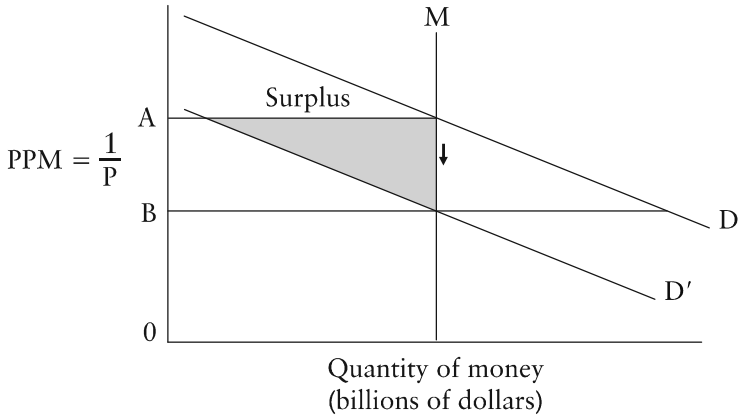


FIGURE 3.7 — A FALL IN THE DEMAND FOR MONEY

The demand for money falls from D to D′. In other words, whatever the price level, people are now, for whatever reason, willing to hold lower cash balances than they did before. At the old equilibrium price level, 0A, people now find that they have a surplus of cash balances burning a hole in their pockets. As they spend the surplus, demand curves for goods rise, driving up prices. But as prices rise, the total supply of cash balances, M, becomes no longer surplus, for it now must do cash balance work at a higher price level. Finally, when prices rise (PPM falls) to 0B, the surplus of cash balance has disappeared and the demand and supply of money has been equilibrated. The same money supply, M, is once again satisfactory despite the fall in the demand for money, because the same M must do more cash balance work at the new, higher price level.

So prices, overall, can change for only two reasons: If the supply of money increases, prices will rise; if the supply falls, prices will fall. If the demand for money increases, prices will fall (PPM rises); if the demand for money declines, prices will rise (PPM falls). The purchasing power of the dollar varies inversely with the supply of dollars, and directly with the demand. Overall prices are determined by the same supply-and-demand forces we are all familiar with in individual prices. *Micro* and *macro* are not mysteriously separate worlds; they are both plain economics and governed by the same laws.

2024

SESSION 2

Potential Solutions

CHAPTER 9

The Cure for Inflation

Compare two rectangles of paper of about the same size. One is mostly green on the back side and has a picture of Abraham Lincoln on the front side, which also has the number 5 on each of its corners and some printing. You can exchange this piece of paper for some quantity of food, clothing, or other goods. People will willingly make the trade.

The other piece of paper, perhaps cut from a glossy magazine, may also have a picture, some numbers, and some printing on its face. It may also be colored green on its back. Yet it is fit only to light the fire.

Whence the difference? The printing on the $5 bill gives no answer. It simply says, "FEDERAL RESERVE NOTE / THE UNITED STATES OF AMERICA / FIVE DOLLARS" and, in smaller print, "THIS NOTE IS LEGAL TENDER FOR ALL DEBTS, PUBLIC AND PRIVATE." Until not very many years ago, the words "WILL PROMISE TO PAY" were included between "THE UNITED STATES OF AMERICA" and "FIVE DOLLARS." That seemed to explain the difference between the two pieces of paper. But it meant only that if you had gone to a Federal Reserve Bank and asked a teller to redeem the promise, he would have given you five identical pieces of paper except that the number 1 took the place of the number 5 and George Washington's picture the place of Abraham Lincoln's. If you had then asked the teller to pay the $1 promised by one of these pieces of paper, he would have given you coins which, if you had melted them down (despite its being illegal to do so), would have sold for less than $1 as metal. The present wording is at least more candid if equally unrevealing. The legal-tender quality means that the government will accept the pieces of paper in discharge of debts and taxes due to itself, and that the courts will regard them as discharging debts stated in dollars. Why should they also be accepted by private persons in private transactions in exchange for goods and services?

The short answer is that each person accepts them because he is confident that others will. The pieces of green paper have value because everybody thinks they have value. Everybody thinks they have value because in his experience they have had value. The United States could not operate at more than a small fraction of its present level of productivity without a common and widely accepted medium of exchange (or at most a small number of such media); yet the existence of a common and widely accepted medium of exchange rests on a convention that owes its existence to the mutual acceptance of what, from one point of view, is a fiction.

The convention or the fiction is no fragile thing. On the contrary, the value of having a common money is so great that people will stick to the fiction even under extreme provocation—whence, as we shall see, comes part of the gain that issuers of the money can derive from inflation and hence the temptation to inflate. But neither is the fiction indestructible: the phrase "not worth a Continental" is a reminder of how that fiction was destroyed for the Continental currency issued in excessive amount by the U.S. Continental Congress to finance the American Revolution.

Though the value of money rests on a fiction, money serves an extraordinarily useful economic function. Yet it is also a veil. The "real" forces that determine the wealth of a nation are the capacities of its citizens, their industry and ingenuity, the resources at their command, their mode of economic and political organization, and the like. As John Stuart Mill wrote more than a century ago: "There cannot, in short, be intrinsically a more insignificant thing, in the economy of society, than money; except in the character of a contrivance for sparing time and labour. It is a machine for doing quickly and commodiously, what would be done, though less quickly and commodiously, without it: and like many other kinds of machinery, it only exerts a distinct and independent influence of its own when it gets out of order."[1]

Perfectly true, as a description of the role of money, provided we recognize that society possesses hardly any other contrivance that can do more damage when it gets out of order.

We have already discussed one example: the Great Depression,

when money got out of order through too sharp a reduction in its quantity. This chapter discusses the opposite and more common way in which money has gotten out of order—through too sharp an increase in quantity.

VARIETIES OF MONEY

An amazing variety of items has been used as money at one time or anoher. The word "pecuniary" comes from the Latin *pecus*, meaning "cattle," one of the many things that have been used as money. Others include salt, silk, furs, dried fish, even feathers, and, on the Pacific island of Yap, stones. Cowrie shells and beads have been the most widely used forms of primitive money. Metals—gold, silver, copper, iron, tin—have been the most widely used forms among more advanced economies before the victory of paper and the bookkeeper's pen.

The one thing all the items used as money have had in common is their acceptance, in the particular place and time, in return for other goods and services in the faith that others would likewise accept them.

The "wampum" that the early settlers of America used in trade with Indians was a form of shell, analogous to the cowrie shells used in Africa and Asia. A most interesting and instructive money used in the American colonies was the tobacco money of Virginia, Maryland, and North Carolina: "The first law passed by the first General Assembly of Virginia, July 31, 1619 [twelve years after Captain John Smith landed and established at Jamestown the first permanent settlement in the New World], was in reference to tobacco. It fixed the price of that staple 'at three shillings the beste, and the second sorte at 18d. the pounde.' . . . Tobacco was already the local currency." [2]

At various periods tobacco was declared the only legal currency. It remained a basic money of Virginia and its neighboring colonies for close to two centuries, until well after the American Revolution. It was the money that the colonists used to buy food, clothing, to pay taxes—even to pay for a bride: "The Rev. Mr. Weems, a Virginian writer, intimates that it would have done a man's heart good to see the gallant young Virginians hastening

to the waterside when a vessel arrived from London, each carrying a bundle of the best tobacco under his arm, and taking back with him a beautiful and virtuous young wife." [3] And another writer, quoting this passage, goes on to remark, "They must have been stalwart, as well as gallant, to hasten with a roll of tobacco weighing 100 to 150 pounds under the arm." [4]

As money goes, so tobacco went. The original price set on it in terms of English money was higher than the cost of growing it, so planters set to with a will and produced more and more. In this case, the money supply grew literally as well as figuratively. As always happens when the quantity of money increases more rapidly than the quantity of goods and services available for purchase, there was inflation. Prices of other things in terms of tobacco rose drastically. Before the inflation ended about half a century later, prices in terms of tobacco had risen fortyfold.

The growers of tobacco were most unhappy about the inflation. Higher prices of other things in terms of tobacco meant that tobacco would command less of those other things. The price of money in terms of goods is the reciprocal of the price of goods in terms of money. Naturally, tobacco growers turned to government for help. One law after another was passed prohibiting certain classes of people from growing tobacco; providing for destroying part of the crop; prohibiting the planting of tobacco for one year. All to no avail. Finally, people took matters into their own hands, banded together, and went around the countryside destroying tobacco plants: "The evil reached such proportions that in April, 1684, the Assembly passed a law declaring that these malefactors had passed beyond the bounds of riot, and that their aim was the subversion of the government. It was enacted that if any persons to the number of eight or more should go about destroying tobacco plants, they should be adjudged traitors and suffer death." [5]

The tobacco currency vividly illustrates one of the oldest laws in economics: Gresham's Law, "Bad money drives out good." The grower of tobacco, who had to pay taxes or other obligations fixed in terms of tobacco, understandably used the poorest quality tobacco to discharge obligations and retained the best quality for export in return for "hard" money, i.e., British sterling. As a re-

sult, only poor quality tobacco tended to circulate as money. Every device of human ingenuity was used to make tobacco appear higher in quality than it was: "Maryland in 1698 found it necessary to legislate against the fraud of packing trash in hogsheads that contained good tobacco on top. Virginia adopted a similar measure in 1705, but apparently it did not offer relief." [6]

The quality problem was somewhat alleviated when "[i]n 1727 tobacco notes were legalized. These were in the nature of certificates of deposit issued by the inspectors. They were declared by law current and payable for all tobacco debts within the warehouse district where they were issued." [7] Despite numerous abuses of the system, "[s]uch receipts performed the office of currency right to the eve of the 19th century." [8]

That was not the last use of tobacco as money. During World War II cigarettes were widely used as a medium of exchange in German and Japanese prison camps. After World War II cigarettes were widely used as money in Germany during the period when the occupation authorities enforced ceilings on prices in legal currency that were well below the levels that would have cleared the market. The result was to destroy the usefulness of the legal money. People resorted to barter and to the use of cigarettes as a medium of exchange for small transactions, and cognac for large ones—by all odds the most liquid currency of which we have record. Ludwig Erhard's monetary reform ended that instructive—and destructive—episode.[9]

The general principles illustrated by tobacco money in Virginia remain relevant in the modern era, though paper money issued by government and bookkeeping entries called deposits have replaced commodities or warehouse receipts for commodities as the basic money of the society.

It remains as true now as it was then that a more rapid increase in the quantity of money than in the quantity of goods and services available for purchase will produce inflation, raising prices in terms of that money. It does not matter why the quantity of money increases. In Virginia the quantity of tobacco money grew and produced an inflation of prices in terms of tobacco because the cost of producing tobacco in terms of labor and other resources fell drastically. In Europe in the Middle Ages, silver and

gold were the dominant money, and inflation of prices in terms of gold and silver occurred because precious metals from Mexico and South America flooded Europe via Spain. In the mid-nineteenth century inflation of prices in terms of gold occurred around the world because of gold discoveries in California and Australia; later, from the 1890s to 1914, because of the successful commercial application of the cyanide process to the extraction of gold from low-grade ore, primarily in South Africa.

Today, when the commonly accepted media of exchange have no relation to any commodity, the quantity of money is determined in every major country by government. Government and the government alone is responsible for any rapid increase in the quantity of money. That very fact has been the major source of confusion about the cause and the cure of inflation.

THE PROXIMATE CAUSE OF INFLATION

Inflation is a disease, a dangerous and sometimes fatal disease, a disease that if not checked in time can destroy a society. Examples abound. Hyperinflations in Russia and Germany after World War I—when prices sometimes doubled and more than doubled from one day to the next—prepared the ground for communism in the one country and nazism in the other. The hyperinflation in China after World War II eased Chairman Mao's defeat of Chiang Kai-shek. Inflation in Brazil, where it reached about 100 percent a year in 1954, brought military government. A far more extreme inflation contributed to the overthrow of Allende in Chile in 1973 and of Isabel Perón in Argentina in 1976, followed in both countries by the assumption of power by a military junta.

No government is willing to accept responsibility for producing inflation, even in less virulent degree. Government officials always find some excuse—greedy businessmen, grasping trade unions, spendthrift consumers, Arab sheikhs, bad weather, or anything else that seems even remotely plausible. No doubt, businessmen are greedy, trade unions are grasping, consumers are spendthrifts, Arab sheikhs have raised the price of oil, and weather is often bad. All these can produce high prices for individual items; they

cannot produce rising prices for goods in general. They can cause temporary ups or downs in the rate of inflation. But they cannot produce continuing inflation for one very simple reason: none of the alleged culprits possesses a printing press on which it can turn out those pieces of paper we carry in our pockets; none can legally authorize a bookkeeper to make entries on ledgers that are the equivalent of those pieces of paper.

Inflation is not a capitalist phenomenon. Yugoslavia, a communist country, has experienced one of the most rapid rates of inflation of any country in Europe; Switzerland, a bastion of capitalism, one of the lowest. Neither is inflation a communist phenomenon. China had little inflation under Mao; Italy, the United Kingdom, Japan, the United States—all largely capitalist countries—have experienced substantial inflation in the past decade. In the modern world, inflation is a printing press phenomenon.

The recognition that substantial inflation is always and everywhere a monetary phenomenon is only the beginning of an understanding of the cause and cure of inflation. The more basic question is, why do modern governments increase the quantity of money too rapidly? Why do they produce inflation when they understand its potential for harm?

Before turning to that question, it is worth dwelling a bit longer on the proposition that inflation is a monetary phenomenon. Despite the importance of that proposition, despite the extensive historical evidence supporting it, it is still widely denied—in large part because of the smoke screen with which governments try to conceal their own responsibility for inflation.

If the quantity of goods and services available for purchase—output, for short—were to increase as rapidly as the quantity of money, prices would tend to be stable. Prices might even fall gradually as higher incomes led people to want to hold a larger fraction of their wealth in the form of money. Inflation occurs when the quantity of money rises appreciably more rapidly than output, and the more rapid the rise in the quantity of money per unit of output, the greater the rate of inflation. There is probably no other proposition in economics that is as well established as this one.

Output is limited by the physical and human resources available and by the improvement in knowledge and capacity to use them. At best, output can grow only fairly slowly. Over the past century, output in the United States grew at the average rate of about 3 percent per year. Even at the height of the rapid growth of Japan after World War II, output grew about 10 percent per year. The quantity of commodity money is subject to similar physical limits, though, as the examples of tobacco, precious metals from the New World, and gold in the nineteenth century illustrate, commodity money has at times grown far more rapidly than output in general. Modern forms of money—paper and bookkeeping entries—are subject to no physical limits. The nominal quantity, that is, the number of dollars, pounds, marks, or other monetary units, can grow at any rate, and at times has grown at fantastic rates.

During the German hyperinflation after World War I, for example, hand-to-hand money grew at the *average* rate of more than 300 percent a *month* for more than a year, and so did prices. During the Hungarian hyperinflation after World War II, hand-to-hand money rose at the *average* rate of more than 12,000 percent per *month* for a year, and prices at the even higher rate of nearly 20,000 percent a month.[10]

During the far more moderate inflation in the United States from 1969 to 1979, the quantity of money rose at the average rate of 9 percent per year and prices at the average rate of 7 percent per year. The difference of two percentage points reflects the 2.8 percent average rate of growth of output over the same decade.

As these examples show, what happens to the quantity of money tends to dwarf what happens to output; hence our reference to inflation as a *monetary* phenomenon, without adding any qualification about output. These examples also show that there is not a precise one-to-one correspondence between the rate of monetary growth and the rate of inflation. However, to our knowledge there is no example in history of a substantial inflation that lasted for more than a brief time that was not accompanied by a roughly correspondingly rapid increase in the quantity of money; and no example of a rapid increase in the quantity of

money that was not accompanied by a roughly correspondingly substantial inflation.

A few charts (Figures 1–5) show the persistence of this relation in recent years. The solid line on each chart is the quantity of money per unit of output for the country in question, year by year from 1964 through 1977. The other line is the consumer price index. In order to make the two series comparable, both have been expressed as percentages of their average values over the period as a whole (1964–1977 = 100 for both lines). The two lines necessarily have the same average level, but there is nothing in the arithmetic that requires the two lines to be the same for any single year.

The two lines for the United States on Figure 1 are almost indistinguishable. As the remaining figures show, that is not special to the United States. Though the two lines differ more for some of the other countries than they do for the United States, for every country the two lines are remarkably similar. The different countries experienced very different rates of monetary growth. In every case, that difference was matched by a different rate of inflation. Brazil is the most extreme (Figure 5). It experienced more rapid monetary growth than any of the others, and also more rapid inflation.

Which causes which? Does the quantity of money grow rapidly because prices increase rapidly, or vice versa? One clue is that on most of the charts the number plotted for the quantity of money is for a year ending six months *earlier* than the year to which the matching price index corresponds. More decisive evidence is provided by examination of the institutional arrangements that determine the quantity of money in these countries and by a large number of historical episodes in which it is crystal clear which is cause and which is effect.

One dramatic example comes from the American Civil War. The South financed the war largely by the printing press, in the process producing an inflation that averaged 10 percent a month from October 1861 to March 1864. In an attempt to stem the inflation, the Confederacy enacted a monetary reform: "In May, 1864, the currency reform took hold, and the stock of money was reduced. Dramatically, the general price index dropped . . . in

Figure 1. MONEY AND PRICES UNITED STATES (1964-1977)

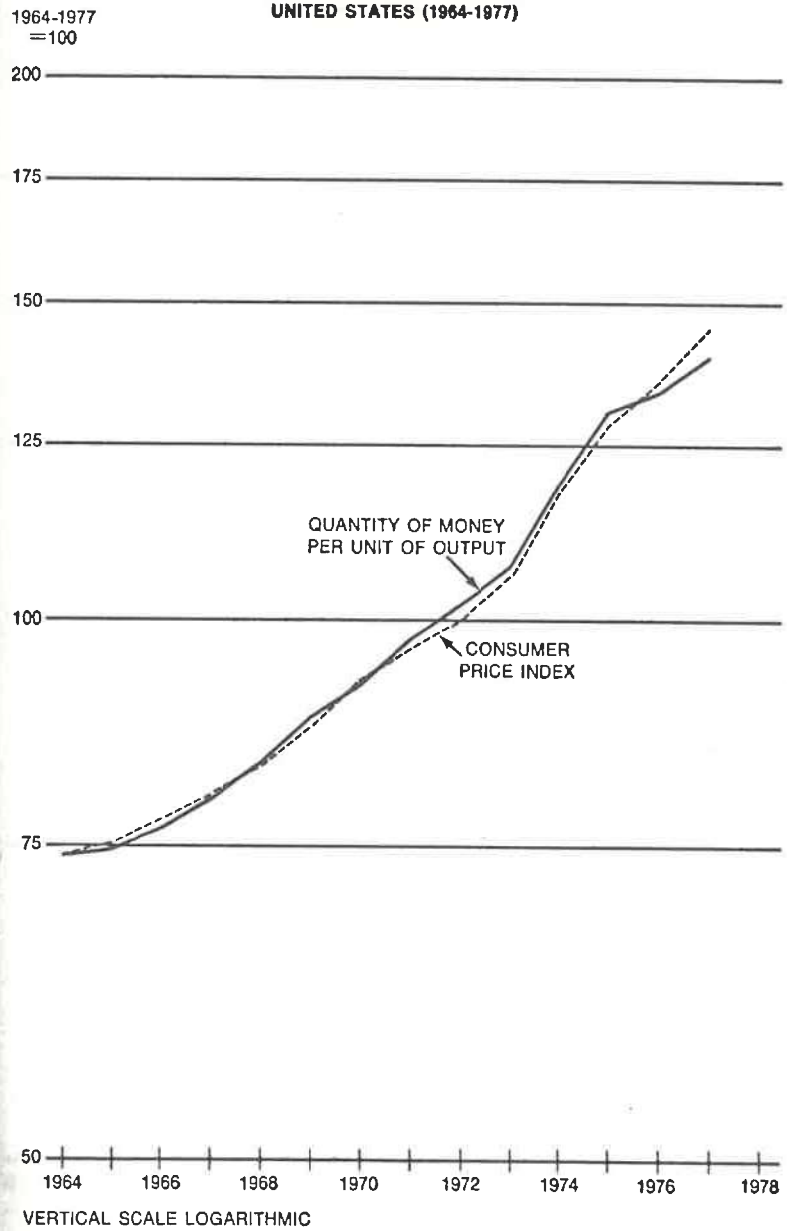


Figure 2. MONEY AND PRICES GERMANY (1964-1977)

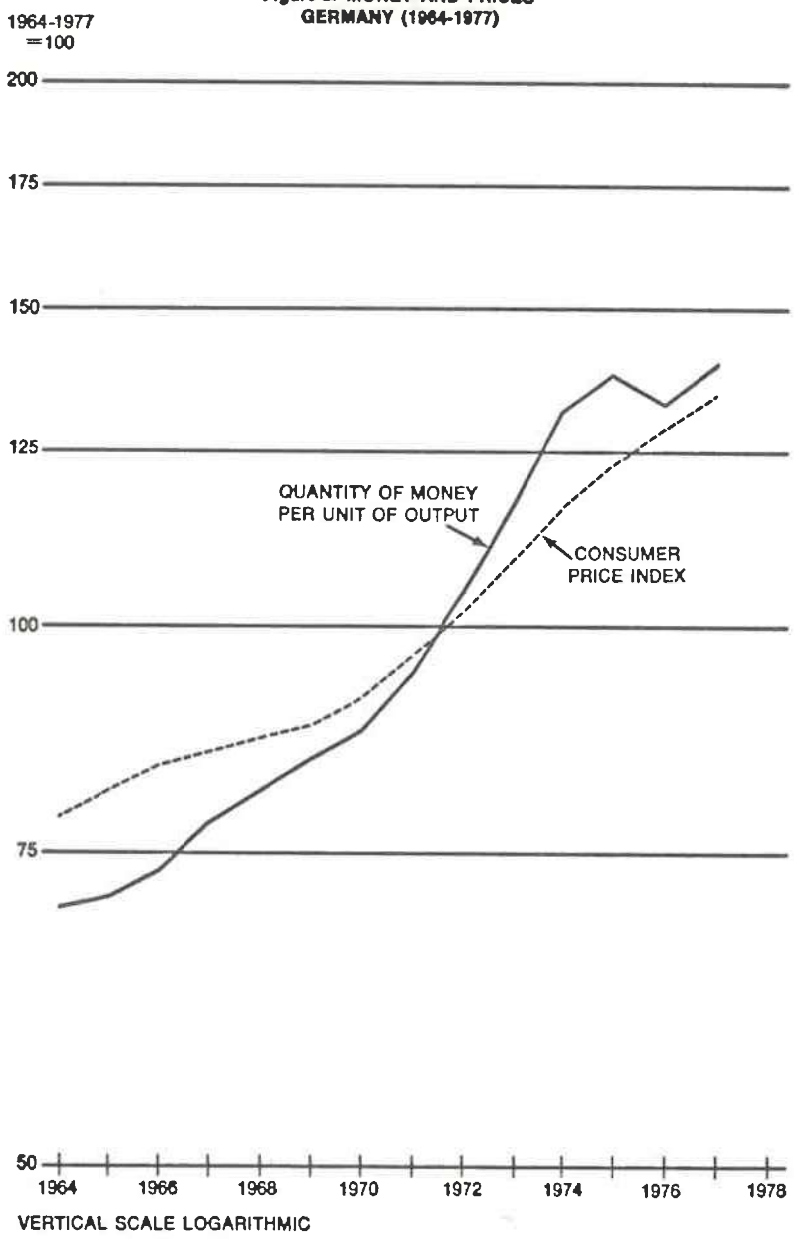


Figure 3. MONEY AND PRICES JAPAN (1964-1977)

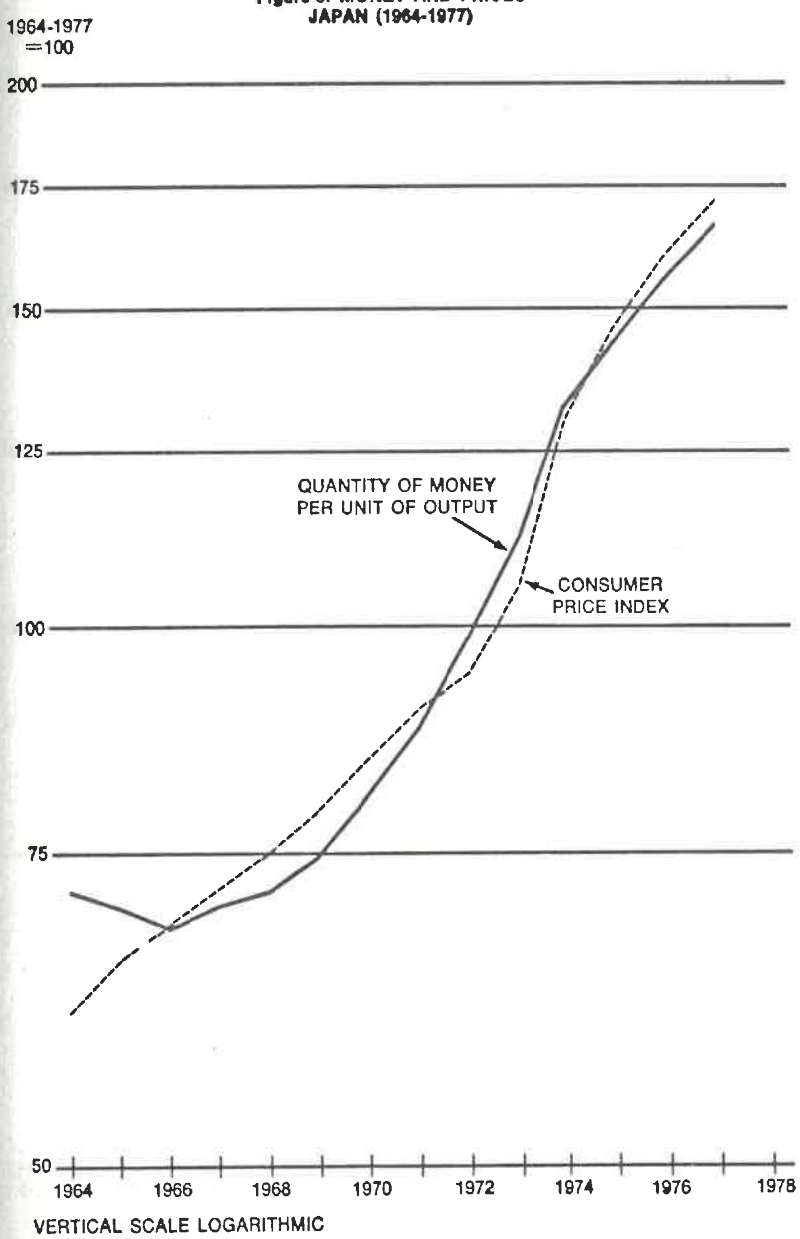


Figure 4. MONEY AND PRICES UNITED KINGDOM (1964-1977)

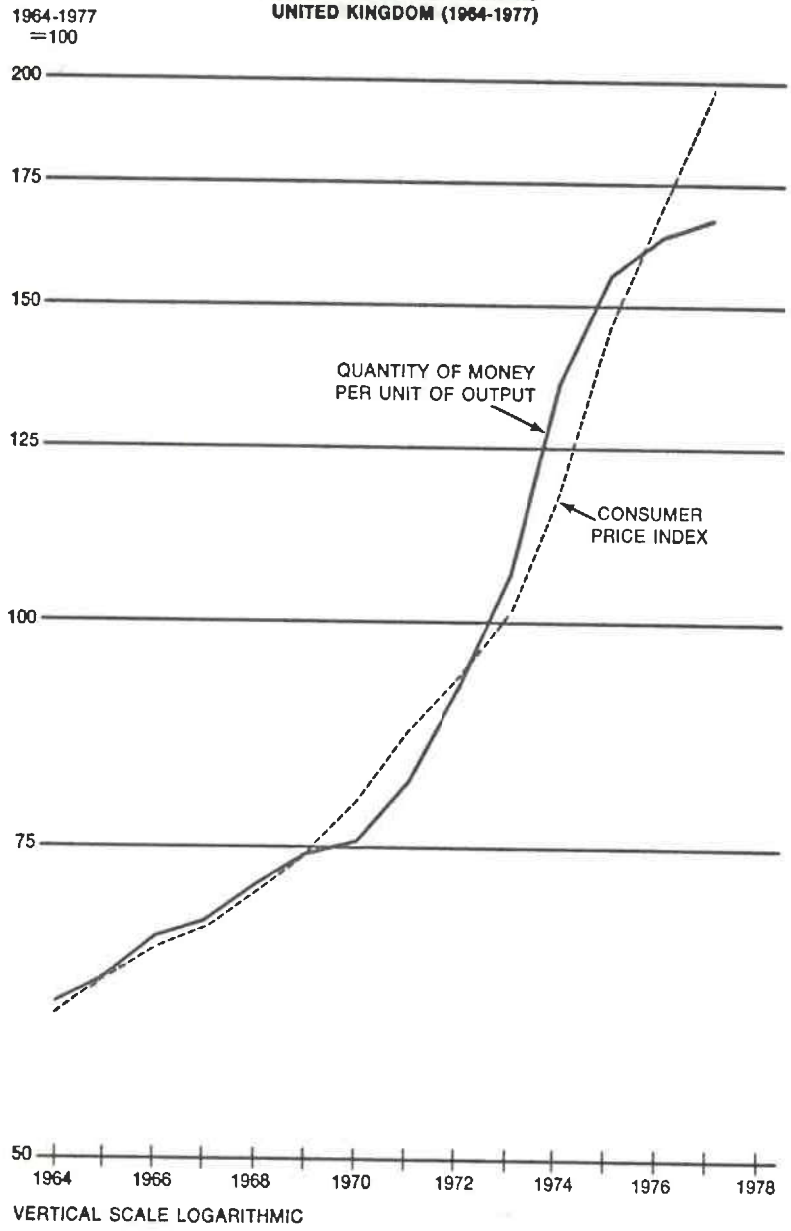
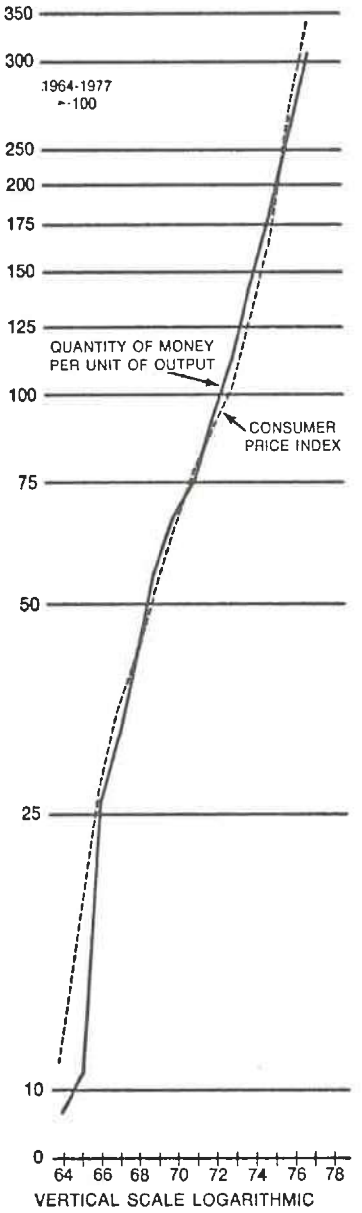


Figure 5. MONEY AND PRICES BRAZIL (1964-1977)



spite of invading Union armies, the impending military defeat, the reduction in foreign trade, the disorganized government, and the low morale of the Confederate army. Reducing the stock of money had a more significant effect on prices than these powerful forces." [11]

These charts dispose of many widely held explanations of inflation. Unions are a favorite whipping boy. They are accused of using their monopoly power to force up wages, which drive up costs, which drive up prices. But then how is it that the charts for Japan, where unions are of trivial importance, and for Brazil, where they exist only at the sufferance and under the close control of the government, show the same relation as the charts for the United Kingdom, where unions are stronger than in any of the other nations, and for Germany and the United States, where unions have considerable strength? Unions may provide useful services for their members. They may also do a great deal of harm by limiting employment opportunities for others, but they do not produce inflation. Wage increases in excess of increases in productivity are a result of inflation, rather than a cause.

Similarly, businessmen do not cause inflation. The rise in the prices they charge is a result or reflection of other forces. Businessmen are surely no more greedy in countries that have experienced much inflation than in countries that have experienced little, no more greedy at one period than another. Why then is inflation so much greater in some places and at some times than in other places and at other times?

Another favorite explanation of inflation, particularly among government officials seeking to shift blame, is that it is imported from abroad. That explanation was often correct when the currencies of the major countries were linked through a gold standard. Inflation was then an international phenomenon because many countries used the same commodity as money and anything that made the quantity of that commodity money grow more rapidly affected them all. But it clearly is not correct for recent years. If it were, how could the rates of inflation be so different in different countries? Japan and the United Kingdom experienced inflation at the rate of 30 percent or more a year in the early 1970s, when inflation in the United States was around 10 percent

and in Germany under 5 percent. Inflation is a worldwide phenomenon in the sense that it occurs in many countries at the same time—just as high goverment spending and large government deficits are worldwide phenomena. But inflation is not an international phenomenon in the sense that each country separately lacks the ability to control its own inflation—just as high government spending and large government deficits are not produced by forces outside each country's control.

Low productivity is another favorite explanation for inflation. Yet consider Brazil. It has experienced one of the most rapid rates of growth in output in the world—and also one of the highest rates of inflation. True enough, what matters for inflation is the quantity of money per unit of output, but as we have noted, as a practical matter, changes in output are dwarfed by changes in the quantity of money. Nothing is more important for the long-run economic welfare of a country than improving productivity. If productivity grows at 3.5 percent per year, output doubles in twenty years; at 5 percent per year, in fourteen years—quite a difference. But productivity is a bit player for inflation; money is center stage.

What about Arab sheikhs and OPEC? They have imposed heavy costs on us. The sharp rise in the price of oil lowered the quantity of goods and services that was available for us to use because we had to export more abroad to pay for oil. The reduction in output raised the price level. But that was a once-for-all effect. It did not produce any longer-lasting effect on the rate of inflation from that higher price level. In the five years after the 1973 oil shock, inflation in both Germany and Japan declined, in Germany from about 7 percent a year to less than 5 percent; in Japan from over 30 percent to less than 5 percent. In the United States inflation peaked a year after the oil shock at about 12 percent, declined to 5 percent in 1976, and then rose to over 13 percent in 1979. Can these very different experiences be explained by an oil shock that was common to all countries? Germany and Japan are 100 percent dependent on imported oil, yet they have done better at cutting inflation than the United States, which is only 50 percent dependent, or than the United Kingdom, which has become a major producer of oil.

We return to our basic proposition. Inflation is primarily a *monetary phenomenon*, produced by a more rapid increase in the quantity of money than in output. The behavior of the quantity of money is the senior partner; of output, the junior partner. Many phenomena can produce temporary fluctuations in the rate of inflation, but they can have lasting effects only insofar as they affect the rate of monetary growth.

WHY THE EXCESSIVE MONETARY GROWTH?

The proposition that inflation is a monetary phenomenon is important, yet it is only the beginning of an answer to the causes of and cures for inflation. It is important because it guides the search for basic causes and limits possible cures. But it is only the beginning of an answer because the deeper question is why excessive monetary growth occurs.

Whatever was true for tobacco money or money linked to silver and gold, with today's paper money, excessive monetary growth, and hence inflation, is produced by governments.

In the United States the accelerated monetary growth during the past fifteen years or so has occurred for three related reasons: first, the rapid growth in government spending; second, the government's full employment policy; third, a mistaken policy pursued by the Federal Reserve System.

Higher government spending will not lead to more rapid monetary growth and inflation *if* additional spending is financed either by taxes or by borrowing from the public. In that case, government has more to spend, the public has less. Higher government spending is matched by lower private spending for consumption and investment. However, taxing and borrowing from the public are politically unattractive ways to finance additional government spending. Many of us welcome the additional government spending; few of us welcome additional taxes. Government borrowing from the public diverts funds from private uses by raising interest rates, making it both more expensive and more difficult for individuals to get mortgages on new homes and for businesses to borrow money.

The only other way to finance higher government spending is

by increasing the quantity of money. As we noted in Chapter 3, the U.S. government can do that by having the U.S. Treasury—one branch of the government—sell bonds to the Federal Reserve System—another branch of the government. The Federal Reserve pays for the bonds either with freshly printed Federal Reserve Notes or by entering a deposit on its books to the credit of the U.S. Treasury. The Treasury can then pay its bills with either the cash or a check drawn on its account at the Fed. When the additional high-powered money is deposited in commercial banks by its initial recipients, it serves as reserves for them and as the basis for a much larger addition to the quantity of money.

Financing government spending by increasing the quantity of money is often extremely attractive to both the President and members of Congress. It enables them to increase government spending, providing goodies for their constituents, without having to vote for taxes to pay for them, and without having to borrow from the public.

A second source of higher monetary growth in the United States in recent years has been the attempt to produce full employment. The objective, as for so many government programs, is admirable, but the results have not been. "Full employment" is a much more complex and ambiguous concept than it appears to be on the surface. In a dynamic world, in which new products emerge and old ones disappear, demand shifts from one product to another, innovation alters methods of production, and so on without end, it is desirable to have a good deal of labor mobility. People change from one job to another and often are idle for a time in between. Some people leave a job they do not like before they have found another. Young people entering the labor force take time to find jobs and experiment with different kinds of jobs. In addition, obstacles to the free operation of the labor market—trade union restrictions, minimum wages, and the like—increase the difficulty of matching worker and job. Under these circumstances, what average number of persons employed corresponds to full employment?

As with spending and taxes, there is here, too, an asymmetry. Measures that can be represented as adding to employment are politically attractive. Measures that can be represented as adding

to unemployment are politically unattractive. The result is to impart a bias to government policy in the direction of adopting unduly ambitious targets of full employment.

The relation to inflation is twofold. First, government spending can be represented as adding to employment, government taxes as adding to unemployment by reducing private spending. Hence, the full employment policy reinforces the tendency for government to increase spending and lower taxes, and to finance any resulting deficit by increasing the quantity of money rather than by taxes or borrowing from the public. Second, the Federal Reserve System can increase the quantity of money in ways other than financing government spending. It can do so by buying outstanding government bonds, paying for them with newly created high-powered money. That enables the banks to make a larger volume of private loans, which can also be represented as adding to employment. Under pressure to promote full employment, the Fed's monetary policy has had the same inflationary bias as the government's fiscal policy.

These policies have not succeeded in producing full employment but they have produced inflation. As Prime Minister James Callaghan put it in a courageous talk to a British Labour party conference in September 1976: "We used to think that you could just spend your way out of a recession and increase employment by cutting taxes and boosting government spending. I tell you, in all candor, that that option no longer exists; and that insofar as it ever did exist, it only worked by injecting bigger doses of inflation into the economy followed by higher levels of unemployment as the next step. That is the history of the past twenty years."

The third source of higher monetary growth in the United States in recent years has been a mistaken policy by the Federal Reserve System. Not only has the Fed's policy had an inflationary bias because of pressures to promote full employment, but that bias has been exacerbated by its attempt to pursue two incompatible objectives. The Fed has the power to control the quantity of money and it gives lip service to that objective. But like Demetrius in Shakespeare's *A Midsummer Night's Dream*, who shuns Helena, who is in love with him, to pursue Hermia, who loves another, the Fed has given its heart not to controlling the quantity of money but to controlling interest rates, something that

it does not have the power to do. The result has been failure on both fronts: wide swings in both money and interest rates. These swings, too, have had an inflationary bias. With memories of its disastrous mistake from 1929 to 1933, the Fed has been much prompter in correcting a swing toward a low rate of monetary growth than in correcting a swing toward a high rate of monetary growth.

The end result of higher government spending, the full employment policy, and the Fed's obsession with interest rates has been a roller coaster along a rising path. Inflation has risen and then fallen. Each rise has carried inflation to a higher level than the preceding peak. Each fall has left inflation above its preceding trough. All the time, government spending has been rising as a fraction of income; government tax receipts, too, have been rising as a fraction of income, but not quite as fast as spending, so the deficit, too, has been rising as a fraction of income.

These developments are not unique to the United States or to recent decades. Since time immemorial, sovereigns—whether kings, emperors, or parliaments—have been tempted to resort to increasing the quantity of money to acquire resources to wage wars, construct monuments, or for other purposes. They have often succumbed to the temptation. Whenever they have, inflation followed close behind.

Nearly two thousand years ago the Roman Emperor Diocletian inflated by "debasing" the coinage—that is, replacing silver coins by look-alikes that had less and less silver and more and more of a worthless alloy until they became "no more than base metal washed over with silver." [12] Modern governments do so by printing paper money and making entries on books—but the ancient method has not entirely disappeared. The once full-bodied silver coins of the United States are now copper coins washed over, not even with silver, but with nickel. And a small-size Susan B. Anthony dollar coin has been introduced to replace what was once a full-bodied silver coin.

GOVERNMENT REVENUE FROM INFLATION

Financing government spending by increasing the quantity of money looks like magic, like getting something for nothing. To

take a simple example, government builds a road, paying for the expenses incurred with newly printed Federal Reserve Notes. It looks as if everybody is better off. The workers who build the road get their pay and can buy food, clothing, and housing with it. Nobody has paid higher taxes. Yet there is now a road where there was none before. Who has paid for it?

The answer is that all holders of money have paid for the road. The extra money raises prices when it is used to induce the workers to build the road instead of engage in some other productive activity. Those higher prices are maintained as the extra money circulates in the spending stream from the workers to the sellers of what they buy, from those sellers to others, and so on. The higher prices mean that the money people previously held will now buy less than it would have before. In order to have on hand an amount of money that can buy as much as before, they will have to refrain from spending all of their income and use part of it to add to their money balances.

The extra money printed is equivalent to a tax on money balances. If the extra money raises prices by 1 percent, then every holder of money has in effect paid a tax equal to 1 percent of his money holdings. The extra pieces of paper he now must hold (or book entries he must make) in order to have the same purchasing power in the form of money as before are indistinguishable from the other pieces of paper in his pocket or safe deposit box (or from book entries), but they are in effect receipts for taxes paid.

The physical counterpart to these taxes is the goods and services that could have been produced by the resources that built the road. The people who spent less than their income in order to maintain the purchasing power of their money balances have given up these goods and services in order that the government could get the resources to build the road.

You can see why John Maynard Keynes, in discussing the inflations after World War I, wrote: "There is no subtler, no surer means of overturning the existing basis of society than to debauch the currency. The process engages all the hidden forces of economic law on the side of destruction, and does it in a manner which not one man in a million is able to diagnose." [13]

The additional currency printed and the additional deposits

entered on the books of the Federal Reserve Bank correspond to only part of the revenue that government gets from inflation.

Inflation also yields revenue indirectly by automatically raising effective tax rates. As people's dollar incomes go up with inflation, the income is pushed into higher brackets and taxed at a higher rate. Corporate income is artificially inflated by inadequate allowance for depreciation and other costs. On the average, if income rises by 10 percent simply to match a 10 percent inflation, federal tax revenue tends to go up by more than 15 percent—so the taxpayer has to run faster and faster to stay in the same place. That process has enabled the President, Congress, state governors and legislatures to pose as tax cutters when all they have done is to keep taxes from going up as much as they otherwise would have gone up. Each year, there is talk of "cutting taxes." Yet there has been no reduction in taxes. On the contrary, taxes correctly measured have gone up—at the federal level from 22 percent of national income in 1964 to 25 percent in 1978; at the state and local level from 11 percent in 1964 to 15 percent in 1978.

Still a third way inflation yields revenue to the government is by paying off—or repudiating, if you will—part of the government's debt. Government borrows in dollars and pays back in dollars. But thanks to inflation, the dollars it pays back can buy less than the dollars it borrowed. That would not be a net gain to the government if in the interim it had paid a high enough interest rate on the debt to compensate the lender for inflation. But for the most part it did not. Savings bonds are the clearest example. Suppose you had bought a savings bond in December 1968, had held it to December 1978, and then cashed it in. You would have paid $37.50 in 1968 for a ten-year bond with a face value of $50 and you would have received $64.74 when you cashed it in 1978 (because the government raised the interest rate in the interim to make some allowance for inflation). By 1978 it took $70 to buy as much as $37.50 would have bought in 1968. Yet not only would you have gotten back only $64.74, you would have had to pay income tax on the $27.24 difference between what you received and what you paid. You would have ended up paying for the dubious privilege of lending to your government.

Paying off the debt by inflation has meant that although the

federal government has run large deficits year after year and its debt in terms of dollars has gone up, the debt has gone up far less in terms of purchasing power and has actually fallen as a percentage of the national income. In the decade from 1968 through 1978, the federal government had a cumulative deficit of more than $260 billion, yet the debt amounted to 30 percent of national income in 1968, to 28 percent in 1978.

THE CURE FOR INFLATION

The cure for inflation is simple to state but hard to implement. Just as an excessive increase in the quantity of money is the one and only important cause of inflation, so a reduction in the rate of monetary growth is the one and only cure for inflation. The problem is not one of knowing what to do. That is easy enough. Government must increase the quantity of money less rapidly. The problem is to have the political will to take the measures necessary. Once the inflationary disease is in an advanced state, the cure takes a long time and has painful side effects.

Two medical analogies suggest the problem. One is about a young man who had Buerger's disease, a disease that interrupts the blood supply and can lead to gangrene. The young man was losing fingers and toes. The cure was simple to state: stop smoking. The young man did not have the will to do so; his addiction to tobacco was simply too great. His disease was in one sense curable, in another not.

A more instructive analogy is between inflation and alcoholism. When the alcoholic starts drinking, the good effects come first; the bad effects only come the next morning when he wakes up with a hangover—and often cannot resist easing the hangover by taking "the hair of the dog that bit him."

The parallel with inflation is exact. When a country starts on an inflationary episode, the initial effects seem good. The increased quantity of money enables whoever has access to it—nowadays, primarily governments—to spend more without anybody else having to spend less. Jobs become more plentiful, business is brisk, almost everybody is happy—at first. Those are the good effects. But then the increased spending starts to raise prices; workers find that their wages, even if higher in dollars, will buy less; businessmen find that their costs have risen, so that the extra sales are not as profitable as they anticipated, unless they can raise their prices even faster. The bad effects start to emerge: higher prices, less buoyant demand, inflation combined with stagnation. As with the alcoholic, the temptation is to increase the quantity of money still faster, which produces the roller coaster we have been on. In both cases, it takes a larger and larger amount—of alcohol or money—to give the alcoholic or the economy the same "kick."

The parallel between alcoholism and inflation carries over to the cure. The cure for alcoholism is simple to state: stop drinking. It is hard to take because, this time, the bad effects come first, the good effects come later. The alcoholic who goes on the wagon suffers severe withdrawal pains before he emerges in the happy land of no longer having an almost irresistible desire for another drink. So also with inflation. The initial side effects of a slower rate of monetary growth are painful: lower economic growth, temporarily high unemployment, without, for a time, much reduction of inflation. The benefits appear only after one or two years or so, in the form of lower inflation, a healthier economy, the potential for rapid noninflationary growth.

Painful side effects are one reason why it is difficult for an alcoholic or an inflationary nation to end its addiction. But there is another reason, which, at least in the earlier stage of the disease, may be even more important: the lack of a real desire to end the addiction. The drinker enjoys his liquor; he finds it hard to accept that he really is an alcoholic; he is not sure he wants to take the cure. The inflationary nation is in the same position. It is tempting to believe that inflation is a temporary and mild matter produced by unusual and extraneous circumstances, and that it will go away of its own accord—something that never happens.

Moreover, many of us enjoy inflation. We would naturally like to see the prices of the things we *buy* go down, or at least stop going up. But we are more than happy to see the prices of the things we *sell* go up—whether goods we produce, our labor services, or houses or other items we own. Farmers complain about inflation but congregate in Washington to lobby for higher prices for their products. Most of the rest of us do the same in one way or another.

One reason inflation is so destructive is because some people

benefit greatly while other people suffer; society is divided into winners and losers. The winners regard the good things that happen to them as the natural result of their own foresight, prudence, and initiative. They regard the bad things, the rise in the prices of the things they buy, as produced by forces outside their control. Almost everyone will say that he is against inflation; what he generally means is that he is against the bad things that have happened to him.

To take a specific example, almost every person who has owned a home during the past two decades has benefited from inflation. The value of his home has risen sharply. If he had a mortgage, the interest rate was generally below the rate of inflation. As a result the payments called "interest," as well as those called "principal," have in effect been paying off the mortgage. To take a simple example, suppose both the interest rate and inflation rate were 7 percent in one year. If you had a $10,000 mortgage on which you paid only interest, a year later the mortgage would correspond to the same buying power as $9,300 would have a year earlier. In real terms you would owe $700 less—just the amount you paid as interest. In real terms you would have paid nothing for the use of the $10,000. (Indeed, because the interest is deductible in computing your income tax, you would actually benefit. You would have been paid for borrowing.) The way this effect becomes apparent to the homeowner is that his equity in the house goes up rapidly. The counterpart is a loss to the small savers who provided the funds that enabled savings and loan associations, mutual savings banks, and other institutions to finance mortgage loans. The small savers had no good alternative because government limits narrowly the maximum interest rate that such institutions can pay to their depositors—supposedly to protect the depositors.

Just as high government spending is one reason for excessive monetary growth, so lower government spending is one element that can contribute to reducing monetary growth. Here, too, we tend to be schizophrenic. We would all like to see government spending go down, provided it is not spending that benefits us. We would all like to see deficits reduced, provided it is through taxes imposed on others.

As inflation accelerates, however, sooner or later it does so much damage to the fabric of society, creates so much injustice and suffering, that a real public will develops to do something about inflation. The level of inflation at which that occurs depends critically on the country in question and its history. In Germany it came at a low level of inflation because of Germany's terrible experiences after World War I and II; it came at a much higher level of inflation in the United Kingdom and Japan; it has not yet come in the United States.

SIDE EFFECTS OF A CURE

We read over and over again that higher unemployment and slow growth are cures for inflation, that the alternatives we must face are more inflation *or* higher unemployment, that the powers that be are reconciled to, or are positively promoting, slower growth and higher unemployment in order to cure inflation. Yet over the past several decades, the growth of the U.S. economy has slowed, the average level of unemployment has risen, and at the same time, the rate of inflation has moved higher and higher. We have had both more inflation and more unemployment. Other countries have had the same experience. How come?

The answer is that slow growth and high unemployment are not *cures* for inflation. They are *side effects* of a successful cure. Many policies that impede economic growth and add to unemployment may, at the same time, increase the rate of inflation. That has been true of some of the policies we have adopted—sporadic price and wage control, increasing government intervention into business, all accompanied by higher and higher government spending, and a rapid increase in the quantity of money.

Another medical example will perhaps make clear the difference between a *cure* and a *side effect*. You have acute appendicitis. Your physician recommends an appendectomy but warns you that after the operation you will be confined to bed for an interval. You refuse the operation but take to your bed for the indicated interval as a less painful *cure*. Silly, yes, but faithful in every detail to the confusion between unemployment as a side effect and as a cure.

The side effects of a cure for inflation are painful so it is important to understand why they occur and to seek means to mitigate them. The basic reason why the side effects occur has already been pointed out in Chapter 1. They occur because variable rates of monetary growth introduce static into the information transmitted by the price system, static that is translated into inappropriate responses by the economic actors, which it takes time to overcome.

Consider, first, what happens when inflationary monetary growth starts. The higher spending financed by the newly created money is no different to the seller of goods or labor or other services from any other spending. The seller of pencils, for example, finds that he can sell more pencils at the former price. He does so initially without changing his price. He orders more pencils from the wholesaler, the wholesaler from the manufacturer, and so on down the line. *If* the demand for pencils had increased at the expense of some other segment of demand, say at the expense of the demand for ball-point pens, rather than as a result of inflationary monetary growth, the increased flow of orders down the pencil channel would be accompanied by a decreased flow down the ball-point pen channel. Pencils and later the materials used to make them would tend to rise in price; pens and the materials used to make them would tend to fall in price; but there would be no reason for prices *on the average* to change.

The situation is wholly different when the increased demand for pencils has its origin in newly created money. The demand for pencils and pens and most other things can then go up simultaneously. There is more spending (in dollars) in total. However, the seller of pencils does not know this. He proceeds as before, initially holding the price at which he sells constant, content to sell more until, as he believes, he will be able to restock. But now the increased flow of orders down the pencil channel is accompanied by an increased flow down the pen channel, and down many other channels. As the increased flow of orders generates a greater demand for labor and materials to produce more, the initial reaction of workers and producers of materials will be like that of the retailers—to work longer and produce more and also charge more in the belief that the demand for what they have

been providing has gone up. But this time there is no offset, there are no declines in demand roughly matching the increases in demand, no declines in prices matching the increases. Of course, this will not at first be obvious. In a dynamic world demands are always shifting, some prices going up, some going down. The general signal of increasing demand will be confused with the specific signals reflecting changes in relative demands. That is why the initial side effect of faster monetary growth is an appearance of prosperity and greater employment. But sooner or later the signal will get through.

As it does, workers, manufacturers, retailers will discover that they have been fooled. They reacted to higher demand for the small number of things they sell in the mistaken belief that the higher demand was special to them and hence would not much affect the prices of the many things they buy. When they discover their mistake, they raise wages and prices still higher—not only to respond to higher demand but also to allow for the rises in the prices of the things they buy. We are off on a price-wage spiral that is itself an effect of inflation, not a cause. If monetary growth does not speed up further, the initial stimulus to employment and output will be replaced by the opposite; both will tend to go down in response to the higher wages and prices. A hangover will succeed the initial euphoria.

It takes time for these reactions to occur. On the average over the past century and more in the United States, the United Kingdom, and some other Western countries, roughly six to nine months have elapsed before increased monetary growth has worked its way through the economy and produced increased economic growth and employment. Another twelve to eighteen months have elapsed before the increased monetary growth has affected the price level appreciably and inflation has occurred or speeded up. The time delays have been this long for these countries because, wartime aside, they were long spared widely varying rates of monetary growth and inflation. On the eve of World War II wholesale prices in the United Kingdom averaged roughly the same as two hundred years earlier, and in the United States, as one hundred years earlier. The post–World War II inflation is a new phenomenon in these countries. They have experienced

many ups and downs but not a long movement in the same direction.

Many countries in South America have had a less happy heritage. They experience much shorter time delays—amounting at most to a few months. If the United States does not cure its recent propensity to indulge in widely varying rates of inflation, the time delays will shorten here as well.

The sequence of events that follows a slowing of monetary growth is the same as that just outlined except in the opposite direction. The initial reduction in spending is interpreted as a reduction in demand for specific products, which after an interval leads to a reduction in output and employment. After another interval inflation slows, which in turn is accompanied by an expansion in employment and output. The alcoholic is through his worst withdrawal pains and on the road to contented abstinence.

All of these adjustments are set in motion by *changes* in the rates of monetary growth and inflation. If monetary growth were high and steady, so that, let us say, prices tended to rise year after year by 10 percent, the economy could adjust to it. Everybody would come to anticipate a 10 percent inflation; wages would rise by 10 percent a year more than they otherwise would; interest rates would be 10 percentage points higher than otherwise—in order to compensate the lender for inflation; tax rates would be adjusted for inflation, and so on and on.

Such an inflation would do no great harm, but neither would it serve any function. It would simply introduce unnecessary complexities in arrangements. More important, such a situation, if it ever developed, would probably not be stable. If it were politically profitable and feasible to generate a 10 percent inflation, the temptation would be great, when and if inflation ever settled there, to make the inflation 11 or 12 or 15 percent. Zero inflation is a politically feasible objective; a 10 percent inflation is not. That is the verdict of experience.

MITIGATING THE SIDE EFFECTS

We know no example in history in which an inflation has been ended without an interim period of slow economic growth and

higher than usual unemployment. That is the basis in experience for our judgment that there is no way to avoid side effects of a cure for inflation.

However, it is possible to mitigate those side effects, to make them milder.

The most important device for mitigating the side effects is to slow inflation *gradually but steadily* by a policy announced in advance and adhered to so it becomes credible.

The reason for gradualness and advance announcement is to give people time to readjust their arrangements—and to induce them to do so. Many people have entered into long-term contracts—for employment, to lend or borrow money, to engage in production or construction—on the basis of *anticipations* about the likely rate of inflation. These long-term contracts make it difficult to reduce inflation rapidly and mean that trying to do so will impose heavy costs on many people. Given time, these contracts will be completed or renewed or renegotiated, and can then be adjusted to the new situation.

One other device has proved effective in mitigating the adverse side effects of curing inflation—including an automatic adjustment for inflation in longer-term contracts, what are known as escalator clauses. The most common example is the cost-of-living adjustment clause that is included in many wage contracts. Such a contract specifies that the hourly wage shall increase by, say, 2 percent plus the rate of inflation or plus a fraction of the rate of inflation. In that way, if inflation is low, the wage increase in dollars is low; if inflation is high, the wage increase in dollars is high; but in either case the wage has the same purchasing power.

Another example is for contracts for the rental of property. Instead of being stated as a fixed number of dollars, the rental contract may specify that the rent shall be adjusted from year to year by the rate of inflation. Rental contracts for retail stores often specify the rent as a percentage of the gross receipts of the store. Such contracts have no explicit escalator clause but implicitly they do, since the store's receipts will tend to rise with inflation.

Still another example is for a loan. A loan is typically for a fixed dollar sum for a fixed period at a fixed annual rate of interest, say, $1,000 for one year at 10 percent. An alternative is

to specify the rate of interest not at 10 percent but, say, 2 percent plus the rate of inflation, so that if inflation turns out to be 5 percent, the interest rate will be 7 percent; if inflation turns out to be 10 percent, the interest rate will be 12 percent. An alternative that is roughly equivalent is to specify the amount to be repaid not as a fixed number of dollars but as a number of dollars adjusted for inflation. In our simple example the borrower would owe $1,000 increased by the rate of inflation plus interest at 2 percent. If inflation turned out to be 5 percent, he would owe $1,050; if 10 percent, $1,100; in both cases plus interest at 2 percent.

Except for wage contracts, escalator clauses have not been common in the United States. However, they are spreading, especially in the form of variable interest mortgages. And they have been common in just about all countries that have experienced both high and variable rates of inflation over any extensive period.

Such escalator clauses reduce the time delay between slowing down monetary growth and the subsequent adjustment of wages and prices. In that way they shorten the transition period and reduce the interim side effects. However, useful though they are, escalator clauses are far from a panacea. It is impossible to escalate *all* contracts (consider, for example, paper money), and costly to escalate many. A major advantage of using money is precisely the ability to carry on transactions cheaply and efficiently, and universal escalator clauses reduce this advantage. Far better to have no inflation and no escalator clauses. That is why we advocate resort to escalator clauses in the private economy only as a device for easing the side effects of curing inflation, not as a permanent measure.

Escalator clauses are highly desirable as a permanent measure in the federal government sector. Social Security and other retirement benefits, salaries of federal employees, including the salaries of members of Congress, and many other items of government spending are now automatically adjusted for inflation. However, there are two glaring and inexcusable gaps: income taxes and government borrowing. Adjusting the personal and corporate tax structure for inflation—so that a 10 percent price rise would raise taxes in dollars by 10 percent, not, as it does now, by something over 15 percent on the average—would eliminate the imposition of higher taxes without their having been voted. It would end this taxation without representation. By so doing, it would also reduce the incentive for the government to inflate, since the revenue from inflation would be reduced.

The case for inflation-proofing government borrowing is equally strong. The U.S. government has itself produced the inflation that has made the purchase of long-term government bonds such a poor investment in recent years. Fairness and honesty toward citizens on the part of their government require introducing escalator clauses into long-term government borrowing.

Price and wage controls are sometimes proposed as a cure for inflation. Recently, as it has become clear that controls are not a cure, they have been urged as a device for mitigating the side effects of a cure. It is claimed that they will serve this function by persuading the public that the government is serious in attacking inflation. That, in turn, is expected to lower the anticipations of future inflation that are built into the terms of long-term contracts.

Price and wage controls are counterproductive for this purpose. They distort the price structure, which reduces the efficiency with which the system works. The resulting lower output adds to the adverse side effects of a cure for inflation rather than reducing them. Price and wage controls waste labor, both because of the distortions in the price structure and because of the immense amount of labor that goes into constructing, enforcing, and evading the price and wage controls. These effects are the same whether controls are compulsory or are labeled "voluntary."

In practice, price and wage controls have almost always been used as a substitute for monetary and fiscal restraint, rather than as a complement to them. This experience has led participants in the market to regard the imposition of price and wage controls as a signal that inflation is heading up, not down. It has therefore led them to raise their inflation expectations rather than to lower them.

Price and wage controls often seem effective for a brief period after they are imposed. Quoted prices, the prices that enter into index numbers, are kept down because there are indirect ways of raising prices and wages—lowering the quality of items produced, eliminating services, promoting workers, and so on. But then, as

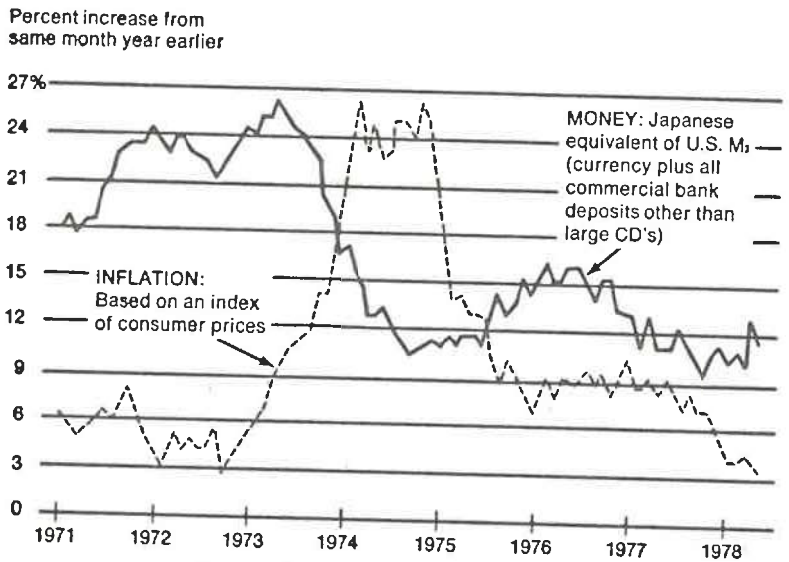
the easy ways of avoiding the controls are exhausted, distortions accumulate, the pressures suppressed by the controls reach the boiling point, the adverse effects get worse and worse, and the whole program breaks down. The end result is more inflation, not less. In light of the experience of forty centuries, only the short time perspective of politicians and voters can explain the repeated resort to price and wage controls.[14]

A CASE STUDY

Japan's recent experience provides an almost textbook illustration of how to cure inflation. As Figure 6 shows, the quantity of money in Japan began growing at higher and higher rates in 1971, and by mid-1973, it was growing more than 25 percent a year.[15]

Inflation did not respond until about two years later, in early 1973. The subsequent dramatic rise in inflation produced a funda-

Figure 6. INFLATION FOLLOWS MONEY: THE CASE OF JAPAN



Source: Japanese Economic Planning Agency

mental change in monetary policy. Emphasis shifted from the external value of the yen—the exchange rate—to its internal value—inflation. Monetary growth was reduced sharply, from more than 25 percent a year to between 10 and 15 percent. It was kept there, with minor exceptions, for five years. (Because of Japan's high rate of economic growth, monetary growth in this range would produce roughly stable prices. The comparable rate for the United States is 3 to 5 percent.)

About eighteen months after monetary growth started declining, inflation followed suit, but it took two and a half years before inflation fell below double digits. Inflation then held roughly constant for about two years—despite a mild upturn in monetary growth. Inflation then started moving rapidly toward zero in response to a new decline in monetary growth.

The numbers on inflation in the chart are for consumer prices. Wholesale prices did even better. They actually declined after mid-1977. The postwar shift of workers in Japan from low-productivity sectors to high-productivity sectors, such as automobiles and electronics, has meant that prices of services have risen sharply relative to prices of commodities. As a result, consumer prices have risen relative to wholesale prices.

Japan experienced lower growth and higher unemployment after it slowed monetary growth, particularly during 1974 before inflation started to respond appreciably to the slower monetary growth. The low point was reached at the end of 1974. Output then began recovering and grew thereafter—more modestly than in the boom years of the 1960s but at a highly respectable rate nonetheless: more than 5 percent per year.

Price and wage controls were not imposed at any time during the tapering down of inflation. And the tapering down occurred at the same time that Japan was adjusting to higher prices for crude oil.

CONCLUSIONS

Five simple truths embody most of what we know about inflation:

1. Inflation is a monetary phenomenon arising from a more

rapid increase in the quantity of money than in output (though, of course, the reasons for the increase in money may be various).

2. In today's world government determines—or can determine —the quantity of money.

3. There is only one cure for inflation: a slower rate of increase in the quantity of money.

4. It takes time—measured in years, not months—for inflation to develop; it takes time for inflation to be cured.

5. Unpleasant side effects of the cure are unavoidable.

The United States has embarked on rising monetary growth four times during the past twenty years. Each time the higher monetary growth has been followed first by economic expansion, later by inflation. Each time the authorities have slowed monetary growth in order to stem inflation. Lower monetary growth has been followed by an inflationary recession. Later still, inflation has declined and the economy has improved. So far the sequence is identical with Japan's experience from 1971 to 1975. Unfortunately, the crucial difference is that we have not displayed the patience Japan did by continuing monetary restraint long enough. Instead, we have overreacted to the recession by accelerating monetary growth, setting off on another round of inflation, and condemning ourselves to higher inflation plus higher unemployment.

We have been misled by a false dichotomy: inflation or unemployment. That option is an illusion. The real option is only whether we have higher unemployment as a result of higher inflation or as a temporary side effect of curing inflation.

CHAPTER 10

The Tide Is Turning

The failure of Western governments to achieve their proclaimed objectives has produced a widespread reaction against big government. In Britain the reaction swept Margaret Thatcher to power in 1979 on a platform pledging her Conservative government to reverse the socialist policies that had been followed by both Labour and earlier Conservative governments ever since the end of World War II. In Sweden in 1976, the reaction led to the defeat of the Social Democratic party after more than four decades of uninterrupted rule. In France the reaction led to a dramatic change in policy designed to eliminate government control of prices and wages and sharply reduce other forms of government intervention. In the United States the reaction has been manifested most dramatically in the tax revolt that has swept the nation, symbolized by the passage of Proposition 13 in California, and realized in a number of states in constitutional amendments limiting state taxes.

The reaction may prove short-lived and be followed, after a brief interval, by a resumption of the trend toward ever bigger government. The widespread enthusiasm for reducing government taxes and other impositions is not matched by a comparable enthusiasm for eliminating government programs—except programs that benefit other people. The reaction against big government has been sparked by rampant inflation, which governments can control if they find it politically profitable to do so If they do, the reaction might be muted or disappear.

We believe that the reaction is more than a response to transitory inflation. On the contrary, the inflation itself is partly a response to the reaction. As it has become politically less attractive to vote higher taxes to pay for higher spending, legislators have resorted to financing spending through inflation, a hidden tax that can be imposed without having been voted, taxation without rep-

GOLD AND ECONOMIC FREEDOM
Alan Greenspan, 1967

An almost hysterical antagonism toward the gold standard is one issue which unites statists of all persuasions. They seem to sense-perhaps more clearly and subtly than many consistent defenders of laissez-faire -- that gold and economic freedom are inseparable, that the gold standard is an instrument of laissez-faire and that each implies and requires the other.

In order to understand the source of their antagonism, it is necessary first to understand the specific role of gold in a free society.

Money is the common denominator of all economic transactions. It is that commodity which serves as a medium of exchange, is universally acceptable to all participants in an exchange economy as payment for their goods or services, and can, therefore, be used as a standard of market value and as a store of value, i.e., as a means of saving.

The existence of such a commodity is a precondition of a division of labor economy. If men did not have some commodity of objective value which was generally acceptable as money, they would have to resort to primitive barter or be forced to live on self-sufficient farms and forgo the inestimable advantages of specialization. If men had no means to store value, i.e., to save, neither long-range planning nor exchange would be possible.

What medium of exchange will be acceptable to all participants in an economy is not determined arbitrarily. First, the medium of exchange should be durable. In a primitive society of meager wealth, wheat might be sufficiently durable to serve as a medium, since all exchanges would occur only during and immediately after the harvest, leaving no value-surplus to store. But where store-of-value considerations are important, as they are in richer, more civilized societies, the medium of exchange must be a durable commodity, usually a metal. A metal is generally chosen because it is homogeneous and divisible: every unit is the same as every other and it can be blended or formed in any quantity. Precious jewels, for example, are neither homogeneous nor divisible. More important, the commodity chosen as a medium must be a luxury. Human desires for luxuries are unlimited and, therefore, luxury goods are always in demand and will always be acceptable. Wheat is a luxury in underfed civilizations, but not in a prosperous society. Cigarettes ordinarily would not serve as money, but they did in post-World War II Europe where they were considered a luxury. The term "luxury good" implies scarcity and high unit value. Having a high unit value, such a good is easily portable; for instance, an ounce of gold is worth a half-ton of pig iron.

In the early stages of a developing money economy, several media of exchange might be used, since a wide variety of commodities would fulfill the foregoing conditions. However, one of the commodities will gradually displace all others, by being more widely acceptable. Preferences on what to hold as a store of value, will shift to the most widely acceptable commodity, which, in turn, will make it still more acceptable. The shift is progressive until that commodity becomes the sole medium of exchange. The use of a

single medium is highly advantageous for the same reasons that a money economy is superior to a barter economy: it makes exchanges possible on an incalculably wider scale.

Whether the single medium is gold, silver, seashells, cattle, or tobacco is optional, depending on the context and development of a given economy. In fact, all have been employed, at various times, as media of exchange. Even in the present century, two major commodities, gold and silver, have been used as international media of exchange, with gold becoming the predominant one. Gold, having both artistic and functional uses and being relatively scarce, has significant advantages over all other media of exchange. Since the beginning of World War I, it has been virtually the sole international standard of exchange. If all goods and services were to be paid for in gold, large payments would be difficult to execute and this would tend to limit the extent of a society's divisions of labor and specialization. Thus a logical extension of the creation of a medium of exchange is the development of a banking system and credit instruments (bank notes and deposits) which act as a substitute for, but are convertible into, gold.

A free banking system based on gold is able to extend credit and thus to create bank notes (currency) and deposits, according to the production requirements of the economy. Individual owners of gold are induced, by payments of interest, to deposit their gold in a bank (against which they can draw checks). But since it is rarely the case that all depositors want to withdraw all their gold at the same time, the banker need keep only a fraction of his total deposits in gold as reserves. This enables the banker to loan out more than the amount of his gold deposits (which means that he holds claims to gold rather than gold as security of his deposits). But the amount of loans which he can afford to make is not arbitrary: he has to gauge it in relation to his reserves and to the status of his investments.

When banks loan money to finance productive and profitable endeavors, the loans are paid off rapidly and bank credit continues to be generally available. But when the business ventures financed by bank credit are less profitable and slow to pay off, bankers soon find that their loans outstanding are excessive relative to their gold reserves, and they begin to curtail new lending, usually by charging higher interest rates. This tends to restrict the financing of new ventures and requires the existing borrowers to improve their profitability before they can obtain credit for further expansion. Thus, under the gold standard, a free banking system stands as the protector of an economy's stability and balanced growth.

When gold is accepted as the medium of exchange by most or all nations, an unhampered free international gold standard serves to foster a world-wide division of labor and the broadest international trade. Even though the units of exchange (the dollar, the pound, the franc, etc.) differ from country to country, when all are defined in terms of gold the economies of the different countries act as one -- so long as there are no restraints on trade or on the movement of capital. Credit, interest rates, and prices tend to follow similar patterns in all countries. For example, if banks in one country extend credit too liberally, interest rates in that country will tend to fall, inducing depositors to shift their gold to higher-interest paying banks in other countries. This will immediately cause a

shortage of bank reserves in the "easy money" country, inducing tighter credit standards and a return to competitively higher interest rates again.

A fully free banking system and fully consistent gold standard have not as yet been achieved. But prior to World War I, the banking system in the United States (and in most of the world) was based on gold and even though governments intervened occasionally, banking was more free than controlled. Periodically, as a result of overly rapid credit expansion, banks became loaned up to the limit of their gold reserves, interest rates rose sharply, new credit was cut off, and the economy went into a sharp, but short-lived recession. (Compared with the depressions of 1920 and 1932, the pre-World War I business declines were mild indeed.) It was limited gold reserves that stopped the unbalanced expansions of business activity, before they could develop into the post-World Was I type of disaster. The readjustment periods were short and the economies quickly reestablished a sound basis to resume expansion.

But the process of cure was misdiagnosed as the disease: if shortage of bank reserves was causing a business decline-argued economic interventionists -- why not find a way of supplying increased reserves to the banks so they never need be short! If banks can continue to loan money indefinitely -- it was claimed -- there need never be any slumps in business. And so the Federal Reserve System was organized in 1913. It consisted of twelve regional Federal Reserve banks nominally owned by private bankers, but in fact government sponsored, controlled, and supported. Credit extended by these banks is in practice (though not legally) backed by the taxing power of the federal government. Technically, we remained on the gold standard; individuals were still free to own gold, and gold continued to be used as bank reserves. But now, in addition to gold, credit extended by the Federal Reserve banks ("paper reserves") could serve as legal tender to pay depositors.

When business in the United States underwent a mild contraction in 1927, the Federal Reserve created more paper reserves in the hope of forestalling any possible bank reserve shortage. More disastrous, however, was the Federal Reserve's attempt to assist Great Britain who had been losing gold to us because the Bank of England refused to allow interest rates to rise when market forces dictated (it was politically unpalatable). The reasoning of the authorities involved was as follows: if the Federal Reserve pumped excessive paper reserves into American banks, interest rates in the United States would fall to a level comparable with those in Great Britain; this would act to stop Britain's gold loss and avoid the political embarrassment of having to raise interest rates.

The "Fed" succeeded; it stopped the gold loss, but it nearly destroyed the economies of the world in the process. The excess credit which the Fed pumped into the economy spilled over into the stock market -- triggering a fantastic speculative boom. Belatedly, Federal Reserve officials attempted to sop up the excess reserves and finally succeeded in braking the boom. But it was too late: by 1929 the speculative imbalances had become so overwhelming that the attempt precipitated a sharp retrenching and a consequent demoralizing of business confidence. As a result, the American economy collapsed. Great Britain fared even worse, and rather than absorb the full consequences of her previous

folly, she abandoned the gold standard completely in 1931, tearing asunder what remained of the fabric of confidence and inducing a world-wide series of bank failures. The world economies plunged into the Great Depression of the 1930's.

With a logic reminiscent of a generation earlier, statists argued that the gold standard was largely to blame for the credit debacle which led to the Great Depression. If the gold standard had not existed, they argued, Britain's abandonment of gold payments in 1931 would not have caused the failure of banks all over the world. (The irony was that since 1913, we had been, not on a gold standard, but on what may be termed "a mixed gold standard"; yet it is gold that took the blame.) But the opposition to the gold standard in any form -- from a growing number of welfare-state advocates -- was prompted by a much subtler insight: the realization that the gold standard is incompatible with chronic deficit spending (the hallmark of the welfare state). Stripped of its academic jargon, the welfare state is nothing more than a mechanism by which governments confiscate the wealth of the productive members of a society to support a wide variety of welfare schemes. A substantial part of the confiscation is effected by taxation. But the welfare statists were quick to recognize that if they wished to retain political power, the amount of taxation had to be limited and they had to resort to programs of massive deficit spending, i.e., they had to borrow money, by issuing government bonds, to finance welfare expenditures on a large scale.

Under a gold standard, the amount of credit that an economy can support is determined by the economy's tangible assets, since every credit instrument is ultimately a claim on some tangible asset. But government bonds are not backed by tangible wealth, only by the government's promise to pay out of future tax revenues, and cannot easily be absorbed by the financial markets. A large volume of new government bonds can be sold to the public only at progressively higher interest rates. Thus, government deficit spending under a gold standard is severely limited. The abandonment of the gold standard made it possible for the welfare statists to use the banking system as a means to an unlimited expansion of credit. They have created paper reserves in the form of government bonds which -- through a complex series of steps -- the banks accept in place of tangible assets and treat as if they were an actual deposit, i.e., as the equivalent of what was formerly a deposit of gold. The holder of a government bond or of a bank deposit created by paper reserves believes that he has a valid claim on a real asset. But the fact is that there are now more claims outstanding than real assets. The law of supply and demand is not to be conned. As the supply of money (of claims) increases relative to the supply of tangible assets in the economy, prices must eventually rise. Thus the earnings saved by the productive members of the society lose value in terms of goods. When the economy's books are finally balanced, one finds that this loss in value represents the goods purchased by the government for welfare or other purposes with the money proceeds of the government bonds financed by bank credit expansion.

In the absence of the gold standard, there is no way to protect savings from confiscation through inflation. There is no safe store of value. If there were, the government would have to make its holding illegal, as was done in the case of gold. If everyone decided, for example, to convert all his bank deposits to silver or copper or any other good, and

thereafter declined to accept checks as payment for goods, bank deposits would lose their purchasing power and government-created bank credit would be worthless as a claim on goods. The financial policy of the welfare state requires that there be no way for the owners of wealth to protect themselves.

This is the shabby secret of the welfare statists' tirades against gold. Deficit spending is simply a scheme for the confiscation of wealth. Gold stands in the way of this insidious process. It stands as a protector of property rights. If one grasps this, one has no difficulty in understanding the statists' antagonism toward the gold standard.

Llewellyn H. Rockwell, Jr., ed.
The Gold Standard: An Austrian Perspective
Lexington, MA: D. C. Heath and Co., 1985, pp. 61-79

The "Costs" of a Gold Standard

Roger W. Garrison

I. Introduction

The term "gold standard," whether used in an historical or a theoretical context, can mean many different things. And for each meaning of this term, a reference to the "costs" of a gold standard will not be unambiguous. Yet it is commonly believed, by economists and laypersons alike, that cost considerations eliminate gold as a viable medium of exchange in modern-day economies.

The purpose of this paper is to examine the arguments against the gold standard which are based on considerations of costs. Section II identifies the benefits of a gold standard in order to put the discussion of costs into proper perspective. Section III compares two conflicting views of the gold standard and of the resources devoted to maintaining it. Section IV deals with actual estimates of the resource costs of gold. Sections V and VI employ the more broadly conceived concept of "opportunity costs" to argue the irrelevance of "resource costs" to the comparison of alternative monetary institutions. Section VII calls into question the assumed equivalence of monetary stability and price-level stability. This assumption, which underlies many of the cost estimates, has clouded some fundamental issues in ways that have prejudiced both monetary theorists and policymakers against the gold standard. Section VIII provides a summary assessment.

II. The Gold Standard: Costs and Benefits

Any discussion of the costs of a gold standard and of the controversy that surrounds this issue is, by its very nature, a one-sided discussion. The comparison of alternative standards on the basis of costs will not be meaningful unless the corresponding benefits are brought into view. Spelling out the particular type of gold standard being discussed and identifying its benefits—in comparison to a paper standard—puts the cost comparisons into proper perspective and goes a long way toward justifying the costs.

The term "gold standard" in the present paper is used to denote the outcome of a market process. Using the term in this way serves to consolidate at least three propositions—based on both economic theory and historical insight—about the nature of markets and about the nature of money. (1) Left to its own devices, a market economy will give rise to medium of exchange.(1) (2) The commodity that emerges as the medium of exchange will be one that possesses a certain set of characteristics.(2) (3) This set of characteristics has its clearest and most pronounced manifestation in gold.(3) So conceived, the gold standard—at least in its purest form—neither requires nor permits the State to exercise control over the money supply. And as argued below, the absence of centralized, discretionary monetary control constitutes the primary benefit of the gold standard.

The perception by the layperson that the costs of a gold standard are "too high" is not difficult to understand. Under a gold standard suppliers of goods or of labor services exchange their supplies for gold (or for bank notes redeemable in gold) not because the gold standard is seen as having great merit, but because gold is the customarily accepted medium of exchange. To each market participant gold *per se* has no particular benefits, although the custom of accepting some specific thing does. When consciously pondering the nature of money, the layperson is likely to see the custom in a different context and to see the value that others attach to gold—or that "the market" attaches to gold—as "irrational," as being based on superstition or mythology. Gold in the layperson's view is a "barbarous relic" (to use Maynard Keynes's phrase). Yet individuals in modern economies continue to devote resources to securing this shiny yellow metal. Could not some other custom have the same benefits without having such high costs?

Market-oriented economists adopt a third view of the gold standard—one that differs from the views of both the market participant and the layperson pondering the gold question. The economists see the difficulties—and costs—of replacing an evolved custom with a designed system. The differences among such economists stem from the differing estimates of the nature and magnitude of these difficulties.

Economists who oppose the gold standard may recognize what has to be achieved in order to make a centrally controlled paper standard superior to a decentralized gold standard. Milton Friedman poses the key question: "[H]ow ... can we establish a monetary system that is stable, free from irresponsible tinkering, and incapable of being used as a source of power to threaten economic and political freedom?"(4) How, in other words, can we design a system that we cannot tinker with? While Friedman goes on to suggest how such a system might be designed, economists who support the gold standard argue that this objective is self-contradictory and operationally impossible. Any monetary institution that is designed and implemented by a central authority can be abused by that central authority.

The proponents of gold are not suggesting that irresponsible tinkering is inevitable—whatever the nature of the monetary system; they are instead making the sharp distinction between a designed institution and an evolved institution. A monetary institution that has been consciously designed is much more subject to tinkering than one that simply emerged as a consequence of market processes. F. A. Hayek points to the positivist slogan that "what man has made he can also alter to suit his desires."(5) The positivists were correct so long as they were referring to consciously and deliberately designed institutions such as a paper standard. Of course, it is government officials (not "man") who design the paper standard, and it is government officials who can (and do) tinker with it. Hayek goes on to point out the limits of the positivists' view. The slogan is a "...complete *non-sequitur* if `made' is understood to include what has arisen from man's actions without his design."(6)

A gold standard—one that has emerged as a consequence of market processes—cannot easily be altered to suit the State's purposes. It would be an overstatement (and a matter of historical inaccuracy) to claim that the State cannot even in the long run interfere with the operation of the gold standard. What is true (both theoretically and historically) is that the State can supplant a spontaneously evolved monetary system with a centrally controlled system only after a prolonged struggle in which it must slowly and gradually overcome (through propaganda and the use of coercion) the market's reluctance to abandoned gold. It is the gold standard's substantial immunity from State manipulation and tinkering, and not the associated

superstition and mythology, that recommends gold as a monetary standard. In the words of Ludwig von Mises, "the advantage of the gold standard ... is due solely to the fact that, if once generally adopted in a definite form, and adhered to, it is no longer subject to specific political interference."(7) In the judgment of the proponents of the gold standard, the benefits of gold, immunity from State intervention and the resulting monetary stability, outweigh the resource costs of gold—and any other costs that might be associated with the gold standard—by a comfortable margin.

III. The Resources Devoted to Gold: Too Few and Too Many

Discussions of the gold standard typically gravitate toward a consideration of the amount of resources used up in the maintenance of it. Well recognized market processes will devote a certain amount of resources to the gold-mining industry, sometimes more resources, sometimes fewer, depending upon market conditions in the rest of the economy. Changing market conditions have both price effects and quantity effects that come into play. Consider, for instance, an increase in the demand for money brought about by a desire on the part of market participants for greater liquidity. This demand shift puts downward pressure on prices. Because the actual adjustment in prices is not immediate, the increased monetary demand will have a temporary effect on quantities as well. Excess supplies of goods and of resources—both labor and capital—will develop. In general, the more rapidly the prices adjust, the less pronounced the temporary adjustment in quantities, and conversely, the more slowly they adjust, the more pronounced the adjustment in quantities.

The adjustment process is facilitated in part by changing market conditions in the gold-mining industry and in supporting industries. In these markets, movements in prices and quantities are opposite in direction to the movements in markets for goods that exchange against money. Downward movements of prices in general mean an increased value of the monetary commodity; excess supplies of labor and capital mean an increased availability of resources for mining gold. Both the price and the quantity effects stimulate the production of the monetary commodity and in the process relieve the pressure that gave rise to the stimulation. The final result is that the increased demand for money is accommodated in part by an actual decline in prices and in part by an increased quantity of the monetary commodity. The relative size of the two accommodating factors depends upon the supply conditions in the gold-mining industry.

Much of the dissatisfaction with the gold standard stems from dissatisfactions with the quantity of resources devoted to the extraction and processing of gold. Paradoxically, some opponents of gold believe that too few resources are involved for the gold standard to be viable, while other opponents believe that too many resources are devoted to the mining of gold. Not surprisingly, these opposing opponents of gold are reasoning in markedly different ways.

The first line of reasoning is based on the assumption that prices are extremely "sticky," and hence that all adjustments to changing market conditions are quantity adjustments. An increased demand for money means a decreased demand for goods. Since the goods cannot all be sold at existing prices, surpluses pile up, production is curtailed, and workers become unemployed. An economy-wide depression sets in. The only excess demand in the economy is for the monetary commodity. But because of the nature of gold—its relative scarcity—the gold-mining industry can absorb only a small fraction of the unemployment. The demand for money cannot be fully met at the existing level of prices. If the gold-mining industry absorbed the same amount of

resources that were unemployed as a result of the increase in the demand for money, then the gold standard would perform admirably in this view and would constitute an automatic countercyclical device. Employment could shift from goods to gold or from gold to goods, but the level of employment would remain unchanged. Unfortunately, the gold-mining industry does not employ enough labor and capital resources to provide for such economic stability.(8) This line of reasoning has even caused one monetary reformer to advocate the abandonment of gold and the adoption of the common clay brick as a monetary standard.(9)

The other line of reasoning considers the alternative of a centrally directed system of paper money that can mimic the countercyclical effects of a clay-brick standard but without devoting any resources at all to the production of clay bricks or to the mining of gold. Each increase in the demand for money could be met with a costlessly produced increase in the quantity of money supplied. An economy whose transactions are facilitated by such a managed paper money would never experience an economy-wide downward pressure on prices that could result in resource idleness. Thus, the economy could devote all its resources to the production of real (non-monetary) output. With this possibility in mind the allocation of any of the economy's resources to the production of gold is seen as wasteful, and as constituting too many resources.(10)

Proponents of the gold standard should not feel called upon to argue in the context of either of these two lines of reasoning that the quantity of resources actually devoted to gold is just enough but not too much. Practically any quantity would at the same time be too little and too much—depending upon the opponent's particular point of view. Both viewpoints, however, can be called into question by an examination of the meaning and relevance of key concepts used by each. Particularly critical to the issue of the gold standard are the concepts of costs, resource costs, price stability, and monetary stability. These and related concepts provide a focus for the remainder of the present paper.

IV. The Costs of Gold and the Costs of a Constant Price Level

Estimates of the resource costs of gold depend critically upon the assumed rate of extraction. The actual rate of extraction, as indicated above, would be determined by market conditions. In an expanding economy with given supply conditions for gold, an increasing demand for money would cause additional resources to be committed to gold-mining operations. If competitive forces in the banking industry have given rise to the circulation of redeemable bank notes, the actual shift in the demand for gold caused by the expansion would be significantly reduced. The additional quantity of resources committed to gold would depend upon the elasticity of supply and the magnitude of the demand shift. Gold's relative inelasticity of supply would ensure that the dominant effects of the increase in the demand for gold, whatever its magnitude, would be a price effect rather than a quantity effect. That is, the value of gold would rise, or conversely, the prices of other goods would fall with respect to gold. There would be some increase in the quantity of gold supplied, but due to the price effect, this increase would be small in comparison to the increase in demand. The resource costs of extracting the additional gold would be correspondingly small.

Unfortunately, the most commonly cited estimates of the resource costs are based on the assumptions that there is no circulation of bank notes and that there is no price effect at all. Further, the supply of gold is assumed to be perfectly elastic.(11) Increases in the demand for money, under these assumptions, are met in full with increases in the quantity of gold supplied. The rate of gold extraction, in other words, is assumed to be

sufficiently large to offset totally the downward movement of prices that would otherwise be necessary in an expanding economy. The fact that the supply of gold is actually *in*elastic is simply brushed aside. The resulting estimate of the resource costs, then, is not an estimate of the costs of a gold standard at all but rather an estimate of the costs of maintaining a constant price level by adopting an elastically supplied commodity money.

Not surprisingly, actual estimates that are based on these assumptions show that the costs of a commodity money are quite high. Neglecting changes in the velocity of money, Friedman calculated that for the first half of the twentieth century, the resource costs of a pure gold standard would have amounted to about one-and-one-half percent of national income, or about one half of the annual growth rate of output. (Velocity considerations would increase these figures to two percent of national income and two-thirds of the annual growth rate.)(12) These particular estimates are three decades out of date, but the estimating procedure is in current use. Allan Meltzer cites the Friedman estimate and then updates the figures to reflect the current ratio of money to income. The new calculations indicate that the costs are down from fifty percent of the annual growth rate to something like sixteen percent. But the cost of a gold standard in Meltzer's own judgment "remains high."(13)

The estimating procedure adopted by Friedman and more recently by Meltzer is flawed on both positive and normative grounds. The positive analysis makes use of the classical long-run perspective in which *all* supply curves are perfectly elastic. Friedman notes explicitly that his cost estimate is independent of which commodity is used as the monetary standard.(14) Perfect supply elasticity is a particularly inappropriate assumption when the *gold* standard is at issue. The supply of gold is inelastic in the short run because of the increasing marginal costs of extraction and inelastic in the long run because of the natural scarcity of this particular element. Friedman's and Meltzer's calculations fail to take into account these particular supply considerations which help to qualify gold as a monetary commodity. Indeed, they fail to make any distinction whatsoever between gold and all other commodities.

The normative judgment upon which their cost calculations are based is the judgment that the maintenance of a constant price level over time is an undisputed *desideratum* and the appropriate basis for evaluating alternative monetary arrangements. The significance of a constant price level in this regard is the focus of the penultimate section of the present paper. The following two sections distinguish between two different cost concepts and question the use of resource costs as a criteria for choosing among alternative monetary institutions.

V. Costs, Resource Costs, and the Gold Standard

It was demonstrated above that commonly cited estimates of the resource costs of gold are based on untenable assumptions about the supply conditions in the gold-mining industry and about the desired behavior of the price level. The present section is concerned with the relevance of any estimate of resource costs to the comparison of gold and paper standards. It is argued that the resource costs are doubly irrelevant in assessing the relative merits—and the relative costs—of the two alternative standards. The critical issues are likely to be overlooked if there is a failure to distinguish between (1) the resource costs of gold and (2) the costs of a gold standard. The two cost concepts are totally dissimilar despite the similarity in the verbiage. This section deals with the costs of the gold standard and of paper standards over and above the narrowly conceived resource costs; the following section puts the resource costs of gold into proper perspective.

So-called resource costs are an inadequate proxy for total costs or opportunity costs—unless the former term is defined in such a way as to make it synonymous with the latter two, in which case the modifier "resource" becomes redundant and misleading. The inadequacy is especially pronounced when the issue is the relative costs of alternative institutional arrangements.(15) A penal system that segregated convicted criminals from the rest of society may involve more "resource costs" than one that only slapped the criminals' wrists and turned them back into society. But it would be difficult to argue that the total costs, which would have to take into account the subsequent crimes perpetrated by convicted criminals, are greater for the former institutional arrangement than for the latter.

There is a similar difficulty in the argument that a gold standard costs more than a paper standard. Comparing the resource costs of gold to the resource costs of paper does not settle the issue. The true costs of the paper standard would have to take into account (1) the costs imposed on society by different political factions in their attempts to gain control of the printing press, (2) the costs imposed by special-interest groups in their attempts to persuade the controller of the printing press to misuse its authority (print more money) for the benefit of the special interests, (3) the costs in the form of inflation-induced misallocations of resources that occur throughout the economy as a result of the monetary authority succumbing to the political pressures of the special interests, and (4) the costs incurred by businessmen in their attempts to predict what the monetary authority will do in the future and to hedge against likely, but uncertain, consequences of monetary irresponsibility. With these considerations in mind, it is not difficult to believe that a gold standard costs less than a paper standard. The judgment that a gold standard is the less costly reflects the wisdom in a simile attributed to Alan Greenspan: Allowing the State to create paper money is like putting a penny in the fuse box. The resource costs of the penny may be lower than the resource costs of the fuse, but the total costs, which take into account the likelihood of a destructive fire, are undoubtedly higher.

Some proponents of a paper standard base their counter arguments on their perception of the costs of a gold standard over and above the narrowly defined resource costs. But it is difficult to produce a laundry list of costs that will rival the list that was easily produced for the paper standard. The one cost most commonly cited stems from the fact that the supply of gold is not perfectly elastic, that gold production chronically falls short of real economic growth.(16) This circumstance requires that the price level be continually adjusted downward as the ratio of money to real output steadily declines. And the market process by which individual prices become so adjusted can be time-consuming and costly. Monetary disequilibrium, in effect, gets translated into disequilibrium in all markets throughout the economy.(17) Until equilibrium is eventually restored, the disequilibrium prices will result in the misdirection of some resources and the idleness of others. Under a paper standard, the monetary authority can eliminate the need to continually adjust prices by continually increasing the money supply to keep pace with the economy's real growth.

At one level of abstraction, the opponents of the gold standard have an appealing, if not compelling, argument. But when the analysis penetrates beneath the issue of the changing price levels associated with the gold and paper standards, the argument all but disappears. At least three considerations are relevant here.(18) First, individual prices in a market economy are changing all the time in response to changing market conditions. Some prices are increasing, some decreasing. If the supply of gold is not increasing as fast a real output, the pattern of individual price changes will be altered. Prices that are increasing will not increase quite so much; prices that are falling will

have to fall a little further. Some prices that would have had to be increased a little will not have to be increased at all; some which would have remained unchanged will have to be slightly decreased. As a result of the altered pattern of price changes, the price *level*, the weighted average of all prices, will be lower. It is misleading, though, to associate the cost of price changes with a changing price level. Prices would have had to be changed in any event—though in a slightly different pattern.

Second, even if we allow ourselves to abstract from individual price changes and think in terms of price levels, an elastically supplied currency will not eliminate the need for costly price adjustments. Consider, for instance, a growing economy in which the real rate of interest is declining. Which price level should the monetary authority keep constant: the consumer price level, the factor price level, or the general price level (which includes prices of both consumer goods and factors of production)? If the consumer price level is kept constant, then factor prices will have to be continually adjusted upward as the rate of interest falls; if the factor price level is kept constant, then the prices of consumer goods will have to be continually adjusted downward; if the general price level is kept constant, then the prices of both factors of production and consumer goods will have to be continually adjusted so as to reflect the declining interest rate. There is no price level whose constancy will eliminate the necessity for economy-wide adjustments in individual prices.

Third, the alleged price-adjustment costs of a gold standard are identified by comparing the gold standard as it actually operates with a paper standard as it ideally operates. Such comparisons never provide a sound basis for choosing between alternative institutional arrangements. The comparison assumes away all the relevant costs of a paper standard. If paper standards were administered by angelic monetary authorities whose sole objective was to minimize money-induced disequilibrium, the choice between a gold standard and a paper standard would be much less consequential than it actually is. But actual paper standards have price-adjustment costs too. And as history teaches, the magnitude and costliness of upward price adjustments under a paper standard dwarf the magnitude and costliness of downward price adjustments under a gold standard.(19)

VI. The Unavoidable Resource Costs of Money

In the preceding section the contention was made that the resource costs are doubly irrelevant to the issue of alternative monetary standards. Total costs, which are poorly proxied by resource costs, are the appropriate bases for comparison. This section establishes the double irrelevancy by showing that while resource costs are only a fraction of total costs, they constitute a part of total costs that the economy incurs whether on a gold standard or a paper standard. That is, the resource costs of gold constitute a part of the costs of both standards, but all of the costs of neither standard. These costs, then, cannot be costs that influence the choice between the two monetary standards.

The effectiveness of the resource-cost argument against the gold standard rests on the popular perception that the activities of mining gold, refining it, casting it into bars or minting it into coins, storing it, and guarding it are collectively wasteful activities and the implicit assumption that if the gold standard were supplanted by a paper standard, these activities would cease. But making the implicit assumption explicit is enough to demonstrate its falsity. The imposition of a paper standard does not cause gold to lose its monetary value. To believe otherwise is to hold the naive view that the State can repeal the laws of economics. Gold continues to be mined, refined, cast or minted, stored, and guarded; the resource costs continue to be incurred. In fact, a paper standard administered by an irresponsible monetary authority may drive the monetary

value of gold so high that more resource costs are incurred under the paper standard than would have been incurred under a gold standard. Market processes operating since antiquity have identified gold as the premier monetary commodity, and until the market's adoption of an alternative standard causes the value of gold to fall to a level that reflects only the non-monetary uses of gold, these resource costs cannot be avoided.

There is a certain asymmetry in the cost comparison that turns the resource-cost argument against paper standards. When an irresponsible monetary authority begins to overissue paper money, market participants begin to hoard gold, which stimulates the gold-mining industry and drives up the resource costs. But when new discoveries of gold are made, market participants do not begin to hoard paper or to set up printing presses for the issue of unbacked currency. Gold is a good substitute for an officially instituted paper money, but paper is not a good substitute for an officially recognized metallic money. Because of this asymmetry, the resource costs incurred by the State in its efforts to impose a paper standard on the economy and manage the supply of paper money could be avoided if the State would simply recognize gold as money. These costs, then, can be counted against the paper standard.

As suggested earlier, resource-cost comparisons that favor paper over gold are comparisons between real-world gold standards and fictitious paper standards.(20) Typically, the alternatives considered are strictly noncormformable: They consist of a market *process* that gives rise to the use of gold as the medium of exchange and an outcome that no known process can bring about. Wouldn't the world be a better place to live if there were no monetary value attached to gold (or to silver, copper...) and if the monetary authority were constitutionally bound to increase the issue of paper money at a relatively slow, fixed, and foreknown rate? Wouldn't the world be an even better place to live if there were some other monetary commodity, a commodity which was relatively scarce, which could not be extracted by any known mining technique, but which was costlessly coughed up by nature at a slow and steady rate in locations that were experiencing economic growth? These worlds can be imagined to look just like the one that we actually live in—minus the resource costs of gold. Such imaginations may provide the basis for bad science fiction, but they are no basis at all for devising monetary theories or for choosing among alternative institutional arrangements.

VII. A Constant Price Level vs. Monetary Stability

The assumption of a constant price level has a history that is both long and wide. Over the years theorists representing diverse schools of thought have invoked this assumption in their effort to abstract from monetary influences on the course of economic activity, and have adopted as a self-evident truth the notion that a constant price level is the hallmark of monetary stability. The significance of a constant price level for both theory and policy has been taken to be so obvious and self-evident that the literature is virtually devoid of attempts to defend these common practices. Yet a sampling of the many writers who do not question this assumption—and of the few who do—exposes the assumption as the Achilles' heel of the popular stance against the gold standard and of many other theoretical pronouncements and policy prescriptions.

Hayek noted in the early 1930s that an unaccountable preeminence of the constant price level characterized the writings of such monetary notables as Gustav Cassel and A. C. Pigou.(21) That a country should regulate its currency so as to achieve a constant price level appeared to Cassel as the "simplest assumption." If a country's currency were so regulated, money would exert no influence of its own, according to

Pigou. The idea that an equality between economic growth and monetary growth is "natural" and that money whose growth rate satisfies this equality is "neutral" had become commonplace by the end of the twenties. The general acceptance of this idea eliminated the need for a theoretical justification.

The assumed relevance and desirability of a constant price level are incorporated in later decades in the writings of American economists. In the early 1950s Clark Warburton included in his list of assumptions that underlie monetary theory the need for monetary growth to accommodate real economic growth. "[A]s a result of [population growth, technological developments, and increasing labor productivity] and of the stability of customs (such as the periodicity of income payments) which affect the rate of circulation of money, the economy needs for equilibrium a continuous increase in the quantity of money."(22) Following suit, Friedman assumed "for convenience" that a stable price level of final products is the objective of policy.(23) (It is both revealing and disquieting to note that Friedman's estimate of the resource costs of a gold standard discussed in Section IV above depends critically upon an assumption that was made for the sake of convenience.)

In the late sixties Friedman reaffirmed that he "simply took it for granted, in line with a long tradition and near-consensus in the profession, that a stable level of prices of final products was a desirable objective."(24) The purpose of the article that contains this statement was to replace the assumed optimum of a constant price level with a theoretically derived optimum. After identifying costs and benefits of a changing price level, Friedman, following standard microeconomic procedures, set the marginal costs equal to the marginal benefits and solved for the optimal, or welfare-maximizing, rate of change in the price level. It turned out that with an assumed rate of economic growth of three or four percent per year, a decline of prices of four or five percent per year would maximize economic welfare.(25) At this rate of price deflation, the marginal gains associated with the deflation-induced increase in real-cash holdings would just be offset by the nearly negligible marginal costs of increasing the supply of money. (These results apply to an economy using fiat currency. If gold were used instead, the marginal costs of extraction would cause the optimal rate of price deflation to be somewhat higher.)

From the outset Friedman failed fully to persuade even himself of the merits of his theoretically derived optimum. He ended the article with "A Final Schizophrenic Note" in which he teetered between endorsing a monetary rule which would optimize welfare as suggested by his theory and endorsing a monetary rule which would maintain a constant price level. In retrospect Friedman's calculations can be seen as a curious and contrived exercise in the application of marginalism. But today his arguments ring hollow. The unquestioned assumption of the desirability of a constant price level has regained its former status in discussions of monetary policy.

Economists of the Austrian school have always held the minority view that stable money and a constant price level are two different things.(26) At root their case is a very simple one. It requires only the most cursory consideration of what goes on behind the aggregates and the averages of the more orthodox monetary theory. It is true that productivity gains increase the level of output and thereby exert downward pressure on the level of prices. An offsetting increase in the total quantity of money can exert upward pressure and thereby preserve a constant price level. But productivity gains are themselves not neutral with respect to the composition of output. Economic growth does not consist in an across-the-board increase in the quantity of goods produced. It consists instead of increases in the quantities of some goods and decreases in the quantities of other goods, improvements in the quality of some goods, and the

introduction of new goods. Growth-induced changes in the pattern of output are accompanied by corresponding changes in the pattern of prices. The fact that the price level calculated on the basis of the new pattern is lower than the price level calculated on the basis of the old pattern is strictly incidental. To the extent that each individual change in the pattern of prices can be attributed to non-monetary factors, the issue of monetary non-neutrality does not arise despite the fall in the price level.(27)

The Austrians go on to point out that if an increase in the supply of money is brought about so that the economic growth can be "accommodated," the effects of the monetary injection on prices will be compounding rather than counteracting. Economic growth coupled with monetary growth may allow for a constant price level, but the pattern of prices will be affected in one way by the economic growth and in some other way by the monetary growth. Although it can be imagined that the increase in the supply of money affects only the price level, this lone effect cannot, because of the very nature of money, be actualized. Actual monetary injections, whether in the presence or the absence of economic growth, are always non-neutral.(28) They always have their own relative-price effects which, in turn, have effects on the pattern of output. A constant price level, then, is neither an appropriate assumption for devising monetary theories nor the most appropriate goal of monetary policies.

In current debate the goal of a constant price level enjoys a certain popularity for two reasons. Both have at least some merit, but neither constitutes a telling case against the gold standard. The first reason has to do with political feasibility. Some may argue that the prospects of persuading the central bank to adopt a constant price level as its goal are better than the prospects of persuading it to surrender totally to the dictates of a commodity money. The adoption of such a goal would at least be a step in the right direction, and it would not preclude a further step to a commodity standard if such a step were to prove desirable and feasible. But those who now favor the gold standard do not expect that the central bank will adopt a goal of a constant price level. In fact, they believe that the central bank's unwillingness to do so—or otherwise to behave responsibly—goes a long way towards proving the desirability of a commodity standard. And they believe that the question of which sort of monetary institution is the most desirable should be kept separate from the question of the political feasibility of bringing about the needed institutional change.(29)

The second reason for the popularity of this goal derives from money's role as a unit of value and its relationship in this regard to other units such as units of length and units of weight.(30) The analogy between the need for invariant units of length and weight and the need for an invariant unit of value is appealing. Carpenters would not fare well in their trade if they had to use measuring devices that expand and shrink on their own; truck drivers would experience an increased dread of weigh stations if they had to wonder how heavy a pound is today. The images conjured up by examples of this sort drive the point home. To be serviceable, units of length, weight and value must be invariant over time. The analogy is persuasive and may be just the right medicine for those who advocate inflation or who advocate artificially cheap credit even if the ultimate result is inflation.

But for the advocates of sound money, there is more to be learned from the sense in which the analogy does not hold than the sense in which it does. Invariance can be achieved for units of length and weight but not for units of value. Modern attempts to discover or create an invariant unit of value (in the form of multiple-commodity standards, indexation schemes, and the like) represent a throwback to the old pre-marginalist, pre-subjectivist classical economics. They require that we unlearn the lessons implicit in Ricardo's fruitless search.(31)

This point can be driven home with an analogy of a different sort. (It takes an analogy to beat an analogy.) A monetary commodity is more like a reference commodity, a base point, or bench mark, than like a measuring unit. An immutable reference value for gauging all other values has as its physical analog an immutable reference point in the cosmos. Some might argue that the Earth cannot serve as such a reference point, because the Earth is revolving around the Sun, which is revolving around the center of the Milky Way Galaxy, which is moving through the Universe. An immutable reference point has to be independent of all these movements. Different schemes for locating such a point, which take into account all the relative locations of all the heavenly bodies, might be proposed. But reflection will reveal that the immutable reference point is as useless as it is elusive. The most relevant reference point is the point where cosmic developments have put us. And so it is with the reference value. The most relevant reference commodity is the monetary commodity that market processes have given us. Once gold emerged as the world's monetary commodity, it became irrelevant that certain prices or prices in general may be "unstable" with respect to some other reference value or some index of values. If undisturbed by political schemes, gold should be regarded as a stable money until the market process itself, for whatever reasons, begins to favor some other commodity as a value reference.(32)

The different opponents of the gold standard have radically different reasons for wanting to reject gold as money. Some want to harness the monetary forces and put the reigns in the hands of government; others want to nullify the monetary forces that are inherent in any commodity standard. The former like to think of monetary stability as those monetary arrangements that result in full employment; the latter like to think of monetary stability as those monetary arrangements that result in a constant price level. Proponents of the gold standard hold that neither full employment nor a constant price level is an appropriate goal of government policy. Nor is either of these goals consistent with monetary stability. And achieving the goal of stable money, which may well result in both a fuller employment and a more nearly constant price level than would otherwise be possible, requires only that the government refrain from interfering with the commodity money chosen by the market.

VIII. Concluding Remarks

Opponents of the gold standard calculate the costs of gold in dollars and cents and report their calculations as a percentage of the economy's output. The intended interpretation is clear: But for the costs of gold, the economy would have had an output that much greater. Proponents of the gold standard would be ill-advised to respond with a cost figure of their own. If the true costs of a gold standard could be calculated at all, it would have to take into account the monetary instability associated with alternative standards and the consequent loss of output. But incorporating these considerations would undoubtedly cause the cost figures to turn negative. The gold standard has net benefits, not net costs. An appreciation for these benefits, but not a precise quantitative estimate, can best be gained by comparisons of historical episodes which are illustrative of economic performance under a gold standard and economic performance under a paper standard. The superiority of the former in comparison to the latter constitutes the net benefits of the gold standard.(33)

Ultimately, the cost of any action, commodity, or institution is the alternative action, commodity, or institution forgone. The opportunity cost is the only cost that counts. The cost of one institution is forgoing some other institution; the cost of the gold standard is forgoing a paper standard; the cost of sound money is forgoing unsound money.

Notes

* I would like to acknowledge the helpful comments and criticisms offered by Don Bellante, Don Boudreaux, and Leland B. Yeager of Auburn University and Gerald P. O'Driscoll, Jr. of the Federal Reserve Bank of Dallas.

1. Carl Menger, "On the Origin of Money," *Economic Journal*, vol. 2 (June, 1892), pp. 239-55. Also see Menger, Principles of Economics, trans. and ed. by James Dingwall and Bert F. Hozelitz (Glencoe, IL: Free Press, 1950), pp. 257-62.

2. One early list of such characteristics includes (1) utility and value, (2) portability, (3) indestructability, (4) homogeneity, (5) divisibility, (6) stability of value, and (7) cognizability. William Stanley Jevons, *Money and the Mechanism of Exchange* (New York: D. Appleton and Company, 1882), p. 31. It might be noted that characteristics 1 and 6 of Jevons' list are strongly reinforced as the particular commodity so characterized emerges as the medium of exchange.

3. *Ibid*. p. 41. Jevons lists in order: gold, silver, copper, tin, lead, and iron. Also see Ludwig von Mises, *The Theory of Money and Credit*, trans. by H. E. Batson (New Haven: Yale University Press, 1953), pp. 30-34.

4. Milton Friedman, "Should There Be an Independent Monetary Authority?" in Leland B. Yeager, ed., *In Search of a Monetary Constitution* (Cambridge: Harvard University Press, 1962), p. 228.

5. F. A. Hayek, "The Results of Human Action but not of Human Design," in Hayek, *Studies in Philosophy, Politics and Economics* (New York: Simon and Schuster, Publishers, 1969), p. 104.

6. *Ibid*. Language and money are probably the two most striking examples of social phenomena that have "arisen from man's actions without his design."

7. Ludwig von Mises, *On the Manipulation of Money and Credit*, trans. by Bettina Bien Graves and ed. by Percy L. Graves, Jr. (Dobbs Ferry, New York: Free Market Books, 1978), p. 80. Also see Mises, *Human Action*, 3rd rev. ed. (Chicago: Henry Regnery Company, 1966), pp. 471-76.

8. John Maynard Keynes, *The General Theory of Employment, Interest, and Money* (New York: Harcourt, Brace and World, Inc., 1964), pp. 229-36. Keynes is very clear on this point: "It is interesting to notice that the characteristic which has been traditionally supposed to render gold especially suitable for use as a standard of value, namely, its inelasticity of supply, turns out to be precisely the characteristic which is at the bottom of the trouble." This statement follows the more fanciful account of the trouble with gold: "Unemployment develops, that is to say, because people want the moon;--men cannot be employed when the object of their desire (i.e. money) is something which cannot be produced and the demand for which cannot be readily chocked off.

There is no remedy but to persuade the public that green cheese is practically the same thing and to have a green cheese factory (i.e. a central bank) under public control." *Ibid*. pp. 235-36.

9. For an account of the clay-brick standard first proposed by C. O. Hardy, see James M. Buchanan, "Predictability: the Criterion of Monetary Constitutions" in Yeager, *In Search of a Monetary Constitution*, pp. 155-83.

10. Milton Friedman, "Commodity-Reserve Currency," in *Essays in Positive Economics* (Chicago: University of Chicago Press, 1953), pp. 204-50. In his discussion of the supply elasticities of alternative monetary commodities, Friedman points out the disadvantages of an elasticity greater than zero (pp. 209-10) as well as the disadvantages of an elasticity less than infinity (pp. 210-14). The idea that the use of a monetary commodity whose supply elasticity is greater than zero involves "waste" has a long history and was endorsed by the classical economists including Smith and Ricardo.

11. Ibid. p. 210. Also, see Milton Friedman, *A Program for Monetary Stability* (New York: Fordham University Press, 1959), p. 5.

12. These estimates are from Friedman's "Commodity-Reserve Currency," which was originally published in 1951. The high resource costs were to become even higher by the end of the decade when Friedman offered a new estimate in his *Program for Monetary Stability*. On the basis of a slightly higher rate of

economic growth, he estimated that about two-and-a-half percent of the economy's output, or $8 billion dollars per year, would have to be devoted to the procurement of additional quantities of the monetary commodity.

13. Allan H. Meltzer, "Monetary Reform in an Uncertain Environment," *Cato Journal*, vol. 3, no. 1 (Spring, 1983), pp. 93-112. Meltzer's updated estimate is found on page 105. This particular estimating procedure, in which the estimate of the resource costs of gold is always some multiple of the economy's growth rate, gives rise to an interesting paradox. If mismanaged paper money causes so much economic discoordination that the growth rate drops to zero, the estimated resource costs of maintaining a gold standard would also drop to zero; but if the adoption of the now-inexpensive gold standard creates economic stability and fosters new growth, maintaining the gold standard would once again become too costly.

14. Friedman, "Commodity-Reserve Currency," p. 210; *A Program for Monetary Stability*, p. 5.

15. This point has been emphasized by Ronald Coase: "It would seem desirable to use [an opportunity-cost] approach when dealing with questions of economic policy and to compare the total product yielded by alternative social arrangements." Coase, "The Problem of Social Cost," *Journal of Law and Economics*, vol. 3 (October, 1960), p. 43.

16. It would be a gross misrepresentation, of course, to add the costs that are incurred as a result of an assumed inelasticity of the supply of gold to the resource costs that are based on an assumed perfect elasticity.

17. Arguments of this sort have their roots in the writings of Clark Warburton. See Warburton, "The Monetary Disequilibrium Hypothesis," in Warburton, *Depression, Inflation, and Monetary Policy, Selected papers, 1945-1953*, (Baltimore: The Johns Hopkins Press, 1966), pp. 25-35. Also, see Leland B. Yeager, "Stable Money and Free-Market Currencies," *Cato Journal*, vol. 3, no. 1 (Spring, 1983), pp. 305-26.

18. A more thorough discussion of the relevance of a constant price level is offered in Section VII below.

19. For historical perspectives on gold and paper, see Murray N. Rothbard, *What has Government Done to Our Money*? (A Pine Tree Publication, 1963), pp. 27-49; Ron Paul and Lewis Lehrman, *The Case for Gold: A Minority Report of the U. S. Gold Commission* (Washington, D. C.: Cato Institute, 1982), pp 17-142; and Alan Reynolds, "Why Gold?" *Cato Journal*, vol. 3, no. 1 (Spring, 1983), pp. 211-32.

20. Again, it was Ronald Coase who sensitized the profession to "the usual treatment of [problems of social cost in which] the analysis proceeds in terms of a comparison between a state of laissez faire and some kind of ideal world." He went on to point out that "very little analysis is required to show that an ideal world is better than a state of laissez faire unless the definitions of a state of laissez faire and an ideal world happen to be the same." Coase, "The Problem of Social Costs," p. 43. Harold Demsetz has elaborated on the differences between the "nirvana approach" (which is implicitly adopted by many critics of the gold standard) and the "comparative-institutions approach" (which is adopted in the present paper). Demsetz, "Information and Efficiency: Another Viewpoint," *Journal of Law and Economics*, vol. 12 (April, 1969), pp. 1-22.

21. F. A. Hayek, *Prices and Production*, 2nd ed. (New York: Augustus M. Kelley, Publishers, 1967), p. 107.

22. Warburton, "The Monetary Disequilibrium Hypothesis," p. 28. Warburton notes that this assumption "pervaded so much economic literature of the nineteenth century and early part of the twentieth that supporting documentary references seem superfluous." *Ibid*., p. 29.

23. Friedman, "Commodity-Reserve Currency," p. 210.

24. Milton Friedman, "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays* (Chicago: Aldine Publishing Company, 1969), p. 48.

25. Ibid., p. 46.

26. Hayek, *Prices and Production*, pp. 105-31; Mises, *Human Action*, pp. 219-28, and *passim*; Mises, *On the Manipulation of Money and Credit*, pp. 1-49; and Mises, *Theory of Money and Credit*, pp. 108-69.

27. Mises, *Theory of Money and Credit*, p. 123. Mises maintains a first-order distinction between "two sorts of determinants of the exchange-value that connects money and other economic goods; those that exercise their effect on the money side of the ratio and those that exercise their effect on the commodity side." Only the effects of the first-mentioned sort are relevant to the issue of monetary neutrality.

28. Mises, *Human Action*, p. 418. "The notion of a neutral money is no less contradictory than that of a money of stable purchasing power. Money without a driving force of its own would not, as people assume, be a perfect money; it would not be money at all." The fact that money is not neutral is the common denominator of the monetary theories of Mises, Hayek, and others who have written in the Austrian tradition.

29. Friedman has enunciated this maxim but has not always abided by it. "The role of the economist in discussions of public policy seems to me to be to prescribe what should be done in the light of what can be done, politics aside, and not to predict what is "politically feasible" and then to recommend it." Milton Friedman, "Comments on Monetary Policy," in *Essays in Positive Economics*, (Chicago: University of Chicago Press, 1953), p. 264. William Hutt, who is critical of the Friedman maxim, recommends that the economist take political considerations into account, but only if such considerations are made explicit. This, he argues, is the business of old-style "political economy." William H. Hutt, *Politically Impossible...?* (London:Institute of Economic Affairs, 1971), pp. 22-27.

30. Yeager, "Stable Money and Free-Market Currencies," pp. 305-308.

31. "The Search for a stable unit of account is ultimately the search for an invariant standard of value, the quixotic goal of classical political economy." Gerald P. O'Driscoll, Jr., "A Free-Market Money: Comment on Yeager," *Cato Journal*, vol. 3, no. 1 (Spring, 1983), p. 328.

32. Economic calculation does not require monetary stability in the sense in which this term is used by the champions of the stabilization movement. The fact that rigidity in the monetary unit's purchasing power is unthinkable and unrealizable does not impair the methods of economic calculation. What economic calculation requires is a monetary system whose functioning is not sabotaged by government interference." Mises, *Human Action*, p. 223. It should be noted that the dramatic fluctuations in the value of gold in recent years do not imply that gold is no longer a viable monetary commodity. Quite to the contrary, volatile shifts in the demand for hard money are a reflection of the instability of our present monetary institution. The paper money is losing *its* viability. See Mises, *On the Manipulation of Money and Credit*, pp. 76-77.

33. Historical studies are cited in note 19, above.

---

x

2024

SESSION 3

Hayekian Solutions

Gwartney Institute

I. THE PRACTICAL PROPOSAL

The concrete proposal for the near future, and the occasion for the examination of a much more far-reaching scheme, is that

> *the countries of the Common Market, preferably with the neutral countries of Europe (and possibly later the countries of North America) mutually bind themselves by formal treaty not to place any obstacles in the way of the free dealing throughout their territories in one another's currencies (including gold coins) or of a similar free exercise of the banking business by any institution legally established in any of their territories.*

This would mean in the first instance the abolition of any kind of exchange control or regulation of the movement of money between these countries, as well as the full freedom to use any of the currencies for contracts and accounting. Further, it would mean the opportunity for any bank located in these countries to open branches in any other on the same terms as established banks.

Free trade in money

The purpose of this scheme is to impose upon existing monetary and financial agencies a very much needed discipline by making it impossible for any of them, or for any length of time, to issue a kind of money substantially less reliable and useful than the money of any other. As soon as the public became familiar with the new possibilities, any deviations from the straight path of providing an honest money would at once lead to the rapid displacement of the offending currency by others. And the individual countries, being deprived of the various dodges by which they are now able temporarily to conceal the effects of their actions by 'protecting' their currency, would be constrained to keep the value of their currencies tolerably stable.

Proposal more practicable than utopian European currency

This seems to me both preferable and more practicable than the utopian scheme of introducing a new European currency, which would ultimately only have the effect of more deeply entrenching the source and root of all monetary evil, the government monopoly of the issue and control of money. It would also seem that, if the countries were not prepared to

adopt the more limited proposal advanced here, they would be even less willing to accept a common European currency. The idea of depriving government altogether of its age-old prerogative of monopolising money is still too unfamiliar and even alarming to most people to have any chance of being adopted in the near future. But people might learn to see the advantages if, at first at least, the currencies of the governments were allowed to compete for the favour of the public.

Though I strongly sympathise with the desire to complete the economic unification of Western Europe by completely freeing the flow of money between them, I have grave doubts about the desirability of doing so by creating a new European currency managed by any sort of supra-national authority. Quite apart from the extreme unlikelihood that the member countries would agree on the policy to be pursued in practice by a common monetary authority (and the practical inevitability of some countries getting a worse currency than they have now), it seems highly unlikely, even in the most favourable circumstances, that it would be administered better than the present national currencies. Moreover, in many respects a single international currency is not better but worse than a national currency if it is not better run. It would leave a country with a financially more sophisticated public not even the chance of escaping from the consequences of the crude prejudices governing the decisions of the others. The advantage of an international authority should be mainly to protect a member state from the harmful measures of others, not to force it to join in their follies.

Free trade in banking

The suggested extension of the free trade in money to free trade in banking is an absolutely essential part of the scheme if it is to achieve what is intended. First, bank deposits subject to cheque, and thus a sort of privately issued money, are today of course a part, and in most countries much the largest part, of the aggregate amount of generally accepted media of exchange. Secondly, the expansion and contraction of the separate national superstructures of bank credit are at present the chief excuse for national management of the basic money.

On the effects of the adoption of the proposal all I will add at this point is that it is of course intended to prevent national monetary and financial authorities from doing many things

politically impossible to avoid so long as they have the power to do them. These are without exception harmful and against the long-run interest of the country doing them but politically inevitable as a temporary escape from acute difficulties. They include measures by which governments can most easily and quickly remove the causes of discontent of particular groups or sections but bound in the long run to disorganise and ultimately to destroy the market order.

Preventing government from concealing depreciation

The main advantage of the proposed scheme, in other words, is that it would prevent governments from 'protecting' the currencies they issue against the harmful consequences of their own measures, and therefore prevent them from further employing these harmful tools. They would become unable to conceal the depreciation of the money they issue, to prevent an outflow of money, capital, and other resources as a result of making their home use unfavourable, or to control prices—all measures which would, of course, tend to destroy the Common Market. The scheme would indeed seem to satisfy all the requirements of a common market better than a common currency without the need to establish a new international agency or to confer new powers on a supra-national authority.

The scheme would, to all intents and purposes, amount to a displacement of the national circulations only if the national monetary authorities misbehaved. Even then they could still ward off a complete displacement of the national currency by rapidly changing their ways. It is possible that in some very small countries with a good deal of international trade and tourism, the currency of one of the bigger countries might come to predominate, but, assuming a sensible policy, there is no reason why most of the existing currencies should not continue to be used for a long time. (It would, of course, be important that the parties did not enter into a tacit agreement not to supply so good a money that the citizens of the other nations would prefer it! And the presumption of guilt would of course always have to lie against the government whose money the public did not like!)

I do not think the scheme would prevent governments from doing anything they ought to do in the interest of a well-functioning economy, or which in the long run would benefit

any substantial group. But this raises complex issues better discussed within the framework of the full development of the underlying principle.

II. THE GENERALISATION OF THE UNDERLYING PRINCIPLE

If the use of several concurrent currencies is to be seriously considered for immediate application in a limited area, it is evidently desirable to investigate the consequences of a general application of the principle on which this proposal is based. If we are to contemplate abolishing the exclusive use within each national territory of a single national currency issued by the government, and to admit on equal footing the currencies issued by other governments, the question at once arises whether it would not be equally desirable to do away altogether with the monopoly of government supplying money and to allow private enterprise to supply the public with other media of exchange it may prefer.

The questions this reform raises are at present much more theoretical than the practical proposal because the more far-reaching suggestion is clearly not only much too strange and alien to the general public to be considered for present application. The problems it raises are evidently also still much too little understood even by the experts for anyone to make a confident prediction about the precise consequences of such a scheme. Yet it is clearly possible that there is no necessity or even advantage in the now unquestioned and universally accepted government prerogative of producing money. It may indeed prove to be harmful and its abolition a great gain, opening the way for very beneficial developments. Discussion therefore cannot begin early enough. Though its realisation may be wholly impracticable so long as the public is mentally unprepared for it and uncritically accepts the dogma of the necessary government prerogative, this should no longer be allowed to act as a bar to the intellectual exploration of the fascinating theoretical problems the scheme raises.

Competition in currency not discussed by economists

It is an extraordinary truth that competing currencies have

until quite recently never been seriously examined.[1] There is no answer in the available literature to the question why a government monopoly of the provision of money is universally regarded as indispensable, or whether the belief is simply derived from the unexplained postulate that there must be within any given territory one single kind of money in circulation—which, so long as only gold and silver were seriously considered as possible kinds of money, might have appeared a definite convenience. Nor can we find an answer to the question of what would happen if that monopoly were abolished and the provision of money were thrown open to the competition of private concerns supplying different currencies. Most people seem to imagine that any proposal for private agencies to be allowed to issue money means that they should be allowed to issue the *same* money as anybody else (in token money this would, of course, simply amount to forgery) rather than *different* kinds of money clearly distinguishable by different denominations among which the public could choose freely.

Initial advantages of government monopoly in money

Perhaps when the money economy was only slowly spreading into the remoter regions, and one of the main problems was to teach large numbers the art of calculating in money (and that was not so very long ago), a single easily recognisable kind of money may have been of considerable assistance. And it may be argued that the exclusive use of such a single uniform sort of money greatly assisted comparison of prices and therefore the growth of competition and the market. Also, when the genuineness of metallic money could be ascertained only by a difficult process of assaying, for which the ordinary person had neither the skill nor the equipment, a strong case could be made for guaranteeing the fineness of the coins by the stamp of some generally recognised authority which, outside the great commercial centres, could be only the government. But today these initial advantages, which might have served as an excuse for governments to appropriate the exclusive right of issuing metallic money, certainly do not outweigh the

[1] But, though I had independently arrived at the realisation of the advantages possessed by independent competing currencies, I must now concede intellectual priority to Professor Benjamin Klein, who, in a paper written in 1970 and published in 1975 [35], until recently unknown to me, had clearly explained the chief advantage of competition among currencies.

disadvantages of this system. It has the defects of all monopolies: one must use their product even if it is unsatisfactory, and, above all, it prevents the discovery of better methods of satisfying a need for which a monopolist has no incentive.

If the public understood what price in periodic inflation and instability it pays for the convenience of having to deal with only one kind of money in ordinary transactions, and not occasionally to have to contemplate the advantage of using other money than the familiar kind, it would probably find it very excessive. For this convenience is much less important than the opportunity to use a reliable money that will not periodically upset the smooth flow of the economy—an opportunity of which the public has been deprived by the government monopoly. But the people have never been given the opportunity to discover this advantage. Governments have at all times had a strong interest in persuading the public that the right to issue money belongs exclusively to them. And so long as, for all practical purposes, this meant the issue of gold, silver and copper coins, it did not matter so much as it does today, when we know that there are all kinds of other possible sorts of money, not least paper, which government is even less competent to handle and even more prone to abuse than metallic money.

III. THE ORIGIN OF THE GOVERNMENT PREROGATIVE OF MAKING MONEY

For more than 2,000 years the government prerogative or exclusive right of supplying money amounted in practice merely to the monopoly of minting coins of gold, silver or copper. It was during this period that this prerogative came to be accepted without question as an essential attribute of sovereignty—clothed with all the mystery which the sacred powers of the prince used to inspire. Perhaps this conception goes back to even before King Croesus of Lydia struck the first coins in the sixth century BC, to the time when it was usual merely to punch marks on the bars of metal to certify its fineness.

At any rate, the minting prerogative of the ruler was firmly established under the Roman emperors.[1] When, at the begin-

[1] W. Endemann [15], Vol. II, p. 171.

ning of the modern era, Jean Bodin developed the concept of sovereignty, he treated the right of coinage as one of the most important and essential parts of it.[1] The *regalia*, as these royal prerogatives were called in Latin, of which coinage, mining, and custom duties were the most important, were during the Middle Ages the chief sources of revenue of the princes and were viewed solely from this angle. It is evident that, as coinage spread, governments everywhere soon discovered that the exclusive right of coinage was a most important instrument of power as well as an attractive source of gain. From the beginning the prerogative was neither claimed nor conceded on the ground that it was for the general good but simply as an essential element of governmental power.[2] The coins served, indeed, largely as the symbols of might, like the flag, through which the ruler asserted his sovereignty, and told his people who their master was whose image the coins carried to the remotest parts of his realm.

Government certificate of metal weight and purity

The task the government was understood to assume was of course initially not so much to make money as to certify the weight and fineness of the materials that universally served as money,[3] which after the earliest times were only the three metals, gold, silver, and copper. It was supposed to be a task rather like that of establishing and certifying uniform weights and measures.

The pieces of metal were regarded as proper money only if

[1] J. Bodin [5], p. 176. Bodin, who understood more about money than most of his contemporaries, may well have hoped that the governments of large states would be more responsible than the thousands of minor princelings and cities who, during the later part of the Middle Ages, had acquired the minting privilege and sometimes abused it even more than the richer princes of large territories.

[2] The same applies to the postal monopoly which everywhere appears to provide a steadily deteriorating service and of which in Great Britain (according to *The Times*, 25 May, 1976) the General Secretary of the Union of Post Office Workers (!) said recently that 'Governments of both political complexions have reduced a once great public service to the level of a music-hall joke'. *Politically* the broadcasting monopoly may be even more dangerous, but *economically* I doubt whether any other monopoly has done as much damage as that of issuing money.

[3] Cf. Adam Smith [54, p. 40]: '... those public offices called mints: institutions exactly of the same nature with those of the aulnagers and stampmasters of woollen and linen cloth'.

they carried the stamp of the appropriate authority, whose duty was thought to be to assure that the coins had the proper weight and purity to give them their value.

During the Middle Ages, however, the superstition arose that it was the act of government that conferred the value upon the money. Although experience always proved otherwise, this doctrine of the *valor impositus*[1] was largely taken over by legal doctrine and served to some extent as justification of the constant vain attempts of the princes to impose the same value on coins containing a smaller amount of the precious metal. (In the early years of this century the medieval doctrine was revived by the German Professor G. F. Knapp; his *State Theory of Money* still seems to exercise some influence on contemporary legal theory.)[2]

There is no reason to doubt that private enterprise would, if permitted, have been capable of providing as good and at least as trustworthy coins. Indeed occasionally it did, or was commissioned by government to do so. Yet so long as the technical task of providing uniform and recognisable coins still presented major difficulties, it was at least a useful task which government performed. Unfortunately, governments soon discovered that it was not only useful but could also be made very profitable, at least so long as people had no alternative but to use the money they provided. The seignorage, the fee charged to cover the cost of minting, proved a very attractive source of revenue, and was soon increased far beyond the cost of manufacturing the coin. And from retaining an excessive part of the metal brought to the government mint to be struck into new coins, it was only a step to the practice, increasingly common during the Middle Ages, of recalling the circulating coins in order to recoin the various denominations with a lower gold or silver content. We shall consider the effect of these debasements in the next Section. But since the function of government in issuing money is no longer one of merely certifying the weight and fineness of a certain piece of metal, but involves a deliberate determination of the quantity of money to be issued, governments have become wholly inadequate for the task and, it can be said without qualifications, have incessantly and everywhere abused their trust to defraud the people.

[1] Endemann [15], p. 172.

[2] Knapp [36], and compare Mann [41].

The appearance of paper money

The government prerogative, which had originally referred only to the issue of coins because they were the only kind of money then used, was promptly extended to other kinds of money when they appeared on the scene. They arose originally when governments wanted money which they tried to raise by compulsory loans, for which they gave receipts that they ordered people to accept as money. The significance of the gradual appearance of government paper money, and soon of bank notes, is for our purposes complicated because for a long time the problem was not the appearance of new kinds of money with a different denomination, but the use as money of paper claims on the established kind of metallic money issued by government monopoly.

It is probably impossible for pieces of paper or other tokens of a material itself of no significant market value to come to be gradually accepted and held as money unless they represent a claim on some valuable object. To be accepted as money they must at first derive their value from another source, such as their convertibility into another kind of money. In consequence, gold and silver, or claims for them, remained for a long time the only kinds of money between which there could be any competition; and, since the sharp fall in its value in the 19th century, even silver ceased to be a serious competitor to gold. (The possibilities of bimetallism[1] are irrelevant for our present problems.)

Political and technical possibilities of controlling paper money

The position has become very different, however, since paper money established itself everywhere. The government monopoly of the issue of money was bad enough so long as metallic money predominated. But it became an unrelieved calamity since paper money (or other token money), which can provide the best and the worst money, came under political control. A money deliberately controlled in supply by an agency whose self-interest forced it to satisfy the wishes of the *users* might be the best. A money regulated to satisfy the demands of group interests is bound to be the worst possible (Section XVIII).

The value of paper money obviously can be regulated according to a variety of principles—even if it is more than

[1] Section VII, below, pp. 43-45.

doubtful that any democratic government with unlimited powers can ever manage it satisfactorily. Though historical experience would at first seem to justify the belief that only gold can provide a stable currency, and that all paper money is bound to depreciate sooner or later, all our insight into the processes determining the value of money tells us that this prejudice, though understandable, is unfounded. The *political* impossibility that governments will achieve it does not mean there is reason to doubt that it is *technically* possible to control the quantity of any kind of token money so that its value will behave in a desired manner, and that it will for this reason retain its acceptability and its value. It would therefore now be possible, if it were permitted, to have a variety of essentially different monies. They could represent not merely different quantities of the same metal, but also different abstract units fluctuating in their value relatively to one another. In the same way, we could have currencies circulating concurrently throughout many countries and offering the people a choice. This possibility appears, until recently, never to have been contemplated seriously. Even the most radical advocates of free enterprise, such as the philosopher Herbert Spencer[1] or the French economist Joseph Garnier,[2] seem to have advocated only private coinage, while the free banking movement of the mid-19th century agitated merely for the right to issue notes in terms of the standard currency.[3]

Monopoly of money has buttressed government power

While, as we shall see presently, government's exclusive right to issue and regulate money has certainly not helped to give us a better money than we would otherwise have had, and probably a very much worse one, it has of course become a chief instrument for prevailing governmental policies and profoundly assisted the general growth of governmental power. Much of contemporary politics is based on the assumption that government has the power to create and make people accept any amount of additional money it wishes. Governments will for this reason strongly defend their traditional rights. But for the same reason it is also most important that they should be taken from them.

A government ought not, any more than a private person, to

[1] Herbert Spencer [57]. [2] Joseph Garnier [21]. [3] Vera C. Smith [55].

be able (at least in peace-time) to take whatever it wants, but be limited strictly to the use of the means placed at its disposal by the representatives of the people, and to be unable to extend its resources beyond what the people have agreed to let it have. The modern expansion of government was largely assisted by the possibility of covering deficits by issuing money—usually on the pretence that it was thereby creating employment. It is perhaps significant, however, that Adam Smith [54, p. 687] does not mention the control of the issue of money among the 'only three duties [which] according to the system of natural liberty, the sovereign has to attend to'.

IV. THE PERSISTENT ABUSE OF THE GOVERNMENT PREROGATIVE

When one studies the history of money one cannot help wondering why people should have put up for so long with governments exercising an exclusive power over 2,000 years that was regularly used to exploit and defraud them. This can be explained only by the myth (that the government prerogative was necessary) becoming so firmly established that it did not occur even to the professional students of these matters (for a long time including the present writer[1]) ever to question it. But once the validity of the established doctrine is doubted its foundation is rapidly seen to be fragile.

We cannot trace the details of the nefarious activities of rulers in monopolising money beyond the time of the Greek philosopher Diogenes who is reported, as early as the fourth century BC, to have called money the politicians' game of dice. But from Roman times to the 17th century, when paper money in various forms begins to be significant, the history of coinage is an almost uninterrupted story of debasements or the continuous reduction of the metallic content of the coins and a corresponding increase in all commodity prices.

History is largely inflation engineered by government

Nobody has yet written a full history of these developments. It would indeed be all too monotonous and depressing a story,

[1] F. A. Hayek [29], pp. 324 *et seq.*

but I do not think it an exaggeration to say that history is largely a history of inflation, and usually of inflations engineered by governments and for the gain of governments—though the gold and silver discoveries in the 16th century had a similar effect. Historians have again and again attempted to justify inflation by claiming that it made possible the great periods of rapid economic progress. They have even produced a series of inflationist theories of history[1] which have, however, been clearly refuted by the evidence: prices in England and the United States were at the end of the period of their most rapid development almost exactly at the same level as two hundred years earlier. But their recurring rediscoverers are usually ignorant of the earlier discussions.

Early Middle Ages' deflation local or temporary

The early Middle Ages may have been a period of deflation that contributed to the economic decline of the whole of Europe. But even this is not certain. It would seem that on the whole the shrinking of trade led to the reduction of the amount of money in circulation, not the other way round. We find too many complaints about the dearness of commodities and the deterioration of the coin to accept deflation as more than a local phenomenon in regions where wars and migrations had destroyed the market and the money economy shrank as people buried their treasure. But where, as in Northern Italy, trade revived early, we find at once all the little princes vying with one another in diminishing the coin—a process which, in spite of some unsuccessful attempts of private merchants to provide a better medium of exchange, lasted throughout the following centuries until Italy came to be described as the country with the worst money and the best writers on money.

But though theologians and jurists joined in condemning these practices, they never ceased until the introduction of paper money provided governments with an even cheaper method of defrauding the people. Governments could not, of course, pursue the practices by which they forced bad money upon the people without the cruellest measures. As one legal treatise on the law of money sums up the history of punishment for merely refusing to accept the legal money:

[1] Especially Werner Sombart [56] and before him Archibald Alison [1] and others. Cf. on them Paul Barth [4], who has a whole chapter on 'History as a function of the value of money', and Marianne von Herzfeld [32].

'From Marco Polo we learn that, in the 13th century, Chinese law made the rejection of imperial paper money punishable by death, and twenty years in chains or, in some cases death, was the penalty provided for the refusal to accept French *assignats*. Early English law punished repudiation as *lese-majesty*. At the time of the American revolution, non-acceptance of Continental notes was treated as an enemy act and sometimes worked a forfeiture of the debt.'[1]

Absolutism suppressed merchants' attempts to create stable money

Some of the early foundations of banks at Amsterdam and elsewhere arose from attempts by merchants to secure for themselves a stable money, but rising absolutism soon suppressed all such efforts to create a non-governmental currency. Instead, it protected the rise of banks issuing notes in terms of the official government money. Even less than in the history of metallic money can we here sketch how this development opened the doors to new abuses of policy.

It is said that the Chinese had been driven by their experience with paper money to try to prohibit it for all time (of course unsuccessfully) before the Europeans ever invented it.[2] Certainly European governments, once they knew about this possibility, began to exploit it ruthlessly, not to provide people with good money, but to gain as much as possible from it for their revenue. Ever since the British Government in 1694 sold the Bank of England a limited monopoly of the issue of bank notes, the chief concern of governments has been not to let slip from their hands the power over money, formerly based on the prerogative of coinage, to really independent banks. For a time the ascendancy of the gold standard and the consequent belief that to maintain it was an important matter of prestige, and to be driven off it a national disgrace, put an effective restraint on this power. It gave the world the one long period—200 years or more—of relative stability during which modern industrialism could develop, albeit suffering from periodic crises. But as soon as it was widely understood some 50 years ago that the convertibility into gold was merely a method of controlling the *amount* of a currency, which was the real factor determining its

[1] A. Nussbaum [50], p. 53.

[2] On the Chinese events, see W. Vissering [61] and G. Tullock [58], who does not, however, allude to the often recounted story of the 'final prohibition'.

value, governments became only too anxious to escape that discipline, and money became more than ever before the plaything of politics. Only a few of the great powers preserved for a time tolerable monetary stability, and they brought it also to their colonial empires. But Eastern Europe and South America never knew a prolonged period of monetary stability.

But, while governments have never used their power to provide a decent money for any length of time, and have refrained from grossly abusing it only when they were under such a discipline as the gold standard imposed, the reason that should make us refuse any longer to tolerate this irresponsibility of government is that we know today that it is possible to control the quantity of a currency so as to prevent significant fluctuations in its purchasing power. Moreover, though there is every reason to mistrust government if not tied to the gold standard or the like, there is no reason to doubt that private enterprise whose business depended on succeeding in the attempt could keep stable the value of a money it issued.

Before we can proceed to show how such a system would work we must clear out of the way two prejudices that will probably give rise to unfounded objections against the proposal.

V. THE MYSTIQUE OF LEGAL TENDER

The first misconception concerns the concept of 'legal tender'. It is not of much significance for our purposes, but is widely believed to explain or justify government monopoly in the issue of money. The first shocked response to the proposal here discussed is usually 'But there must be a legal tender', as if this notion proved the necessity for a single government-issued money believed indispensable for the daily conduct of business.

In its strictly legal meaning, 'legal tender' signifies no more than a kind of money a creditor cannot refuse in discharge of a debt due to him in the money issued by government.[1] Even so, it is significant that the term has no authoritative definition

[1] Nussbaum [50], Mann [41] and Breckinridge [6].

in English statute law.[1] Elsewhere it simply refers to the means of discharging a debt contracted in terms of the money issued by government or due under an order of a court. In so far as government possesses the monopoly of issuing money and uses it to establish one kind of money, it must probably also have power to say by what kind of objects debts expressed in its currency can be discharged. But that means neither that all money need be legal tender, nor even that all objects given by the law the attribute of legal tender need to be money. (There are historical instances in which creditors have been compelled by courts to accept commodities such as tobacco, which could hardly be called money, in discharge of their claims for money.[2])

The superstition disproved by spontaneous money

The term 'legal tender' has, however, in popular imagination come to be surrounded by a penumbra of vague ideas about the supposed necessity for the state to provide money. This is a survival of the medieval idea that it is the state which somehow confers value on money it otherwise would not possess. And this, in turn, is true only to the very limited extent that government can force us to accept whatever it wishes in place of what we have contracted for; in this sense it can give the substitute the same value for the debtor as the original object of the contract. But the superstition that it is necessary for government (usually called the 'state' to make it sound better) to declare what is to be money, as if it had created the money which could not exist without it, probably originated in the naive belief that such a tool as money must have been 'invented' and given to us by some original inventor. This belief has been wholly displaced by our understanding of the spontaneous generation of such undesigned institutions by a process of social evolution of which money has since become the prime

[1] Mann [41], p. 38. On the other hand, the refusal until recently of English Courts to give judgement for paying in any other currency than the pound sterling has made this aspect of legal tender particularly influential in England. But this is likely to change after a recent decision (Miliangos v. George Frank Textiles Ltd [1975]) established that an English Court can give judgement in a foreign currency on a money claim in a foreign currency, so that, for instance, it is now possible in England to enforce a claim from a sale in Swiss francs. (*Financial Times*, 6 November, 1975: the report is reproduced in F. A. Hayek [31], pp. 45-6).

[2] Nussbaum [50], pp. 54-5.

paradigm (law, language and morals being the other main instances). When the medieval doctrine of the *valor impositus* was in this century revived by the much admired German Professor Knapp it prepared the way for a policy which in 1923 carried the German Mark down to $\frac{1}{1{,}000{,}000{,}000{,}000}$ of its former value!

Private money preferred

There certainly can be and has been money, even very satisfactory money, without government doing anything about it, though it has rarely been allowed to exist for long.[1] But a lesson is to be learned from the report of a Dutch author about China a hundred years ago who observed of the paper money then current in that part of the world that '*because it is not legal tender* and because it is no concern of the State it is generally accepted as money'.[2] We owe it to governments that within given national territories today in general only one kind of money is universally accepted. But whether this is desirable, or whether people could not, if they understood the advantage, get a much better kind of money without all the to-do about legal tender, is an open question. Moreover, a 'legal means of payment' (*gesetzliches Zahlungsmittel*) need not be specifically designated by a law. It is sufficient if the law enables the judge to decide in what sort of money a particular debt can be discharged.

The commonsense of the matter was put very clearly 80 years ago by a distinguished defender of a liberal economic policy, the lawyer, statistician and high civil servant Lord Farrer. In a paper written in 1895[3] he contended that if nations

> 'make nothing else but the standard unit [of value they have adopted] legal tender, there is no need and no room for

[1] Occasional attempts by the authorities of commercial cities to provide a money of at least a constant metallic content, such as the establishment of the Bank of Amsterdam, were for long periods fairly successful and their money used far beyond the national boundaries. But even in these cases the authorities sooner or later abused their quasi-monopoly positions. The Bank of Amsterdam was a state agency which people had to use for certain purposes and its money even as exclusive legal tender for payments above a certain amount. Nor was it available for ordinary small transactions or local business beyond the city limits. The same is roughly true of the similar experiments of Venice, Genoa, Hamburg and Nuremberg.

[2] Willem Vissering [61].

[3] Lord Farrer [17], p. 43.

the operation of any special law of legal tender. The ordinary law of contract does all that is necessary without any law giving special function to particular forms of currency. We have adopted a gold sovereign as our unit, or standard of value. If I promised to pay 100 sovereigns, it needs no special currency law of legal tender to say that I am bound to pay 100 sovereigns, and that, if required to pay the 100 sovereigns, I cannot discharge the obligation by anything else.'.

And he concludes, after examining typical applications of the legal tender conception, that

> '*Looking to the above cases of the use or abuse of the law of legal tender other than the last* [i.e. that of subsidiary coins] *we see that they possess one character in common—viz. that the law in all of them enables a debtor to pay and requires a creditor to receive something different from that which their contract contemplated.* In fact it is a forced and unnatural construction put upon the dealings of men by arbitrary power'.[1]

To this he adds a few lines later that 'any Law of Legal Tender is in its own nature "suspect" '.[2]

Legal tender creates uncertainty

The truth is indeed that legal tender is simply a legal device to force people to accept in fulfilment of a contract something

[1] *Ibid.*, p. 45. The *locus classicus* on this subject from which I undoubtedly derived my views on it, though I had forgotten this when I wrote the First Edition of this *Paper*, is Carl Menger's discussion in 1892 [43a] of legal tender under the even more appropriate equivalent German term *Zwangskurs*. See pp. 98-106 of the reprint, especially p. 101, where the *Zwangskurs* is described as 'eine Massregel, die in der überwiegenden Zahl der Fälle den Zweck hat, gegen den Willen der Bevöklerung, zumindest durch einen Missbrauch der Münzhoheit oder des Notenregals entstandene pathologische (also exceptionelle[?]) Formen von Umlaufsmitteln, durch einen Missbrauch der Justizhoheit dem Verkehr aufzudrängen oder in demselben zu erhalten'; and p. 104 where Menger describes it as 'ein auf die Forderungsberechtigten geübter gesetzlicher Zwang, bei Summenschulden (bisweilen auch bei Schulden anderer Art) solche Geldsorten als Zahlung anzunehmen, welche dem ausdrücklich oder stillschweigend vereinbarten Inhalte der Forderungen nicht entsprechen, oder dieselben sich zu einem Wert aufdrangen zu lassen, der ihrem Wert im freien Verkehr nicht entspricht'. Especially interesting also is the first footnote on p. 102 in which Menger points out that there had been fairly general agreement on this among the liberal economists of the first half of the 19th century, while during the second half of that century, through the influence of the (presumably German) lawyers, the economists were led erroneously to regard legal tender as an attribute of perfect money.

[2] *Ibid.*, p. 47.

they never intended when they made the contract. It becomes thus, in certain circumstances, a factor that intensifies the uncertainty of dealings and consists, as Lord Farrer also remarked in the same context,

> 'in substituting for the free operation of voluntary contract, and a law which simply enforces the performance of such contracts, an artificial construction of contracts such as would never occur to the parties unless forced upon them by an arbitrary law'.

All this is well illustrated by the historical occasion when the expression 'legal tender' became widely known and treated as a definition of money. In the notorious 'legal tender cases', fought before the Supreme Court of the United States after the Civil War, the issue was whether creditors must accept at par current dollars in settlement of their claims for money they had lent when the dollar had a much higher value.[1] The same problem arose even more acutely at the end of the great European inflations after the First World War when, even in the extreme case of the German Mark, the principle 'Mark is Mark' was enforced until the end—although later some efforts were made to offer limited compensation to the worst sufferers.[2]

Taxes and contracts

A government must of course be free to determine in what currency taxes are to be paid and to make contracts in any currency it chooses (in this way it can support a currency it issues or wants to favour), but there is no reason why it should not accept other units of accounting as the basis of the assessment of taxes. In non-contractual payments such as damages or compensations for torts, the courts would have to decide the currency in which they have to be paid, and might for this purpose have to develop new rules; but there should be no need for special legislation.

[1] Cf. Nussbaum [50], pp. 586-592.

[2] In Austria after 1922 the name 'Schumpeter' had become almost a curse word among ordinary people, referring to the principle that 'Krone is Krone', because the economist J. A. Schumpeter, during his short tenure as Minister of Finance, had put his name to an order of council, merely spelling out what was undoubtedly valid law, namely that debts incurred in crowns when they had a higher value could be repaid in depreciated crowns, ultimately worth only a 15,000th part of their original value.

There is a real difficulty if a government-issued currency is replaced by another because the government has disappeared as a result of conquest, revolution, or the break-up of a nation. In that event the government taking over will usually make legal provisions about the treatment of private contracts expressed in terms of the vanished currency. If a private issuing bank ceased to operate and was unable to redeem its issue, this currency would presumably become valueless and the holders would have no enforceable claim for compensation. But the courts may decide that in such a case contracts between third parties in terms of that currency, concluded when there was reason to expect it to be stable, would have to be fulfilled in some other currency that came to the nearest presumed intention of the parties to the contract.

VI. THE CONFUSION ABOUT GRESHAM'S LAW

It is a misunderstanding of what is called Gresham's law to believe that the tendency for bad money to drive out good money makes a government monopoly necessary. The distinguished economist W. S. Jevons emphatically stated the law in the form that better money cannot drive out worse precisely to prove this. It is true he argued then against a proposal of the philosopher Herbert Spencer to throw the coinage of gold open to free competition, at a time when the only different currencies contemplated were coins of gold and silver. Perhaps Jevons, who had been led to economics by his experience as assayer at a mint, even more than his contemporaries in general, did not seriously contemplate the possibility of any other kind of currency. Nevertheless his indignation about what he described as Spencer's proposal

> 'that, as we trust the grocer to furnish us with pounds of tea, and the baker to send us loaves of bread, so we might trust Heaton and Sons, or some of the other enterprising firms of Birmingham, to supply us with sovereigns and shillings at their own risk and profit',[1]

[1] W. S. Jevons [34], p. 64, as against Herbert Spencer [57].

VIII. PUTTING PRIVATE TOKEN MONEY INTO CIRCULATION

I shall assume for the rest of this discussion that it will be possible to establish a number of institutions in various parts of the world which are free to issue notes in competition and similarly to carry cheque accounts in their individual denominations. I shall call these institutions simply 'banks', or 'issue banks' when necessary to distinguish them from other banks that do not choose to issue notes. I shall further assume that the name or denomination a bank chooses for its issue will be protected like a brand name or trade mark against unauthorised use, and that there will be the same protection against forgery as against that of any other document. These banks will then be vying for the use of their issue by the public by making them as convenient to use as possible.

The private Swiss 'ducat'

Since readers will probably at once ask how such issues can come to be generally accepted as money, the best way to begin is probably to describe how I would proceed if I were in charge of, say, one of the major Swiss joint stock banks. Assuming it to be legally possible (which I have not examined), I would announce the issue of non-interest bearing certificates or notes, and the readiness to open current cheque accounts, in terms of a unit with a distinct registered trade name such as 'ducat'. The only legal obligation I would assume would be to redeem these notes and deposits on demand with, at the option of the holder, either 5 Swiss francs or 5 D-marks or 2 dollars per ducat. This redemption value would however be intended only as a floor below which the value of the unit could not fall because I would announce at the same time my intention to regulate the quantity of the ducats so as to keep their (precisely defined) purchasing power as nearly as possible constant. I would also explain to the public that I was fully aware I could hope to keep these ducats in circulation only if I fulfilled the expectation that their real value would be kept approximately constant. And I would announce that I proposed from time to time to state the precise commodity equivalent in terms of which I intended to keep the value of the ducat constant, but that I reserved the right, after announcement, to alter the composition of the commodity standard as

experience and the revealed preferences of the public suggested.

*

It would, however, clearly be necessary that, though it seems neither necessary nor desirable that the issuing bank legally commits itself to maintain the value of its unit, it should in its loan contracts specify that any loan could be repaid either at the nominal figure in its own currency, or by corresponding amounts of any other currency or currencies sufficient to buy in the market the commodity equivalent which at the time of making the loan it had used as its standard. Since the bank would have to issue its currency largely through lending, intending borrowers might well be deterred by the formal possibility of the bank arbitrarily raising the value of its currency, that they may well have to be explicitly reassured against such a possibility. **[1]

These certificates or notes, and the equivalent book credits, would be made available to the public by short-term loans or sale against other currencies. The units would presumably, because of the option they offered, sell from the outset at a premium above the value of any one of the currencies in which they were redeemable. And, as these governmental currencies continued to depreciate in real terms, this premium would increase. The real value at the price at which the ducats were first sold would serve as the standard the issuer would have to try to keep constant. If the existing currencies continued to depreciate (and the availability of a stable alternative might indeed accelerate the process) the demand for the stable currency would rapidly increase and competing enterprises offering similar but differently-named units would soon emerge.

The sale (over the counter or by auction) would initially be the chief form of issue of the new currency. After a regular market had established itself, it would normally be issued only in the course of ordinary banking business, i.e. through short-term loans.

Constant but not fixed value

It might be expedient that the issuing institution should from the outset announce precisely the collection of commodities in terms of which it would aim to keep the value of the 'ducat'

[1] [To assist readers of the First Edition to identify major additions, we have inserted a single asterisk at the beginning and double asterisks at the end of substantial, self-contained new passages. – ED.]

constant. But it would be neither necessary nor desirable that it tie itself legally to a particular standard. Experience of the response of the public to competing offers would gradually show which combination of commodities constituted the most desired standard at any time and place. Changes in the importance of the commodities, the volume in which they were traded, and the relative stability or sensitivity of their prices (especially the degree to which they were determined competitively or not) might suggest alterations to make the currency more popular. On the whole I would expect that, for reasons to be explained later (Section XIII), a collection of raw material prices, such as has been suggested as the basis of a commodity reserve standard,[1] would seem most appropriate, both from the point of view of the issuing bank and from that of the effects of the stability of the economic process as a whole.

Control of value by competition

In most respects, indeed, the proposed system should prove a more practicable method of achieving all that was hoped from a commodity reserve standard or some other form of 'tabular standard.' At the same time it would remove the necessity of making it fully automatic by taking the control from a monopolistic authority and entrusting it to private concerns. The threat of the speedy loss of their whole business if they failed to meet expectations (and how any government organisation would be certain to abuse the opportunity to play with raw material prices!) would provide a much stronger safeguard than any that could be devised against a government monopoly. Competition would certainly prove a more effective constraint, forcing the issuing institutions to keep the value of their currency constant (in terms of a stated collection of commodities), than would any obligation to redeem the currency in those commodities (or in gold). And it would be an infinitely cheaper method than the accumulation and the storing of valuable materials.

The kind of trust on which private money would rest would not be very different from the trust on which today all private banking rests (or in the United States rested before the governmental deposit insurance scheme!). People today trust that a bank, to preserve its business, will arrange its affairs so that it

[1] Cf. Hayek [30], pp. 318-320.

will at all times be able to exchange demand deposits for cash, although they know that banks do not have enough cash to do so if everyone exercised his right to demand instant payment at the same time. Similarly, under the proposed scheme, the managers of the bank would learn that its business depended on the unshaken confidence that it would continue to regulate its issue of ducats (etc.) so that their purchasing power remained approximately constant.

Is the risk in the venture therefore too big to justify entry by men with the kind of conservative temper its successful conduct probably requires?[1] It is not to be denied that, once announced and undertaken, the decision on how large the commitment was to grow would be taken out of the hands of the issuing institution. To achieve its announced aim of maintaining the purchasing power of its currency constant, the amount would have to be promptly adapted to any change of demand, whether increase or decrease. Indeed, so long as the bank succeeded in keeping the value of its currency constant, there would be little reason to fear a sudden large reduction of the demand for it (though successful competitors might well make considerable inroads on its circulation). The most embarrassing development might be a rapid growth of demand beyond the limits a private institution likes to handle. But we can be fairly sure that, in the event of such success, new competition would soon relieve a bank of this anxiety.

The issuing bank could, at first, at no prohibitive cost keep in cash a 100 per cent reserve of the currencies in terms of which it had undertaken to redeem its issue and still treat the premiums received as freely available for general business. But once these other currencies had, as the result of further

[1] On the question of its attractiveness the discussion by S. Fischer [18] of the notorious reluctance of enterprise to issue indexed bonds is somewhat relevant. It is true that a gradual increase of the value of the notes issued by a bank in terms of other concurrent currencies might produce a situation in which the aggregate value of its outstanding notes (*plus* its liabilities from other sources) would exceed its assets. The bank would of course not be legally liable to redeem its notes at this value, but it could preserve this business only if it did in fact promptly buy at the current rate any of its notes offered to it. So long as it succeeded in maintaining the real value of its notes, it would never be called upon to buy back more than a fraction of the outstanding circulation. Probably no one would doubt that an art dealer who owns the plates of the engravings of a famous artist could, so long as his works remained in fashion, maintain the market value of these engravings by judiciously selling and buying, even though he could never buy up all the existing prints. Similarly, a bank could certainly maintain the value of its notes even though it could never buy back all the outstanding ones.

inflation, substantially depreciated relative to the ducat, the bank would have to be prepared, in order to maintain the value of the ducat, to buy back substantial amounts of ducats at the prevailing higher rate of exchange. This means that it would have to be able rapidly to liquidate investments of very large amounts indeed. These investments would therefore have to be chosen very carefully if a temporary rush of demand for its currency were not to lead to later embarrassment when the institution that had initiated the development had to share the market with imitators. Incidentally, the difficulty of finding investments of an assured stable value to match similar obligations would not be anything like as difficult for such a bank as we are considering as present-day bankers seem to find it: all the loans made in its own currency would of course represent such stable assets. The curious fact that such an issuing bank would have claims and obligations in terms of a unit the value of which it determined itself, though it could not do so arbitrarily or capriciously without destroying the basis of its business, may at first appear disturbing but should not create real difficulties. What may at first appear somewhat puzzling accounting problems largely disappear when it is remembered that such a bank would of course keep its accounts in terms of its own currency. The outstanding notes and deposits of such a bank are not claims on it in terms of some other unit of value; it determines itself the value of the unit in terms of which it has debts and claims and keeps its books. This will cease to seem shocking when we remember that this is precisely what practically all central banks have been doing for nearly half a century—their notes were of course redeemable in precisely nothing. But notes which may appreciate relatively to most other capital assets may indeed present to accountants problems with which they never before had to deal. Initially the issuing bank would of course be under a legal obligation to redeem its currency in terms of the other currencies against which it was at first issued. But after it has existed for some time their value may have shrunk to very little or they may have altogether disappeared.[1]

[1] A real difficulty could arise if a sudden large increase in the demand for such a stable currency, perhaps due to some acute economic crisis, had to be met by selling large amounts of it against other currencies. The bank would of course have to prevent such a rise in the value and could do so only by increasing its supply. But selling against other currencies would give it assets likely to depreciate

[*Contd. on page 51*]

IX. COMPETITION BETWEEN BANKS ISSUING DIFFERENT CURRENCIES

It has for so long been treated as a self-evident proposition that the supply of money cannot be left to competition that probably few people could explain why. As we have seen, the explanation appears to be that it has always been assumed that there must be only *one* uniform kind of currency in a country, and that competition meant that its amount was to be determined by several agencies issuing it independently. It is, however, clearly not practicable to allow tokens with the same name and readily exchangeable against each other to be issued competitively, since nobody would be in a position to control their quantity and therefore be responsible for their value. The question we have to consider is whether competition between the issuers of clearly distinguishable kinds of currency consisting of *different* units would not give us a better kind of money than we have ever had, far outweighing the inconvenience of encountering (but for most people not even having to handle) more than one kind.

In this condition the value of the currency issued by one bank would not necessarily be affected by the supplies of other currencies by different institutions (private or governmental). And it should be in the power of each issuer of a distinct currency to regulate its quantity so as to make it most acceptable to the public—and competition would force him to do so. Indeed, he would know that the penalty for failing to fulfil the expectations raised would be the prompt loss of the business. Successful entry into it would evidently be a very profitable venture, and success would depend on establishing the credibility and trust that the bank was able and determined to carry out its declared intentions. It would seem that in this situation sheer desire for gain would produce a better money than government has ever produced.[1]

[*Contd. from page 50*]
in terms of its own currency. It probably could not increase its short-term lending very rapidly, even if it offered to lend at a very low rate of interest—even though in such a situation it would be safer to lend even at a small negative rate of interest than to sell against other currencies. And it would probably be possible to grant long-term loans at very low rates of interest against negotiable securities (in terms of its own currency) which it should be easy to sell if the sudden increase of demand for its currency should be as rapidly reversed.

[1] Apart from notes and cheque deposits in its distinctive currency, an issuing bank would clearly also have to provide fractional coins; and the availability of

Effects of competition

It seems to me to be fairly certain that

(a) *a money generally expected to preserve its purchasing power approximately constant would be in continuous demand so long as the people were free to use it;*

(b) *with such a continuing demand depending on success in keeping the value of the currency constant one could trust the issuing banks to make every effort to achieve this better than would any monopolist who runs no risk by depreciating his money;*

(c) *the issuing institution could achieve this result by regulating the quantity of its issue;* and

(d) *such a regulation of the quantity of each currency would constitute the best of all practicable methods of regulating the quantity of media of exchange for all possible purposes.*

Clearly a number of competing issuers of different currencies would have to compete in the quality of the currencies they offered for loan or sale. Once the competing issuers had credibly demonstrated that they provided currencies more suitable to the needs of the public than government has ever provided, there would be no obstacle to their becoming generally accepted in preference to the governmental cur-

[*Contd. from page 51*]

convenient fractional coins in that currency might well be an important factor in making it popular. It would also probably be the habitual use of one sort of fractional coins (especially in slot machines, fares, tips, etc.) which would secure the predominance of one currency in the retail trade of one locality. The effective competition between different currencies would probably be largely confined to inter-business use, with retail trade following the decisions about the currency in which wages and salaries were to be paid.

Certain special problems would arise where present sales practices are based on the general use of uniform coins of a few relatively small standard units, as, e.g., in vending machines, transportation or telephones. Probably even in localities in which several different currencies were in general use, one set of small coins would come to dominate. If, as seems probable, most of these competing currencies were kept at practically the same value, the technical problem of the use of coins might be solved in any one of various ways. One might be that one institution, e.g. an association of retailers, specialised in the issue of uniform coins at slightly fluctuating market prices. Tradesmen and transport and communication undertakings of a locality might join to sell, at market prices and probably through the banks, a common set of tokens for all automats in the locality. We can certainly expect commercial inventiveness rapidly to solve such minor difficulties. Another possible development would be the replacement of the present coins by plastic or similar tokens with electronic markings which every cash register and slot machine would be able to sort out, and the 'signature' of which would be legally protected against forgery as any other document of value.

rencies—at least in countries in which government had removed all obstacles to their use. The appearance and increasing use of the new currencies would, of course, decrease the demand for the existing national ones and, unless their volume was rapidly reduced, would lead to their depreciation. This is the process by which the unreliable currencies would gradually all be eliminated. The condition required in order that this displacement of the government money should terminate before it had entirely disappeared would be that government reformed and saw to it that the issue of its currency was regulated on the same principles as those of the competing private institutions. It is not very likely that it would succeed, because to prevent an accelerating depreciation of its currency it would have to respond to the new currencies by a rapid contraction of its own issue.

'A thousand hounds': the vigilant press

The competition between the issuing banks would be made very acute by the close scrutiny of their conduct by the press and at the currency exchange. For a decision so important for business as which currency to use in contracts and accounts, all possible information would be supplied daily in the financial press, and have to be provided by the issuing banks themselves for the information of the public. Indeed, a thousand hounds would be after the unfortunate banker who failed in the prompt responses required to ensure the safeguarding of the value of the currency he issues. The papers would probably print a

TABLE I

ILLUSTRATION OF POSSIBLE CURRENCY PRICE DEVIATIONS

Currency	*Deviation from* *Announced Standard* %	*Our Test Standard* %
Ducats (SGB)	−0·04	−0·04
Florins (FNB)	+0·02	+0·03
Mengers (WK)	+0·10	+0·10
Piasters (DBS)	−0·06	−0·12
Reals (CNB)	**−1·02**	**−1·01**
Shekels (ORT)	−0·45	−0·45
Talents (ATBC)	+0·26	+0·02

table daily, not only of the current rates of exchange between the currencies but also of the current value, and the deviation of each of the currencies likely to be used by their readers from the announced standard of value in terms of commodities. These tables might look something like Table I (with the initials of the issuing institution given after the name of the currency it issues).

Nothing would be more feared by the bankers than to see the quotation of their currency in heavy type to indicate that the real value had fallen below the standard of tolerance set by the paper publishing the table.

Three questions

This sketch of the competition between several private issuing institutions presupposes answers to a number of questions we shall have to examine in more detail in succeeding sections.

—The first is whether a competing institution issuing its distinctive currency will always be able to regulate its value by controlling its quantity so as to make it more attractive to people than other currencies, and how far other issuers of currencies can by their policy interfere with these efforts.

—The second is which value (or other attribute of a currency) the public will prefer if different banks announce that it is their intention (and demonstrate their ability) to keep announced values of their currency constant.

—A third and no less important question is whether the kind of money most people will individually prefer to use will also best serve the aims of all. Though one might at first think that this must necessarily be so, it is not inevitably true. It is conceivable that the success of people's efforts will depend not only on the money they themselves use but also on the effects of the money others use, and the benefits they derive for themselves from using a particular kind of money may conceivably be more than offset by the disturbances caused by its general use. I do not believe this to be the case in the present instance, but the question certainly requires explicit consideration.

Before we can discuss further the interaction between currencies it will be expedient to devote a section to precisely what we mean by money or currency and its different kinds, and the various ways in which they may differ from one another.

2024

SESSION 4

Applications

The Economic Organisation of a P.O.W. Camp

R. A. Radford

Economica, New Series, Volume 12, Issue 48 (Nov., 1945), 189-201.

Stable URL:
http://links.jstor.org/sici?sici=0013-0427%28194511%292%3A12%3A48%3C189%3ATEOOAP%3E2.0.CO%3B2-D

The Economic Organisation of a P.O.W. Camp

By R. A. Radford

Introduction

After allowance has been made for abnormal circumstances, the social institutions, ideas and habits of groups in the outside world are to be found reflected in a Prisoner of War Camp. It is an unusual but a vital society. Camp organisation and politics are matters of real concern to the inmates, as affecting their present and perhaps their future existences. Nor does this indicate any loss of proportion. No one pretends that camp matters are of any but local importance or of more than transient interest, but their importance there is great. They bulk large in a world of narrow horizons and it is suggested that any distortion of values lies rather in the minimisation than in the exaggeration of their importance. Human affairs are essentially practical matters and the measure of immediate effect on the lives of those directly concerned in them is to a large extent the criterion of their importance at that time and place. A prisoner can hold strong views on such subjects as whether or not all tinned meats shall be issued to individuals cold or be centrally cooked, without losing sight of the significance of the Atlantic Charter.

One aspect of social organisation is to be found in economic activity, and this, along with other manifestations of a group existence, is to be found in any P.O.W. camp. True, a prisoner is not dependent on his exertions for the provision of the necessaries, or even the luxuries of life, but through his economic activity, the exchange of goods and services, his standard of material comfort is considerably enhanced. And this is a serious matter to the prisoner: he is not "playing at shops" even though the small scale of the transactions and the simple expression of comfort and wants in terms of cigarettes and jam, razor blades and writing paper, make the urgency of those needs difficult to appreciate, even by an ex-prisoner of some three months' standing.

Nevertheless, it cannot be too strongly emphasised that economic activities do not bulk so large in prison society as they do in the larger world. There can be little production; as has been said the prisoner is independent of his exertions for the provision of the necessities and luxuries of life; the emphasis lies in exchange and the media of exchange. A prison camp is not to be compared with the seething crowd of higglers in a street market, any more than it is to be compared with the economic inertia of a family dinner table.

Naturally then, entertainment, academic and literary interests, games and discussions of the "other world" bulk larger in everyday life than they do in the life of more normal societies. But it would be

wrong to underestimate the importance of economic activity. Everyone receives a roughly equal share of essentials; it is by trade that individual preferences are given expression and comfort increased. All at some time, and most people regularly, make exchanges of one sort or another.

Although a P.O.W. camp provides a living example of a simple economy which might be used as an alternative to the Robinson Crusoe economy beloved by the text-books, and its simplicity renders the demonstration of certain economic hypotheses both amusing and instructive, it is suggested that the principal significance is sociological. True, there is interest in observing the growth of economic institutions and customs in a brand new society, small and simple enough to prevent detail from obscuring the basic pattern and disequilibrium from obscuring the working of the system. But the essential interest lies in the universality and the spontaneity of this economic life; it came into existence not by conscious imitation but as a response to the immediate needs and circumstances. Any similarity between prison organisation and outside organisation arises from similar stimuli evoking similar responses.

The following is as brief an account of the essential data as may render the narrative intelligible. The camps of which the writer had experience were Oflags and consequently the economy was not complicated by payments for work by the detaining power. They consisted normally of between 1,200 and 2,500 people, housed in a number of separate but intercommunicating bungalows, one company of 200 or so to a building. Each company formed a group within the main organisation and inside the company the room and the messing syndicate, a voluntary and spontaneous group who fed together, formed the constituent units.

Between individuals there was active trading in all consumer goods and in some services. Most trading was for food against cigarettes or other foodstuffs, but cigarettes rose from the status of a normal commodity to that of currency. RMk.s existed but had no circulation save for gambling debts, as few articles could be purchased with them from the canteen.

Our supplies consisted of rations provided by the detaining power and (principally) the contents of Red Cross food parcels—tinned milk, jam, butter, biscuits, bully, chocolate, sugar, etc., and cigarettes. So far the supplies to each person were equal and regular. Private parcels of clothing, toilet requisites and cigarettes were also received, and here equality ceased owing to the different numbers despatched and the vagaries of the post. All these articles were the subject of trade and exchange.

The Development and Organisation of the Market

Very soon after capture people realised that it was both undesirable and unnecessary, in view of the limited size and the equality of supplies,

to give away or to accept gifts of cigarettes or food. "Goodwill" developed into trading as a more equitable means of maximising individual satisfaction.

We reached a transit camp in Italy about a fortnight after capture and received ¼ of a Red Cross food parcel each a week later. At once exchanges, already established, multiplied in volume. Starting with simple direct barter, such as a non-smoker giving a smoker friend his cigarette issue in exchange for a chocolate ration, more complex exchanges soon became an accepted custom. Stories circulated of a padre who started off round the camp with a tin of cheese and five cigarettes and returned to his bed with a complete parcel in addition to his original cheese and cigarettes; the market was not yet perfect. Within a week or two, as the volume of trade grew, rough scales of exchange values came into existence. Sikhs, who had at first exchanged tinned beef for practically any other foodstuff, began to insist on jam and margarine. It was realised that a tin of jam was worth ½ lb. of margarine plus something else; that a cigarette issue was worth several chocolate issues, and a tin of diced carrots was worth practically nothing.

In this camp we did not visit other bungalows very much and prices varied from place to place; hence the germ of truth in the story of the itinerant priest. By the end of a month, when we reached our permanent camp, there was a lively trade in all commodities and their relative values were well known, and expressed not in terms of one another—one didn't quote bully in terms of sugar—but in terms of cigarettes. The cigarette became the standard of value. In the permanent camp people started by wandering through the bungalows calling their offers—"cheese for seven" (cigarettes)—and the hours after parcel issue were Bedlam. The inconveniences of this system soon led to its replacement by an Exchange and Mart notice board in every bungalow, where under the headings "name", "room number", "wanted" and "offered" sales and wants were advertised. When a deal went through, it was crossed off the board. The public and semi-permanent records of transactions led to cigarette prices being well known and thus tending to equality throughout the camp, although there were always opportunities for an astute trader to make a profit from arbitrage. With this development everyone, including non-smokers, was willing to sell for cigarettes, using them to buy at another time and place. Cigarettes became the normal currency, though, of course, barter was never extinguished.

The unity of the market and the prevalence of a single price varied directly with the general level of organisation and comfort in the camp. A transit camp was always chaotic and uncomfortable: people were overcrowded, no one knew where anyone else was living, and few took the trouble to find out. Organisation was too slender to include an Exchange and Mart board, and private advertisements were the most that appeared. Consequently a transit camp was not one

market but many. The price of a tin of salmon is known to have varied by two cigarettes in 20 between one end of a hut and the other. Despite a high level of organisation in Italy, the market was morcellated in this manner at the first transit camp we reached after our removal to Germany in the autumn of 1943. In this camp—Stalag VIIA at Moosburg in Bavaria—there were up to 50,000 prisoners of all nationalities. French, Russians, Italians and Jugo-Slavs were free to move about within the camp: British and Americans were confined to their compounds, although a few cigarettes given to a sentry would always procure permission for one or two men to visit other compounds. The people who first visited the highly organised French trading centre, with its stalls and known prices, found coffee extract—relatively cheap among the tea-drinking English—commanding a fancy price in biscuits or cigarettes, and some enterprising people made small fortunes that way. (Incidentally we found out later that much of the coffee went "over the wire" and sold for phenomenal prices at black market cafés in Munich: some of the French prisoners were said to have made substantial sums in RMk.s. This was one of the few occasions on which our normally closed economy came into contact with other economic worlds.)

Eventually public opinion grew hostile to these monopoly profits—not everyone could make contact with the French—and trading with them was put on a regulated basis. Each group of beds was given a quota of articles to offer and the transaction was carried out by accredited representatives from the British compound, with monopoly rights. The same method was used for trading with sentries elsewhere, as in this trade secrecy and reasonable prices had a peculiar importance, but as is ever the case with regulated companies, the interloper proved too strong.

The permanent camps in Germany saw the highest level of commercial organisation. In addition to the Exchange and Mart notice boards, a shop was organised as a public utility, controlled by representatives of the Senior British Officer, on a no profit basis. People left their surplus clothing, toilet requisites and food there until they were sold at a fixed price in cigarettes. Only sales in cigarettes were accepted—there was no barter— and there was no higgling. For food at least there were standard prices: clothing is less homogeneous and the price was decided around a norm by the seller and the shop manager in agreement; shirts would average say 80, ranging from 60 to 120 according to quality and age. Of food, the shop carried small stocks for convenience; the capital was provided by a loan from the bulk store of Red Cross cigarettes and repaid by a small commission taken on the first transactions. Thus the cigarette attained its fullest currency status, and the market was almost completely unified.

It is thus to be seen that a market came into existence without labour or production. The B.R.C.S. may be considered as "Nature" of the

text-book, and the articles of trade—food, clothing and cigarettes—as free gifts—land or manna. Despite this, and despite a roughly equal distribution of resources, a market came into spontaneous operation, and prices were fixed by the operation of supply and demand. It is difficult to reconcile this fact with the labour theory of value.

Actually there was an embryo labour market. Even when cigarettes were not scarce, there was usually some unlucky person willing to perform services for them. Laundrymen advertised at two cigarettes a garment. Battle-dress was scrubbed and pressed and a pair of trousers lent for the interim period for twelve. A good pastel portrait cost thirty or a tin of " Kam ". Odd tailoring and other jobs similarly had their prices.

There were also entrepreneurial services. There was a coffee stall owner who sold tea, coffee or cocoa at two cigarettes a cup, buying his raw materials at market prices and hiring labour to gather fuel and to stoke ; he actually enjoyed the services of a chartered accountant at one stage. After a period of great prosperity he overreached himself and failed disastrously for several hundred cigarettes. Such large-scale private enterprise was rare but several middlemen or professional traders existed. The padre in Italy, or the men at Moosburg who opened trading relations with the French, are examples : the more subdivided the market, the less perfect the advertisement of prices, and the less stable the prices, the greater was the scope for these operators. One man capitalised his knowledge of Urdu by buying meat from the Sikhs and selling butter and jam in return : as his operations became better known more and more people entered this trade, prices in the Indian Wing approximated more nearly to those elsewhere, though to the end a " contact " among the Indians was valuable, as linguistic difficulties prevented the trade from being quite free. Some were specialists in the Indian trade, the food, clothing or even the watch trade. Middlemen traded on their own account or on commission. Price rings and agreements were suspected and the traders certainly co-operated. Nor did they welcome newcomers. Unfortunately the writer knows little of the workings of these people : public opinion was hostile and the professionals were usually of a retiring disposition.

One trader in food and cigarettes, operating in a period of dearth, enjoyed a high reputation. His capital, carefully saved, was originally about 50 cigarettes, with which he bought rations on issue days and held them until the price rose just before the next issue. He also picked up a little by arbitrage ; several times a day he visited every Exchange or Mart notice board and took advantage of every discrepancy between prices of goods offered and wanted. His knowledge of prices, markets and names of those who had received cigarette parcels was phenomenal. By these means he kept himself smoking steadily—his profits—while his capital remained intact.

Sugar was issued on Saturday. about Tuesday two of us used to visit Sam and make a deal; as old customers he would advance as much of the price as he could spare then, and entered the transaction in a book. On Saturday morning he left cocoa tins on our beds for the ration, and picked them up on Saturday afternoon. We were hoping for a calendar at Christmas, but Sam failed too. He was left holding a big black treacle issue when the price fell, and in this weakened state was unable to withstand an unexpected arrival of parcels and the consequent price fluctuations. He paid in full, but from his capital. The next Tuesday, when I paid my usual visit he was out of business.

Credit entered into many, perhaps into most, transactions, in one form or another. Sam paid in advance as a rule for his purchases of future deliveries of sugar, but many buyers asked for credit, whether the commodity was sold spot or future. Naturally prices varied according to the terms of sale. A treacle ration might be advertised for four cigarettes now or five next week. And in the future market "bread now" was a vastly different thing from "bread Thursday". Bread was issued on Thursday and Monday, four and three days' rations respectively, and by Wednesday and Sunday night it had risen at least one cigarette per ration, from seven to eight, by supper time. One man always saved a ration to sell then at the peak price: his offer of "bread now" stood out on the board among a number of "bread Monday's" fetching one or two less, or not selling at all—and he always smoked on Sunday night.

The Cigarette Currency

Although cigarettes as currency exhibited certain peculiarities, they performed all the functions of a metallic currency as a unit of account, as a measure of value and as a store of value, and shared most of its characteristics. They were homogeneous, reasonably durable, and of convenient size for the smallest or, in packets, for the largest transactions. Incidentally, they could be clipped or sweated by rolling them between the fingers so that tobacco fell out.

Cigarettes were also subject to the working of Gresham's Law. Certain brands were more popular than others as smokes, but for currency purposes a cigarette was a cigarette. Consequently buyers used the poorer qualities and the Shop rarely saw the more popular brands: cigarettes such as Churchman's No. 1 were rarely used for trading. At one time cigarettes hand-rolled from pipe tobacco began to circulate. Pipe tobacco was issued in lieu of cigarettes by the Red Cross at a rate of 25 cigarettes to the ounce and this rate was standard in exchanges, but an ounce would produce 30 home-made cigarettes. Naturally, people with machine-made cigarettes broke them down and re-rolled the tobacco, and the real cigarette virtually disappeared from the market. Hand-rolled cigarettes were not homogeneous and prices could no longer be quoted in them with safety: each cigarette was examined before it was accepted and thin

ones were rejected, or extra demanded as a make-weight. For a time we suffered all the inconveniences of a debased currency.

Machine-made cigarettes were always universally acceptable, both for what they would buy and for themselves. It was this intrinsic value which gave rise to their principal disadvantage as currency, a disadvantage which exists, but to a far smaller extent, in the case of metallic currency;—that is, a strong demand for non-monetary purposes. Consequently our economy was repeatedly subject to deflation and to periods of monetary stringency. While the Red Cross issue of 50 or 25 cigarettes per man per week came in regularly, and while there were fair stocks held, the cigarette currency suited its purpose admirably. But when the issue was interrupted, stocks soon ran out, prices fell, trading declined in volume and became increasingly a matter of barter. This deflationary tendency was periodically offset by the sudden injection of new currency. Private cigarette parcels arrived in a trickle throughout the year, but the big numbers came in quarterly when the Red Cross received its allocation of transport. Several hundred thousand cigarettes might arrive in the space of a fortnight. Prices soared, and then began to fall, slowly at first but with increasing rapidity as stocks ran out, until the next big delivery. Most of our economic troubles could be attributed to this fundamental instability.

Price Movements

Many factors affected prices, the strongest and most noticeable being the periodical currency inflation and deflation described in the last paragraphs. The periodicity of this price cycle depended on cigarette and, to a far lesser extent, on food deliveries. At one time in the early days, before any private parcels had arrived and when there were no individual stocks, the weekly issue of cigarettes and food parcels occurred on a Monday. The non-monetary demand for cigarettes was great, and less elastic than the demand for food: consequently prices fluctuated weekly, falling towards Sunday night and rising sharply on Monday morning. Later, when many people held reserves, the weekly issue had no such effect, being too small a proportion of the total available. Credit allowed people with no reserves to meet their non-monetary demand over the week-end.

The general price level was affected by other factors. An influx of new prisoners, proverbially hungry, raised it. Heavy air raids in the vicinity of the camp probably increased the non-monetary demand for cigarettes and accentuated deflation. Good and bad war news certainly had its effect, and the general waves of optimism and pessimism which swept the camp were reflected in prices. Before breakfast one morning in March of this year, a rumour of the arrival of parcels and cigarettes was circulated. Within ten minutes I sold a treacle ration, for four cigarettes (hitherto offered in vain for three), and many similar deals went through. By 10 o'clock the rumour was denied, and treacle that day found no more buyers even at two cigarettes.

More interesting than changes in the general price level were changes in the price structure. Changes in the supply of a commodity, in the German ration scale or in the make-up of Red Cross parcels, would raise the price of one commodity relative to others. Tins of oatmeal, once a rare and much sought after luxury in the parcels, became a commonplace in 1943, and the price fell. In hot weather the demand for cocoa fell, and that for soap rose. A new recipe would be reflected in the price level: the discovery that raisins and sugar could be turned into an alcoholic liquor of remarkable potency reacted permanently on the dried fruit market. The invention of electric immersion heaters run off the power points made tea, a drug on the market in Italy, a certain seller in Germany.

In August, 1944, the supplies of parcels and cigarettes were both halved. Since both sides of the equation were changed in the same degree, changes in prices were not anticipated. But this was not the case: the non-monetary demand for cigarettes was less elastic than the demand for food, and food prices fell a little. More important however were the changes in the price structure. German margarine and jam, hitherto valueless owing to adequate supplies of Canadian butter and marmalade, acquired a new value. Chocolate, popular and a certain seller, and sugar, fell. Bread rose; several standing contracts of bread for cigarettes were broken, especially when the bread ration was reduced a few weeks later.

In February, 1945, the German soldier who drove the ration waggon was found to be willing to exchange loaves of bread at the rate of one loaf for a bar of chocolate. Those in the know began selling bread and buying chocolate, by then almost unsaleable in a period of serious deflation. Bread, at about 40, fell slightly; chocolate rose from 15; the supply of bread was not enough for the two commodities to reach parity, but the tendency was unmistakable.

The substitution of German margarine for Canadian butter when parcels were halved naturally affected their relative values, margarine appreciating at the expense of butter. Similarly, two brands of dried milk, hitherto differing in quality and therefore in price by five cigarettes a tin, came together in price as the wider substitution of the cheaper raised its relative value.

Enough has been cited to show that any change in conditions affected both the general price level and the price structure. It was this latter phenomenon which wrecked our planned economy.

Paper Currency—Bully Marks

Around D-Day, food and cigarettes were plentiful, business was brisk and the camp in an optimistic mood. Consequently the Entertainments Committee felt the moment opportune to launch a restaurant, where food and hot drinks were sold while a band and variety turns performed. Earlier experiments, both public and private, had pointed the way, and the scheme was a great success. Food was bought at

market prices to provide the meals and the small profits were devoted to a reserve fund and used to bribe Germans to provide grease-paints and other necessities for the camp theatre. Originally meals were sold for cigarettes but this meant that the whole scheme was vulnerable to the periodic deflationary waves, and furthermore heavy smokers were unlikely to attend much. The whole success of the scheme depended on an adequate amount of food being offered for sale in the normal manner.

To increase and facilitate trade, and to stimulate supplies and customers therefore, and secondarily to avoid the worst effects of deflation when it should come, a paper currency was organised by the Restaurant and the Shop. The Shop bought food on behalf of the Restaurant with paper notes and the paper was accepted equally with the cigarettes in the Restaurant or Shop, and passed back to the Shop to purchase more food. The Shop acted as a bank of issue. The paper money was backed 100 per cent. by food; hence its name, the Bully Mark. The BMk. was backed 100 per cent. by food: there could be no over-issues, as is permissible with a normal bank of issue, since the eventual dispersal of the camp and consequent redemption of all BMk.s was anticipated in the near future.

Originally one BMk. was worth one cigarette and for a short time both circulated freely inside and outside the Restaurant. Prices were quoted in BMk.s and cigarettes with equal freedom—and for a short time the BMk. showed signs of replacing the cigarette as currency. The BMk. was tied to food, but not to cigarettes: as it was issued against food, say 45 for a tin of milk and so on, any reduction in the BMk. prices of food would have meant that there were unbacked BMk.s in circulation. But the price of both food and BMk.s could and did fluctuate with the supply of cigarettes.

While the Restaurant flourished, the scheme was a success: the Restaurant bought heavily, all foods were saleable and prices were stable.

In August parcels and cigarettes were halved and the Camp was bombed. The Restaurant closed for a short while and sales of food became difficult. Even when the Restaurant reopened, the food and cigarette shortage became increasingly acute and people were unwilling to convert such valuable goods into paper and to hold them for luxuries like snacks and tea. Less of the right kinds of food for the Restaurant were sold, and the Shop became glutted with dried fruit, chocolate, sugar, etc., which the Restaurant could not buy. The price level and the price structure changed. The BMk. fell to four-fifths of a cigarette and eventually farther still, and it became unacceptable save in the Restaurant. There was a flight from the BMk., no longer convertible into cigarettes or popular foods. The cigarette re-established itself.

But the BMk. was sound! The Restaurant closed in the New Year with a progressive food shortage and the long evenings without lights due to intensified Allied air raids, and BMk.s could only be spent in

the Coffee Bar—relict of the Restaurant—or on the few unpopular foods in the Shop, the owners of which were prepared to accept them. In the end all holders of BMk.s were paid in full, in cups of coffee or in prunes. People who had bought BMk.s for cigarettes or valuable jam or biscuits in their heyday were aggrieved that they should have stood the loss involved by their restricted choice, but they suffered no actual loss of market value.

Price Fixing

Along with this scheme came a determined attempt at a planned economy, at price fixing. The Medical Officer had long been anxious to control food sales, for fear of some people selling too much, to the detriment of their health. The deflationary waves and their effects on prices were inconvenient to all and would be dangerous to the Restaurant which had to carry stocks. Furthermore, unless the BMk. was convertible into cigarettes at about par it had little chance of gaining confidence and of succeeding as a currency. As has been explained, the BMk. was tied to food but could not be tied to cigarettes, which fluctuated in value. Hence, while BMk. prices of food were fixed for all time, cigarette prices of food and BMk.s varied.

The Shop, backed by the Senior British Officer, was now in a position to enforce price control both inside and outside its walls. Hitherto a standard price had been fixed for food left for sale in the shop, and prices outside were roughly in conformity with this scale, which was recommended as a " guide " to sellers, but fluctuated a good deal around it. Sales in the Shop at recommended prices were apt to be slow though a good price might be obtained : sales outside could be made more quickly at lower prices. (If sales outside were to be at higher prices, goods were withdrawn from the Shop until the recommended price rose : but the recommended price was sluggish and could not follow the market closely by reason of its very purpose, which was stability.) The Exchange and Mart notice boards came under the control of the Shop : advertisements which exceeded a 5 per cent. departure from the recommended scale were liable to be crossed out by authority : unauthorised sales were discouraged by authority and also by public opinion, strongly in favour of a just and stable price. (Recommended prices were fixed partly from market data, partly on the advice of the M.O.)

At first the recommended scale was a success : the Restaurant, a big buyer, kept prices stable around this level : opinion and the 5 per cent. tolerance helped. But when the price level fell with the August cuts and the price structure changed, the recommended scale was too rigid. Unchanged at first, as no deflation was expected, the scale was tardily lowered, but the prices of goods on the new scale remained in the same relation to one another, owing to the BMk., while on the market the price structure had changed. And the modifying influence of the Restaurant had gone. The scale was moved

up and down several times, slowly following the inflationary and deflationary waves, but it was rarely adjusted to changes in the price structure. More and more advertisements were crossed off the board, and black market sales at unauthorised prices increased: eventually public opinion turned against the recommended scale and authority gave up the struggle. In the last few weeks, with unparalleled deflation, prices fell with alarming rapidity, no scales existed, and supply and demand, alone and unmellowed, determined prices.

Public Opinion

Public opinion on the subject of trading was vocal if confused and changeable, and generalisations as to its direction are difficult and dangerous. A tiny minority held that all trading was undesirable as it engendered an unsavoury atmosphere; occasional frauds and sharp practices were cited as proof. Certain forms of trading were more generally condemned; trade with the Germans was criticised by many. Red Cross toilet articles, which were in short supply and only issued in cases of actual need, were excluded from trade by law and opinion working in unshakable harmony. At one time, when there had been several cases of malnutrition reported among the more devoted smokers, no trade in German rations was permitted, as the victims became an additional burden on the depleted food reserves of the Hospital. But while certain activities were condemned as anti-social, trade itself was practised, and its utility appreciated, by almost everyone in the camp.

More interesting was opinion on middlemen and prices. Taken as a whole, opinion was hostile to the middleman. His function, and his hard work in bringing buyer and seller together, were ignored; profits were not regarded as a reward for labour, but as the result of sharp practices. Despite the fact that his very existence was proof to the contrary, the middleman was held to be redundant in view of the existence of an official Shop and the Exchange and Mart. Appreciation only came his way when he was willing to advance the price of a sugar ration, or to buy goods spot and carry them against a future sale. In these cases the element of risk was obvious to all, and the convenience of the service was felt to merit some reward. Particularly unpopular was the middleman with an element of monopoly, the man who contacted the ration wagon driver, or the man who utilised his knowledge of Urdu. And middlemen as a group were blamed for reducing prices. Opinion notwithstanding, most people dealt with a middleman, whether consciously or unconsciously, at some time or another.

There was a strong feeling that everything had its "just price" in cigarettes. While the assessment of the just price, which incidentally varied between camps, was impossible of explanation, this price was nevertheless pretty closely known. It can best be defined as the price usually fetched by an article in good times when cigarettes were

plentiful. The "just price" changed slowly; it was unaffected by short-term variations in supply, and while opinion might be resigned to departures from the "just price", a strong feeling of resentment persisted. A more satisfactory definition of the "just price" is impossible. Everyone knew what it was, though no one could explain why it should be so.

As soon as prices began to fall with a cigarette shortage, a clamour arose, particularly against those who held reserves and who bought at reduced prices. Sellers at cut prices were criticised and their activities referred to as the black market. In every period of dearth the explosive question of "should non-smokers receive a cigarette ration?" was discussed to profitless length. Unfortunately, it was the non-smoker, or the light smoker with his reserves, along with the hated middleman, who weathered the storm most easily.

The popularity of the price-fixing scheme, and such success as it enjoyed, were undoubtedly the result of this body of opinion. On several occasions the fall of prices was delayed by the general support given to the recommended scale. The onset of deflation was marked by a period of sluggish trade; prices stayed up but no one bought. Then prices fell on the black market, and the volume of trade revived in that quarter. Even when the recommended scale was revised, the volume of trade in the Shop would remain low. Opinion was always overruled by the hard facts of the market.

Curious arguments were advanced to justify price fixing. The recommended prices were in some way related to the calorific values of the foods offered: hence some were overvalued and never sold at these prices. One argument ran as follows:—not everyone has private cigarette parcels: thus, when prices were high and trade good in the summer of 1944, only the lucky rich could buy. This was unfair to the man with few cigarettes. When prices fell in the following winter, prices should be pegged high so that the rich, who had enjoyed life in the summer, should put many cigarettes into circulation. The fact that those who sold to the rich in the summer had also enjoyed life then, and the fact that in the winter there was always someone willing to sell at low prices were ignored. Such arguments were hotly debated each night after the approach of Allied aircraft extinguished all lights at 8 p.m. But prices moved with the supply of cigarettes, and refused to stay fixed in accordance with a theory of ethics.

Conclusion

The economic organisation described was both elaborate and smooth-working in the summer of 1944. Then came the August cuts and deflation. Prices fell, rallied with deliveries of cigarette parcels in September and December, and fell again. In January, 1945, supplies of Red Cross cigarettes ran out: and prices slumped still further: in February the supplies of food parcels were exhausted and the depression became a blizzard. Food, itself scarce, was almost given away in

order to meet the non-monetary demand for cigarettes. Laundries ceased to operate, or worked for £s or RMk.s: food and cigarettes sold for fancy prices in £s, hitherto unheard of. The Restaurant was a memory and the BMk. a joke. The Shop was empty and the Exchange and Mart notices were full of unaccepted offers for cigarettes. Barter increased in volume, becoming a larger proportion of a smaller volume of trade. This, the first serious and prolonged food shortage in the writer's experience, caused the price structure to change again, partly because German rations were not easily divisible. A margarine ration gradually sank in value until it exchanged directly for a treacle ration. Sugar slumped sadly. Only bread retained its value. Several thousand cigarettes, the capital of the Shop, were distributed without any noticeable effect. A few fractional parcel and cigarette issues, such as one-sixth of a parcel and twelve cigarettes each, led to momentary price recoveries and feverish trade, especially when they coincided with good news from the Western Front, but the general position remained unaltered.

By April, 1945, chaos had replaced order in the economic sphere: sales were difficult, prices lacked stability. Economics has been defined as the science of distributing limited means among unlimited and competing ends. On 12th April, with the arrival of elements of the 30th U.S. Infantry Division, the ushering in of an age of plenty demonstrated the hypothesis that with infinite means economic organisation and activity would be redundant, as every want could be satisfied without effort.

Search...

AIER | AMERICAN INSTITUTE for ECONOMIC RESEARCH

(https://www.aier.org)

Yes, This Time We'll Have Inflation, and Here's Why

James D. Gwartney (https://www.aier.org/people/james-d-gwartney/) – February 25, 2021

(/custom-search-results/?post_date=02252021) Reading Time: 4 minutes

After several decades of relatively low rates of inflation, it is easy to think that we will continue to see little change in prices. But the seeds of inflation have been planted.

Purchases of financial assets, primarily Treasury securities, are the primary tool the Fed uses to control the money supply. When the Fed purchases Treasury securities, it provides the federal government with spendable funds. When these funds are spent, the money supply increases. Essentially, money is created out of nothing.

Since 2008, the Fed has expanded these purchases far more rapidly than in the past. Fed holdings of financial assets quadrupled during 2008-2019, expanding from $900 billion to $4.1 trillion. During the past 12 months, these purchases have surged another 80 percent, soaring to $7.47 trillion In February 2021.

Until now, however, the inflation rate has remained relatively low. Two factors have combined to keep inflation in check. Starting in October 2008, the Fed began paying banks interest on their deposits held with the Fed. These interest payments encourage banks to hold larger Fed deposits, rather than undertake investments and extend loans. During 2008-2019, the Fed used these interest payments to induce banks to hold a larger share of their assets as reserves, dampening the money growth and inflationary effects of the Fed's huge expanded purchases of financial assets.

In addition, worldwide interest rates declined to historic low levels during the decade following the Great Recession. A variety of factors caused this. One that has been largely ignored was a dramatic demographic shift in high-income countries, as the number of people in the lending phase of life (roughly age 50 to 75) increased relative to those in the borrowing phase (under age 50). The resulting low and declining interest rates reduced the opportunity cost of holding money, causing the velocity or turnover rate of money to plummet. As the result of this combination of factors, the Fed's huge increase in purchases of financial assets and money creation has, to this point, exerted only a minimal impact on inflation.

However, this favorable scenario is about to reverse course. Three major factors underlie the reversal.

First, the Fed's current money creation dwarfs those of recent decades. Propelled by the $3.6 trillion Covid-19 spending financed by borrowing from the Fed, the narrow measure of the money supply known as M1 has expanded from $4.0 trillion to $6.8 trillion during the past 12 months, a 70 percent increase. By way of comparison, the 12-month increases in M1 during the inflationary 1970s never exceeded 10 percent. The largest previous single-year M1 increase in recent decades was a 21 percent figure in the aftermath of the Great Recession. The story is the same for the broader M2 measure of money, which has increased by 25 percent during the past year. The next largest 12-month expansions in M2 during the past 75 years were the 1975-1976 increase of 13.8 percent during the double-digit inflation of the 1970s and the 10.3 percent increase during 2011.

One has to go all the way back to World War II to find anything comparable to the money supply increases of the past 12 months. Moreover, even these gigantic figures understate the current monetary surge. The Treasury is currently holding more than a trillion dollars of committed funds in its bank account, which will be added to the money supply when they are spent in the next few months. Congress is expected to add additional fuel to the fire with the $1.9 trillion spending package currently under consideration.

Second, the inflation triggered by the huge monetary expansion will increase the expected rate of inflation and nominal interest rates. In turn, the higher nominal interest rates will cause the velocity of money to increase, unleashing additional inflationary pressures from the rapid money growth of 2008–2019. Rising inflation rates and higher nominal interest rates are two peas in the same pod. When the former occurs the latter will follow. As inflation pushes nominal interest rates upward, the recent reductions in velocity will reverse, adding to the inflationary pressure.

Third, the mandated shutdowns have resulted in a huge pent-up demand. Once a sizable share of the population has received the vaccine and the virus is brought under control, spending will increase substantially, providing an

additional boost to both demand and the general level of prices. While the shutdown imposed large costs on small businesses and employees who lost their jobs as the result of business closures, the income of another sizable share of Americans continued as usual. In fact, the incomes of many in this group received an additional boost from the government's direct payments to households in the Covid relief packages. Restrictions on travel, tourism, sporting events, and other entertainment curtailed spending and the personal savings rate more than doubled, jumping from 8 to 17 percent during 2020. Now, many people who have money have been cooped up for too long, and when they can, they are going to spend "big time."

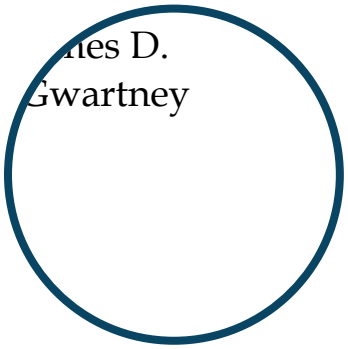
The current situation is similar to that of World War II and its aftermath. As during the pandemic, the surge in government spending during the war was financed mostly by debt and money creation. Similarly, spending options were severely limited during both of these events. The performance of the economy after the war provides insight on our likely future. Propelled by pent-up demand and wartime savings, the post-war recovery was stronger than expected. But it was also characterized by inflation. The CPI and GDP deflator (two measures of inflation) both increased at double-digit rates during 1946 and 1947.

The next two or three years are likely to be similar. It is a virtual certainty that inflation will rise, perhaps to double-digit levels. Demand will be strong and real GDP is likely to grow, albeit at a sluggish rate. Currently, the political forces supportive of anti-growth policies such as trade restrictions, higher minimum wages, perverse energy regulations, and cronyism appear to be on the rise, and they will dampen future growth. These policies, along with the uncertainties accompanying inflation and the burden of financing the larger outstanding debt will slow real growth. But inflation is going to be the big story of the post-pandemic economy. Get ready for an inflationary ride.

The author would like to thank Jane Shaw Stroup, Joseph Connors, and David Macpherson for their assistance in the preparation of this article.

READ MORE (/article)

James D. Gwartney (https://www.aier.org/people/james-d-gwartney/)



James D. Gwartney is professor of economics and policy sciences at Florida State University. He is an expert on such economic issues as taxation, labor policy, and the economic analysis of government.

His research has focused on the measurement and determination of factors that influence cross-country differences in income levels and growth rates. Dr. Gwartney is the co-author of the annual report, Economic Freedom of the World, which provides information on the consistency of institutions and policies with economic freedom for more than 150 countries.

Get notified of new articles from James D. Gwartney and AIER.

SUBSCRIBE (/email-signup)

otherwise be absent. Disequilibrium money prices signal the existence of a faulty pattern of resource use and they provide entrepreneurs with an incentive to eliminate this discoordination. '[D]isequilibrium prices provide through profits a *feedback mechanism* for their own correction that makes them a more sophisticated informational device than they may seem when concentrating only on their equilibrium role' (Thomsen 1992: 3; emphasis added). The advantage of this feedback mechanism is that entrepreneurs do not have to identify instances of imperfect coordination between market transactions per se; they simply have to identify price discrepancies for the same good.

The existence of money and money prices enhances the self-efficacy of entrepreneurs because it enables them to engage in 'economic calculation', which is an advanced cognitive operation. Money enhances the capacity of the entrepreneurial mind to deal with problems of a higher order. '[T]he very possibility of purposive action within the framework of social division of labor depends on the faculty of the human intellect to conceive cardinal numbers and manipulate them in arithmetic operations' (Salerno 1990b: 27).

Money strengthens the ability of entrepreneurs to make *rational* judgemental decisions about reallocating scarce productive resources in light of the higher-valued uses that they have discovered. '[M]onetary economic calculation is the intellectual basis of the market economy' (Mises 1966: 259). In an economy based on a complex division of labour, economic calculation is pivotal to the ability of entrepreneurs to make rational economic decisions and systematic production plans under uncertainty. 'Every single step of entrepreneurial activities is subject to scrutiny by monetary calculation' (Mises 1966: 229).

So what exactly is economic calculation? Economic calculation is an individual's numerical computation, in terms of money prices, of the consequences of his or her actions in the market (Mises 1966: 201–213). It includes the *ex ante* computation of the expected outcomes of planned actions (e.g. the expected costs and expected revenues of a project) and the retrospective computation of the results of past transactions (i.e. entrepreneurial success and failure, profit and loss).

Money is necessary for economic calculation because computation requires a common denominator (a common unit of calculation), to which exchange ratios can be reduced. 'The whole structure of the calculations of the entrepreneur and the consumer rests on the process of valuing commodities in money. Money has thus become an aid that the human mind is no longer able to dispense with in making economic calculations' (Mises 1971: 48–49).

The numbers that entrepreneurs employ in economic calculation are money prices – that is, the numerical ratios of exchange between money and other economic goods and services. '[P]rices are not measured in money; they consist in money' (Mises 1966: 217). Money prices are the mental tools of rational entrepreneurial planning and economic calculation.

The bounded rationality of entrepreneurs means that, in the absence of money prices, it would be impossible for them to monitor directly all the changes in market conditions and to determine the corresponding alterations in exchange ratios between numerous pairs of goods and services. If there were n commodities in a direct barter economy, and each commodity could be exchanged with each of the other $n - 1$ commodities, then an entrepreneur participating regularly in the market would have to monitor $\frac{1}{2}n(n - 1)$ separate exchange ratios. If, however, one of the commodities were to become a generally accepted medium of exchange and a unit of account, then the exchange matrix would be reduced dramatically to $n - 1$ exchange ratios (Brunner and Meltzer 1971: 787). By establishing money prices for every good and service, money enables market participants to compare prices easily: 'Instead of a myriad of isolated markets for each good and every other good, each good exchanges for money, and the exchange ratios between every good and every other good can easily be estimated by observing their money prices' (Rothbard 1993: 203).

Psychological research on numerical cognition supports the suggestion that money reduces the cost of obtaining, comparing and remembering exchange ratios. By supplying a common unit of calculation (i.e. a common denominator), money reduces the *problem difficulty* effect (Ashcraft 1992; Dehaene 1992). Without money, individuals would have to engage in more complex cognitive processes. The absence of a common denominator reduces the efficiency of direct memory retrieval as the predominant strategy for performing simple arithmetic in price comparisons. It also greatly increases the need to use relatively slow calculation procedures, and it affects more than just the encoding of numerical stimuli. The reaction time to perceive and quantify price differences, and the scope for error, will all be greater under barter than for comparable transactions involving exchange ratios expressed in a common unit of account.

Money prices are set by unconditional market exchanges of goods and services against money. It follows that entrepreneurs can only engage in economic calculation within an institutional framework that supports private ownership of the means of production and the exchange of goods and services against money. 'Economic calculation cannot comprehend things which are not sold and bought against money' (Mises 1966: 214).

If fallible entrepreneurs are to have the competency (efficacy) to distinguish, from among the myriad technologically feasible projects, those ventures which are profitable from those that are not, they must have recourse to calculative 'aids' of the 'human mind', such as that provided by money, in assessing the possible consequences of planned actions. Monetary calculation provides 'a guide amid the bewildering throng of economic possibilities' (Mises 1981: 101). Moreover, economic calculation does not require perfect-knowledge surrogates. In spite of the 'imperfect configuration of disequilibrium relative prices' that occurs in a world of out-of-equilibrium trading, economic calculation can still aid entrepreneurs

in separating out the profitable investment projects from those that are not (Lavoie 1985a: 57).

Before embarking upon a full-scale operation, entrepreneurs are able to assess alternative methods of production by means of monetary calculations, to evaluate them symbolically and to reject or tentatively accept them on the grounds of their predicted consequences. 'Through the medium of symbols people can solve problems without having to enact all the various alternative solutions' (Bandura 1977b: 13). The cognitive capacity for economic calculation is thus embedded in symbolic ability. For example, a business plan with projections of money income and money outlay is an objective symbolic depiction of the entrepreneur's ideas. Because economic calculation enables entrepreneurs to sift their way early on through possible projects whose implementation would consume a great deal of time and effort, it quickens the pace of their learning.

Economic calculation is not only required at the *ex ante* stage of forming an entrepreneurial plan, however. It also facilitates implementation of plans and their evaluation *ex post*. '[I]n a world of partial ignorance, there is much more to effective decision-making than the selection of the correct alternative from the choice set' (Loasby 1976: 88). Because entrepreneurial decisions are made in conditions that are liable to lead to error, economic calculation is also required in assessing the success of the decisions that they have made.

Once entrepreneurs have chosen a project, they may engage in monetary calculation and profit and loss accounting so as to gauge the progress of their ventures. Feedback that conveys successful accomplishment and competence strengthens entrepreneurs' perceptions of their self-efficacy. But even negative feedback may facilitate entrepreneurial learning and the development of entrepreneurial capabilities and perceived self-efficacy. Feedback in the form of a discrepancy between actual performance and the planned decision is an important source of information in the formulation of entrepreneurs' efficacy perceptions (cf. Bandura and Cervone 1986). The frequency and immediacy of these signals help entrepreneurs to discover their errors earlier than they would otherwise in the course of carrying out their plans. In the absence of money prices, entrepreneurs would not be able to avoid errors that they can now avoid because of the condensed and detailed (albeit imperfect) information made available to them by price signals.

A whole host of abstract pecuniary concepts and practices has evolved from the use of money and economic calculation. These products of the human mind are essential tools of entrepreneurial thought and action and support what Mitchell (1937: 160) calls the 'whole countinghouse attitude toward economic activities'. For example, entrepreneurs rely upon the system of double-entry bookkeeping and the process of capital budgeting. These methods enable entrepreneurs to establish business hierarchies in which they delegate subordinate entrepreneurial tasks to their managers, thereby enabling themselves to focus on the big picture. The system of business accounting, in conjunction with market-based transfer prices for factor inputs, enables entrepreneurs to calculate the profit or loss imputed to different divisions of the enterprise and to determine their contribution to overall performance. Lewin (1998) adds the twist that business hierarchies are also *necessary* for the smooth functioning of market processes because they facilitate economic calculation by supplying a cognitive backdrop and set of procedures for the attribution of input costs.

A less obvious point is that economic calculation *always* pertains to the future. The introduction of money and monetary calculation supports 'futurity', which Lane defines as 'longer-term purposiveness, a teleological orientation that necessarily points to the future' (1991: 87). Entrepreneurs engage in economic calculation to handle changes in market conditions and to anticipate the future. They use past money prices as a starting point for forming their expectations of the future structure of market prices for particular goods and services (what Mises calls entrepreneurial 'appraisement') (Mises 1966: 331–332).[2] Entrepreneurs only take past money prices into account in economic calculation to the extent that it assists them in adjusting their actions to their current expectations of the future. But because future prices cannot be deduced from past prices, economic calculation does not supply entrepreneurs with certain knowledge of the future constellation of the market or definitive knowledge of a profit opportunity. '[I]t does not deprive action of its speculative character' (Mises 1966: 214). Even retrospective assessments of past action may rely upon speculative anticipations of future prices (e.g. the prices that will be paid on the market for assets acquired in previous transactions).

Money and entrepreneurs' locus of control beliefs

The discussion so far has focused upon the impact of money upon entrepreneurs' perceptions of self-efficacy as a determinant of alertness. The introduction of money might also have a tendency to strengthen entrepreneurs' beliefs that economic outcomes are contingent upon action rather than external forces beyond their control. That is, it might reinforce a sense of internal locus of control – the other component of personal agency beliefs that enhances alertness.

A system of direct exchange or barter is one in which all market transactions involve 'the exchange of one useful good for another, each for purposes of direct use by the party to the exchange' (Rothbard 1993: 160). There is no universally used medium of exchange; money does not exist. Goods and services are directly traded on the market against other goods and services. Each party acquires a good either for the direct satisfaction of his or her wants or for the services it renders directly to the production of other goods.

Barter is a very cumbersome and high-transaction-cost system in which every entrepreneurial transaction requires a 'double coincidence' of wants. At the very least, the entrepreneur must find two individuals, each of whom

But it suffices to emphasise that not only can money emerge spontaneously out of a barter economy through the interaction of market participants, but also money can *only* evolve organically through private market exchanges and cannot be consciously created by the state through central planning (Menger 1892; 1994). It is 'epistemologically impossible for the State to create a common medium of exchange outside the context of exchange practice' (Boettke 2001: 255). The state by itself does not have the power to transform a commodity into a generally accepted medium of exchange.[6] Although the legal order of a society can have an effect on the money character of commodities, it is only the common commercial practice of all the individuals who participate in the market that can create money. (This also implies that the institutions of property and contract must exist before money can emerge.) Furthermore, as a result of changes in economic conditions, business customs and the marketability of various assets, a commodity that once served as money may eventually be displaced over time by another more liquid commodity. 'Advanced transactions technologies, liquid spot and futures markets, and the development of financial intermediaries all contribute to the replacement of one set of exchange media by another' (Cowen and Kroszner 1994: 596–597).

Although, as a Big Player, the state or its central bank does not have the power to create a medium of exchange, it does have the power to change the value of the monetary unit and to discoordinate market processes. A Big Player is defined as 'anyone who habitually exercises *discretionary power* to influence the market while himself remaining wholly or largely immune from the discipline of profit and loss' (Koppl and Yeager 1996: 368; emphasis added). It is clear that as a Big Player, a central bank is an archetype of a 'powerful other' in Rotter's (1966) psychological sense of the term. When a central bank pursues inflationary monetary policies or intervenes to influence short-term interest rates or to defend an exchange rate, its actions distort the structure of relative prices and reduce the information content of market signals – money prices less accurately reflect underlying market fundamentals (Butos and Koppl 1999: 269).

A change in the epistemic quality of money prices brought about by discretionary monetary policies inhibits entrepreneurial alertness to market opportunities through its stifling effects on entrepreneurial self-efficacy and locus of control beliefs. Big-Player intervention in the monetary sphere reduces entrepreneurs' sense of self-efficacy because it reduces their ability to obtain useful information about market developments. 'Big Players weaken the epistemic foundations of entrepreneurial action' (Butos and Koppl 1999: 272). Discretionary policy reduces the capacity of entrepreneurs to know their environment. Entrepreneurs become less confident in the reliability of their expectations, and they feel less able to respond to changes in the data of the market.

In addition, the discretionary actions of a central bank diminish the internality of entrepreneurs' locus of control beliefs because they weaken the perceived contingency of events upon what entrepreneurs discover and do in the market.[7] Economic outcomes (profit and loss) become more dependent upon external forces, that is the actions of key central bank officials and politicians. 'Entrepreneurial success now becomes more closely tied to discovering or anticipating the behavior of the Big Player' (Butos and Koppl 1999: 270). As mentioned earlier, money is the linking pin in market transactions, the other side of all market exchanges in a monetary economy. This implies that if state intervention alters the value of the monetary unit, it also changes the pattern of exchanges in the market, the constellation of relative prices, the structure of profit opportunities and therefore the allocation of real productive resources among competing uses. When a Big Player enters the game, observed market events and the pattern of relative prices are no longer as tightly linked to shifts in the data of the market, such as changes in consumer preferences, technology and resource endowments, because they depend more and more upon political factors. '[D]iscretionary policy attenuates the link between action and the economic environment by making the underlying reality [of the market] less important. ... In short, Big Players introduce free parameters into the environment that may change in unpredictable and arbitrary ways' (Butos and Koppl 1999: 270). As a result of the declining contingency of entrepreneurial outcomes upon their actions, entrepreneurs shift their attention away from market opportunities towards political developments and the capricious actions of the monetary authority.

As long as there is a central bank and government monopoly of money, political tinkering with the money supply is likely to continue (Wagner 1989a). The state monopoly over money gives decision-makers who control or influence government an additional means for pursuing their political interests. The incumbent political party can try to buy votes through monetary manipulations that change the structure of relative prices and the distribution of income for the benefit of those whose political support is desired. Greater economic instability is thus the inevitable by-product of the rational pursuit of political gain (electoral success) in a democracy in which the state has a monetary monopoly. Cartelisation of the banking industry is also a possibility:

> Instead of being the agency for the provision of a public good, a central bank seems more reasonably seen as an agent for cartelizing a banking system that otherwise would be competitive. The member banks gain from the formation of this cartel, as the members of any cartel gain from the cartel's formation. It is the government that makes this cartel possible, and which enforces the cartel, so it too would share in the gains from the monopolization of money and credit.
>
> (Wagner 1980: 14)

One approach to constraining discretionary monetary interventions is to denationalise money and to introduce competition into the monetary sphere.

In addition to providing a sounder banking system and more effective administration of the money supply, a free banking system of competitive note issue is likely to enhance the alertness of entrepreneurs both within and outside the banking sector.[8] Under free banking, entrepreneurs in the market for monetary services are free from interference by monetary authorities, such as a central bank or a government deposit insurance agency. They have the freedom to issue bank notes bearing their brand name, set interest rates and introduce new types of loans and deposits, subject only to the general laws of contract. Entrepreneurs recognise that there is no official lender of last resort that will provide them with emergency loans if they make poor business decisions. Eliminating the central bank and privatising the money supply credibly signals to all potential entrepreneurs that the government is committed to a limited role in economic affairs. It takes away the ability of the government to finance its expenditures through inflation, and it thereby assures entrepreneurs that government officials will not manipulate the value of the monetary unit and distort the structure of relative prices. All in all, free banking is likely to strengthen entrepreneurs' sense of personal agency and to heighten their alertness to profit opportunities.

Political decentralisation

By itself, the rule of law, even in Hayek's thoroughgoing version, does not secure a system of personal liberty and vigorous entrepreneurship. The rule of law, and the formal requirements it imposes (certainty, generality and equality), does not guarantee an effective bulwark against the discretionary power of government. It is a mistake to present the rule of law as a *sufficient* condition for individual freedom and the unimpeded operation of spontaneous market forces, when it is possibly only a *necessary* condition (Hamowy 1971: 375). The same could be said of well-defined property rights and freedom of contract.

The fundamental problem is one of *credible political commitment* to maintaining markets and protecting economic liberties. If entrepreneurs are going to discover lucrative business opportunities and engage in the innovation that creates wealth, a political infrastructure is needed that credibly restricts the power of the state to expropriate entrepreneurial profits and other people's property.

Federalism

A sound political foundation for market processes requires political decentralisation of the authority to determine economic policy, which so far has been best achieved by a federal constitutional order. A federal system of government is one of the institutional configurations most conducive to economic freedom and development (de Tocqueville 1990a: 172). It is a way

of minimising the potential for political coercion by injecting the principles of the market into the political structure (Buchanan 1995; 1995/96). The processes of entry, exit and intergovernmental competition that are essential features of federalist structures serve as a constitutional limitation on governmental power. By protecting the autonomy of private decision-makers, federalism strengthens people's beliefs in their ability to exert power over what happens in their lives, and it raises their general level of attentiveness to market opportunities. 'In a federation economic policy will have to take the form of providing a rational permanent framework within which *individual initiative will have the largest possible scope* and will be made to work as beneficently as possible' (Hayek 1948: 268; emphasis added).

So what exactly is federalism, and what form of federalism is most conducive to a strong sense of agency and heightened entrepreneurial alertness? If the overall objective is to maximise people's entrepreneurial propensity to discover opportunities in the dynamic world in which they live, then it is relatively straightforward to define the ideal of competitive federalism:

> A central government authority should be constitutionally restricted to the enforcement of *openness of the whole nexus of economic interaction*. Within this scope, the central authority must be strong, but it should not be allowed to extend beyond the limits constitutionally defined. Other political-collective activities should be carried out, if at all, by separate state-provincial units that exist side-by-side, as competitors of sorts, in the inclusive polity.
>
> (Buchanan 1995/96: 265; emphasis added)

Riker's (1994) definition of federalism captures the main elements common to most conceptions of this polycentric political order. He defines federalism as a political system that comprises a *hierarchy of governments*, each of which has its own clearly specified *sphere of authority*, and that institutionalises the *autonomy* of each government by means of *self-enforcing limitations* on political discretion.

Though necessary, these conditions are not sufficient for activating acute levels of entrepreneurial alertness and promoting wealth creation. Not all forms of federal organisation generate thriving market processes and vibrant economic development. A case in point is the *de jure* federal system in Argentina, once the world's fourth-richest country in the 1930s. Thus, extra criteria are needed for pinpointing the subset of federal systems that effectively supports competitive markets. To this end, I draw upon Weingast's (1995) concept of '*market-preserving* federalism'.

It is predicted that market-preserving federalism is the political structure that is most conducive to internal locus of control beliefs, strong self-efficacy and heightened alertness on the part of economic actors. To qualify as market preserving, according to Weingast (1995), a particular federal system must