

**ECONOMIC RESEARCH REPORTS**

THE EVOLUTION OF A  
FREE BANKING SYSTEM

by

George A. Selgin  
and  
Lawrence H. White

R.R. #85-38

November 1985

**C. V. STARR CENTER  
FOR APPLIED ECONOMICS**



**NEW YORK UNIVERSITY  
FACULTY OF ARTS AND SCIENCE  
DEPARTMENT OF ECONOMICS  
WASHINGTON SQUARE  
NEW YORK, N.Y. 10003**

THE EVOLUTION OF A FREE BANKING SYSTEM

George A. Selgin  
Department of Economics  
George Mason University

Lawrence H. White  
Department of Economics  
New York University

October, 1985

The authors are indebted to the Institute for Humane Studies for the opportunity to work together on this article, and to Chris Fauvelais, David Glasner, Israel Kirzner, and Mario Rizzo for useful comments. White's research is supported by the Scaife Foundation. The usual disclaimer applies.

## THE EVOLUTION OF A FREE BANKING SYSTEM

In recent years monetary theorists have produced a substantial literature on the properties of a completely unregulated monetary system.<sup>1</sup> They have made diverse assumptions concerning the institutional features of such a system, ranging for example from the proliferation of numerous competing private fiat currencies at one extreme to the complete disappearance of money at the other. While these assumptions have helped generate clear-cut conclusions, their plausibility or realism in light of historical experience is open to serious doubt. This doubtfulness may unfortunately be taken to suggest that any discussion of an unregulated monetary system (or "free banking" system) must be tenuous and highly speculative. This paper intends to show to the contrary, that general institutional features can be theoretically derived for a free banking model and need not be arbitrarily posited. The model can then yield plausible conclusions regarding the likely implications of deregulation in the real world.

This paper attempts to identify the predictable institutional features of free banking system by constructing a conjectural history of the evolution of an unregulated monetary system. This procedure provides insight into the arrangements that would actually have arisen had state intervention never been present. It therefore helps to identify the degree to which various features of current monetary and banking institutions are rooted in market forces and the degree to which they have grown out of regulatory intervention. Such information is surely relevant, though of course not sufficient, for drawing conclusions about the ways in which future deregulation would modify institutions. The investigation shows that sophisticated and orderly monetary arrangements, whose institutional features are described, emerge in the absence

of regulation. No strong claims are advanced here about the welfare properties of these arrangements.

The evolutionary account presented here is conjectural because of necessity it cannot rest solely on inferences from actual events. There have been few historical episodes of free banking, and none realizing it in a pure form. An effort is made to avoid generalizing from the peculiarly local features of any actual episode and from features attributable to the presence of regulation. Actual historical examples are cited in order to illustrate the theoretical processes under discussion. The primary aim, however, is to present a theoretical account rather than to do economic history in the usual sense. The analysis deliberately avoids basing conjectures on the assumption that agents consciously aim at erecting a banking system with desirable features. Instead it views them as motivated purely by private goals and beliefs. The institutional patterns derived are thus the unintended consequences of self-interested, individual actions, or in other words are products of "invisible-hand" processes.<sup>2</sup>

The evolution of a free banking system, following the emergence of standardized commodity money, is seen as proceeding through three stages. These are, first, the development of basic money-transfer services which substitute for the physical transportation of specie; second, the emergence of easily assignable and negotiable bank demand liabilities (inside money); and third, the development of arrangements for the routine exchange ("clearing") of inside monies among rival banks. The historical time that might separate these stages is not crucial. The path of development, rather than being one of steady progress as will be pictured here, may in practice involve dead ends or creative leaps along the way. What is essential, though, is that each stage of progress is a logical outgrowth of the circumstances that preceded it. In other words,

each step in this process of evolution can be shown to originate in individuals' discovery of new ways in which to promote their self-interest, even though the final result is an arrangement of far greater complexity and importance than any of them can have been contemplated.

### Commodity Money

Since the use of money logically and historically precedes the emergence of banking firms, an account of the evolution of a free banking system appropriately begins with an account of the origin of money. By starting at this earliest stage we can show that a state intervention, which might influence later developments, is not necessary in order to establish a monetary system's base money or unit of account. Carl Menger (1892) has furnished an invisible-hand explanation, consistent with the weight of historical and anthropological evidence, of how money originated as a product of undesigned or spontaneous evolution. Menger's theory also serves as a useful prototype for the rest of the discussion of the evolution of a free banking system, since the purpose of this discussion is to show how complex banking institutions can evolve in a similarly spontaneous fashion.

Menger (1892, p. 250) argues that money, rather than being invented or adopted by some conscious collective decision, emerges as "the spontaneous outcome, the unpremeditated resultant, of particular, individual efforts of the members of a society."<sup>3</sup> In a pre-monetary society agents relying upon barter offer goods in exchange only for other goods having some direct use-value to them. The number of bargains struck this way must be very small, owing to the improbability of what Jevons termed a "double coincidence of wants." Before long some individual, perhaps one frustrated by his inability to find a taker for his goods among those who offer what he wants to buy, realizes that he can

increase his opportunities for successful barter by adopting a two-stage procedure. He can trade his goods for some--any--more saleable good, regardless of its value to him, which will more easily find a taker among those selling what he wants. It follows that the earliest media of exchange are simply the most "barterable" of goods: scarce items for which there is widespread want. The increased demand for these things owing to their use as media of exchange further enhances the saleability of the chosen commodities. Other agents eventually recognize the gains achieved by those using indirect exchange, and emulate them, even though they may be unaware of the reason for the advantages from using a medium of exchange. This emulation further enhances the acceptance of the most widely accepted media. Thus a "widening chasm" comes to separate one or several goods from all others with respect to their relative saleability. The snowballing of saleability results in the spontaneous appearance of generally-accepted media of exchange. Eventually agents throughout an economically integrated region converge on a single good as the most generally accepted medium of exchange. This is money.

Historical evidence on primitive monies indicates that cattle were often the most frequently exchanged commodity, and that a standardized "cow" was the earliest unit of account. Cattle were, however, ill-suited to serve as a general medium of exchange because of their relative non-transportability and non-uniformity.<sup>4</sup> It was the discovery of metals and of methods for working them that finally allowed the use of money to replace barter on a widespread basis.<sup>5</sup>

The emergence of coinage can also be explained as a spontaneous development. Like the idea of money itself, that of coinage need not "flash out upon the world" or spring full-blown into the mind of any ruler (Burns 1927a, p. 285). Rather it is an unplanned result of attempts by merchants to minimize the

necessity for assessing and weighing amounts of crude commodity money received in exchange. Arthur R. Burns (1927a, pp. 297-304; 1927b, p. 59) has illustrated this process with evidence from ancient Lydia, where electrum (a naturally occurring silver-gold alloy) coins came into early use. The merchants of Lydia at first adopted the practice of marking irregular electrum blobs after having assessed their quality. Then a merchant recognizing his own mark or the mark of another merchant with whom he was familiar could take advantage of this to avoid the trouble and cost of re-assessment. Marking soon gave way to stamping or punching, which eventually led to specialists' making coins in their modern form. Techniques for covering the entire surface of coins with type provided a safeguard against clipping and sweating, and so allowed coinage to serve as a guarantee of weight as well as of quality.

States seem to have monopolized their coinage early in history, but not because public authorities were the best makers of coin or because coinage is a natural monopoly.<sup>6</sup> Rather, the facts suggest that state coinage monopolies were regularly established by means of legal compulsion and for reasons of prestige and monopoly profit. Once political authorities erected a state coinage they would prohibit private issues. State-minted coins thereby functioned both as a symbol of rule and as a source of profits from shaving, clipping, and seignorage. In this way by the end of the seventh century coinage throughout the Greek world became a state function.<sup>7</sup>

In the absence of state interference, coinage would be a private industry encompassing various competing brands. Less reliable or underweight coins would either circulate at a discount or be driven from circulation. Competition would, given the demand for readily exchangeable coins, promote the emergence of commodity money pieces of consistent weight and fineness, valued according to bullion content plus a premium equal to the marginal cost of mintage.<sup>8</sup>

### The Development of Banking Firms

The use of metallic money entails considerable inconveniences in counting and transporting coin. Traders, particularly those frequently making large or distant exchanges, will naturally adopt lower-cost means of transferring ownership of money where such means are developed. One likely locus for their development is the market where local coins are exchanged for foreign coins. (Standard coins may differ interlocally either because of local state interventions or more generally because of geographic diseconomies in reputation-building for mints. A coin-exchange market then naturally arises with interlocal trade.) A trader who uses a money-changer finds that he must count and carry in local coin each time he wants to acquire foreign coin, or vice versa. He may then discover cost savings in establishing a standing account balance, which he can build up at his convenience and draw upon as desired. The money-changer's inventories equip him to provide such accounts, which constitute demand deposits, and even to make overdrafts available. These deposits may initially be non-transferrable. But it will soon be apparent, in cases where one customer plans to withdraw coins in order to pay a recipient who intends to redeposit the coins with the same exchange-banker, that the transfer is more easily made by meeting at the banker's place of business, or most easily yet by persuading the banker to make the transfer on his books without any coins ever being handled.

Thus trading individuals come by steps to the practice of keeping money balances with agencies which can make payments by means of ledger-account transfers. Money-transfer services of this sort, provided by money-changers and bill-brokers in twelfth-century Genoa and at the medieval trade fairs in Champagne, mark the earliest forms of banking for which records exist.<sup>9</sup> In time



there were "transfer banks", as Raymond de Roover (1974, p. 184) has called them, in all the major European trading centers.

Money-transfer services may also develop in connection with deposits which are initially made not for purpose of trade but for safekeeping. The well-known story of the origins of goldsmith-banking in seventeenth-century England illustrates such a development. Persons of substantial wealth may temporarily place sums of commodity money in the strongboxes of scrivenors, goldsmiths, mintmasters, and other reputable tradespeople accustomed to having and protecting valuable property.<sup>10</sup> Coin and bullion thus lodged for safekeeping must be physically withdrawn and transferred when its owner desires to use it as a means of payment. Exchanges in which the recipient of coin will redeposit it in the same vault offer (just as redeposits with a money-changer or bill-broker do) obvious advantages in making the transfer at the vault, or better yet in simply notifying the vault's custodian to make the transfer on his books. The English goldsmiths evidently became transfer bankers in this manner during the seventeenth century, when they "began to keep a 'running cash' for the convenience of merchants and country gentlemen" (De Roover 1974, pp. 183-4).

Transfer banking is not connected with intermediation between borrowers and lenders when the banker acts strictly as a warehouseman, giving deposit receipts which are regular warehouse dockets.<sup>11</sup> Thus the strict warehouse-banker is a bailee rather than a debtor to his depositors, and can make loans only out of his personal wealth. Two conditions make it possible, however, to take advantage of the interest income available from lending out depositors' balances, even while satisfying depositors' desire to have their funds withdrawable on demand: (1) money is fungible, which allows a depositor to be repaid in coin and bullion not identical to the coin and bullion he had originally brought in, and (2) the law of large numbers with random withdrawals

and deposits makes a fractional reserve sufficient to meet actual withdrawal demands even though any single account may be removed without notice.<sup>12</sup> The lending of depositors' balances is a significant innovation that taps a vast new source of loanable funds and alters fundamentally the relationship of the banker to his depositor-customers. Historically in England, according to Richards (1929, pp. 223), "the bailee . . . developed into the debtor of the depositor; and the depositor became an investor who loaned his money . . . for a consideration." Thus money "warehouse receipts" became merely ready promissory notes. W. R. Bisschop (1910, p. 50n.) reports that English warehouse-bankers had begun to function as intermediaries by the time of Charles II (1660~85): "any deposit made in any other shape than ornament was looked upon by them as a free loan." Competition for deposits prompted the payment of interest on deposits, and the attractiveness of interest on safe and accessible deposits in turn made the practice of depositing widespread among all ranks of people (Powell 1966, pp. 56~57). Under these circumstances the effective money supply obviously becomes greater than the existing stock of specie alone.

#### Assignability and Negotiability

At this stage, though, the most important banking procedures and devices have yet to develop. Many purchases are still made with actual coin, which means that substantial wealth is tied up in the circulating medium. Bank depositors, in order to satisfy changing needs for money at hand, make frequent withdrawals from and deposits into their bank balances. These actions, though they may in the aggregate largely cancel out through the law of large numbers, require the banks to hold greater precautionary commodity money reserves than they need to hold when payments practices become more sophisticated. Greater sophistication comes with the emergence of negotiable instruments able to pass

easily in exchange from one person to another which replace non-negotiable deposit receipts and coin in transactions balances. The development of more efficient means for the bank-mediated transfer of liabilities represented by book entries also serves to reduce the need for coin in circulation.

The development of assignability and negotiability may proceed through several steps. Assume that initially deposited money (whether "warehoused" or entrusted to the banker for lending at interest) may be assigned by the depositor to another party by oral transfer only--a procedure requiring the presence of all three parties to the exchange or their attorneys.<sup>13</sup> Money "warehouse receipts" or promissory notes cannot be assigned by the mere endorsement of their owner and without the banker acting as witness. An important innovation is then the development of bank-issued promissory notes transferable by endorsement. These assignable notes in turn give way to fully negotiable bank notes assigned to no one in particular but instead payable to the bearer on demand. A parallel development is that of the non-negotiable check which can be used to transfer deposits represented in book-entries. Eventually non-negotiable checks give way to negotiable ones.<sup>14</sup> Thus the modern forms of inside money--redeemable bearer bank notes and checkable deposits--are established. Once this stage is reached it is not difficult for bankers to conceive what Hartley Withers (1920, p. 24) calls "the epoch-making notion"--on our view it is merely an incremental step--of giving inside money not only to depositors of metal, but also to those who come to borrow money.

The use of inside money enhances both the convenience of bank customers and the profits of banks, so that only the possible reluctance of courts to enforce obligations represented by assigned or bearer paper stands in the way of its rapid development.<sup>15</sup> In England the step of recognizing bearer notes was taken

during the reign of Charles II, about the time when warehouse banking was giving way to fractional-reserve transfer banking. At first the courts gave their grudging approval to the growing practice of repeated endorsement of promissory notes. Then after some controversy, fully negotiable notes were recognized by Act of Parliament. In France, Holland, and Italy during the 16th century merchants' checks "drawn in blank" circulated within limited circles and may have cleared the way for the appearance of bank notes. (Usher 1943, p. 189; Richards 1965, pp, 46, 225).

#### Regular Note\*Exchange

Further economies in the use of commodity money require, first, more complete use of inside money in place of commodity money in circulation and, second, more complete development of bank note and check clearing facilities so that there is less need for commodity money reserves. It is relatively straightforward to show that bankers and other agents pursuing their self-interest are indeed led to improve the acceptability of inside money and the efficiency of banking operations.

At the stage we have now reached, although bank notes are less cumbersome than coin, and checkable deposits are both convenient for certain transactions and (in an unregulated system) interest-paying, some coin will still remain in circulation. Consumers can better trust a local bank's notes than a distant bank's notes because consumers are more aware of the former's likelihood of being honored and more familiar with their appearance, hence less prone to accepting forgeries. It follows that the cost to a bank of building a positive reputation for its issues\*particularly in the matter of note convertibility\*is higher in places further removed from the place of issue and redemption. In the early stages of banking development the par circulation of every bank's notes is

geographically relatively limited.<sup>16</sup> People who generally hold the inside money of a local bank but who wish to do business in distant towns must either undertake the bothersome step of redeeming some of their holdings for gold and incur the added inconvenience of transporting coin, or suffer a loss in value on their notes by carrying them to a distant locale where they are accepted only at a discount if at all. (The alternative practice of keeping on hand stocks of notes from each of the localities with which they deal is likely to be prohibitively costly in terms of foregone interest.) In general, inside money will (when it first comes into use) only be convenient for transactions between persons located in the vicinity of its issuer, and coin will continue to be held alongside notes of like denomination.

The use of commodity money in circulation necessitates the holding of commodity money bank reserves greater than those required by the use of inside money in its place, because the withdrawal of commodity money for spending generates more volatile reserve outflows than the spending of notes or deposits. In this situation, however, profit opportunities happen to exist which sponsor actions leading to more general acceptance of particular inside monies. The discounting of notes outside of the immediate neighborhood of the issuing bank's office creates an arbitrage opportunity if the par value of notes (i.e., their face redemption value in commodity money) exceeds the price at which they can be purchased for commodity money or local issues in a distant town plus the cost of transaction and transportation. "Note brokers" can make a business, just as retail foreign currency brokers do today, of buying discounted non-local notes and either transporting them to their par circulation areas or reselling them to travelers bound for those areas.

Competition eventually reduces note discounts to the level of transaction and transportation costs plus a factor for risk whenever notes of an unreliable

issuer are involved. By offering to accept the notes of unfamiliar banks at minimal commission rates, brokers increase the general acceptability of notes, and help to promote their use instead of commodity money in persons' balances.

To this point it has implicitly been assumed that banks refuse to accept one another's notes. This is not an unreasonable assumption, since banks have as many reasons as other individuals do to refuse notes which are unfamiliar to them or difficult to redeem. They have in addition a further incentive for refusing to accept notes from rival banks, which is that by doing so they help to limit the acceptability of these notes, thereby enhancing the demand for their own issues.<sup>17</sup>

Nevertheless note brokerage presents opportunities for profit to bankers. Banks can out-compete other brokers because, unlike other brokers, the banks can issue their own notes (or deposit balances) to purchase "foreign" notes and need not hold costly till money. Each bank has an additional incentive to accept rival notes: the new issue is a source of interest earnings if the notes acquired are redeemed sooner than the notes issued, because interest-earning assets can be held in the interim. This profit from "float" can be continually renewed. In other words, a bank can maintain a permanently larger circulation of its own notes by continually replacing other notes with its own, and correspondingly can hold a larger portfolio of earning assets than it otherwise could.<sup>18</sup> Where transaction and transportation costs and risk are low enough, competition for circulation will narrow the brokerage fee to zero, that is, will lead the banks to general acceptance of one another's notes at par. The development of par acceptance, it should be emphasized, does not require that the banks explicitly and mutually agree to such a policy.

An alternative scenario, which assumes strategic behavior by the banks,

leads to the same result. A bank may aggressively deploy the foreign notes it collects by suddenly returning large quantities of notes to their issuers for redemption in commodity money, hoping to force an unprepared issuer to suspend payments. The aggressor may thereby gain market share by damaging a rival's reputation or even forcing it into liquidation. This tactic, historically known as "note duelling", naturally leads the other issuers to respond in kind. Collecting and redeeming the first bank's notes not only returns the damage, but helps replenish their own reserves. Par acceptance of its rival's notes allows a bank to collect them in greater quantities, and may therefore be adopted. (Arbitrage redemption precludes acceptance above par.) In the long run, nonaggression between any two banks should be expected to emerge, being less costly for both sides.<sup>19</sup> But again, no explicitly negotiated pact is necessary. It only takes a single bank acting without any cooperation from other banks to nudge the rest towards par acceptance (zero brokerage fees) as a defensive measure to maintain their reserves and circulation.

In New England at the beginning of the 19th century it was the Boston banks that gave the nudge that put the whole region--with its multitude of "country" banks of issue far removed from the city--on a par acceptance basis (Trivoli 1979). In Scotland it was the Royal Bank of Scotland, which when it opened for business in 1727 immediately began accepting at par the notes of the Bank of Scotland, at that time its only rival, and instigated a note duel that lasted for a few months (Checkland 1975, p. 17). In both places established banks, even after they had begun accepting each other's notes at par, sometimes refused to take the notes of new entrants. They soon changed their policies because the new banks that accepted and redeemed their notes were draining their reserves, whereas the established banks were not offsetting this by engaging in the same practice in reverse.

In the long run, banks that accept other banks' notes at par improve the market for their own notes and, unintentionally, for the notes that they accept.<sup>20</sup> Statistics from Boston dramatically illustrate this: from 1824 to 1833 the note circulation of the Boston banks increased 57%, but the Boston circulation of country banks increased 148%, despite the Boston banks' intent to drive the country banks out of business (Lake 1947, p. 186; Trivoli 1979, pp. 10-12). There is room for all banks to gain because the spread of par acceptance makes inside money more attractive to hold relative to commodity money. Since notes from one town are now accepted in a distant town at par, there is no longer good reason for people to lug around cumbersome amounts of commodity money. As par note acceptance developed in Scotland, Canada, and New England—places where note issue was least restricted—during the 19th century, gold virtually disappeared from circulation.<sup>21</sup> In England and the rest of the United States, where banking (and note issue in particular) were less free, considerable amounts of gold remained in circulation.

Even the complete displacement of commodity money in circulation by inside money does not, however, exhaust the possibilities for economizing on the use of commodity money. A large part of the specie formerly used in circulation to settle exchanges outside the banks may still be needed to settle clearings among them. The banks can substantially reduce their prudentially required holdings of commodity money by participating in regular note exchanges which allow them opportunities to offset their mutual obligations. Only net clearings rather than gross clearings are then settled in commodity money. The probability of any given-sized reserve loss in a given period is accordingly reduced (by the law of large numbers) and each bank can prudently reduce its ratio of reserves to demand liabilities.



The gains to be had from rationalization of note-exchange are illustrated by the provincial Scottish banks before 1771, which practiced par acceptance without regular exchange. Note duelling among these banks was not uncommon.<sup>22</sup> In order for a bank to guard against redemption raids it would have to keep substantial reserves. Munn (1981, pp. 23-24, 141) reports that one of the Scottish provincial banking companies kept reserves equal at one point to 61.2% of its inside-money liabilities in order to protect itself against the raids of rivals. More typically reserves during this period were in the neighborhood of 10% of total liabilities. This contrasts with reserve ratios of around 2% that were typical under the Scottish free banking system after note clearings became routine.

The advantages of regular note exchange are great and obvious enough to secure its eventual adoption. This has been the case in every historical instance of relatively unregulated plural note issue. Instead of actively seeking to buy notes in the marketplace, which is not only costly but an unreliable means of promoting circulation, banks soon find it convenient to accept the notes of their rivals only as these are brought to them for deposit or exchange. Notes thus accepted are taken to their issuers promptly in order to be exchanged for commodity-money reserves which can be profitably employed. Note-duelling, as explained above, soon ceases to be profitable. Finally, once banks realize the benefits from offsetting note debits with one another, they reach formal agreements to engage in regular note exchange.

### Clearinghouses

The most readily made arrangements for note exchange are simply bilateral. In a system of more than two issuers, however, banks will be led to discover and exploit the reserve-holding and other economies made possible by a multilateral

note-exchange system. The reserve-holding economies may be illustrated by means of a simple numerical example. Suppose that there are three banks, A, B, and C, and that A has collected \$20,000 of B's notes, B has collected \$20,000 of C's notes, and C has collected \$10,000 of A's notes.<sup>23</sup> If these banks settle their obligations bilaterally, they need to have a total of \$20,000 to \$40,000 of commodity-money reserves on hand among them, depending on the chronological sequence of their exchange. (Even \$40,000 is a savings compared to the \$50,000 of reserves that would be needed if even bilateral offsetting of debts did not take place.) If they settle their balances multilaterally, on the other hand, they only need \$10,000 of reserves among them: A's net balance to B and C combined is +\$10,000; B's net balance to A and C combined is \$0; and C's net balance to A and B combined is \$10,000. Hence all three balances can be settled merely by a transfer of \$10,000 from C to A. In addition to the reduction in reserve needs, multilateral clearing allows savings in time and transportation costs by allowing all debts to be settled in one place and during one meeting rather than in numerous scattered meetings.

The institutional embodiment of multilateral note (and deposit) exchange, the clearinghouse, need not spring into existence full blown because of any sudden and universal perception of its advantages. Instead it may evolve gradually from simpler note-exchange arrangements. For example, the note-exchange agents of banks A and B may accidentally meet each other at the counter of bank C. The greater the number of banks exchanging bilaterally, the less likely it is that such an encounter could be avoided. It would be natural for these two agents to recognize the savings in simple time and shoe-leather costs from settling their own exchange then and there, and from agreeing to do it again next time out, and then regularly. From a set of three pairwise settlements around one table it is not a large step toward the computation and

settlement of clearing balances on a combined net basis. Once the advantages of this become clear to management, particularly the reserve holding economies which may not have concerned the note-porters, the institution will spread. Fourth, fifth, and subsequent banks may be invited individually to join later meetings. Or similar regular few-sided exchanges may be formed among other groups of banks, either independently or at the instigation of one of the first three banks, whose meetings are later combined with the meetings of the original group. Eventually all the banks within an economically integrated region will be connected through one or a small number of clearinghouses.

The histories of the best-known early clearinghouses, in London, Edinburgh, and New York, all conform to this general pattern of development, though geographical, economic, and legislative differences affected the particular shape clearinghouses took in each of these cities. The circumstances leading to the establishment of the New York Clearinghouse in 1853, recounted by Gibbons (1858, p. 292), were typical. The impetus for change from numerous bilateral exchanges to combined multilateral exchange came from note-porters who "crossed and re-crossed each other's footsteps constantly; they often met in companies of five or six at the same counter, and retarded each other, and they were fortunate to reach their respective banks at the end of one or two hours." The porters finally hit upon the idea of meeting at a convenient spot, outside of any bank, to combine and reconcile their claims. Nearly identical events took place in London approximately three-quarters of a century earlier, although here it was mainly checks rather than notes which had to be exchanged because the Bank of England was the only note-issuing bank in the area. Bisschop (1910, p. 160) relates that among the porters who originally travelled among the banks to settle accounts unilaterally or bilaterally, "occasional encounters developed

into daily meetings at a certain fixed place. At length the bankers themselves resolved to organize these meetings on a regular basis in a room specially reserved for this purpose." Thus the London clearinghouse came into being.

The motive force in establishing clearinghouses is the desire for economical exchange and settlement of banks' obligations to one another. Once established, however, a clearinghouse may serve a variety of other purposes. It can become, in the words of James G. Cannon (1908, p.97), "a medium for united action among the banks in ways that did not exist even in the imagination of those who were instrumental in its inception." One of the most common tasks that clearinghouses may take on is to serve as credit information bureaus for their members: by pooling their records, banks can learn whether loan applicants have had bad debts in the past or are overextended to other banks at present, and can then take appropriate precautions.<sup>24</sup> Through the clearinghouse banks can also share information concerning bounced checks, forgeries, and the like.

Another task clearinghouses may take on is policing the soundness of each member bank in order to assure the other member banks that the acceptance of notes and deposits for clearing does not present undue risk. As part of this function, banks may have their books audited by clearinghouse examiners. Cannon (1910, pp. 138-139) reports that the Chicago clearinghouse, beginning in 1906, carried out fairly comprehensive examinations to determine its members' financial condition. Others, such as the Suffolk Bank and the Edinburgh clearinghouse, took their bearings mainly from the trends of members' clearing balances and the traditional canons of sound banking practice. Those two clearinghouses enjoyed such high repute as certifying agencies that to be taken off their lists of members in good standing meant a serious loss in reputation and hence business for the offending bank.<sup>25</sup>

It is possible that a clearinghouse may attempt to organize collusive agreements on interest rates, exchange rates, and fee schedules for its members. However, insofar as the rates set were inconsistent with the results of competition they would tend to break down under unregulated conditions, for the standard reason that secretly underbidding a cartel has concentrated benefits and largely external costs. A clear example of this comes from Scottish experience (Checkland 1975, pp. 391-27). The Edinburgh banks set up a committee in 1828 to collude on borrowing and lending rates. The Glasgow banks joined a new version of the committee in 1836, so at the time it represented the preponderance of Scottish banks in number and in total assets. Though not a clearinghouse association itself, the committee had much the same membership as the Edinburgh clearinghouse. In spite of repeated formal agreements, though, the committee could not hold members to its recommended interest rates. Not until after entry to the industry was closed in 1844 did the agreements become at all effective.

Perhaps the most interesting of all the roles a clearinghouse may perform are its roles in assisting its members in times of crisis.<sup>26</sup> If a bank or group of banks is temporarily unable to pay its clearing balances, or if it experiences a run on its commodity-money reserves, the clearinghouse can serve as a medium through which more liquid banks lend to less liquid ones. It provides the framework for an intermittent, short term credit market similar to the continuous Federal Funds market from which reserve-deficient American banks presently borrow. Another possible emergency function of clearinghouses is that of note-issue. This function is called for when the note issues of member banks are for some reason artificially restricted, so that the banks are not able independently to fulfill all of their depositors' requests for hand-to-hand

means of payment. Currency shortages occurred frequently in the United States during the second half of the 19th century, and clearinghouses helped to fill the void caused by deficient note-issues of the National banks.<sup>27</sup>

### The Mature Free Banking System

We have now reached the stage of mature development of a stylized free banking system, at least in so far as theoretical conjecture and historical evidence allow us to picture it. We may now set forth as a general pattern, lacking in specific detail, the structural and operational characteristics of the system at maturity. The evidence on industry structure from Scotland, Canada, and elsewhere suggests that unregulated development does not produce natural monopoly, but rather an industry consisting of several rivalrously competing banking firms, most having numerous branch facilities, all of which are joined through one or more clearinghouses.<sup>28</sup> The banks issue inside money in the shape of paper notes and checkable deposits<sup>29</sup> that are redeemable upon demand for commodity money. Each bank's notes bear distinct brand-name identification marks and are issued in the denominations the public is most willing to hold. Because of the costs that would be involved, interest is not likely to be paid to holders of bank notes.<sup>30</sup> Demand deposits, however, provide a competitive yield reflecting rates available on interest-earning assets issued outside the banking system.

No significant quantity of commodity money remains in circulation, since all of it has been offered to the banks in exchange for inside money. Much of what was once used for monetary purposes is now devoted to industrial and other uses. Commodity money in bank reserves constitutes only a small fraction of inside-money liabilities. Reserve ratios are determined by banks' precautionary liquidity needs, there being no statutory reserve requirements. Since the

public does not demand full bodied coin for transactions purposes, and assuming that redemption runs due to lost confidence are rare,<sup>31</sup> the banks' liquidity needs depend mainly on the volume and volatility of clearing balances they must settle.<sup>32</sup>

The assets of unregulated banks would presumably include short-term commercial paper, bonds of corporations and government agencies, and loans on various types of collateral. Without particular information on the assets available in the economy, the structure of asset portfolios cannot be characterized in any detail, except to say that the banks presumably strive to maximize their risk-adjusted interest earnings net of operating and liquidity costs. The declining probability of larger liquidity needs, and the trade-off at the margin between liquidity and interest yield, suggest a spectrum of assets ranging from commodity-money reserves, which are perfectly liquid, to interest-earning investments which are highly liquid (these constitute a "secondary reserve"), to higher-earning assets which are less liquid. Thus far, because the focus has been on monetary arrangements, the only bank liabilities discussed have been notes and demand deposits. Of course unregulated banks would almost certainly diversify on the liability side by offering various sorts of time deposits and possibly traveler's checks. It is likely that some banks would also become involved in such related lines of business as the production of full-bodied commodity money and token fractional coins, issue of credit cards, and even management of mutual funds. Such banks would fulfill the contemporary ideal of the "financial supermarket", with the additional feature of issuing bank notes.

This picture of an unregulated banking system differs significantly in its institutional features from the visions presented in some of the recent literature on competitive payments systems. The system described here does

involve assets fitting standard definitions of "money". Banks hold high-powered commodity money (outside money) reserves for settlement of clearing balances, and issue debt liabilities (inside money) with which payments are generally made. These features contrast with the situation envisioned by Black (1970) and Fama (1980), in which "banks" hold no reserve assets and the payments mechanism operates by transferring equities or mutual fund shares.

Bank reserves do not disappear in the evolution of a free banking system, as analyzed here, because the existence of bank liabilities that are promises to pay presupposes some more fundamental means of payment that is the thing promised. Individuals may forego actual redemption of promises, preferring to hold them instead of more cumbersome commodity money, so long as they believe that they will receive high-powered money if they ask for it. Banks, on the other hand, have a competitive incentive to redeem one another's liabilities regularly. So long as net clearing balances have a positive probability of being non-zero, commodity money will continue to be held in bank reserves. The scarcity of reserves moreover serves to limit the quantity of inside money. In a moneyless system it is not clear what forces limit the expansion of payment media.

Our analysis indicates that commodity and commodity-based money would persist in the absence of intervention, for the reason that the supreme saleability of the particular money good is self-reinforcing. This result contradicts the recent views that associate complete deregulation with the replacement of monetary exchange by a sophisticated form of barter.<sup>33</sup> In a commodity-money economy, prices are stated in terms of a unit of the money commodity, so the question of using an abstract unit of account does not arise as it does in a sophisticated barter setting.<sup>34</sup> Our account furthermore



suggests no basis for the spontaneous emergence of a multi-commodity monetary unit or of any pure fiat monetary standard, such as contemplated in works by Hall (1982), Klein (1974), and Hayek (1978). In short the present study views unregulated banking on the whole as much less radically unconventional, and much more akin to existing financial institutions, than does other recent literature on the topic.

One important contemporary financial institution is nonetheless conspicuously missing from our account, namely the central bank. We find no market forces leading to the spontaneous emergence of a central bank, in contrast to Goodhart (1984). Goodhart argues, as other economists have, that the development of a central bank is "natural" because of "natural forces leading to the centralization of reserves" (sec. 1, p. 9). But even on his own account the forces that historically promoted centralization of reserves were not "natural" in any laissez-faire sense. They stemmed crucially from legal favoritism in each country toward a particular state-owned or state-chartered bank, particularly the monopoly of note-issue conferred on that bank, or from legal restrictions on commercial banks that directly or indirectly promote unit banking. Where no legislation inhibits the growth of branched banking firms with direct access to the economy's financial center, it is not at all apparent that profit-seeking compels any more centralization of reserves than is accomplished by the holdings of each bank's head office. Walter Bagehot (1873, pp. 66-68) argues persuasively that "the natural system—that which would have sprung up if Government had let banking alone—is that of many banks of equal or not altogether unequal size" and that in such a system no bank "gets so much before the others that the others voluntarily place their reserves in its keeping." Certainly neither Scotland nor Canada showed any tendency toward complete centralization of reserves or any other tendency toward spontaneous

emergence of a central bank.<sup>35</sup>

## NOTES

1. See for example Black (1970), Klein (1974), Hayek (1978), Fama (1980), Greenfield and Yeager (1983), Wallace (1983), White (1984b), O'Driscoll (1985), and Yeager (1985).
2. On the role of invisible-hand explanations in social science see Ullmann-Margalit (1978). On the prediction of institutional patterns see Hayek (1969). As these articles make obvious, the method of conjectural history is by no means original with the present study. It should also be stated clearly that the method is not employed here in the belief that it is the only valid method for theoretically constructing (or reconstructing) institutional arrangements. The more standard method of building explicit transactions costs or informational imperfections or asymmetries into an optimization model has unquestionably been useful in the task of explaining why banks exist (see Santomero 1984, pp. 577-580, for a survey of this literature). Nor can this paper's conjectures claim to establish the only possible path of monetary evolution; instead they aim to establish the most credible path. Economists who find different institutional outcomes plausible for an unregulated system should similarly try to explain how those would come about.
3. See also Menger (1981 pp. 260-262). The same view appears in Carlisle (1901, p. 5), and Ridgeway (1892, p. 47).
4. See Menger (1981, pp. 263-266); Ridgeway (1892, pp. 6-11); and Burns (1927a, pp. 286-288).
5. On some alleged non-metallic monies of primitive peoples see Quiggin (1963). According to Melitz (1974, p. 95) one must be skeptical concerning attributions of moneyness to primitive media, especially non-metallic "monies" (with the exception of cowries in China), since many of these media do not meet any reasonably strict definition of money. For example, the Yap stones of Melanesia, often cited as a curious form of money, did not in fact serve as general media of exchange.
6. Burns (1927b, p. 78) comments: "There is no evidence that, as a general rule, the officers of the royal or public treasury, by their greater honesty, made the royal or civic seal a mark more reliable than the seals of the [merchants]."
7. Again see Burns (1927a, p. 308).
8. This result may appear to contradict Gresham's Law, at least in its popular version according to which "bad money drives out good." But properly understood Gresham's Law applies only when legal tender or other laws fix an artificial exchange rate between two monies. In that situation payments are naturally tendered in the overvalued ("bad") money, while the undervalued ("good") money is hoarded or exported to a market where it commands full value. See Giffin (1904) or Hayek (1978, pp. 37-39).
9. See Usher (1943), de Roover (1974, chs. 4-5), and Lopez (1979). De Roover

(1974, p. 184) says that "Deposit grew out of [money-changing] activity, because the money-changers developed a system of local payments by book transfer." In our scenario the taking of deposits, at least on a small scale, precedes the development of book-transfer methods of payment.

10. In England, scrivener were the earliest pioneers in the banking trade; in Stuart times they were almost entirely displaced by goldsmith-bankers. The confiscation by Charles II of gold deposited for safekeeping at the royal mint ended that institution's participation in the process of banking development. Nevertheless we conjecture that, if they had been permitted, private mints would have been logical sites for the undertaking of banking activities.
11. The provision of loanable funds was presumably no part of the intention of depositors at early money-changing transfer banks, either, but on the other hand warehousing was not the (principal) service being provided.
12. These conditions may also be met in the warehousing of standard-quality grain, so that fractional-reserve "banking" can likewise develop there. For an interesting historical case see Williams (1984).
13. According to Usher (1943, pp. 7-8) such transactions continued to occur in Europe "long after the bill of exchange was well established" whereas "the use of checks and orders of payment seems to have been incidental and sporadic before 1500".
14. The use of non-negotiable checks in early modern Europe, although it allowed customers to avoid some visits to the bank, did not allow banks to reduce their reserve needs by extending the scope of clearing transactions. "Each check would have to become the basis of specific journal and ledger entries so that to the bookkeeper it made no difference whether the entry originated in an oral order or a written order. [Negotiable checks] made it possible to establish clearing houses in which all transactions among a group of banks were liquidated by drawing up general debit and credit balances between each bank and the clearing house." (Usher 1943 p. 23.) The development of clearinghouses is treated below.
15. "There can be no true note until the law is willing to recognize the right of the bearer to sue in his own name without any supplementary proof of bona fide possession." (Usher 1943, p. 177.)
16. What is said here about notes applies equally to checkable deposits. For expository convenience, and also because unrestricted note issue distinguishes free banking from current arrangements, we will sometimes refer explicitly to bank notes only. See White (1984a, pp. 84-85) for nineteenth-century views on geographic diseconomies in note circulation.
17. To cite just one historical illustration of this, the Bank of Scotland and the Royal Bank of Scotland—the two then existing chartered banks of issue located in Edinburgh—refused to accept the notes of unchartered "provincial" banks of issue for a number of years. See Checkland (1975, p. 126).

18. If other banks are simultaneously replacing Bank A's notes with their own, there may be no absolute increase in A's circulation as compared with the situation in which no bank accepts rival notes. But there will be an increase compared to Bank A not accepting, given whatever policies rivals are following, so that the incentive remains. (We argue below that in fact an indirect consequence of other banks' par acceptance of Bank A notes will be a an absolute increase in A-note\*holding in place of specie\*holding.)
19. This would be a example of the "tit for tat" strategy proving dominant in a repeated\*game setting; on this strategy see Robert Axelrod (1984).
20. This makes a third scenario possible: if two banks both understand these circulation gains, they may explicitly enter a mutual\*par\*acceptance arrangement. Others will emulate them, leading to general par acceptance. This explanation, offered by White (1984, pp. 19-21), assumes slightly more knowledge on the part of banks than the other two. Historical evidence of such explicit arrangements in Scotland is provided by Munn (1975).
21. Small amounts of gold coin were still used in these places at least in part because of restrictions upon the issue of "token" coin and of small denomination notes. In an entirely free system, such restrictions would not exist.
22. See Leslie (1950, pp. 8-9). On note-duelling in the Suffolk system see Magee (1923, pp. 440-444).
23. We use "\$" for expository convenience only. \$1 here is some fixed amount of commodity money.
24. See Cannon (1910, p. 135). This is the most comprehensive work on the clearinghouses.
25. See Trivoli (1979, p. 20); Graham (1911, p. 59).
26. "The clearing house, which was begun simply as a labor\*saving device, has united the banking interests in various communities in closer bonds of sympathy and union and has developed into a marvelous instrumentality for the protection of the community from the evil effects of panics and of bad banking." Cannon (1910, p. 24).
27. See Cannon (1908); Andrew (1908); Timberlake (1984); and Gorton (1985).
28. In Scotland there were nineteen banks of issue in 1844, the final year of free entry. The largest four banks supplied 46.7 percent of the note circulation. In addition to their head offices the banks had 363 branch offices, 43.5 percent of which were owned by the largest (measured again by note issue) four banks. These figures are based on data in White (1984a, p. 37).
29. The development of "paperless" or electronic means of transfer does not imply any change in the nature of checkability.
30. See White (1984a, pp. 8-9; 1985). Wallace (1983) has argued that non-

interest-bearing paper currency is inconsistent with laissez faire in banking.

31. Abstracting from redemption runs is not inconsistent with historical experience under virtually free banking, since general panics were unknown and even runs on individual banks were rare during Scotland's free banking period. Panics and runs were also infrequent in Canadian banking during the era in which Canada had plural note issue and few banking regulations. Runs in the United States and England may be attributable to regulations which weakened banks by restricting their capitalization and branching.
32. On the determination of optimal reserves in the absence of reserve requirements see Baltensperger (1980, pp. 449) and Santomero (1984, pp. 584-586).
33. In addition to Black (1970) and Fama (1980), see Greenfield and Yeager (1983) and Yeager (1985).
34. This point is emphasized by White (1984c). For additional criticism of the Black-Fama-Yeager literature see O'Driscoll (1985), Hoover (1985), and McCallum (1984).
35. On the actual emergence of central banking in Canada see Bordo and Redish (1985). The Scottish system was very gradually subordinated to the English system, beginning with the legislation of 1844 and 1845. On the events and arguments surrounding the appearance of central banks in several nations see Smith (1936).

## REFERENCE

- Andrew, A. Piatt. 1908. "Substitutes for Cash in the Panic of 1907." Quarterly Journal of Economics 22 (August): 497-516.
- Axelrod, Robert. 1984. The Evolution of Cooperation. New York: Basic Books.
- Baltensperger, Ernst. 1980. "Alternative Approaches to the Theory of the Banking Firm." Journal of Monetary Economics 6 (January): 1-37.
- Black, Fischer. 1970. "Banking and Interest Rates in a World Without Money: The Effects of Uncontrolled Banking." Journal of Bank Research 1 (Autumn): 9-20.
- Bisschop, W. R. 1910. The Rise of the London Money Market, 1640-1826. London: P. S. King & Son.
- Bordo, Michael, and Redish, Angela. 1985. "Why Did the Bank of Canada Emerge in 1935?" Unpublished ms.
- Burns, A. R. 1927a. "Early Stages in the Development of Money and Coins." In T. E. Gregory and Hugh Dalton, eds., London Essays in Economics in Honour of Edwin Cannan. London: George Routledge & Sons.
- Burns, A. R. 1927b. Money and Monetary Policy in Early Times. New York: Alfred E. Knopf.
- Cannon, James G. 1908. "Clearing Houses and the Currency." In The Currency Problem and the Present Financial Situation. New York: Columbia University Press.
- Cannon, James Graham. 1910. Clearing Houses. Washington: Government Printing Office.
- Carlisle, William. 1901. The Evolution of Modern Money. London: Macmillan.
- Checkland, S. G. 1975. Scottish Banking: A History, 1695-1973. Glasgow: Collins.
- de Roover, Raymond. 1974. Business, Banking, and Economic Thought in Late Medieval and Early Modern Europe, ed. Julius Kirshner. Chicago: University of Chicago Press.
- Fama, Eugene. 1980. "Banking in the Theory of Finance." Journal of Monetary Economics 6 (January): 39-57.
- Gibbons, J. S. 1858. The Banks of New York, Their Dealers, the Clearing House, and the Panic of 1857. New York: D. Appleton Co.
- Giffen, Robert. 1904. "Note on the Gresham Law." In Economic Inquiries and Studies, vol. 2. London: George Bell and Sons.

- Goodhart, C. A. E. 1984. "The Evolution of Central Banks: A Natural Development?" Unpublished ms., Bank of England.
- Gorton, Gary. 1985. "Clearinghouses and the Origin of Central Banking in the United States". Journal of Economic History 45 (June): 277-283.
- Graham, William. 1911. The One Pound Note in the History of Banking in Great Britain, 2nd ed. Edinburgh: James Thin.
- Greenfield, Robert L., and Yeager, Leland B. 1983. "A Laissez Faire Approach to Monetary Stability." Journal of Money, Credit and Banking 15 (August): 302-315.
- Hall, Robert. 1982. "Explorations in the Gold Standard and Related Policies for Stabilizing the Dollar." In Hall, ed., Inflation Causes and Effects. Chicago: University of Chicago Press for the National Bureau of Economic Research.
- Hayek, F. A. 1969. "The Theory of Complex Phenomena." In Studies in Philosophy, Politics, and Economics. New York: Simon and Schuster.
- Hoover, Kevin D. 1985. "Causality and Invariance in the Money Supply Process." Unpublished doctoral dissertation, Oxford University.
- Klein, Benjamin. 1974. "The Competitive Supply of Money". Journal of Money, Credit, and Banking 6 (November): 423-453.
- Lake, Wilfred S. 1947. "The End of the Suffolk System." Journal of Economic History 7 (November): 183-207.
- Leslie, J. O. 1950. The Note Exchange and Clearing House Systems. Edinburgh: William Blackwood.
- Lopez, Robert S. 1979. "The Dawn of Medieval Banking." In The Dawn of Modern Banking. New Haven: Yale University Press.
- Magee, James D. 1923. "Historical Analogy to the Fight Against Par Collection." Journal of Political Economy 31 (June): 433-445.
- McCallum, Bennett T. 1984. "Bank Deregulation, Accounting Systems of Exchange, and the Unit of Account: A Critical Review." Unpublished ms., Carnegie Mellon University.
- Melitz, Jacques. 1974. Primitive and Modern Money. Reading, MA: Addison-Wesley.
- Menger, Carl. 1892. "The Origin of Money". Economic Journal 2 (June): 239-255.
- Menger, Carl. 1981. Principles of Economics [1871]. New York: New York University Press.



- Munn, Charles W. 1975. "The Origins of the Scottish Note Exchange." Three Banks Review 107: 45-60.
- Munn, Charles W. 1981. The Scottish Provincial Banking Companies, 1747-1864. Edinburgh: John Donald.
- O'Driscoll, Gerald P., Jr. 1985. "Money in a Deregulated Financial System." Economic Review, Federal Reserve Bank of Dallas (May): 1-12.
- Powell, Ellis T. 1966. The Evolution of the Money Market, 1385-1915. New York: Augustus M. Kelley.
- Quiggin, A. Hingston. 1963. A Survey of Primitive Money: The Beginning of Currency. London: Methuen.
- Richards, R. D. 1965. The Early History of Banking in England. New York: Augustus M. Kelley.
- Ridgeway, William. 1892. The Origin of Metallic Currency and Weight Standards. Cambridge: Cambridge University Press.
- Santomero, Anthony M. 1984. "Modeling the Banking Firm: A Survey." Journal of Money, Credit, and Banking. (November): 576-602.
- Smith, Vera C. 1936. The Rationale of Central Banking. London: P. S. King.
- Timberlake, R. H. 1984. "The Central Banking Role of Clearinghouse Associations". Journal of Money, Credit, and Banking 16 (February): 1-15.
- Trivoli, George. 1979. The Suffolk Bank: A Study of a Free-Enterprise Clearing System. London: Adam Smith Institute.
- Ullman-Margalit, Edna. 1978. "Invisible-Hand Explanations." Synthese 39 (October): 263-291.
- Usher, Abbott Payson. 1943. The Early History of Deposit Banking in Mediterranean Europe. Cambridge: Harvard University Press.
- Wallace, Neil. 1983. "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy." Federal Reserve Bank of Minneapolis Quarterly Review (Winter): 1-7.
- White, Lawrence H. 1984a. Free Banking in Britain: Theory, Experience, and Debate, 1800-1845. Cambridge: Cambridge University Press.
- White, Lawrence H. 1984b. "Free Banking as an Alternative Monetary System." In Barry N. Siegel, ed., Money in Crisis. San Francisco: Pacific Institute.
- White, Lawrence H. 1984c. "Competitive Payments Systems and the Unit of Account." American Economic Review 74 (September): 699-712.