Transaction Costs, Money and Units of Account

Bengtsson, Ingemar

2005

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Transaction Costs, Money and Units of Account

by

Ingemar Bengtsson

Department of Economics,

Lund University

Keywords:
Medium of Exchange; Money; Payment Techniques; Quantity Theory; Transaction Costs; Unit of Account

JEL Classification:
B52;D02;D23;E31;E41E42;E51

¹ Ph.D. Ingemar Bengtsson, Department of Economics, Lund University, P.O. Box 7082, SE-220 07 LUND, Sweden, email: ingemar.bengtsson@nek.lu.se
Transaction Costs, Money and Units of Account

Abstract

In the paper, an analogy with length measurement is applied in order to explore the nature of the unit for value measurement, i.e. the unit of account. As the meter is defined as the length traveled by light in vacuum during 1/299 792 458 of a second, the unit of account krona is defined as the purchasing power of the medium of exchange krona. However, one should be cautious when drawing conclusions from this analogy. Our unit of account is defined in our medium of exchange, but it is meaningful only because we can observe prices on real goods expressed in it. As it would be pointless to define the meter as the length traveled by light in vacuum during 1/299 792 458 of a second if we could not compare this length with anything else, it would be pointless to define our unit of account in something that is not priced.

In the paper it is explained how different payment techniques help to overcome transaction costs in the market. In particular, following Alchian (1977), it is argued that to reap the full benefit from the use of payment techniques, it has to be combined with the use of both a unit of account and specialist middlemen. The use of payment techniques helps to reduce costs due to sequential payment, but to reduce costs due to sequential quality evaluation, you need unit of account as well as reputable middlemen.

* This paper is prepared from my Ph.D. thesis Central bank power: a matter of coordination rather than money supply. Some of the feedback that I have received on various parts of the thesis concerns the material in this paper, and I am thus grateful for remarks and suggestions from Tyler Cowen, Kevin Dowd, Benjamin Friedman, Charles Goodhart and Michael Woodford. Naturally, I am solely responsible for all remaining errors and obscurities. I am also grateful for the financial support from Torsten och Ragnar Söderbergs stiftelser, which enabled me finishing the thesis.
1 Outline

In this paper, we will analyze the use of payment techniques and nominal prices/units of account as techniques to reduce the need to engage in sequential transactions during attempts to exploit the advantages of a division of labor. We will look at different kinds of payment techniques that have been used throughout history and pay special attention to the interdependence between the use of payment techniques and units of account. This will help us understand how these entities are related and perhaps more importantly how they are not.

2 Money in economic theory

From Coase (1937), we know that the neo-classical general equilibrium model in the tradition of Walras-Arrow-Debreu (WAD) involves no firms. In this paper, we will follow Coase and analyze another absent friend, namely money. There is nothing surprising about the non-existence of money in the WAD model; it is a description of a static world, where all exchanges take place once-for-all. Money has no function in such a world, since the primary function of money is to solve problems of sequential transactions. The use of money is a way to solve the problem of making credible payment commitments in transactions where simultaneous exchange is impossible.

The inconsistency between mainstream microeconomics and the actual use of money in the observable world has forced economists to make some peculiar conjectures about the nature of money. Some have built models in which agents enjoy happiness from the sound of rustling notes and jingling coins.1 Others hold that there are certain goods or services that can only be paid for in cash, perhaps inspired by the market for illegal drugs.2 The peculiar conjectures underlying the money-in-the-utility-function model and the cash-in-advance model illustrate how the inappropriateness of applying the WAD framework to an analysis of money. Within the WAD framework, one can adequately analyze a world with a smoothly working price mechanism, or in other words, a market economy with zero transaction costs. The problem is that with zero transaction costs, money would have no purpose to fill.3 In the real world, on the other hand, the market economy consists of a wide range of institutions, which have as their sole purpose to bring down transaction costs. Hence, it should be clear that the WAD framework is not suitable for an analysis of market institutions such as e.g. money. Steven G. Medema and Richard O. Zerbe (1998:217) put it nicely: “A blackboard theory that assumes away

---

1 Cf. e.g. Sidrauski (1967).
2 Cf. e.g. Clower (1967).
transaction costs will have predictive value, but only in those instances in which
transaction costs are not determinative.” That is, while there are issues that can be
adequately addressed within the WAD model, issues about money are certainly not
among them.

Here, we will analyze money in a framework where the driving force behind economic
evolution is the division of labor facilitated by a system of voluntary exchanges, which
are associated with certain transaction costs. Although the analysis is not intended to
describe the actual history of the evolution of money, it relies on particular facts from
history during the discussion of the characteristics of different payment techniques.
Thus, by acknowledging actual circumstances, we attempt to avoid the risk of analyzing
only an approximation of a society. Geoff M. Hodgson (1998:33) refers to this in the
following passage:

By confining itself to allegedly universal and ahistorical concepts, mainstream economics fails to
become rooted in any specific socio-economic system. Its very generality provides limited means for
an understanding of capitalism or other specific systems. Instead of attempting to confront a
particular economy, or real object, it becomes confined to a remotely abstract and artificial idea of an
economy, the economy in general.

The risk of ending up with a theory that can not be applied to reality when an
ahistorical approach is used, is exemplified in some of the more recent attempts to
incorporate money into the WAD framework. We will discuss this briefly below. Many
traditional discussions on money have emphasized its function as a medium of exchange
and especially its role in overcoming the double coincidence of wants problem associated
with pure barter exchanges (William S. Jevons, 1875). The focus is often on the intrinsic
properties of objects that make them more or less a natural medium of exchange,
including properties such as relatively low storage or exchange costs (Carl Menger,
(1892)). In recent years, steps have been taken to incorporate money into mainstream
microeconomics along this line, and some of these intuitively appealing ideas have been
formalized by the use of search-theoretic equilibrium models of the exchange process.
Nobuhiro Kiyotaki and Randall Wright (1989) show how an indirect exchange with a few
commodities used as money may evolve because of the usual transaction costs
associated with pure barter.4 In later papers (1991,1993), Kiyotaki and Wright attempt to
explain also the holding of fiat money. In these models, fiat money arises endogenously
as a medium of exchange, leading to reductions in the search and transaction costs
associated with pure barter. However, since neither the possibility of money with
intrinsic value nor convertible paper debt are considered, these models does not explain

4 Cf. Ostroy and Starr, 1990, for a survey of earlier work in the tradition of incorporating monetary theory into the general
equilibrium theory of value.
why fiat money is accepted in the actual world. Besides, a common problem with this class of models is the interpretation of fiat money. In Kiyotaki and Wright (1993:64), for example, fiat money is described as “a collection of pieces of paper or certain types of seashells, for example, with no intrinsic value.” This is a troublesome interpretation of fiat money, since this kind of money has never been used in the kind of transactions the authors are studying, i.e. transactions between strangers. In cases when seashells have been used as a medium of exchange between strangers, it has been in cultures where seashells had a consumption value and they are therefore principally not different from gold coins. To my knowledge, pure token money has only been used within non-stranger environments as a simple bookkeeping device. Fiat paper money, in practice, has never been entirely inconvertible; as a last resort, one can always pay taxes with them (as we will argue later). Consequently, since the authors define fiat money as something different to the objects in reality that we normally call fiat money, the predictions of these models have uncertain value outside the rather special economy they describe.

The absence of a double coincidence can also be interpreted as a problem of asymmetric information about trading histories. Robert M. Townsend (1989) describes a model with private information, spatial separation and limited communication, where a currency-like object – a token – and other forms of credits can be distinguished. Credits can be used among agents in a persisting relationship, i.e. among agents with known trading histories, whereas tokens are needed among relative strangers. Tokens play the role of a bookkeeping device among strangers. Townsend shows that under certain conditions, tokens will exist alongside normal credits. In his analysis, however, it is unclear why indirect exchange would not be used. Put in a historical context, it appears that his paper compares a situation of autarchy with a situation of a modern market economy with fiat money. Historically, there are in fact thousands of years between the two, and indirect exchange as well as intrinsically valuable money and convertible money have been commonly used as money in the meantime. Thus, it is not at all clear from the model if the fiat money would be held in equilibrium were the possibility of other payment techniques considered. Hence, although the model makes valuable contributions to our understanding of the record-keeping function of money (in an abstract sense), it does not explain the use of token money. Again, since the situation depicted in the model has no actual counterpart, we can not say for certain what it teaches us about reality. Richard N. Langlois (1984:34) has claimed, in a different context, that the neoclassical logic of explanation is generally inapplicable to issues about market institutions:
Admittedly, this is an odd sort of explanatory mechanism: rather than literally proposing a process by which the efficient result is achieved, it relies simply on showing that the efficient result is logically possible given the assumptions.

These examples of attempts to incorporate money into the WAD framework illustrate what Hodgson and Langlois warn against: ahistorical discussions about allegedly universal concepts without reference to the actual institutional settings are prone to lead our thoughts in the wrong direction.

3 Money - a payment technique

Money is used to settle debts incurred by e.g. purchases or a liability to pay damages; i.e. to use money is to apply a kind of payment technique. I prefer to talk about the more general term, *payment technique*, rather than money. The reason is, of course, that the term money is far from clear-cut; The nature of money has been subject to many different interpretations: *numeraire, medium of exchange, store of value, medium of account, unit of account, standard of deferred payment, standard of value* and so on. However, to avoid confusion, one should be more precise about what exactly is referred to with the term money. Although I would prefer to dispose with the messy concept of money altogether, this is hardly possible and I will instead try to clarify what money will be supposed to refer to, before I start my analysis.

1.1 The different roles of money?

In this section, we will consider in some depth the concept *unit of account* and its relation to the concept of a *medium of exchange*. Money is often supposed to be not only the medium of exchange but also the unit of account. However, is that actually true? The fact that it has been difficult to define money unambiguously is explained in the following way by Robert Clower (1995:525):

> Few writers seem able to avoid references to “money” that are metaphorical: comments that seem on the surface to refer to money “objects” but refer in truth to an unspecified complex of institutions associated with monetary economies.

In his account of what money has been thought to be, Clower (1995:526) states that:

> “from Aristotle in the 4th century B.C., to John Hicks in 1967, no evident progress was made towards rational understanding of the nature of the ‘thing’ called ‘money’.” Clower then quotes Hicks5, who sums up the conventional view of money as being defined by its functions. It is a functional definition, since it defines money from its perceived functions, which Hicks refers to as threefold: “to act as a unit of account, as a medium of exchange, and as a store of value.” One way or another, most definitions resemble the idea that money does perform these functions, most importantly the unit of account and
the medium of exchange functions. It is obvious that if money in fact does not perform both these functions, then any attempt to define money under the presumption that it does fulfil both functions would be purposeless. I think this lies at the heart of the question why definitions of money have been widely unsuccessful; economists have been unwilling to reconsider in an unprejudiced fashion the idea of money as a unit of account. As the analysis proceeds, we will see that the unit of account is quite a different institution than what can reasonably be called money.

First, we need to decide what the unit of account is. Cash, for example, is undoubtedly money, but it is the unit of account no more than a measuring stick is the standard of measurement of length. I think that an analogy with physics is helpful in this case. The unit of account, e.g. the Swedish krona, performs a function similar to that of, for example, the meter, i.e. it is a standard of measurement of value, in the same sense as the meter is a standard of measurement of length. That is, we choose to define something as our unit of measurement as we defined the standard meter bar, and since 1983 the length traveled by light in vacuum during 1/299 792 458 of a second, as our unit of measurement of length. The purpose of such a standard of measurement is that it simplifies comparisons of the length, or value, of different objects. In neither case is it possible to measure in an absolute way. To say that something is x meters long is just a statement about its length in relation to other objects. Royall Brandis (1966:120) has explored the analogy in some detail:

We could do without a standard of length measurement although it would be very inconvenient to do so for the length of any particular distance would then have to be expressed as ratios of the lengths of all the other distances in the physical universe. This is analogous to the measure of relative values in a pure barter economy without a numeraire which is an equally inconvenient arrangement for the same reason. Thus we establish a standard of length measurement which serves the same function as a numeraire in an economic system. Our length standard does not measure absolute length but only relative length and its own length is unmeasurable. The question - how long is the standard meter bar? - is a meaningless question. Our monetary unit does not measure absolute value but only relative value and its own value is similarly unmeasurable.

That is, it is meaningless to ask what the krona’s value is, since krona is the unit we have created in order to be able to measure the value of other things. Nevertheless, I would like to add that although we can not measure the standard of measurement with the things that it is intended to measure, we understand it in some way like that. Most people will never be able to observe the length traveled by light in vacuum during 1/299 792 458 of a second, so how do they know the (relative) length of the meter? Obviously, if you have a measuring stick that is one meter, it helps you understand how long a meter is. Otherwise, if you know that you are 1,70 meters tall, that would also be helpful to understand the length of a meter. Therefore, we should not be surprised to find that

---

5 Hicks, J, (1967) "Critical Essays".
early standards of measurement were connected to objects everybody was reasonably familiar with, such as an inch or foot. The definition in itself, no matter how impeccable, is not enough to make the concept useful; we must also be able to relate it to the reality we know.

Although the analogy to measurements in physics is useful, it is not perfect. The unit of account krona is not explicitly defined in the same sense as e.g. the meter, and thus it is not quite obvious how it is defined. We do not have a formal definition of a standard of measurement for value, but that does not prevent us from finding a functional definition. We should look for our definition among all those relations that couple the unit of account with values. Such coupling is present in every contract stating a price on an item, and thus, we should expect to find our definition among them. In analogy with the case of length measurement, it would be logical to focus on the one price that does not change, i.e. the price of the medium of exchange. That is, if a tree that used to be one meter now has grown to 1,10 meters, we would say that it is now 1,10 meters, not that the meter is now longer. If, on the other hand, the length traveled by light in vacuum during $\frac{1}{299\,792\,458}$ of a second has become shorter, we would, perhaps, say that the meter now represents a shorter distance than before and that this distance still is one meter. Similarly, if the value of the medium of exchange decreases in relation to the value of all other things, the price of the medium of exchange would still be one. Thus, we could state that the unit of account krona is implicitly defined as having the same value as the medium of exchange krona. As for the meter, the krona becomes meaningful first when we have measured some familiar phenomena with it. It makes sense to treat our unit of account krona as defined by its relation to our medium of exchange krona, because a contractual obligation to make a certain payment that is specified in the unit of account could always be fulfilled by paying with the medium of exchange. Although not perfect, the essence of the analogy still holds; we choose the value (length) of some phenomenon in reality to be our standard of measurement of value (length). Every distance that is just as long as the length traveled by light in vacuum during $\frac{1}{299\,792\,458}$ of a second is said to be one meter long and everything that is possible to buy with a one-krona coin is said to have the value one krona.

To explore the full significance of our analogy, we will consider its implications under the different regimes of gold-convertible money and fiat money respectively. From our definition of the unit of account, it is trivially true that a one-krona coin is worth one krona, in the same sense as it is trivially true that the length traveled by light in vacuum during $\frac{1}{299\,792\,458}$ of a second is one meter. Furthermore, if it is also stated that a one-krona coin always will be worth x gram of gold of a certain grade, it would also be
true in practice that the unit of account krona is equal to the value of that quantity of gold. Although not formally correct, we could for all practical matters say that the standard of measurement of value, the krona, is x gram of gold. This implies that if the value of gold decreases in relation to a basket of goods, the measured value of this basket becomes higher. We can now notice a difference of major practical importance between measurements in physics and economics: while we are quite accustomed to a standard of value that changes in relation to everything else, we do not expect our standard of length to change from one day to the next. While the law of physics changes very slowly, our appreciation of goods changes more or less continually. This difference is also the main reason why we no longer have a standard of value defined in the same way as our standard of length. If the length of particular distances fluctuated as much as the value of particular items, we would perhaps define our standard of measurement of length differently too.

Let us now consider how the definition of the standard of value works under a fiat money regime. In this case, we do not have a permanent link between the medium of exchange and a certain commodity corresponding to the link between the medium of exchange and gold under a convertible money regime. As long as we have not measured any goods with our standard of value, the definition of the standard is meaningless. This is an important difference to the case of length measurement: While it may be difficult to observe the length traveled by light in vacuum during 1/299 792 458 of a second, it is still just a technical problem. In the case of a fiat standard of the measurement of value, the definition is entirely empty until we have measured some values with it. That is, if no prices were quoted in our unit of account, we could not measure any value with it. The definition alone is not sufficient for the fiat unit of account to be a usable standard of measurement, we need to measure at least one value first. It seems that we are stuck in a circle: How can we possibly make our necessary first measurement? The point is that we can not, and indeed, have not. How, then, is it possible that the fiat type of the unit of account is completely dominant today? The answer is that these units have all inherited their meaning from previous, already established, standards of measurement. In practice it means that all items are immediately given an initial value measured in the new unit of account, proportional to their value measured in the old unit of account. The introduction of the euro is a good example. Thus, the problem of the first measurement is overcome. The primary benefit with the fiat unit of account is that it does not fluctuate with the relative value of a single good.

In accordance with what has been said, we will use the term “money” to refer to a subset of all media of exchange, namely those that are deliberately designed to be a
payment technique, i.e. paper notes and gold coins are money, but not gold as such. The broader class of media of exchange, in turn, is a subset of all payment techniques, namely those that have a physical representation. The last category of payment techniques includes those that can be characterized as services and they will be treated under the heading of middlemen. While other classifications are possible, I think this provides a suitable basis for a sound understanding of the role of different payment techniques in different societies and subsets of societies.

4 Payment techniques as a substitute for trust

The division of labor, understood as the specialization in production associated with extensive trade, is one of the most fundamental factors behind economic progress, and it is a decisive factor behind the rise of the market economy. To harvest the advantages of specialized production, it is necessary to trade different goods and services for each other. To carry out trade one must, among a host of other things, be able to make reliable payments. In the rare situations where an immediate exchange of goods with immediately recognizable quality is possible, payment is no issue. However, these situations are very rare indeed, and thus, reliable payment is a crucial issue to the growth of a market economy. Man will try to organize a division of labor in order to obtain potential benefits from it. The complex of institutions - e.g. bookkeeping techniques, payment techniques, measuring techniques and firms - that we refer to as the market economy has emerged as one way to organize the division of labor.

Let us return to the two kinds of relationships that were mentioned before, strangers and non-strangers. In non-stranger relationships, there is no scope for opportunism, either because fraud is observable and would be punished, or because the agents belong to the same social unit as e.g. a family or a tribe and therefore are quite willing to exchange beneficial acts. A functional definition of a transaction follows naturally from this division of relationships: a transaction is an exchange between strangers. Exchanges between strangers are associated with transaction costs because of potentially opportunistic behavior. Neither self-interest nor reciprocity considerations are necessarily sufficient to ensure that the agents will comply with the terms of the contract. Transactions thus involves transaction costs, for example the cost of drawing up enforceable contracts, and if required, maintain one’s rights with the help of the power. The costs come partly from efforts spent on activities to secure a payment and partly from efforts spent on an evaluation of the value of traded goods. Transaction costs,

---

6 Under such circumstances, trading life would be one long repeated game with an indefinite end. In such games, it is reasonable to believe that a co-operative strategy will emerge as norm.
thus comprehended, come with the division of labor, since they are costs that exist in a
catatallactic economy but not in an autarchy. Now we can interpret the market as the
complex of institutions that has emerged as a way to economize on transaction costs. All
transaction costs are not directly related to the problem of trusting a stranger. There are
also considerable transaction costs of a purely practical nature. With the specialization of
production comes an increased need for transportation of both information about goods
and the goods themselves.

Rules are not, however, the only way to deal with the problem of transaction costs. As
discussed in Bengtsson (2003: 50-53), another way is to neutralize the problem by
technical means, e.g. by transforming sequential transactions to simultaneous
transactions. We will now focus on one such technique, namely the method of settling
debts with some kind of payment technique. The application of specialized payment
techniques comes so natural to us today that it is hard to see the fundamental problem
that it solves. Therefore, I think it is appropriate with a moment of contemplation on this
issue. One has to imagine what trade would have looked like if no method of payment
were available, not even indirect exchange. In that case, one would either have to
exchange goods directly in a pure barter deal or engage in a sequential transaction,
possibly involving several persons.

Different payment techniques will be discussed under three headings: pre-monetary
exchange including barter and indirect exchange, money including gold coins, convertible
and inconvertible debt notes and cash cards, and lastly middlemen, including different
kinds of bookkeeping records. The purpose of using the term middlemen is to emphasize
its character of service rather than object. The state of pure barter will function as a
frame of reference and the different payment techniques will be discussed in relation to
it. Each payment technique will be analyzed according to three main issues: its relation
to the payment issue, its relation to the quality evaluation issue and its relation to the
unit of account. The interactive nature of the evolution of payment techniques on the one
hand, and the unit of account on the other, will be highlighted. Finally, acknowledging
that the evolution of society is history-dependent, the institutional prerequisites for each
payment technique will also be discussed.

5 Pre-monetary exchange

There is a fundamental difference between societies using a deliberately designed
medium of exchange and societies, which do not. As we will see, the use of a deliberately
designed medium of exchange requires a higher level of trust in society. This may seem
counterintuitive but is based on the fact that media of exchange not deliberately
designed as such always trade at their consumption value, i.e. their value in their
second-best use is almost the same as their exchange value. This is not necessarily true for deliberately designed media of exchange, which trade for more than their value in their second-best use. Even gold coins often varied somewhat in gold content, and so an element of trust in the issuer is present, something which of course is ever more relevant when we consider paper money. Under this heading, we will treat payment techniques that do not involve deliberately designed media of exchange.

1.2 Barter

The basic form of exchange is pure barter, basic because it only involves goods, the features of which form the basis for the mutually beneficial exchange. Therefore, barter is potentially more utility-improving than any exchange involving an intermediary payment technique, since the use of a payment technique will always use up some of the advantages from the exchange. From this we can conclude that there must be some rather substantial costs involved in barter, since almost all exchanges are in fact conducted through the use of an intermediary payment technique. Some of these problems are often discussed under the label of ‘a double coincidence of wants’ problem, roughly meaning that through barter, goods can not be allocated as efficiently as under a Walrasian auctioneer. There are several reasons why this is the case, one being that goods can not be efficiently allocated since bilateral pure barter can not achieve all possible allocations. There are many other problems as well: how and where to find a potential trading partner is a substantial problem; limited divisibility is another; still another is the fundamental impossibility of immediate exchange of some services – a barber can not cut the dentist’s hair while simultaneously receiving dental care from him.

These are all payment problems following the division of labor. There is one more type of problems following the division of labor, namely problems concerning quality evaluation. During autarchy, each agent has a good idea of the quality of the goods he consumes, simply because he has produced them himself. When production gets specialized, this familiarity with the goods decreases, partly because new goods become available, but also because the familiarity with the ‘old’ goods gradually disappears. We see now that pure barter not only requires a double coincidence of wants in its broadest sense, but also a double coincidence of familiarity with the goods.

---

7 I say ‘roughly’, since the notion of the double coincidence of wants is much older than the concept of a Walrasian market.
1.3 Indirect exchange

The custom of indirect exchange mitigates two of the problems associated with pure barter: Firstly, it increases trading opportunities, since the double coincidence of wants problem is reduced to a single coincidence of wants ditto. A single coincidence is to find someone who has the goods I am looking for; a double coincidence is to find someone who has the goods I am looking for and who desires the goods I have to offer. Secondly, in many cases, it also resolves the problem of securing a payment, since it enables simultaneous transactions instead of sequential transactions. Essentially, these are two sides of the same problem. The double coincidence of wants problem would be much less problematic if sequential transactions were not problematic. If sequential transactions were not costly to handle, you would be able to make a purchase only by finding someone who supplies the goods you desire, i.e. what I have called the single coincidence problem. You and your trading partner could set up a contract that says that you will pay him when you have sold your own production. However, in the real world, sequential transactions are costly to handle; you either have to trust your counterpart and thus face the risk of being cheated, or you have to spend resources on drawing up and enforcing a contract.

In more valuable transactions, a richer choice of institutional devices are available and may be used to solve the problem, such as writing detailed contracts or using the legal system to monitor the parties.\(^8\) One party could also offer some collateral as hostage.\(^9\) In many transactions, however, the cost of enforcing them would outweigh the gains from the exchange. In the case of pure barter, the mutual and simultaneous deliverance of goods solves the problem of securing a payment. However, it requires not only a double coincidence of wants but also a double coincidence of exchange, i.e. that the exchange in its entirety can be performed instantaneously. That is, most services can not be exchanged in a pure barter fashion, since they often take some time to fulfil. The impossibility of instantaneous exchange is most obvious in a transaction where two producers of different services are to exchange services. Recall our earlier discussion of a barber and a dentist, for them it is physically impossible to exchange services without creating a debt/debtor relation; one of them has to perform his side of the transaction first and then hope that the other will fulfil his part.

All kinds of payment technique have this one thing in common: they transform sequential transactions into simultaneous transactions. They make it possible for both sides of a transaction to perform simultaneously, and thus help reduce the number of

---

transactions that give rise to debt/debtor relations. The practice of indirect exchange is one such payment technique. It means that the buying side of a transaction uses some intrinsically valuable and tolerably durable, divisible and portable good as payment. The selling side accepts the payment although he does not want to consume it at the moment. However, he decides that he will either consume it later or be able to use it as payment in another transaction. Hence, by transforming the sequential transaction to a simultaneous transaction where trust is not required, the agent solves the original problem of making credible commitment to comply with the, implicit or explicit, contract. The technique of indirect exchange can be seen as the simplest form of payment technique. Unlike all other payment techniques, it does not involve any, for payment purposes, deliberately designed goods or services.

Initially, the medium of exchange would most likely be goods that the seller already has a stock of and that the buyer accepts without having decided yet if he will use it as medium of exchange, or perhaps consume it himself. A good example is the Aztec’ use of cacao beans as a medium of exchange. This particular example also illustrates a fact that seems partly forgotten today: What we call money with intrinsic value is only intrinsically valuable within a specific cultural context. In the case of cacao beans, it is illustrated by the alleged reaction of the first European pirates who captured a ship carrying cacao beans: they thought the cargo was rabbit droppings and threw it overboard. This should be kept in mind since it reminds us that the line of demarcation between intrinsically valuable money and intrinsically worthless money is less clear-cut than it may seem at first sight.

Over time, a few goods will be discerned as the most salable, as described by Menger (1892: 250-252), in a self-reinforcing process. A salable good should not only be appreciated as valuable in a society, but also divisible, durable and portable. Durable and salable in combination means that it is a suitable store of value. In addition, it should present a modest ‘lemons’ problem; i.e. its quality should be relatively easy to evaluate.

The emergence of a unit of account can be told as a corollary to the story of how different payment techniques evolved. Before indirect exchange, in the pure barter state, no explicit unit of account is employed, relative prices are agreed upon in every transaction. When indirect exchange has become customary, a vast majority of

---

10 Cf. Weatherford (1997) for a description of cacao beans as a medium of exchange in the Aztec culture and for further references on the matter.
transactions involve a medium of exchange, such as gold or silver. This implies that all other relative-price relations gradually will disappear from people’s consciousness. Hence, relative prices become prices expressed in goods accepted as media of exchange, out of which the same number of a unit of account may be distinguished. The emergence of a unit of account stimulated by the emergence of indirect exchange has in turn repercussions on the medium of exchange. The habit of expressing prices in the medium of exchange will provide incentives to further decrease the number of commonly used media of exchange, since traders would then need to know fewer prices. Thus, to reduce the problem of securing payment in an ideal way, there would be very few, maybe only one, medium of exchange, and the unit of account would be a specified amount of the medium of exchange, which hence would serve as the medium of account (MOA). We can see the final stage as a state where pieces of gold, silver and copper are employed as media of exchange and where there are units of account specified as a certain weight of each of these metals. However, even in this ideal state of indirect exchange, there would still be a considerable problem of evaluating the quality of the traded goods. It fact, it would be greater than in pure barter, as demonstrated by Alchian (1977). As long as one has to pay in order to evaluate the quality of the medium of exchange, the use of it would add to the total cost of evaluation. More on this will be said in the next section.

6 Money

The problem concerning the evaluation of the overall quality of the offered good or service is a problem of asymmetric information. The asymmetry in the distribution of information between the agents arises because of the fact that the seller is an expert at evaluating the value of his own goods, whereas the buyer is not. This asymmetry is an unavoidable consequence of the division of labor. Thus, there may be an incentive to sellers to produce low-quality goods and attempt to cheat uninformed buyers, assuming that low-quality goods are cheaper to produce. Therefore, the buyer has to take on value-consuming examinations of the seller’s goods. This applies both to the primary goods and to the medium of exchange. The examination of the quality of a gold nugget, for instance, was associated with great effort and used up a great part of the profit from the trade.

The common feature of different kinds of money, i.e. deliberately designed media of exchange, is that they are denominated in integer numbers of the already established unit of account and that the problem of evaluating the value of the medium of exchange itself is thereby reduced. A gold coin e.g. is struck with a number or symbol intended to indicate its gold content, with gold serving as MOA and a specific weight of gold serving as unit of account. In relation to the custom of indirect exchange, money brings no
further benefits to the double coincidence of wants problem, i.e. what I have called the single coincidence of wants problem remains. Neither the problem of finding a prospective trading partner nor the problem of finding out the quality of the primary goods is resolved by the use of money by itself.

1.4 A standardized medium of exchange with intrinsic value, and convertible debt notes

Gold, silver and copper coins are examples of what we call standardized media of exchange with intrinsic value. The decisive characteristics of them are (a) that they contain a valuable metal and (b) that the content of that metal, regarding weight and pureness, is guaranteed through a stamp on them. Hence, it is clear that their purpose is to overcome the problem of evaluating the quality (and quantity) of the metal itself. In order for the trader to make full use if its benefits, however, the coin must be designed in such a way that it is hard to tamper with and its issuer must be trusted. To refer to our discussion above, the coin is intrinsically valuable at face value only within a certain social context. That is, only those who trust the issuer (or someone else who guarantees the coin’s value) will accept it at face value.

Similar to the payment technique of indirect exchange, the use of intrinsically valuable money is associated with an opportunity cost because the metal has a consumption value. Payments with this technique are consummated when the buyer of the goods or services hands over coins to the seller. It is a very straightforward and simple payment system, but as trade grows, it becomes increasingly expensive. This is because gold coins are costly to store and handle since they are heavy to carry around and exposed to theft.12

Convertible debt notes, or paper money, are similar to gold coins in many ways: the unit of account is the same, they solve the double coincidence of wants problem to the same degree and they do not, by themselves, settle the problem of finding out the quality of the traded goods. The difference between them lies on another level, i.e. in the way they obtain their value. Gold coins obtain their value from a trust in the issuer regarding the gold content, combined with the fact that gold itself is desirable. Convertible debt notes obtain their value from a trust in the issuer regarding the possibility of redeeming the notes in gold. This difference has to do with the institutional settings for each payment technique. Gold coins and other coins were minted by the State. The ancestors of King Croesus of Lydia are believed to have produced the first coins around 640–630 B.C.13 According to Robert M. Cook (1958), these coins were introduced to pay

---

12 See Dowd (1996:10).
mercenaries. Colin M. Kraay (1964) and Hicks (1969), propose that governments minted coins to pay mercenaries only in order to create a medium of exchange for taxes.\textsuperscript{14} Without a convenient medium of exchange, a wide range of production would be difficult to tax. To accept tax payments in kind necessarily results in taxes floating in at irregular intervals – and some of them will be perishable goods.

I think these assessments come close to the core of the issue of coinage and we will find the answer by asking ourselves what is so special about States. The special feature we are looking for here is that all members of a society are actually or potentially indebted to the State, because of the State’s possibility to levy taxes. Presumably, you are more likely to accept, at face value, coins minted by someone you owe money to, than coins minted by someone else. You do not have to worry about whether anyone else would accept them, because you presume that the issuer of the coins will accept them to settle your debt. This idea is reinforced by the fact that for many centuries, produced coins varied between five and ten per cent in weight, and nonetheless, they were accepted as of equivalent value.\textsuperscript{15} This emphasizes that the stamp on the coin did not guarantee the actual content of gold but rather at which price the issuer was willing to accept it as a (tax) payment. Again, we can see that the line of demarcation between media of exchange with intrinsic value and those without is not that sharp.

While coinage was a governmental activity, the introduction of debt notes was a private sector enterprise. As trade grew, so did the possibility of basing commercial relationships on trust due to the hostage effect. With money came an early form of marketplaces where merchants settled down.\textsuperscript{16} According to John Weatherford, marketplaces, with numerous small retailers, appeared for the first time in Sardes, Lydia at the end of the 6\textsuperscript{th} century B.C.

Presumably, a merchant who works permanently in the same location could charge a higher price for goods, the quality of which it takes some time to assess, as compared to a traveling merchant. The point is the repeated nature of his transactions with the inhabitants in the region. The character of the relation between buyer and seller starts to change towards a no-stranger relation and thereby social norms based on reciprocity become increasingly relevant.\textsuperscript{17} With time, by routinely assessing the trustworthiness of retailers in their daily life, the public will become accustomed to recognizing signs of trustworthiness. People will develop a tacit knowledge for judging trustworthiness, which

\textsuperscript{14} Cf. also Goodhart (1998b) and Redish (1992).
\textsuperscript{15} Cf. Melitz (1974).
\textsuperscript{17} By the way, we could observe that, in modern society, the alienation of buyer and seller has increased again and reciprocity-based trust is replaced by sunk cost investments in brand names and goodwill.
in turn paves the way for trust-intensive money. We could say that the ever-growing experience of buyer/seller relations helps people develop what Donald (Deirdre) N. McCloskey (1994) has called bourgeois virtues. Now, in a society where people know how to distinguish between those who are trustworthy and those who are not, there is probably a fairly small cost for trusting a paper note issuer. Paper note issuers will use similar signs of trustworthiness as those used by successful merchants, for example investments in permanent facilities. They will also build personal relations, and, of course, earn a reputation of honesty by actually making honest business. Personal relations are very important since we as human beings have an intuitive tendency to regard a person with a familiar face as someone inside our reciprocity sphere. We try to identify our relatives based on their social relation to us.

Eventually, people increasingly switch from gold coins to paper notes that represent legal claims to gold coins. When the trust problem is resolved, or at least considerably reduced, paper notes offer the user the same advantages as gold coins and the additional benefit that they are easier to handle and store. To the supplier, however, notes offer the decisive advantage that they are considerable cheaper to produce – and increasingly so as trade grows. This advantage benefits the public also, since they can place gold holdings with a banker in exchange for convertible paper notes and hence earn interest on their savings.18 The banker is willing to pay interest since he, because of the law of large numbers, does not have to keep 100 percent in reserves and thus is able to provide credit facilities with a higher interest rate.

1.5 Nominal debt notes and cash cards

Nominal debt notes, or fiat money, are intrinsically worthless pieces of paper representing a claim for its nominal value to the issuer. The formal difference between a convertible debt note and a nominal debt note is that while the former represents a legal claim to a certain commodity, i.e. gold coins, the latter only represents a legal claim to a nominal value. The issuer of nominal debt notes does not have to redeem them in anything but new notes of the same kind. Instead of being legal claim to something in particular, they are legal tender. However, as we know, e.g. from the former Soviet Union, legal tender is not enough to buy all goods. In civil society, it is not possible to force an unwilling seller to accept a certain payment technique only by referring to its status as legal tender. It is the special combination of a certain payment technique having legal tender status and the fact that most citizens are indebted to the issuer who makes
nominal debt notes a generally accepted payment technique. Nominal debt notes are always issued either directly by the government or by a subsidiary to it, as a central bank. Hence, holders of the money expect the State to accept their own debt notes as payment of taxes and we can hence easily understand why they accept these notes in exchange for real goods. It is only required that the holders expect the State to accept the notes for at least as long as they intend to keep them.

Regarding the problem of securing payment and evaluating product quality, nominal debt notes entail no difference compared to convertible debt notes. The reason for its introduction must be sought elsewhere and will probably be found in the extra seignorage that it allows the government to obtain.\(^\text{19}\)

Regarding the unit of account, however, the transition to nominal debt notes implied an important change. As long as gold convertibility was retained, the unit of account was tied to gold by the law of no arbitrage. However, it is important to understand that, although the unit of account derived its value from a fixed relation to gold, people did not use this relation to judge if the price of a good was fair. When we walk around in a grocery store and are confronted with a new brand of olive oil, we do not evaluate its value for money by comparing it to the amount of gold coins we could get for the same price. Rather, we compare its price to the price of brands of olive oil with which we are familiar. The point is that at this moment it is much more important that our debt notes are redeemable at fixed rates for all the different items in the grocery, than that they are redeemable for a certain amount of gold coins. This has two reasons: first, because it is more relevant to the actual choice we face, i.e. in a grocery store, our intention is to buy food, not gold; secondly, because the debt notes’ convertibility into gold is only valuable to the extent that we believe that gold can buy other goods, such as food. Most people are not good at evaluating the value of gold in its best opportunity use, i.e. its use in jewelry or its industrial uses. The average consumer knows the value of gold because he knows the prices of staple commodities as expressed in gold. This implies that people know the value of the unit of account through prices on items in their shopping basket, rather than through the value of gold.

We can generalize what we have just said: as long as there are fixed prices on items that you are familiar with, the fixed relation between the unit of account and gold brings no additional information about the real value of the unit of account. On the other hand, as a theoretical matter, when you negotiate a long-term contract, you may know very

---

\(^{18}\) Dowd (1996) suggests that initially, goldsmith bankers would charge a fee for storing and protecting gold coins and issue receipts, which gave the depositors the right to demand their gold back. These receipts gradually started to circulate as money and thus, a proportion of the gold was never circulated. This lay the foundation for fractional reserve banking.

\(^{19}\) Cf. e.g. Goodhart (1998b).
little about the real value of the unit of account at the end of the contract and hence be forced to use your expectations on the relative price between gold and the goods in question. Thus, it is only in long-term contracts that the unit of account is determined differently under the fiat money regime than under the convertible money regime. In Bengtsson (2003: 86-119), I discuss at length how long-term expectations under such circumstances are determined.

Although the determination issue is not at the center of our attention now, there is one point that must be stated here regarding the relation between the unit of account and central bank liabilities. A popular idea among monetary theorists is that central bank liabilities determine the value of the unit of account. For instance, Woodford (2000) writes:

From whence could any special role of the central bank in equilibrium determination derive? The answer is that the unit of account in a purely fiat system is defined in terms of the liabilities of the central bank.

Moreover, on the next page he accentuates the idea by stating:

But the market value of a dollar deposit in such an account [settlement account at the central bank] cannot be anything other than a dollar --- because this defines the meaning of a "dollar"!

How could that be? What would a promise to pay back a certain number of dollars be worth if no prices were quoted in dollars? No one would ever get the idea of issuing nominal debts in terms of dollars, pounds, krona or whatever, if the unit had not already been established as a unit of account. Paper money, and generally all kinds of money with a face value that is higher than its intrinsic value, presupposes an already established unit of account. This is most obvious in the case of inconvertible money: if the dollar were not already established as a unit of account, how would it be possible to put these notes into circulation? Legal tender would not be enough, because it has no meaning when no prices are quoted in that unit. It should be clear that to issue inconvertible money, it is essential that there are other contracts which determine the 'conversion rate' between the unit and its real value. Paper money that is convertible into gold coins would never have been issued if gold were not already recognized as a measure of value. Similarly, fiat paper notes would never have been issued were not the nominal unit of account recognized as a measure of value.

Close substitutes to paper notes, as cash cards and different kinds of e-wallets are equivalent to cash in most, economically relevant, aspects – the differences are mainly of technical nature. Obviously, cash cards require a much more advanced state of electronic development and they are associated with higher costs because they presuppose that the payee has the relevant equipment. On the other hand, they promise
lower costs for shops and banks because of the reduced risk for robbery. Regarding our main issues, securing of payment, evaluation of quality and the unit of account, cash cards are equivalent to cash.

7 Middlemen

Payment techniques do not have to involve a physical medium of exchange. Some of the most important payment techniques, both historically and in contemporary society, take the shape of services rather than goods. In fact, the single coincidence part of the double coincidence of wants problem can not be resolved without the service of middlemen. The same is true for the quality evaluation problem. This is the central message in Alchian’s paper “Why Money” from 1977 – perhaps the single most important writing on money in the twentieth century.

1.6 Merchants

From our discussion, we can see that the problems of a single coincidence of wants and quality evaluation are not resolved through the use of any medium of exchange so far discussed. In his paper, Alchian (1977:133) demonstrates the conjunct function of money and middlemen in order to overcome the problem of asymmetric information regarding both of these problems.

Ignorance of availability of goods and of their terms of trade and attributes will provoke efforts to reduce that ignorance in order to achieve more trade. Several institutions have evolved to reduce costs of reducing that ignorance: money; specialist middlemen who are expert in assessing attributes of goods, who carry inventories, and whose reliability of assurance is high; specialized marketplaces; and even unemployment. This paper concentrates on the way in which that ignorance leads to the use of money and how money requires concurrent exchange with specialist, expert, highly reputable middlemen.

We have already mentioned the role of merchants and marketplaces to the introduction of convertible debt notes. The main function of these middlemen was not, however, to provide paper money, but to overcome the single coincidence of wants problem and the problem of evaluating product quality, i.e. to reduce the ignorance of the availability of goods and of their terms of trade and attributes. As Alchian shows in his paper, the combination of reputable expert middlemen and an easily recognizable medium of exchange considerably reduces the transaction costs stemming from an asymmetric distribution of information about product quality. Still, it is not self-evident that the same agent should perform both functions. The middleman must primarily overcome the single coincidence problem by becoming an expert on finding buyers and sellers and keeping inventory. Secondary, he can offer the additional service of quality assurance, which will enable him to charge a higher price. However, the producer could take on this role himself by investing in a brand name – this is a better description of
many of today’s more alienated markets. Nevertheless, it is probably correct to assume that the first middlemen performed both functions and that it was not until later that producers took on the quality assurance role. As long as business is sufficiently small-scale, the buyer/seller relationship was not a pure stranger/stranger relation and we could therefore expect reciprocity to be an important factor in all transactions.

Alchian does not discuss different kinds of payment techniques and the only hint regarding what he refers to is the following statement: “We mean by money a commodity used in all, or a dominant number of exchanges.” (1977:133). While this statement appears to point in the direction of gold bullion – commodity – , the paper’s argumentation rather points to gold coins. From the low inspection costs he ascribes to his money, one may conclude that it can not be gold bullion. On the other hand, in the summarizing paragraph below, it is clear that Alchian imagines his money to evolve with middlemen to overcome the costs of identifying quality (1977:139).

Costs of identifying qualities of a good are what count. If costs for some good are low and generally low across members of society, the good will become a medium through which information costs can be reduced and exchange made more economical. But it will rise only with the rise of chains of experts in various goods and commodities, who know the goods cheaply, whose reputation for reliable evaluation is high, and who, because of that knowledge and the low cost of assuring buyer, become specialist middlemen in the good both as inventory carriers and buying and selling agents.

In order to incorporate Alchian’s analysis into the framework of this paper, we have to reinterpret it slightly. More precisely, we need to reinterpret it in terms of the different payment techniques discussed. In an ideal state of indirect exchange, one commodity, such as gold bullion, is used as a medium of exchange and prices are stated in terms of that commodity, i.e. the unit of account is a certain amount of gold. By itself, it helps to reduce the double coincidence of wants problem, and with “chains of experts in various goods and commodities, who know the goods cheaply, whose reputation for reliable evaluation is high”, it helps to reduce the single coincidence of wants problem as well as the quality evaluation problem. We can see that an additional service – which reduces transaction costs – arises from the simultaneous existence of a common unit of account and expert middlemen.

Standardized media of exchange, such as gold coins, further reduce transaction costs by greatly decreasing the identifying costs of the medium of exchange itself. No quality evaluation is required since it is sufficient to read the stamp on it to know what it is worth. This, in turn, is possible because the custom of indirect exchange with only one commodity acting as the medium of exchange has made people used to thinking of prices in terms of a unit of account, rather than in terms of relative prices. This is important because it is the habit of stating prices in a unit of account and recording debt in a unit of account that together with expert middlemen enable a society to reduce the
transaction costs that arise due to an asymmetric distribution of information about product quality. Thus, physical money is not necessarily required, since there are other ways to record debt.

Another very important reduction in transaction costs comes from the combination of a common unit of account and middlemen acting as market makers. This enables the establishment of market prices, something which considerably reduces the cost for assessing one’s opportunity set.

An important thing to learn from this analysis is that one can not understand all benefits from a monetized economy by studying the payment technique in isolation. The benefits of reduced costs for identifying the product quality do not appear if there are not also middlemen.

1.7 Debt-recording services

Under certain circumstances, there is no need for a payment technique represented by a common medium of exchange to overcome the problem of securing a payment, since there are other ways to fulfil a debt-recording function. At a medieval trade fair in Flanders for instance, all transactions were recorded throughout the trading period and the remaining debts after clearing, were settled only at the end of the trading period, as seen in the passage below about the fairs of Champagne (De Liebaart (2001)).

The grand fairs of Champagne clearly aimed at the international businessman. The organisation of a grand fair was strict and well defined. The first week was spent setting up trading stalls along the town streets. This was followed by a ten-day cloth sale, an eleven-day leather sale and nineteen days when various other goods were allowed to change ownership. A number of days devoted to the settling and closing of all accounts ended each fair.

This method of payment, involving a high degree of sequentiality, was successful since it was easy to assess if a person behaved fraudulent within the fair and since each participant had to take part in the fair to be profitable. Although the propensity for reciprocity may be important when such a system is initiated, eventually, it was the threat of being excluded from future trade that prevented the participants from cheating and made the system stable. This example illustrates how a monetary system that essentially is a bookkeeping system could survive within an entity with sufficient internal control. The decisive factor is the transparency of actions within the particular society or part of society. As mentioned before, if an agent’s performance in transactions could be identified without any costs, long-term self-interest would motivate the agent to fulfil his obligations in various transactions. The payment technique sometimes referred to as

---

20 Cf. also Pohl (1994:47). For another account of the long existence of cashless subsocieties, see Origo (1957). The scope of the merchant’s business had nothing to do with his stock of “base money”. The decisive factors were his trustworthiness (perceived solidity) and the supply of profitable business opportunities.
bank money in literature makes use of this property. It can be checks, off-line debit cards or giro systems. The middleman always offers a payment service that, compared to nominal debt notes, provides lower opportunity costs and a smaller risk of theft. The middleman specializes in the particular technique required and in monitoring the customer’s payment performance. By using the middleman’s service, the customer makes his actions transparent to the middleman. One of the characteristic features of a middleman is that he is powerful enough to be able to enforce the contract in most cases, and to survive losses from possible unsolved cases. This is probably the reason why middlemen in practice act as jobbers rather than as brokers; i.e. the middlemen take on the risk instead of the payee. Furthermore, the long-term benefits from being able to use the service bring most customers’ self-interest in line with an honest behavior.

Nowadays, an increasing proportion of payments is made through on-line debit cards. Theoretically, this payment technique is quite different from those using off-line debit cards. While the latter can be characterized as trust for hire, the former is a pure debt-recording function. Payment is completed simultaneously as the goods are handed over. In the trust for hire business, the middleman is a specialist both in dealing with risks of non-performing debtors and in providing the required technique. With the on-line payment technique, the middleman specializes only in the technique, since the payments involve no risk. (There are, of course, other risks involved. Payment services are often combined with credit facilities. This gives rise to another kind of risk, but that is a different issue.) What is particularly interesting about the on-line payment technique is that it highlights the fundamental payment problem that money solves, i.e. how to know if the transaction counterpart is trustworthy.

The core of payment techniques based on debt-recording, on-line or off-line, is that there is a middleman who has specialized in providing the debt-recording service to overcome problems concerning how to secure a payment. Combined with expert middlemen and a unit of account, these payment techniques overcome the problem of value evaluation, too.

8 Conclusions

In the paper, an analogy with length measurement is applied, in order to explore the nature of the unit for value measurement, i.e. the unit of account. As the meter is defined as the length traveled by light in vacuum during 1/299 792 458 of a second, the unit of account krona is defined as the purchasing power of the medium of exchange krona. However, one should be cautious when drawing conclusions from this analogy. Our unit of account is defined in our medium of exchange, but it is meaningful only because we can observe prices on real goods expressed in it. As it would be pointless to
define the meter as the length traveled by light in vacuum during $1/299\,792\,458$ of a second if we could not compare this length with anything else, it would be pointless to define our unit of account in something that is not priced.

In the paper it is explained how different payment techniques help to overcome transaction costs in the market. In particular, following Alchian (1977), it is argued that to reap the full benefit from the use of payment techniques, it has to be combined with the use of both a unit of account and specialist middlemen. The use of payment techniques helps to reduce costs due to sequential payment, but to reduce costs due to sequential quality evaluation, you need unit of account as well as reputable middlemen.
References


Jevons, W. S. (1875), Money and the Mechanism of Exchange, Appleton.


