CEU Notes on the Theory of Money and Banking

David Andolfatto

October 2008

1 The Basic Environment

There are $N \geq 3$ dates, labeled $t = 1, 2, ..., N$ and N nonstorable time-dated goods, labeled $\{y_t\}_{t=1}^N$. There are also N types of agents indexed by t and a continuum of each type of agent indexed by $i \in [0, 1]$. A type (t, i) agent has the ability to produce output at date t and has a desire to consume output at date $t + 1$. Preferences for a type t agent are given by,

$$U_t(i) = -y_t(i) + u(y_{t+1}(i)); \tag{1}$$

for $t = 1, 2, ..., N$ (modulo N). Assume that $u'' < 0 < u'$, $u(0) = 0$, $u(\overline{y}) = \overline{y}$ for some $\overline{y} > 0$; and that there exists a $y^* \in [0, \overline{y}]$ satisfying $u'(y^*) = 1$.

The structure of this economy can be depicted diagrammatically as follows.

Type\Date	1	2	3	$\cdots$	N
1	$-y_1$	$u(y_2)$			
2		$-y_2$	$u(y_3)$		
3			$-y_3$		
$\vdots$				$\ddots$	
N	$u(y_1)$				$-y_N$

Note that there are no **bilateral** gains to trade; i.e., there is a complete **lack of double-coincidence** of wants. Despite this fact, there are clearly **multilateral** gains to trade; achievable, for example, by the coalitions consisting of at least one agent of each type.

2 Method of Analysis

In what follows, I study the problem of resource allocation from the perspective of a benevolent social planner, whose powers will be limited in specific ways depending on the nature of the physical environment. The solution to a planning problem is useful because it provides us with a benchmark. This benchmark describes the upper limit, in terms of efficiency, that any well-designed institutional framework can ever hope to achieve. At the same time, it identifies the properties that any efficient (or constrained-efficient) allocation will have to satisfy irrespective of how trade actually takes place.

The first benchmark is derived under the assumption of an all-powerful planner. Let $\{c_{t+1}(i), y_t(i)\}$ denote an allocation for a type (t,i) agent; where $c_{t+1}(i)$ denotes his consumption (at date $t+1$) and $y_t(i)$ denotes his required production (at date t). Let $T \equiv \{1, 2, ..., N\}$. An **allocation** for this economy then is given by,

$$\{c_t(i), y_t(i) : (t,i) \in T \times [0,1]\}. \tag{2}$$

A **resource feasible** allocation must satisfy,

$$\int c_t(i)di \leq \int y_t(i)di \tag{3}$$

for all $t \in T$ (modulo N). Even the most powerful planner must respect this resource constraint.

Assume that the planner assigns an equal Pareto-weight to each agent; the social welfare function is given by $\sum_{t=1}^{N} \int U_t(i)di$. The solution to this problem clearly entails setting $y_t(i) = y^*$ for all (t,i). Efficiency dictates all the production of type t agents is allocated for the purpose of consumption to type $t-1$ agents; for all $t \in T$ (modulo N). In what follows, I call this the **first-best allocation**.

Under standard assumptions, the first-best allocation can be implemented as the equilibrium of a competitive market structure. In an Arrow-Debreu market, for example, each agent is imagined to issue securities representing sure claims against his output and using this wealth to purchase claims against the output he desires. In equilibrium, all securities will trade at par; and all agents will end up consuming and producing y^*. There are many other institutional settings, however, that can potentially achieve the same result.

3 Limited Commitment

In this section, I weaken the planner's power by assuming that–with the exception of type N agents–people cannot be forced to produce output. Instead, type $t \in T \backslash N$ agents must be given the proper incentive to produce at each point in time; their actions must be **sequentially rational**. Another way of thinking of this is that this latter group of agents cannot commit beforehand to any recommended production plan; and that only type N agents can commit in this manner. It is in this sense that commitment is limited in this environment.

How might the planner handle this situation? Imagine that the planner recommends that each agent produce $y_t(i) = y \in [0, \overline{y}]$ units of output. Type N

agents are required to say 'yes' or 'no' to this proposal at the beginning of time; if they answer in the affirmative, they are committed to following through with their promise.[1]

What incentive do agents have in following through with the planner's recommendation? To induce compliance, the planner must devise a sequentially rational incentive scheme. Here is something that will work:

$$c_t(i) = \begin{cases} y & \text{if} \quad y_{t-1}(i) \geq y; \\ 0 & \text{if} \quad y_{t-1}(i) < y; \end{cases} \tag{4}$$

for all (t, i). In words, the planner promises to reward each person with consumption y if they produced at least y units of output in the past.[2] If agents fail to comply with the planner's recommendation, they are punished in the worse possible way (they are disentitled to any consumption).

The allocation rule (4) is clearly resource feasible and it respects sequential rationality as $u(y) - y \geq 0$ for all $y \in [0, \overline{y}]$. An allocation that respects both resource and incentive constraints is called **incentive-feasible**. Given the allocation rule (4), it follows that any allocation $y \in [0, \overline{y}]$ is incentive-feasible. And of course, as y^* belongs to this set, the first-best allocation is implementable.

Alternatively, the planner may not make any recommendations at all. Instead, he might ask all agents to submit their production plans while at the same time announcing the following allocation rule:

$$c_t(i) = y_{t-1}(i), \tag{5}$$

for all (t, i). That is, with the exception of type N agents, all agents are entitled to consume one unit of output at date t for every unit of output they contributed at date $t-1$. A type N agent is entitled to consume one unit of output at date 1 for every unit of output he *promises* to deliver at date N (remember that type N agents alone can commit to such promises).

In formulating their production (consumption) plans, agents take the allocation rule (5) as given. Their choice problem can therefore be stated as follows,

$$\max \left\{ -y_t(i) + u(c_{t+1}(i)) : c_{t+1}(i) = y_t(i) \right\}. \tag{6}$$

The solution to this problem is a production plan $y_t^s(i)$ satisfying,

$$u'(y_t^s(i)) = 1. \tag{7}$$

Evidently, this implies that $y_t^s(i) = y^*$ for all (t, i). The collection of such plans is clearly feasible; the resulting allocation corresponds to the first-best allocation.

[1] Alternatively, we might instead just assume that the planner announces that he will forcibly extract y units of output from type N agents at date N.

[2] Of course, type N agents receive y units of output at date 1 in exchange for their promise of delivering y units of output at date N.

3.1 Discussion

The analysis above demonstrates the incentive problems introduced by limited commitment can be dealt with by designing an appropriate incentive scheme. In this simple environment, an incentive scheme can be designed to circumvent the limited commitment problem entirely; but this is not a general result. In general, the best-designed incentive scheme might only implement a constrained-efficient allocation. In either case, however, a well-designed incentive scheme will share a common property; namely, that punishments and rewards should be dispersed on the basis of **individual trading histories**.

Of course, the viability of this incentive scheme depends on a number of things. First, it assumes that individual trading histories can be observed and recorded. Second, it assumes that the requisite punishments and rewards can be administered in a credible manner. Both of these assumptions seem natural enough if we view the planner as a sort of intermediary that physically collects and distributes output; while at the same time maintaining a record of contributions and disbursements. In fact, an intermediary that combines both of these functions for a coalition of agents is possibly the correct way in which to define the concept of a *firm*.

4 Private Information and Record-Keeping

Imagine that our planner (or a conglomerate firm, if you prefer) can identify people at each date by type, but not necessarily by their trading history. That is, at a point in time, the planner can see whether someone desires consumption, but cannot remember the agent's past production. This would not pose any problem if agents could simply be relied upon to tell the truth about their personal histories; but such an expectation would be rather naive. In fact, there is good reason to believe that, if agents could profitably exploit the opportunity to make false claims, they would.

In general, the way economists deal with private information is to restrict attention to **incentive-compatible** allocations. In the present context, this would entail having the planner asking each consumer at every date to submit a *report* of their trading history. The consumer is then allocated consumption depending on the report they submit. Frequently, the allocation can be designed in a manner that gives agents an incentive to tell the truth; at least, assuming that everyone else in the economy tells the truth as well. Unfortunately, when agents are free to fabricate their own personal histories at every date, agents are in effect **anonymous**; i.e., allocations cannot depend on personal trading histories.

The only solution to this problem then is to introduce a **record-keeping technology**. Perhaps the planner cannot identify anyone today with what they did in the past, and perhaps people cannot be expected to communicate this information truthfully; but there are other ways in which this information might be stored or communicated. Precisely how this might be done depends on assumptions concerning the set of available record-keeping technologies.

Consider the following technology. The planner has access to a computer (or a set of books). Each producer who wishes to make a deposit is assigned a unique passcode. An electronic account is opened and a deposit of y units of output is recorded as a credit measured in some unit of account; e.g., $m = py$ dollars (where $p > 0$ is an arbitrary constant). Future access to this account is limited to those who present the correct passcode. Of course, this opens up the possibility of **identity theft**. For example, some unscrupulous person may hack into the computer system and guess your passcode correctly. Or you may be mugged and forced to hand over your passcode. On the other hand, you might simply lose your passcode if it is embedded in a plastic card, for example. For now, I abstract from such possibilities.

Let me now describe how things might work here. At date 1, a type N agent approaches the planner and asks for a loan of m dollars of output. As collateral for this loan, a type N agent can offer a security consisting of a promise to deliver m dollars of output at date N. At this moment then, a type N agent has m dollars in assets and m dollars in liabilities. At the same time, a type 1 agent deposits m dollars of output. The latter's account is credited by m dollars and the former's is debited by m dollars; as the deposited output is subsequently transferred to a type N agent. At the end of date 1, a type 1 agent has m dollars in assets and a type N agent has m dollars in liabilities.

Now, at date 2, a type 1 agent desires consumption. But to be entitled to consumption, he will have to provide **verifiable evidence** that he produced in the past. This evidence now exists in the form of credits that are sitting in his account (he is in no position to fabricate his account balances). Moreover, only the true owner of the account can access it with the requisite passcode (others are not in a position to take undue credit for services rendered in the past). Hence, at date 2, type 2 agents produce and deposit their output y for credit m. The output is then transferred to those type 1 agents who wish to make a withdrawal; if they do, their credit balance falls to zero.

This whole process can then just repeat itself until date N. At this final date, type N agents are obliged to produce y units of output to settle their debt with the planner (society). Of course, it is essential here that either the planner can forcibly extract this output (assume that he can identify type N agents and their histories at date N), or that type N agents are simply endowed with the commitment to make good on their promises.

4.1 Discussion

At the present level of generality, the planning mechanism can be interpreted in a variety of ways. One can almost see a rudimentary theory of banking emerging from this analysis. In particular, note that the planner begins business by accepting a security as collateral for a loan, simultaneously creating an accounting instrument (money) of equal value. The accounting instrument then 'circulates' from account to account as agents buy and sell goods over time. The accounting object is extinguished once the loan is repaid. The planner's key role in this process is to offer society an accurate and secure record-keeping service.

There are at least two defects of this model as a theory of banking. The first is that the transformation $m = py$ is entirely innocuous; which is to say that while there is no harm in doing so, it is not necessary either. As $p > 0$ is an arbitrary number, we may as well set $p = 1$, so that $m = y$. In this case, one dollar is by definition one unit of output; and whether value is measured is in units of dollars or output is the same thing.[3] The implication of this is that the planner may as well use the securities deposited by type N agents as the unit of account (there is no specific need to transform units into some other measure). But then this begs the question of why type N agents cannot simply use their own securities as money. They are prevented from doing so here by the assumption that all trade is intermediated through the planner. This then is related to second defect of the model as a theory of banking; namely, that it is not clear whether intermediated exchange is essential either. Is it possible, for example, to implement the first-best allocation as the equilibrium of a game or of a competitive market?

Whatever the defects of the model here as a theory of banking, it seems a little more satisfactory as a theory of money. The key insight appears to be that money is a record-keeping device; or, equivalently, a communication device. The logic of the argument is as follows. When commitment is limited, efficiency dictates that agents have private incentives to behave in the social interest. These incentives must be based on individual trading histories. If individual trading histories can be fabricated with impunity, then society must devise or make use of a record-keeping technology. Demonstrated acts of good will (like producing output for others) must be given credit upon which the contributor can ultimately draw on as a reward. As agents cannot be relied upon to communicate their own past contributions truthfully when receiving a reward, their sacrifice must be recorded at the time it takes place. This record must be made accurate and secure.

One might add that these records are fiat objects; they are intrinsically useless and constitute no enforceable claim against anything of intrinsic value at any point in time–except for the final period. In equilibrium, these accounting objects are exchanged for output at each point in time; but exchange is entirely voluntary (sequentially rational) in all but the final period.

[3] Note that the term *dollar* originally referred to a specific quantity of gold or silver.

4.2 Tangible Money

In the literature, the definition of money is sometimes limited to objects that are tangible in nature; e.g., paper notes or gold coins. This is in contrast to the intangible book-entry accounting objects described above. Let me refer to tangible money as **cash**.

When is money in cash form necessary? In the previous section, I assumed that the planner had access to a computer or a set of books; i.e., a technology for keeping records. Imagine that such a technology is prohibitively expensive to operate, or that it is simply unavailable. Since some form of record-keeping is necessary, an alternative (cheaper) technology must be available if any trade is to take place. Imagine then that the planner has a technology for creating tokens. One might imagine that everyone has a technology for creating tokens as well; we need to assume that only the planner (society) can create **non-counterfeitable** tokens.

Now, when agents make a deposit of output, instead of being issued an account and a passcode, they are simply given tokens (in proportion to the output they deposit). These tokens are, in effect, a receipt; that is, they constitute evidence of a private trading-history that can be credibly communicated to the planner in the future (when the tokens are used to purchase output). In this way, cash is simply another record-keeping technology that substitutes for some other record-keeping technology that is either unavailable or too costly to operate.[4]

Finally, note that just as intangible records may be subject to identity theft, tangible tokens may be subject to theft or counterfeit as well. To put things another way, when somebody steals your money, they are in effect stealing a part of your identity (that part which records your past sacrifice). What this suggests is that the technology of theft may play a significant role in determining the optimal record-keeping technology. For example, imagine that a centralized book-keeping system offers security but for a service fee; while cash entails no cost, but is subject to theft. Depending on the cost of using the centralized service and the probability of theft, one or the other method of payment will be desirable.

[4] Money (cash) is frequently motivated in the literature by assuming the absence of a record-keeping technology. What is really meant by this is that money is (by assumption) a superior record-keeping technology.

Money (cash) is sometimes also motivated in the literature by assuming that private trading histories are private information. The analysis above suggests a subtle distinction; i.e., that money is essential, but cash need not be (depending on the best available record-keeping technology).

5 Intermediated versus Unintermediated Trade

Despite the added frictions of limited commitment and private information (over individual trading histories), it is not immediately clear that an intermediated structure like a bank (our planner) is necessary. In particular, can we not exploit the fact that type N agents are endowed with commitment and allow them to issue their own money that subsequently circulates as a means of payment in a competitive market?

At date 1, a type N agent wishes to purchase y units of output from a type 1 agent. The only thing the former can offer in payment is a security entitled the *bearer* to y units of output at date N. If society has access to the computerized record-keeping technology described above, these securities can be issued in an intangible form; otherwise, they can be issued in paper form. In fact, paper form is likely to be preferable if (for whatever reason) not all members of society have immediate access to a computer network at all times. In any case, it is essential that these securities cannot be counterfeited. It is also essential that agents accepting these securities as payment are aware of their collateral value.

With these assumptions in place, one might imagine the existence of a set of sequential debt markets. Let R_{t+1} denote the (gross) real rate of return on a security from date t to date $t+1$. In a competitive market, each agent takes as given the price sequence $\{R_{t+1}\}_{t=1}^{N-1}$ when formulating their consumption/saving decisions.

For a representative type N agent, the choice problem can be stated as

$$\max\left\{u(c_1) - y_N : c_1 \leq R^{-1}y_N\right\}; \tag{8}$$

where $R \equiv \Pi_{t=1}^{N-1}R_{t+1}$. For all other agents $t \neq N$, the choice problem can be stated as

$$\max\left\{u(c_{t+1}) - y_t : c_{t+1} \leq R_{t+1}y_t\right\}. \tag{9}$$

Given the symmetry embedded in the environment, it is easy to verify that the competitive equilibrium allocation and price system is given by $y_t = y^*$ and $R_t^* = 1$ for all t.

5.1 Discussion

The equilibrium described above has a fraction $(1/N)$ of the population issuing their own securities (each trading at par) that subsequently circulate as a medium of exchange. In the model, these are the agents that can back their securities with collateral.

One might question whether empirical evidence supports this scenario. In fact, absent legal restrictions prohibiting the practice of small note issue, this type of behavior is prevalent throughout recent history. For example, Bodenhorn (1993) quotes an Italian General Secretary of the Banco D'Italia how, prior to 1874,

> "everyone was issuing notes, even individuals and commercial firms; the country was overrun with little notes of 50, 25, and 20 centimes issued by everyone who liked to do so."

The author also notes that when state legislation banned U.S. banks from issuing notes of less than $5, railroad companies, public houses, merchants and even churches filled the void with their own notes. Countless other examples are available.

On the other hand, the quote above suggests that there may have been a problem with many small and independent note-issuers. In particular, he claims that the country was "overrun" with notes–suggesting that notes were being "over-issued" in some sense. Of course, one should check the historical evidence to see whether this was true or not. But regardless of the evidence, there is a definite impression that a small number of currencies is to be preferred to a large number. Presumably, the idea is that it is easier to counterfeit any given note, if it is just one of several hundred or more. Or perhaps the idea is that for small transactions, it is not worth the bother to check the quality of the note (i.e., its collateral value)–making it easier for scoundrels to issue "unsound" money. None of these problems are present in our model (although, it would be interesting to introduce them).

One might also add that in any well-developed economy, the vast bulk of the money supply is created by intermediaries called banks. This was true even in the absence of any government money; for example, as in the U.S. free-banking era (1836-63). In the antebellum U.S., there were hundreds of banks issuing their own dollar notes, mad redeemable in specie (gold and silver coins). The same holds true today; except that bank liabilities now take an electronic form and are redeemable in government paper money (as when you make a withdrawal at an ATM).

The model above is not rich enough to explain why a bank (a money-issuing intermediary) may be essential, but it does give us some hints. Imagine, for example, that it is difficult for agents to access the quality of the collateral offered by type N agents. Implicitly, I am assuming that the identity of all agents is also private information (so that anyone can make promises by pretending to be type N agents). If this is the case, then the securities that are offered by legitimate type N agents will be **illiquid** (they cannot be disposed of quickly; some time and effort will have to be used to access the quality of the collateral).

Now imagine that the planner (bank) has a comparative advantage in accessing the quality of collateral assets. For simplicity, assume that the bank can costlessly identify type N agents. Then the bank could ask legitimate type N agents to deposit their securities with the bank; with the bank then issuing a new security that is backed by its pool of (illiquid) collateral assets. If other agents trust the bank in its capacity to perform such an activity, then the bank's security can circulate as a means of payment.

In fact, in the real world, this is precisely what banks do; they are specialized in screening loan applicants. Successful loan applicants have their accounts credited with money (in this case, private bank notes or electronic money). These agents can then use the bank money to purchase the goods they desire. They offer as collateral for their loan an illiquid security (if it was not illiquid, then they could have used their own security as money). The bank's essential function here is that of an **asset transformer**; i.e., it creates liquid liabilities out of illiquid assets.