Optima in heterogeneous-agent monetary economies

Neil Wallace

April 22, 2014

1 Introduction

What do we know about good policy or optima in monetary economies? Not much. By not much I mean that it is hard to reach conclusions about optima in general monetary economies because optimal policy—even its qualitative aspects—depends on details of the economy that we are unlikely to know much about. This happens because the features that give money a role have at least two consequences: they make it desirable to enhance the return on money (in accord with the so-called Friedman rule) and they inhibit risk-sharing. Moreover, within the class of feasible policies, policies consistent with money having a role, enhancing the return on money and improving risk-sharing are competing goals. Therefore, and not surprisingly, the best way to balance them depends on details of the economy.

2 General monetary economies

A monetary economy is one in which currency-like objects are essential in the sense that their presence allows for the implementation of some good allocations that would otherwise not be implementable. What sort of environments make money essential in this sense? For more than 1,500 years, scholars have pointed to environments with double-coincidence problems. However, we now know that a double-coincidence problem is only one necessary condition for essentiality of money. (Robinson Crusoe and Friday can encounter double-coincidence problems, but few would say that they need money to overcome those problems.) Many of us now think that a second necessary condition has been identified in the work of Ostrov [20], Townsend [24], and Kocherlakota [14]. This necessary condition, which rests on the view that money is nothing but a signal


1If you regard that as an unpleasant message, then too bad. I recall that Milton Friedman said: the problem is not what people don’t know; the problem is what they think they know.
about previous actions, is imperfect monitoring—some lack of common knowledge about previous actions. Therefore, we seem to have two necessary conditions for essentiality—a double-coincidence problem and imperfect monitoring. However, as indicated by recent work on folk theorems, they are far from sufficient. Nevertheless, those conditions can provide a foundation for constructing examples of monetary economies (see, for example, Araujo [5]). I should also mention that my discussion presumes that the background environment is one in which people in the model cannot commit to future actions—the standard repeated- and dynamic-game assumption.

Now we can turn to what general means. A double-coincidence problem requires heterogeneity—as does any model in which there is a motivation for trade. Therefore, let the state of the economy be a joint distribution across the population over types and asset holdings, denoted $\lambda$. And, for now, consider the following schematic law of motion for $\lambda$ in a discrete-time world with a fixed stock of outside (fiat) money and nothing that we would normally identify with policy,

$$\lambda_{t+1} = H(\lambda_t, \text{shocks}_t, \text{trades}_t).$$ (1)

Although I am being vague about the environment, let’s treat the order of the arguments of $H$ as depicting the sequence of actions so that the shocks, idiosyncratic and/or aggregate, are realized before the trades occur. And we should think of consumption and production as occurring simultaneously with or just after the trades. The shocks could be about tastes (see Lucas [19]), endowments (see Levine [18]), or meetings (see Kiyotaki and Wright [13]). General in this context means that there are shocks and that $\lambda_t$ affects the trades at $t$ and that those trades affect $\lambda_{t+1}$—the usual situation in heterogeneous-agent models. Put somewhat differently, the trades at $t$ have two roles: they affect date-$t$ period payoffs of the agents (because they affect consumption and production) and they affect the state at the next date. As a consequence, good trades must represent a compromise between what would be best for date-$t$ period payoffs and what would be best for the date-$t+1$ state.

One good way to highlight this meaning of general is to review a list of well-known models that are not general in that sense. One is an OLG model of two-period-lived people with one good per date. In that model the distribution of money holdings among the old at a date is not a state variable because they are at a corner; they offer all their money at any price of money. Another is Shi [21] and Trejos and Wright [25], similar random matching models with money holdings limited to the set $\{0, 1\}$. In those models, the distribution $\lambda$ is the fraction who hold money at the start of a date, a fraction which is determined entirely by the exogenous stock of money and, therefore, does not depend on the trades that are made. Still another such model is the large-family model in which asset holdings of the members are merged and redistributed in any way across the members (see Shi [22] for an application to money). And, finally, there is Lagos and Wright [17] and its many offshoots in which periods of centralized trade with quasi-linear preferences or other special assumptions eliminate any inherited heterogeneity of asset holdings. While these are evidently knife-edge
specifications, that would be okay if their consequences for good policies are robust. I will try to convince you that that is not the case.

3 Implementability, policy, and the main trade-off

To begin, let’s generalize the law of motion to include policy, so that

$$\lambda_{t+1} = H'(\lambda_t, shocks_t, trades_t, policy_t).$$

(2)

As above, the ordering of arguments depicts the sequence of actions. In accord with that sequencing, I distinguish between the trade-stage and the subsequent policy-stage at each date. The policy stage has neither production nor consumption; it only has taxes and transfers. Moreover, we want those policies to be consistent with the environment and, in particular, with the assumptions that make money essential.

For the moment, everything I say is about weak implementability. (In order to achieve strong implementability (uniqueness), it is necessary to eliminate non-monetary equilibria and I have nothing new to say about how that might be accomplished.) In order to proceed and to discuss examples, which I will do later, it is helpful to delineate for each stage combinations of the information structure and the notion of defection as depicted in the following table,

<table>
<thead>
<tr>
<th>Specifications for each stage</th>
<th>symmetric</th>
<th>asymmetric</th>
</tr>
</thead>
<tbody>
<tr>
<td>individual defection</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>group defection</td>
<td>III</td>
<td>IV</td>
</tr>
</tbody>
</table>

The columns here do not refer to histories. Instead, they refer to tastes, endowments, and portfolios. For example, in a model in which people trade in pairs, there is asymmetric information at the trade-stage if the seller has private information about his endowment or his disutility of production. And at either stage, there is asymmetric information if agents can hide assets. As regards rows, under individual defection, the agents get to choose only from the set \(\{yes, no\}\) in response to a planner suggestion; under group defection they are free to cooperatively defect to any current action that is feasible for the group. Cooperatively is in italics because it is a straightforward notion under symmetric information, but is far from straightforward under asymmetric information.

I favor specifications which permit group defection at the trade stage. Why, after all, should a group be limited to choosing only from \(\{yes, no\}\)? Moreover, such a strategy set is so restrictive that it trivializes puzzles like coexistence of money and higher-return assets. If the planner suggests cash-in-advance trade and agents can only respond from \(\{yes, no\}\), then cash-in-advance trade is almost always implementable. Under symmetric information at the trade stage, group defection at that stage has the attractive property that it exhausts the gains from trade in meetings and does not rationalize cash-in-advance. However,
group defection and asymmetric information at either stage call for the adoption of some notion of the core under private information, or, perhaps, some notion of renegotiation-proofness. So far as I know, existing work on optima in heterogeneous-agent monetary models has not dealt with such specifications.

In models with a large number of agents, symmetric information, and centralized trade (everyone together), group defection implies price-taking trade. In one respect, that is convenient because then the planner has nothing to choose at the trade stage. If trade occurs in small groups under symmetric information, then there is usually a nondegenerate core and the planner can choose among them. At the policy stage, group defection limits nonneutral transfers to be lump-sum transfers. And, if people can hide assets at that stage, then transfers must be weakly increasing in asset holdings.

Before turning to illustrative environments and examples, I can offer a conjecture about circumstances under which there is a role for policy and under which optima depend on details of the economy. First, let’s define a welfare criterion and a notion of first-best. My argument that optima depend on details is only strengthened if I use the simplest welfare criterion; namely, ex ante representative-agent welfare. Therefore, I assume a fixed population of infinitely-lived people and treat them as identical prior to the realizations determined by $\lambda_0$, the initial state. (We may want to treat $\lambda_0$ as arbitrary and given, or we may want to treat the initial asset-holding distribution as something chosen by the planner.) I also assume that the planner can commit to future actions. By first-best, I mean the optimum in the given environment, but with imperfect monitoring replaced by perfect monitoring—but still subject to no-commitment by agents. Now, assume that the environment is a general monetary economy. For such an economy, the conjecture is that the best outcome has a role for policy, falls short of the first-best, and involves a trade-off between raising the return on money and improving risk-sharing that is best resolved in a way that depends on all the details of the environment.

These conclusions are not true in the knife-edge models noted above. In Shi [21] and Trejos and Wright [25], the models with money holdings limited to the set $\{0, 1\}$, there is no scope for policy. In the large-family model, risk-sharing is accomplished within the family; while in Lagos and Wright [17], the quasi-linear preferences or other special assumptions eliminate any role for risk-sharing. Thus, in those models, policy does not confront the trade-off between enhancing the return on money and improving risk-sharing. Hence, in those models the first-best can sometimes be attained (see, for example, Hu et al. [10] and Hu and Rocheteau [11]).

Although I could end here, I want to say a bit about work that bears on the validity of the above conjecture. I start with economies with only spot trade and then turn to economies with both spot trade and credit.
4 Pure quid-pro-quo economies

A pure quid-pro-quo economy usually means one in which only spot trade is possible. Such economies have been studied for a long time (examples are Lucas [19], Bewley [6], and Kiyotaki and Wright [13]). I will mean by a pure quid-pro-quo economy one with no monitoring so that each person’s previous actions are private to that person. Such complete absence of monitoring rules out trigger strategies of any sort to support credit and risk-sharing. When combined with group defection at the trade stage and at least individual defection at the transfer stage, it also eliminates taxation—except via inflation.

In Wallace [26] under the assumption that the planner is restricted to using a single uniform asset labeled money, I described a conjecture about optimal policy in such economies under specification II for the transfer stage (individual defection and hiding of money). In order to fix ideas, I focussed on the following two-parameter class of transfer functions at each date: a function \( \tau_t \) that maps end-of-trade money holdings of a person into a transfer; namely

\[
\tau_t(x) = \max\{0, a_t + b_t x\},
\]

(3)

where \( b_t \geq 0 \). When \( a_t = 0 \), this is a change in units and is equivalent to no transfer. When \( a_t > 0 \), it is equivalent to lump-sum transfers, a scheme with \( b_t = 0 \). When \( a_t < 0 \), the transfer is zero for \( x \leq -(a_t/b_t) \), and proportional to \( x - (a_t/b_t) \) for \( x > -(a_t/b_t) \). (The first use of an \( a_t < 0 \) scheme is in Andolfatto [3]). The conjecture is that schemes with \( a_t \neq 0 \) for most dates are optimal for general monetary economies. The very simple idea is that schemes with \( a_t = 0 \) are equivalent to no intervention and that \( a_t = 0 \) is interior in the set of feasible policies. Therefore, because no-intervention is not first-best in a general monetary economy, some intervention helps. Moreover, \( a_t > 0 \) improves risk-sharing, but lowers the return on money; while \( a_t < 0 \) raises the return on money over some range of holdings while worsening risk-sharing. Which is better depends on all the details of the economy. (Andolfatto [3] works within the general framework of Lagos and Wright [17] in which there is no risk-sharing motive. Therefore, he is able to show a welfare improvement from \( a_t < 0 \).)

In Wallace [26], I studied such a class of interventions in one numerical example borrowed from Kehoe et al. [12], a smooth preference version of Levine [18]. The model is a generalization of Townsend [23], an alternating endowment model; the generalization is that there is an i.i.d. aggregate shock with a two-point support. When the outcome is heads, those who had the high endowment at the last date receive the low endowment, and vice-versa for those who had the low endowment at the last date. Otherwise, there is no such switch. Let \( \pi \) be the probability of heads. With \( a_t \equiv 0 \), there is a monetary two-state equilibrium for all \( \pi \in (\pi_{\min}, 1] \), where \( \pi_{\min} \in (0, 1) \). I showed that there is a local improvement from \( a_t \neq 0 \) for all such \( \pi \). For all sufficiently high \( \pi \), \( a_t < 0 \), which raises the return on money held by those with the high endowment, produces a local improvement. For all other \( \pi \), both \( a_t < 0 \) and \( a_t > 0 \) produce local improvements, where the former raises the return on money
and the latter lowers it by way of lump-sum transfers.\textsuperscript{2}

Here, I want to discuss a richer class of policies, one involving more than one kind of asset. In particular, I want to discuss a possible role for a second asset, which we should think of as a bond, an asset with a higher rate of return than money. There are two challenges relating to policies of that sort. One involves feasibility; in particular, how do we get people to hold both assets? The second is about optimality. As noted above, feasibility is easy if only individual defection is imposed at the trade stage. Under that assumption, Kocherlakota \cite{15} showed that money and bonds are optimal in some economies.

There is work underway at Penn State that deals with both challenges. It starts from Shi \cite{21} and Trejos and Wright \cite{25}, but with a larger set of individual asset holdings. It follows Zhu and Wallace \cite{28} in adding a portfolio-choice stage just prior to the trade stage. Suppose that the planner offers one-period bonds at a discount chosen by the planner. The work assumes III at the trade stage (symmetric information and group defection) and II at the transfer stage (allows the hiding of assets). It also follows Zhu and Wallace \cite{28} in making the division of the gains from trade in a meeting depend on the portfolios held.\textsuperscript{3} In particular, buyers in a meeting who hold only the high return asset get a bad deal. Because having only one asset is a feasible policy, it is not a surprise to find that having two distinct assets with different rates of return is welfare improving.\textsuperscript{4}

5 Monetary economies with credit and taxation

Quid pro quo economies are extreme. Indeed, given the increase in the use of credit and debit cards, there has long been speculation that the world is headed toward being a cashless economy. In order to deal with money and credit, it is desirable to introduce some monitoring—but not so much as to eliminate a role for money. The presence of some monitoring will also allow for implementable outcomes that resemble taxation.

There are many ways of introducing some monitoring. There are models in which monitored status is a feature of meetings, rather than of people (see, for example, Aiyagari and Williamson \cite{2} and Williamson and Wright \cite{27}). There are also models with two separate sub-economies: a monitored part and a nonmonitored part (see Kranton \cite{16} and Antinolfi \textit{et. al.} \cite{4}). In those models, the threat of banishment to the nonmonitored part, which is a quid pro quo sub-

\textsuperscript{2}It may seem strange that both $a_t < 0$ and $a_t > 0$ produce local improvements for some parameters. That may happen because the equilibria are such that low endowment people do not save (and, of course, cannot dissave).

\textsuperscript{3}In a version of Lagos and Wright \cite{17} with money and capital, Hu and Rocheteau \cite{11} use a similar device to support a higher return on capital than that on money. They show that capital overaccumulation can sometimes be eliminated.

\textsuperscript{4}When carrying out this work in a model with an exogenous upper bound on individual asset holdings, it is desirable to specify the trades so that a second asset does not help simply because it helps overcome the bound. It is well-known (see, for example, Aiyagari \textit{et. al.} \cite{1}) that a second asset can play that role.
economy, induces people in the monitored part, which is cashless, to cooperate. Motivated initially by wanting to compare inside (private) money to outside money, Cavalcanti and Wallace [7] also have two segments of the population, one perfectly monitored and the rest not monitored at all, but they have the segments interacting with each other. They work against the background of the random matching models of Shi [21] and Trejos and Wright [25]. Even in versions with money holdings in \( \{0, 1\} \), the distribution of wealth between monitored and nonmonitored agents becomes an endogenous state variable of the model and, for that reason, makes the resulting economy a general monetary economy.

Some preliminary work on that model has turned up results consistent with the claim that optima in heterogeneous-agent models depend on all the details—consistent in the sense that the results are a bit surprising. Deviatov and Wallace [8] studied optimal seasonal policy—there are two seasons, winter and summer—in a version with only outside money and found that optimal seasonal policy has the surprising feature that the planner extends zero-interest loans to monitored people at the start of winter, the low trade season, and is repaid at the start of summer, the high-trade season. I see no reason to think that this is a general property of optima. Deviatov and Wallace [9] study optimal inflation in an inside-money version of the model (without seasons) and find that inflation is optimal. At each date, the monitored people, the issuers of inside money, spend more in meetings with nonmonitored people than they earn in such meetings. The excess disappears via inflation, which is modeled by random disintegration of the money holdings of nonmonitored people. In this case, the assumption that money holdings are in \( \{0, 1\} \) seems to work in favor of such an inflationary outcome. Therefore, some work is underway in order to explore whether that finding survives when the set of money holdings is allowed to be richer, \( \{0, 1, ..., B\} \) for some sizable \( B \).

In none of the above work was there a search over parameters to find what seem in retrospect to be surprising outcomes. But, by itself, that work is too limited to show that optima, even their qualitative features, depend on details. For example, to make that case for the seasonal model, one would want to explore optima over regions of the space of parameters and show that for some region of that space loans are granted at the beginning of summer and otherwise are granted at the beginning of winter. Consistent with the conjecture that intervention is generally beneficial, the possibility that optima have no loans ought to be nongeneric.

6 Concluding remarks

We want models in order to make better decisions. Put more succinctly, if a model does not have normative implications, then it’s nothing. Moreover, it is obvious that normative conclusions should be robust to obvious generalizations of the models that we use. In general monetary economies, policy confronts a trade-off between raising the return on money and enhancing risk-sharing and the best way to resolve the trade-off depends on details. Such economies
are complicated and their implications for optima may be somewhat difficult to learn about. That, however, is not a justification for using special models whose implications are not robust to obvious generalizations.

References


