An alternative to New-Keynesian models for the study of optimal (monetary) policy*

Neil Wallace†

April 21, 2013

Abstract

The alternative, based on Cavalcanti-Wallace 1999, is one way to bridge the gap between two extreme models: an economy with no monitoring (and, therefore, no credit) and an economy with perfect monitoring (and, therefore, no role for money). Here, using an outside-money version of the model, optima for two examples, drawn from earlier work by Deviatov and Wallace, are described. Although the examples are simple, the optima are not simple and depend on all the details of the examples. The results should make users of New Keynesian models wonder why optima in their models are so different.

Key words: money, monitoring, optima.

JEL classification: E52, E58

1 Introduction

Fifty years ago and earlier, the challenge for monetary theory was to integrate money into the theory of value, the general-equilibrium competitive model (see Debreu [6], page 36, note 3). Slowly, via the work of Ostroy [21], Townsend [23], and Kocherlakota [13], we have learned to pose the challenge differently, in a way consistent with a mechanism-design approach to monetary theory: find settings in which money helps to achieve good outcomes—or, in Hahn’s terminology, in which money is essential. The general result that has emerged from this new way of posing the challenge is that imperfect monitoring, some privacy of the history of individual actions, is necessary for essentiality of money (see Wallace [25]). Stated very loosely, money is potentially useful in trade between strangers; it is not needed when everyone knows what everyone else has done in the past—in

*This paper was prepared for presentation at the “Workshop on monetary policy in the presence of micro-founded market and informational frictions,” (Bank of Italy, via Nazionale 91, Roma, 6-7 June 2013). I am indebted to Randy Wright for helpful comments on an earlier version. This is a preliminary version. Please do not cite it without permission of the author.
†Department of Economics, Penn State University <neilw@psu.edu>.
an idealized version of an isolated Amish community, in an Israeli kibbutz, or on Crusoe’s island even after he meets Friday.

There are no general necessary and sufficient conditions for essentiality of money. In particular, imperfect monitoring is not sufficient to give a role for money.\(^1\) Therefore, many monetary models rule out monitoring completely—either implicitly or explicitly.\(^2\) Roughly speaking, the following conditions are sufficient for essentiality of money: no monitoring, discounting (that is not taken to the limit of no discounting), a large number of agents, some background absence-of-double-coincidence, and no durable objects other than money.

While useful for some purposes, such economies—which, following Lucas [19], I label \textit{pure-currency} economies—have no credit and, somewhat less obviously, no taxation. Indeed, they are best viewed as extreme versions of underground economies. The absence of credit seems troublesome because the role of central banks is widely viewed to be intervention involving credit. For that reason and for the sake of generality, we want a setting that has neither perfect monitoring nor is a \textit{pure-currency} economy. However, as is true elsewhere in economics, there are many ways to formulate intermediate situations, situations with some monitoring, and any way of doing so is almost certainly more complicated than either extreme.

About a dozen years ago, Ricardo Cavalcanti and I formulated an intermediate situation by having some exogenous fraction of the population be perfectly monitored, labeled \(m\)-people, and having the rest, labeled \(n\)-people, be not monitored at all (see [5]). The model was designed to compare inside (private money) and outside money as alternative monetary systems. Here, I review some work by Alexei Deviatov and me that uses the outside-money version of that model to study optimal policy. This is the model that I propose as an alternative to New-Keynesian models.

The results that I will report using the model are optima for two quite arbitrary numerical examples: one geared to finding optimal inflation; the other geared to finding optimal seasonal policy. Despite the arbitrariness of those examples, there is a general lesson to be learned: it is not easy to guess at even the qualitative nature of an optimum and the optimum depends on all the details of the model. As I will emphasize later, those features should be present in any model of money and credit that does not make very special, knife-edge, assumptions.

\section{The environment}

The background setting is borrowed from Trejos-Wright [24] and Shi [22], a \textit{pure-currency} economy with pairwise meetings at random. Time is discrete and there is a nonatomic measure of people, each of whom maximizes expected

\(^1\)Any setting in which the \textit{folk theorem} holds is a setting in which money has no role and many such settings have imperfect monitoring.

\(^2\)Whenever borrowing and lending is ruled out—as, for example, in Lucas [19], Bewley [2], and many other papers—the implicit assumption is no monitoring.
discounted utility with discount factor $\beta \in (0, 1)$. Production and consumption occur in pairwise meetings that occur at random in the following way. Just prior to such meetings, each a person looks forward to being a consumer (a buyer) who meets a random producer (seller) with probability $1/K$, looks forward to being a producer who meets a random consumer with probability $1/K$, and looks forward to no pairwise meeting with probability $1-(2/K)$, where integer $K \geq 2$.

The period utility of someone who becomes a consumer and consumes $y \in \mathbb{R}_+$ is $u(y)$, where $u$ is strictly increasing, strictly concave, differentiable, and satisfies $u(0) = 0$. The period utility of someone who becomes a producer and produces $y \in \mathbb{R}_+$ is $-c(y)$, where $c$ is strictly increasing, convex, and differentiable, and satisfies $c(0) = 0$. In addition, $y^* = \arg\max_{y \geq 0} [u(y) - c(y)]$ exists and is positive. Production is perishable; it is either consumed or lost.\footnote{If $K$ exceeds two, then, as is well-known, it can be interpreted as the number of goods and specialization types in Trejos and Wright [24] and Shi [22].}

In addition, either $u$ is bounded above or $c$ is such that $y$ is bounded above, an assumption that allows us to invoke the principle of one-shot deviations.

People in the model are ex ante identical but the fraction $\alpha$ become permanently monitored ($m$-people), while the rest are permanently nonmonitored ($n$-people).\footnote{The interpretation is that the fraction $\alpha$ realize a zero cost of attaining $m$-status and that the rest realize a prohibitively high cost of attaining $m$-status.} For $m$-people, histories and money holdings are common knowledge; for $n$-people, they are private. However, the monitored status and consumer-producer status of people in a pairwise meeting are common knowledge. And, no one except the planner can commit to future actions.

At each date, there are two stages. The first stage has the pairwise meetings just described. There is a second stage at which transfers are made. There is neither production nor consumption at the second stage. Money is uniform and indivisible, and each person’s holding of money is limited to be in the set $\{0, 1\}$ at any time.

The only feasible punishment is permanent banishment of an individual $m$-person to the set of $n$-people. Underlying this assumption about punishment is free exit at any time from the set of $m$-people into the set of $n$-people and the ruling out of global punishments—like the shutting down of all trade in response to individual defections.

### 3 Implementable allocations and the optimum problem

The search for an optimum is limited to allocations that are steady states and are symmetric, where symmetry means that all people in the same situation take the same action, an action that can be a lottery. (In general, lotteries here have the form of a deterministic amount of output exchanged for a probability of getting money.) The state of the economy entering a date is $(\theta_m, \theta_n)$, where $\theta_m \in [0, \alpha]$ is the fraction who are $m$-people with money and $\theta_n \in [0, 1-\alpha]$ is the fraction who are $n$-people with money. With $S = \{m, n\} \times \{0, 1\}$, the
state of a meeting is an element in $S \times S$, where the first component is the state of the producer and the second is that of the consumer. The planner chooses $(\theta_m, \theta_n)$, trades in meetings (as a function of the states of the producer and the consumer in the meeting), and second-stage transfers.

The planner is constrained by the steady-state restriction and by self-selection constraints that follow from the specification of private information and of punishments. The trades that the planner chooses for pairwise meetings are restricted to be both individually rational (IR) and immune to cooperative defection by the pair in any meeting. At the transfer stage, transfers are subject to being IR. In addition, as noted above, $n$-people are permitted to hide money at any time.

Several comments are in order about this notion of implementability. First, defection by an $n$-person has no further consequences for the person. Second, any defection by an $m$-person means permanent loss of $m$-status beginning at the next stage or date. The cooperative defection in meetings is static and assures only that the trade is in the pairwise core for the meeting taking as given the relevant continuation values.

The planner’s objective is ex ante expected utility, where $\alpha$ is the probability of becoming an $m$-person and where the probabilities of starting with money are determined by $(\theta_m, \theta_n)$. This notion of welfare is proportional to a weighted average of the surpluses in meetings; namely,

$$\sum_{s \in S} \sum_{s' \in S} \pi_s \pi_{s'} [u(y_{ss'}) - c(y_{ss'})],$$

where $y_{ss'}$ is production and consumption when the producer is in state $s$ and the consumer is in state $s'$ and where

$$(\pi_{m1}, \pi_{m0}, \pi_{n1}, \pi_{n0}) = (\theta_m, \alpha - \theta_m, \theta_n, 1 - \alpha - \theta_n).$$

The first-best is $u(y^*) - c(y^*)$—namely, output equal to $y^*$ in every (single-coincidence) meeting. Below, we express welfare as a fraction of that first-best welfare.

### 4 Optimal inflation

In models with divisible money, a standard normalization holds the stock of money fixed and represents inflation by a proportional tax on money holdings. The approach taken here is the same, except that the discreteness of money in the model—each person’s money holding is constrained to be in the set \{0, 1\}—dictates that we use a probabilistic version of such a tax: a person who ends up after trade with a unit of money loses it with some probability. This way of modeling inflation, which was first used by Victor Li (see [16] and [17]) and

---

5One could also study Pareto allocations by varying the weight in the welfare criterion attached to different people—for example, $m$-people and $n$-people or even types distinguished by both monitored status and money holdings.
has been used by others, has the same effects on incentives to acquire money as inflation in a model with divisible money. A literal interpretation is that money is made of stuff such that each unit disintegrates at each date with a probability that the planner chooses.

The following example is taken from Deviatov and Wallace [8]:

\[ u(y) = 1 - e^{-10y}, \quad c(y) = y, \quad K = 3, \quad \beta = .59, \quad \alpha = 1/4. \]

All the choices are arbitrary, except that for \( \beta \). It is chosen to satisfy two conditions on optima for \( \alpha \in \{0, 1\} \), the extreme situations with regard to monitoring. First, given the other aspects of the specification, we choose \( \beta \) so that if everyone is an \( m \)-person (\( \alpha = 1 \)), then the first best is implementable. In other words, only the presence of \( n \)-people prevents implementability of the first best. Second, if everyone is an \( n \)-person (\( \alpha = 0 \)), then it would be desirable to pay interest on money if doing so were feasible. Her are the details.

First consider \( \alpha = 1 \). This is an economy with no role for money and one in which punishment is permanent autarky for a defector. In it, the output \( y \) in every single-coincidence meeting is implementable if and only if

\[ c(y) \leq \beta[u(y) - c(y)]/K(1 - \beta) \quad (1) \]

or, equivalently,

\[ c(y) \leq \beta u(y)/[\beta + K(1 - \beta)]. \quad (2) \]

Let \( \bar{y} \) be the largest \( y \) for which (2) holds at equality. Then any \( y \in [0, \bar{y}] \) is implementable. The best implementable \( y \) is \( \min\{y^*, \bar{y}\} \). For our choice of \( u, c, \) and \( K \), the smallest \( \beta \) such that \( \bar{y} \geq y^* \) is \( \beta \approx .51 \). Therefore, for our choice, \( \beta = .59 \), the optimum is the first best if \( \alpha = 1 \).

Now consider \( \alpha = 0 \). Then trade occurs only if the producer has no money and the consumer has money and optimal inflation is zero. The relevant participation constraint is easily shown to be

\[ c(y) \leq \beta u(y)/[\beta + K(1 - \beta)/(1 - \theta_n)]. \quad (3) \]

Also, if \( y = y^* \) and \( \theta_n = 1/2 \) satisfies (3), then the optimum has \( y = y^* \), \( \theta_n = 1/2 \), and ex ante welfare equal to 1/4 of the first best—1/4 because that is the fraction of single-coincidence meetings in which the producer has no money and the consumer has money when \( \theta_n = 1/2 \). (Obviously, \( \theta_n = 1/2 \) maximizes the fraction of single-coincidence meetings in which the producer has no money and the consumer has money.) If not, then the optimum has \( y < y^* \), \( \theta_n < 1/2 \), ex ante welfare less than 1/4 of the first best, and payment of interest on money would be desirable (if it were feasible). For our choice of \( u, c, \) and \( K \), the smallest \( \beta \) for which (3) holds with \( y = y^* \) and \( \theta = 1/2 \) is \( \beta \approx .67 \). Thus, with \( \beta = .59 \) (the mid-point between \( \beta = .51 \) and \( \beta = .67 \)), it would be desirable to pay interest on money because doing so would loosen constraint (3).

With \( \alpha = 1/4 \), the optimum has ex ante welfare equal to 34% of the first best, has \( \theta_m = 1/4 \) (all the \( m \)-people have money), has \( \theta_n = .18 \) (only about one-quarter of the \( n \)-people have money), has a 16% inflation (disintegration) rate, and has no transfers to \( n \)-people at the second stage. There are five meetings in which trade can occur (see Table 1).
Table 1. Optimal trades

<table>
<thead>
<tr>
<th>(producer)(consumer)</th>
<th>output/(money transferred)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n0)(n1)</td>
<td>0.573/(1.0)</td>
</tr>
<tr>
<td>(n0)(m1)</td>
<td>0.573/(1.0)</td>
</tr>
<tr>
<td>(m1)(n0)</td>
<td>0.113/(0)</td>
</tr>
<tr>
<td>(m1)(n1)</td>
<td>0.381/(1.0)</td>
</tr>
<tr>
<td>(m1)(m1)</td>
<td>0.381/(na)</td>
</tr>
</tbody>
</table>

In the table, output is reported as a fraction of \( y^* \) and (*) means a binding a producer IR constraint. When the producer IR constraint is binding, the trade is a take-it-or-leave-it offer by the consumer. Given those binding IR constraints, it easy to see why output is lower in the last two rows than in the first two rows. In the first two rows, the binding defection payoff for the producer is the discounted value of entering the next date in state \((n,0)\); in the last two rows, it is the higher discounted value of entering the next date in state \((n,1)\).\(^6\) And, even though the third-row trade does not have the \(m\)-person on the verge of defecting, a higher output in that meeting would decrease the discounted value of being in state \((m,1)\) and increase that of being in state \((n,0)\), and, thereby, lead to a violation of the IR constraints in all the other meetings.

Because only about one-quarter of the \(n\)-people have money, the inflow of money into holdings by \(n\)-people (the second-row meeting) is roughly three times the outflow (the fourth-row meeting). The inflation reconciles those flows with a constant \(\theta_n\), while transfers at the second stage reconcile those flows with a constant \(\theta_m\). Here is one way to describe the transfers. The \(m\)-people could have a risk-sharing arrangement among themselves that has those who end stage-one with two units money surrender one unit with the proceeds distributed to those who end stage-1 without money. Given the above first-stage trades, net transfers by the planner at each date are needed to reconcile those insurance payments with a constant \(\theta_m\).

5 Optimal seasonal policy

In order to discuss seasonal policy, the setting is modified so that it contains a deterministic seasonal (see Deviatov and Wallace [7]). That is done by having a two-date periodic \(c\) function: at odd dates (winter), the disutility of production is higher than at even dates (summer). In other respects, the model is the same except that we now rule out inflation (there is no disintegration of money) and we look for the best two-date periodic implementable allocation taking the first date to be winter.\(^7\)

\[^6\] That the defection payoff for an \(m\)-person with money is that of an \((n,1)\) person depends on the assumption that money is uniform and that global punishments have been excluded. If the \(m\)-person had a person-specific money, which is one interpretation of inside money, and that person-specific money is worthless after a defection by the holder of that money, then better allocations are implementable.

\[^7\] See Cavalcanti and Nosal [4] for a similar background setting, but with only \(n\)-people. They permit random confiscation of money held by \(n\)-people, which is not allowed according
Our example has $\alpha = 1/4$, $K = 3$, $u(y) = 2y^{1/2}$, $\beta = .95$, and
\[
c_t(y) = \begin{cases} 
  y/(.8) & \text{if } t \text{ is odd (winter)} \\
  y/(1.25) & \text{if } t \text{ is even (summer)} \end{cases}.
\] (4)

For this example, maximum surplus is attained at $y = .8$ in a winter meeting and at $y = 1.25$ in a summer meeting. Also, first-best welfare is proportional to $[.8 + \beta(1.25)]$. If everyone is monitored, then the first-best is implementable. And if no one is monitored, then the optimum is $\theta_n = 1/2$ at every date, output such that the maximum surplus is attained when the consumer has money and the producer does not, and welfare equal to $1/4$ of first-best welfare. With $\alpha = 1/4$, the optimum is described in Tables 2 and 3.

<table>
<thead>
<tr>
<th>Table 2. Optimal quantity of money and welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>$\theta_m$</td>
</tr>
<tr>
<td>$\theta_n$</td>
</tr>
<tr>
<td>welfare/first-best welfare</td>
</tr>
</tbody>
</table>

Again, all the $m$-people have money at the start of each date and there are no transfers to $n$-people at the second stage of either date. Notice that the stock of money is larger at the start of winter than at the start of summer. The trades in meetings appear in Table 3.

<table>
<thead>
<tr>
<th>Table 3. Optimal trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>(prod)(con)</td>
</tr>
<tr>
<td>(n0)(n1)</td>
</tr>
<tr>
<td>(n0)(m1)</td>
</tr>
<tr>
<td>(m1)(n0)</td>
</tr>
<tr>
<td>(m1)(n1)</td>
</tr>
<tr>
<td>(m1)(m1)</td>
</tr>
</tbody>
</table>

In the table, output at each date is expressed as a fraction of the respective first-best output, (*) denotes a binding producer IR constraint, and (†) denotes a binding truth-telling constraint (the $n$-person with money is indifferent between the fourth-row trade and the third-row trade). When money transferred is in $(0, 1)$, as in the second row, it is the probability that the consumer transfers 1 unit to the producer.

In order to interpret optimal intervention, we again focus on inflows into and outflows from money holdings of $n$ people. In winter, the difference is proportional to $(.750 -.312)(.505) -.312(.813)$, which is negative. And, because there

---

8 Although this example does not satisfy the boundedness requirement, it is easily amended to satisfy it in a way that does not affect the optimum. One possibility is to assume that (4) holds only for $y \leq \hat{y}$, where $\hat{y} = 100y^*$ and that $c(y) = \infty$ for $y \geq \hat{y}$.

---

...to the notion of implementability used here. Whether their class of policies could instead be described as positive transfers and inflation is not immediately apparent.
are no transfers to \(n\)-people at the second stage, those flows imply that fewer \(n\)-people have money at the start of summer than at the start of winter (see Table 2). In summer, an exactly offsetting net flow occurs—a consequence of the restriction to two-date periodic allocations without inflation. To reconcile those flows with every \(m\)-person starting with money at every date, \(m\)-people in the aggregate surrender money at the beginning of summer and receive an exactly offsetting amount at the beginning of winter—transfers which can be interpreted as loans. If the \(m\)-people are in an insurance arrangement in which \(m\)-people who earn money surrender it and those who spend money collect money, then without stage-2 intervention there would excess money at the beginning of summer and a shortfall at the beginning of winter. One way to offset those consequences is to have the planner extend zero-interest loans to \(m\)-people at the beginning of winter with repayment made at the beginning of summer.\(^9\)

Of course, real business-cycle enthusiasts will notice that the seasonal example is a special case of a real business cycle model. Nothing in principle prevents changing the model into one with a random process for the disutility of production. Such a model, but with only \(n\)-people, is studied in Cavalcanti and Erosa [3] and in Huang and Igarashi [12].

### 6 Comments on the model

Given that my goal is to present an attractive alternative to New-Keynesian models, I want to discuss some aspects of the model that may seem objectionable. First, there are no sticky prices in the model. Although I find that a virtue, this paper is not the place for a debate on that matter.\(^{10}\) Instead, I want to comment on other features of the model.

People in the model meet in pairs to trade and meetings occur at random. As regards meetings in pairs, even if we set aside all the descriptions of absence-of-double-coincidence situations that presume such meetings, there are good reasons for using such models. Pairwise meetings are, of course, the standard model in labor economics. In addition, they have been used to study the following diverse topics in monetary economics, none of which can be addressed in models in which trade occurs among a large group: float (see Wallace and Zhu [26]), the denomination structure of currency (see Lee et. al. [15]), coexistence

---

\(^9\)One might have guessed that the loans would be extended at the start of summer (when goods are plentiful) and would be repaid at the start of winter (when goods are scarce), as hinted at in the following statement:

> "[For the Bank of England in 1805] knowing the direction of the wind was [important] ... If ... from the east, ships would soon be sailing up the Thames to unload goods in London. The Bank would need to supply lots of money.....
> If a westerly was blowing, the Bank would mop up any excess money..., thereby avoiding inflation." ...Mervyn King, the current governor, told the FT in an interview .... (Winds of change by Chris Giles, Financial Times, May 14, 2007).

\(^{10}\)Even adherents of sticky-price models may not want to assume that prices are sticky relative to a trend or that they are sticky seasonally.
of money and higher return assets (see Zhu and Wallace [27]), and counterfeiting (see, for example, Hu [11]). As regards randomness of meetings, such randomness is simple and could, at a small cost in terms of additional structure, be replaced by heterogeneous taste shocks.

Another matter of concern is money holdings in \{0, 1\}. I have to concede that such holdings prejudice the result toward inflation in our first example. In that example, if \(n\)-people spend more than they earn in meetings with \(m\)-people, then money has to be returned to \(n\)-people as transfers to those who would otherwise have no money. Such transfers have harmful incentive effects on \(n\)-people who are producers in meetings. If, instead, money holdings were richer, say \{0, 1, 2, ..., \(B\)\} with \(B\) large, then such transfers could be paid in a way that approximates payment of interest on money held by \(n\)-people. Nevertheless, I am skeptical that an optimum in such a version would look like the Friedman rule. Spending by \(m\)-people serves several purposes in the model. Therefore, taxing them by having them earn more than they spend in meetings is far from costless in terms of welfare.\(^{11}\)

The assumption that an exogenous fraction are perfectly monitored is a special case of a model with a smooth distribution of costs of getting monitored across the population. In such a model, the planner chooses a cut-off cost subject to people self-selecting in accord with that cut-off. Some examples with such specifications were explored in Deviatov and Wallace [8]. It was found that the extreme version used above is not misleading. More interesting and challenging is a departure from the extreme situation of some perfectly monitored people and others not monitored at all. Some preliminary work on a model of that sort is in Mills [20].

7 Concluding remarks

In terms of the environments and the notion of implementability, the examples presented above are simple. However, the optima are complicated in two senses: they depend on all the details of the examples and even their qualitative properties—inflation in the first example and a larger money supply at the beginning of the low-activity season in the second example—are not obvious and may not be general. Two features of the model account for why the optima are complicated.

The above examples are such that the optima at both extremes—either all \(m\)-people or all \(n\)-people—are easy to describe. That is the case because neither extreme has an endogenous state variable. The version with only \(m\)-people is a simple repeated game. The version with only \(n\)-people is also a repeated game, because, with money holdings in \{0, 1\}, the fraction who end the first stage

\(^{11}\)A similar effect on welfare appears in Antonolfi et al [1]. They have a two-sector model in which \(m\)-people interact solely with each other in a Kehoe-Levine [18] credit market, while \(n\)-people interact solely with each other in a market with spot trade in money. As in our model, \(m\)-people face the threat of being banished into the set of \(n\)-people. In other respects, the models are very different. In our model, as highlighted above, the interactions in equilibrium between \(m\)-people and \(n\)-people are central to the results.
with money is unaffected by the first-stage trades. In contrast, a version with both types has an endogenous state variable; namely, the distribution of money holdings between the two types. In it, the trades at a date affect the current return (in the welfare function) and also help to determine the distribution at the next date which matters for future trades. Such a multiple role for trades and the impossibility of attaining the first-best, which are features of any general model of money, make the optimum complicated.

There are several attractive aspects of the model I have used: it is built up from fundamental ideas about the role of money; it fits well with the rest of economics; and it has endogenous taxation as opposed to having arbitrary assumptions about the allowable set of taxes. At a minimum, it should make users of New Keynesian models wonder why optima in flexible-price versions of their models are so different from optima in my model. One source of the difference is that their models eliminate potential endogenous state variables by having everyone in their models be identical or be a member of a so-called large family (see, for example, Gertler and Kiyotaki [9]). Another is that their agents seem able to commit to future actions, which is more extreme than having everyone in the model I have described be an \( m \)-person. My examples show that when those extreme aspects of their models are dropped, then interesting issues for optimal monetary and aggregative fiscal policy arise even without assuming sticky prices.

References


---

12 Another device for eliminating state variables is that used by Lagos and Wright [14] and its many offshoots. They assume periodic centralized trade with quasi-linear preferences.


