Interest on cash with endogenous fiscal policy*

Alexei Deviatov[†] and Neil Wallace[‡]

April 14, 2010

Abstract

We study optima in a model that has extreme versions of an above-ground economy and an underground economy. The former consists of people who are perfectly monitored, while the latter consists of people who are anonymous and not monitored at all. It follows that the underground economy uses currency and is not subject to direct taxation. Therefore, interest on currency must be financed by taxes on the above-ground economy. However, in most of the examples, it is optimal— according to an ex ante, representative-agent criterion—not to use taxes to raise the rate of return on currency.

1 Introduction

It is well-known that monetary policy and the welfare cost of inflation cannot be studied without some specification of allowable fiscal instruments. Phelps [15] made that point and subsequent work has only reinforced it. The equivalence results in Wallace [20] and Sargent and Smith [16] rest on complete markets and the assumption that lump-sum taxation is possible. Those in Correia *et. al.* [4] also depend on the availability of a rich set of fiscal instruments. The message we should take from this work is that progress depends on using a model in which feasible taxation is an *implication* of the model. It will be if money and monetary policy are studied as aspects of a *mechanism-design* analysis within a model that is explicit about the frictions that generate a role for money.

According to recent work in monetary theory, the crucial friction that generates a role for *money* is imperfect monitoring: incomplete knowledge of the previous actions of at least some people (see, for example, Ostroy [14], Townsend

*This is a preliminary and incomplete version. Please do not cite without permission of the authors.

†New Economic School, Moscow <deviatov@list.ru>.

‡Penn State <neilw@psu.edu>.

[18] and Kocherlakota [11]). That friction—and, perhaps, others—has consequences for feasible taxation. That the consequences may be important is suggested by even the most casual observations concerning the role of *currency* in actual economies.[1]

A pervasive observation is that currency is intensively used to help avoid taxes and in what is labeled the underground economy—the part of the economy in which, by definition, activities are somewhat hidden and, therefore, difficult to tax. Even if the underground economy is benign, as it is in the model we study, it would be surprising if the difficulty of taxing underground activities did not have important implications for the desirability of paying interest on currency.

We conduct our analysis within a variant of Cavalcanti and Wallace [3], a model that has extreme versions of an above-ground economy and an underground economy. The former consists of people who are perfectly monitored, while the latter consists of people who are anonymous and not monitored at all. It follows that the underground economy uses currency and is not subject to direct taxation. Therefore, if interest is to be paid on currency, then it must be financed by taxes on the above-ground economy. However, it turns out that such taxation is scarce in the model—so scarce in most of the examples we study that it is optimal, according to an ex ante, representative-agent criterion, to have a lower rate of return on currency than in the comparable economy that consists entirely of an underground economy.

We set up an economy with the following properties: if the cost of becoming monitored is prohibitive, then it would desirable to pay interest on cash but it is not feasible to do so (because the entire economy is an underground economy). Then we replace the prohibitive monitoring costs by a cross-section distribution of (lower) costs that makes it feasible for some fraction of the population to become monitored. For examples we compute the ex ante optimum for each economy—the one with prohibitive monitoring costs and the one with lower costs. We compare the optima according to a measure of the return on cash held by nonmonitored people. That comparson tells us whether the presence of monitored people, which makes it feasible to raise the return on cash, is used

[1] Although it is unfashionable to identify money with currency, we do so for the following reasons. Currency is the only *outside* component of what is usually labeled *money* in most economies and, therefore, is the only component that calls for a special policy regarding its return. Interest can and is paid on demand deposit balances in many countries and could be paid on cell-phone money and other stored value instruments like cash cards. Although interest rates on some of those instruments are low, that is easy to explain. In the U.S., taxes are due on explicit interest, but not on payments in kind that take the form of free check-writing services. Also, the Diamond-Dybvig model [7] *predicts* a low interest rate on demand deposits.

for that purpose. As noted above, for most of our examples, which include ones with an exogenous fraction of monitored people and ones with endogenous determination of that fraction, the presence of monitored people is not used to raise the return on cash.

Our model is closest in spirit to Antinolfi et al.[1]. They study an economy with two sectors: one is a credit economy modeled as in Kehoe and Levine [10]; the other is a currency economy modeled as in Bewley [2]. However, there is one major difference between their model and ours. In their model, in equilibrium there is no contact between the currency and credit sectors; in our model, the contact between monitored people and nonmonitored people is central to the results. While there is a version of our model without contact between them, in that version the possible *policies* are so limited that the model is uninteresting.

2 The model

Time is discrete with two stages at each date. There is a nonatomic unit measure of people who maximize expected discounted utility with discount factor $\beta \in (0,1)$. The first stage at each date has pairwise meetings and the second stage has a centralized meeting. Just prior to the first stage, a person looks forward to being a consumer who meets a random producer with probability $\frac{1}{K}$, looks forward to being a producer who meets a random consumer with probability $\frac{1}{K}$, and looks forward to no pairwise meeting with probability $1-\frac{2}{K}$, where $K \geq 2$. 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 $c(0)=0$. In addition, $y^* = \arg\max_y [u(y)-c(y)]$ exists and is positive. Production is perishable; it is either consumed or lost.[2] There is no utility associated with actions at the second stage.

People in the model are ex ante identical and make an initial and one-time choice between becoming monitored (an m person) or becoming nonmonitored (an n person). For m people, histories and money holdings are common knowledge; for n people, they are private. However, the monitored status and consumer-producer status (in a pairwise meeting) of each person are common knowledge. And, no one, except the planner, can commit to future actions.

People choose m or n status after receiving a private and independent draw

[2] This formulation is borrowed from Trejos-Wright [19] and Shi [17]. If K is an integer that exceeds two, then, as is well-known, it can be interpreted as the number of goods and specialization types in those models.

from a distribution of an additively separable one-time utility cost of becoming an m person. Let $F : \mathbb{R}_+ \to [0, 1]$, where $F(x)$ is the probability of having a utility cost of becoming monitored no greater than x. We treat F as both the distribution from which individual draws are made and as the realized distribution of costs over the population.

We study two classes of F functions. Let B be a cost so high that a person who realizes cost B would never choose to become an m person. One class of F functions is that used in earlier work (see Deviatov and Wallace [6]); namely,

$$F^e_\alpha(x) = \begin{cases} \alpha \text{ if } x < B \\ 1 \text{ if } x \geq B \end{cases} . \tag{1}$$

Here, α is the exogenous fraction who are m people. The other class has a smoother F. For it, if the allocation ends up with some m people and some n people, then there is an internal cut-off cost.

In order to allow for a discussion of inside money, people and the planner have printing presses capable of turning out identical, indivisible, and somewhat durable objects. Those turned out by the printing press of any one person are, however, distinguishable from those turned out by other peoples' printing presses.

Finally, each person's holding of money (issued by others) is restricted to be in $\{0, 1\}$—both at the start of stage 1 and at the start of stage 2. Because of the assumed restriction on individual money holdings, we assume that the planner can choose a probability with which money disintegrates between stage 1 and stage 2 at each date—whether money is inside money or outside money. As explained further below, that is the model's analogue of inflation.

3 Implementable allocations and the optimum problem

We limit the search for an optimum to allocations that are constant and symmetric. By symmetry, we mean that all people in the same situation take the same action, where that action could be a lottery. (That is, there is no randomization.) We also limit allocations to ones in which all monies issued by m people who have not defected and money issued by the planner are treated as perfect substitutes and in which all monies issued by n people are worthless. (Hence, we simply assume that n people do not issue money.)

The planner chooses m or n status as a function of a person's realized cost of becoming an m person (the person's draw from F); the fraction of m people with a unit of money and the fraction of n people with a unit of money; trades in stage-1 meetings (as functions of the states of the producer and the consumer); the disintegration-of-money probability; and stage 2 transfers.

The sequence of actions at the first date is as follows. People are ex ante identical and the planner's objective is maximization of ex ante expected utility, a representative-agent criterion. First, the planner's choice is announced. Then each person gets a private draw from the monitoring-cost distribution, F, and chooses m or n status, a choice that is observed. Then, initial money holdings are distributed coditional on m or n status. Then, the two stages occur at the first date and all subsequent dates.

The planner's choice is subject to self-selection constraints that follow from our specification of private information and of punishments. We assume that the only punishment is permanent banishment of an individual m-person to the set of n-people, which includes loss of the ability to issue money in the inside-money version of the model. Underlying this assumption is free exit at any time from the set of m-people and the ruling out of global punishments—like the shutting down of all trade in response to individual defections. We allow both individual and cooperative defection of those in a stage-1 meeting, but only individual defection from stage-2 transfers (because there are no static gains from trade at stage 2). The details follow.

3.1 Notation

Let $S = \{m, n\} \times \{0, 1\}$ be the set of individual states, where $s = (s_1, s_2)$ and $s' = (s'_1, s'_2)$ denote generic elements of S. We denote an action x in a stage-1 meeting by $x^{s,s'}$, where the first superscript is the state of the producer and the second is the state of the consumer. Our main notation is summarized in the following table.

Table 1. Notation

$y^{s,s'}$	production by s and consumption by s' in a meeting
$\lambda_p^{s,s'}(i), \lambda_c^{s,s'}(i)$	prob. that end-of-stage-1 money is i; p (c) for producer (consumer)
ξ	prob. that end-of-stage-1 money disintegrates prior to stage 2
$\phi^s(i)$	prob. that end-of-stage-2 money of s is i
θ^s	fraction in state $s \in S$ at the start of a date
v^s	discounted utility for $s \in S$ at the start of a date

The planner chooses the variables in the first five rows subject to the constraints set out below.

3.2 Feasibility and steady state conditions

The λ's and ϕ's must, of course, be lotteries on $\{0, 1\}$. Under outside money, money is not created in stage-1 meetings. Therefore, under outside money, if $s_2 + s_2^{'} = 0$, then $\lambda_j^{s,s'}(0) = 1$; while if $s_2 + s_2^{'} = 1$, then $\lambda_p^{s,s'}(1) \leq \lambda_c^{s,s'}(0)$ and $\lambda_p^{s,s'}(0) \geq \lambda_c^{s,s'}(1)$. Under inside money, money can be created by m people. Therefore, under inside money these constraints apply only when $s_1 = s_1^{'} = n$. Also, these constraints do not rule out disposal of money. This possibility is especially relevant in a meeting between an $(m, 1)$ producer and an $(n, 1)$ consumer. In this case, we do not require that $\lambda_p^{s,s'}(1) = \lambda_c^{s,s'}(1) = 1$. Instead, the allocation can require the consumer to give up money with some probability—even though this requires free disposal (on the part of the $(m, 1)$ producer).

Stage-1 and stage-2 actions imply the transition probabilities of a person's money holding from the start of one date to the start of the next date. The probability that a person in state $(s_1, i) \in S$ transits to state $(s_1, j) \in S$ is

$$t^{s_1}(i,j) = \frac{1}{K} \sum_{s' \in S} \theta\,(s')\,[(\lambda_p^{(s_1,i),s'} + \lambda_c^{s',(s_1,i)} + (K-2)\delta_i)\Psi\Phi^{s_1}(j)], \qquad (2)$$

where $\lambda_p^{s,s'} = (\lambda_p^{s,s'}(0), \lambda_p^{s,s'}(1))$, $\lambda_c^{s,s'} = (\lambda_c^{s,s'}(0), \lambda_c^{s,s'}(1))$, δ_i is the two-element unit vector in direction $i + 1$,

$$\Psi = \begin{bmatrix} 1 & 0 \\ \xi & 1-\xi \end{bmatrix}, \qquad (3)$$

and $\Phi^{s_1}(j) = (\phi^{(s_1,0)}(j), \phi^{(s_1,1)}(j))'$. If T^{s_1} denotes the 2×2 matrix whose (i, j)-th component is $[t^{s_1}(i, j)]$, then the stationarity requirements are

$$(\theta^{(s_1,0)}, \theta^{(s_1,1)})T^{s_1} = (\theta^{(s_1,0)}, \theta^{(s_1,1)}), \qquad (4)$$

for $s_1 \in \{m, n\}$.

3.3 Incentive constraints

It is convenient to first define discounted expected utility at the start of stage 1. For $s \in S$, we have

$$v^s = \frac{1}{K} \sum_{s' \in S} \theta^{s'}[\pi^p(s, s') + \pi^c(s', s) + (K-2)\pi^0(s)], \qquad (5)$$

where

$$\pi^p(s,s') = -c(y^{s,s'}) + \beta\lambda_p^{s,s'}\Psi\Phi^{s_1}v^{s_1}, \tag{6}$$

$$\pi^c(s,s') = u(y^{s,s'}) + \beta\lambda_c^{s,s'}\Psi\Phi^{s'_1}v^{s'_1}, \tag{7}$$

and

$$\pi^0(s) = \beta\delta_{s_2}\Psi\Phi^{s_1}v^{s_1}. \tag{8}$$

Here, δ_{s_2} is the 1×2 unit vector in direction $s_2 + 1$,

$$\Phi^{s_1} = \left[\begin{array}{c} \phi^{(s_1,0)}(0), \phi^{(s_1,0)}(1) \\ \phi^{(s_1,1)}(0), \phi^{(s_1,1)}(1) \end{array} \right], \tag{9}$$

and $v^{s_1} = (v^{(s_1,0)}, v^{(s_1,1)})'$. Given the variables in the first five rows of Table 1, Blackwell's sufficient conditions for contraction imply that v^s exists and is unique. We express the incentive constraints in terms of the v's. This is legitimate because the principle of one-shot deviations applies to this model.

There are truth-telling constraints only for n people with money when they are consumers. They are

$$\pi^c(s,(n,1)) \geq u(y^{s,(n,0)}) + \beta(\xi, 1-\xi)\Phi^n v^n. \tag{10}$$

This potentially binds only when $s_1 = m$, when the producer is an m person.

The individual rationality constraints for stage 1 meetings are

$$\pi^p((s_1,0),s') \geq \beta\phi^{(n,0)}v^n \text{ and } \pi^p((s_1,1),s') \geq \beta(\xi,1-\xi)\Phi^n v^n, \tag{11}$$

$$\pi^c(s,(s'_1,0)) \geq \beta\phi^{(n,0)}v^n \text{ and } \pi^c(s,(s'_1,1) \geq \beta(\xi,1-\xi)\Phi^n v^n, \tag{12}$$

and

$$\pi^0((s_1,0),s') \geq \beta\phi^{(n,0)}v^n \text{ and } \pi^0((s_1,1),s')\beta(\xi,1-\xi)\Phi^n v^n. \tag{13}$$

We also have a constraint which says that m people prefer the stage-2 transfers intended for them to defecting to n-status just prior to those transfers; namely,

$$\phi^{(m,s_2)}v^m \geq \phi^{(n,s_2)}v^n. \tag{14}$$

There is also a constraint that transfers to n people at stage 2 are nonnegative.

Next we consider cooperative defections for people in stage-1 meetings. We start with a definition of the pairwise core for meetings between n people who might trade.

Definition. For $(s, s') = ((n, 0), (n, 1))$, we say that (π^p, π^c) is in the pairwise core if it solves the problem,

$$\max_{y^{s,s'}, \lambda_p^{s,s'}, \lambda_c^{s,s'}} \gamma \pi^p(s, s') + (1 - \gamma)\pi^c(s, s') \tag{15}$$

subject to (11) and (12) for some $\gamma \in [0, 1]$.

The following lemma, the proof of which is routine and relegated to the appendix, fully characterizes this pairwise core.

Lemma 1 *Let $a = \beta\phi^{(n,0)} v^n$ (the lowest possible payoff for the producer in the above problem) and let b be the solution for $\pi^p((n, 0), (n, 1))$ to the above problem for $\gamma = 1$ (the highest possible payoff to the producer in the above problem). Let $\psi : [a, b] \to \mathbb{R}$ be defined by*

$$\psi(x) \equiv \begin{cases} \varsigma - x & \text{if } x \in [a,\ \rho) \\ u[c(\beta(\xi, 1 - \xi)\Phi^n v^n - x)] + a & \text{if } x \in [\rho,\ b] \end{cases}, \tag{16}$$

where

$$\rho = \beta(\xi, 1 - \xi)\Phi^n v^n - c(y^*), \tag{17}$$

and

$$\varsigma = u(y^*) - c(y^*) + \beta[\phi^{(n,0)} + (\xi, 1 - \xi)\Phi^n]v^n. \tag{18}$$

(Notice that ρ is π^p if the producer acquires money with probability 1 and produces the surplus maximizing output, while ς is the sum of π^p and π^c if that output is produced.) For $(s, s') = ((n, 0), (n, 1))$, (π^p, π^c) is in the pairwise core iff

$$(\pi^p, \pi^c) \in \{(x, \psi(x)) : x \in [a, b]\}. \tag{19}$$

Because defection by an m person converts the person to an n person, the payoffs in a stage-1 meeting in which one person is an m person or both are m people must satisfy

$$\pi^c(s, s') \geq \psi(\pi^p(s, s')). \tag{20}$$

Finally, we have the self-selection constraint for the initial choice of m and n status. Let

$$D = (\theta_c^{(m,0)}, \theta_c^{(m,1)})v^m - (\theta_c^{(n,0)}, \theta_c^{(n,1)})v^n, \tag{21}$$

where $\theta_c^{(s_1,s_2)} = \theta^{(s_1,s_2)}/(\theta^{(s_1,0)}+\theta^{(s_1,1)})$, the probability of being in state (s_1, s_2) conditional on being in state s_1. For $F = F_\alpha^e$, we require only that $0 \leq D \leq B$, where the first inequality is implied by the individual rationality constraints and the second holds by the choice of the parameter B. If $(\theta^{(m,0)} + \theta^{(m,1)}) \in (0,1)$ and $F^{-1}(\theta^{(m,0)} + \theta^{(m,1)})$ exists and is continuous at $(\theta^{(m,0)} + \theta^{(m,1)})$, then

$$F^{-1}(\theta^{(m,0)} + \theta^{(m,1)}) = D, \tag{22}$$

the usual cut-off property. If $(\theta^{(m,0)} + \theta^{(m,1)}) \in \{0,1\}$, then the obvious inequalities must hold.

3.4 The planner's problem

The planner chooses the variables in Table 1 to maximize

$$W = \sum_{s \in S} \theta^s v^s - \int_{y=0}^{y=D} y dF(y) \tag{23}$$

subject to all the relevant constraints.

There are two versions of this choice problem: one for outside money and one for inside money. Under outside money, no one except the planner issues money. Under the assumption that monies issued by different people are distinguishable, outside money satisfies all the constraints. In general, however, it leads to a worse outcome than inside money because under inside money, m people do not hold money and, therefore, have a defection payoff which is $v^{(n,0)}$. Under outside money, when m people hold money, their defection payoff is $v^{(n,1)}$, which is larger. We include the outside-money version because there is a government monopoly on currency-issue in most economies.[3]

4 Remarks on the model

There are three frictions in the model: the discount factor, the private cost of getting monitored, and money holdings in $\{0, 1\}$. The last is of particular concern because it seems to make it difficult to vary the return on money.

First, consider the discount factor. In this economy, the best ex ante outcome subject only to physical feasibility is production and consumption at each date

[3] See Wallace [21] for an attempt to rationalize outside money by dropping the assumption that inside monies are perfectly distinguishable by issuer.

in each stage-1 meeting equal to y^*, the output that maximizes *surplus* in a meeting. If everyone were an m person, then that best outcome would be implementable if

$$\frac{u(y^*)}{c(y^*)} \geq 1 + K(1-\beta)/\beta. \tag{24}$$

(This assures that a producer in a meeting weakly prefers producing y^* when others will do so in future meetings to permanent autarky.) Let β^* denote the β for which (24) holds at equality. Because we want to focus on the role of money, a role which arises only in the presence of n people, all our examples will have $\beta \geq \beta^*$. In other words, we assume that only the presence of n people prevents the first-best outcome from being attained in meetings.

Now suppose that there are only n people. Then, with $\{0, 1\}$ money holdings, trade occurs only in *trade* meetings—meetings in which the producer has no money and the consumer has money. The trade is some amount of production in exchange for a probability of a transfer of money from the consumer to the producer. The optimum with only n people and $\{0, 1\}$ money holdings is easy to describe. There is only one relevant constraint in this economy; namely,

$$\frac{u(y)}{c(y)} \geq 1 + \frac{K(1-\beta)/\beta}{1-\theta}, \tag{25}$$

where θ is the fraction with a unit of money and y is the amount produced in a trade meeting. (This says that a producer in a meeting weakly prefers producing y and acquiring money with probability 1—given that others will do that in future meetings—to permanent autarky.) Let $\bar{\beta}$ be the value of β for which (25) holds at equality when $y = y^*$ and $\theta = 1/2$. Obviously, $\bar{\beta} > \beta^*$.

As is well-known, if $\beta \geq \bar{\beta}$, then the optimum is $y = y^*$ and $\theta = 1/2$. If $\beta < \bar{\beta}$, then the optimum has $\theta < 1/2$, $y < y^*$, and (25) at equality. With only n people, having $\theta < 1/2$ is the only way to loosen constraint (25). It does so by reducing the expected number of periods during which money is held before a trading opportunity occurs and comes at the expense of a reduction in the fraction of *trade* meetings. Thus, $\beta < \bar{\beta}$ is the condition under which it would be desirable to pay interest on money if it were feasible and costless to do so. Because most economists believe that the world is such that it would be desirable to pay interest on money if it were feasible and costless to do so, we restrict attention to $\beta \leq \bar{\beta}$.

If there are n people and m people, then there is another way to raise the return on money for n people. A stationary allocation can have m people produce more per unit of money acquired than they consume per unit of money spent in meetings with n people. That is, stationarity does not require that trades be identical in $((m,i)(n,j))$ and $((n,i)(m,j))$ meetings, the former being meetings in which the m person is the producer and the latter being meetings in which the n person is the producer.

How misleading is the restriction to $\{0,1\}$ money holdings? For any specification of allowable money holdings, an n person must receive money (with positive probability) in order to produce. Given the randomness of earning and consumption opportunities, even if large individual holdings of money are allowed, n people who experience a string of consumption opportunities will run out of money, while those who experience a string of earnings opportunities will not want to expend much effort to earn more. Therefore, the assumption that money holdings are in $\{0,1\}$ gives rise to an exaggerated, but not misleading, version of the liquidity problem of n people. Under that restriction on money holdings, n people with money are so rich that they cannot be induced to produce.

But does the restriction to $\{0,1\}$ money holdings limit the possibility of varying the return on money? Suppose, instead, that money is divisible and individual holdings are unconstrained, while still considering stationary allocations, steady states. Suppose first that there are only n people. The inability to tax n people has nothing to do with allowable money holdings. Hence, with only n people, there is no possibility of paying interest on money. There is, however, the possibility of inflating. Nothing prevents positive transfers at each date—helicopter drops of money. As noted above, we emulate inflation by letting the planner choose a probability, denoted ξ, that money carried from one date to the next disintegrates. With only n people, stationarity requires that there is an offsetting probabilistic stage-2 transfer of money. In terms of incentive effects, both the disintegration and the transfer lower the return on acquiring money—just as does an inflation tax that comes about through lump-sum transfers when money is divisible and unbounded.[4]

Now suppose that there are m and n people. With divisible money, consider, without loss of generality, paying explicit interest on money in the form of money. Of course, the planner cannot be creating money to pay interest on it because that would be inflationary and (super) neutral. The interest must be financed by taxes and the taxes must be paid by m people. In addition, in order that average money holdings of n people remain constant, m people must be acquiring more money in their stage-1 trades with n people than they are spending, a net acquisition that must equal the interest payments on money.

[4] It follows that with $\{0,1\}$ money holdings and only n people, $\xi > 0$ is not optimal. That is not true with richer individual holdings of money.(see, for example, Levine [12], Deviatov [5], and Green and Zhou [9]).

Within such a scheme, what do n people see that induces them to be more willing to produce to acquire money? They do not see themselves getting richer relative to other people. And on average their trades with other n people do not give rise to a positive return from acquiring money. Therefore, there are only the same two possibilties that show up in the model with $\{0, 1\}$ money holdings: either the distribution of money is altered so that n people can more quickly spend money or n people get better trades on average when money is spent in a meeting with m people than when money is earned in a meeting with m people.

This leads us to conclude that while $\{0, 1\}$ money holdings exaccerbate the liquidity problems of n people, that specification does not prevent us from studying analogues of the policies we would study in a version with rich individual holdings of money.[5] Moreover, with both m and n people there may be a role for inflating, for $\xi > 0$. A positive ξ can contribute to satisfying the stationary condition which requires that the inflow and outflow from holding of money of n people be equated.

4.1 The rate of return on money for n people

Here we describe how an outside observer of our model would measure the average rate of return on money for n people. This (gross) rate of return is a weighted average over the different discounted returns faced by an $(n, 0)$ person in the different meetings in which the $(n, 0)$ person produces (in order to acquire money). There are three such meetings: with an $(n, 1)$ consumer, with an $(m, 1)$ consumer, and, under inside money, with an $(m, 0)$ consumer. For $s'' \in \{(n, 1), (m, 1), (m, 0)\}$, let

$$R(s'') = \frac{\lambda_p^{(n,0),s''}(1)(1-\xi)\beta\tilde{v}^{(n,1)}}{y^{(n,0),s''}}, \tag{26}$$

where $\tilde{v}^{(n,1)}$, which will be defined in a moment, is the expected discounted quantity of goods obtained when holding a unit money. (That is, the denominator is the quantity of goods surrendered in an $((n, 0), s'')$ meeting and the numerator is the discounted expected acquisition of goods.) Then the average return is

$$R = \frac{\sum_{s''} \theta^{s''} R(s'')}{\sum_{s''} \theta^{s''}}, \tag{27}$$

[5] There is, however, one substantial and simplifying consequence of $\{0, 1\}$ money holdings that we should mention. We allow n people to hide money, which gives rise to two-sided asymmetric information in a meeting between n people. However, with $\{0, 1\}$ money holdings, that asymmetric information plays no role because neither the producer nor the consumer in a trade meeting between n people sees any benefit from hiding money.

where the summations are over $s'' \in \{(n,1),(m,1),(m,0)\}$. Finally, $\tilde{v}^{(n,1)}$ is the (unique) solution to

$$\tilde{v}^{(n,1)} = \frac{1}{K}\sum_{s' \in S} \theta^{s'}[(K-1)(1-\xi)\phi^{(n,1)}(1)\beta\tilde{v}^{(n,1)} + \tilde{\pi}^c(s',(n,1))] \tag{28}$$

and

$$\tilde{\pi}^c(s',(n,1)) = y^{s',(n,1)} + \lambda_c^{s',(n,1)}(1)(1-\xi)\beta\tilde{v}^{(n,1)}. \tag{29}$$

We use the notation $\tilde{v}^{(n,1)}$ because $\tilde{v}^{(n,1)}$ is similar to $v^{(n,1)}$ (see (5)) except that $u(y)$ is replaced by y and $v^{(n,0)}$ is replaced by 0. Also, because future quantities of goods are discounted, we should think of R as being measured relative to β^{-1}. Thus, for example, if producing y implied consuming y/β with probability one at the next date (an outcome consistent with the Friedman rule), then we would find $R = 1$.

One comparison we make using R is to that for an economy with no monitored people—to the R in the optimum for an economy with $F(x) = 0$ for all $x < B$. That tells us whether the presence of m people is used to raise the return on money for n people.

5 Examples

In order to learn a bit about the properties of optima, we compute optima for some examples. Throughout, we assume $u(y) = 1 - e^{-10y}$, $c(y) = y$, and $K = 3$. (This form for u has a finite $u^{'}(0)$, which somewhat simplifies the programming.) This specification for u and c implies that $y^* = \ln 10/10$ and that $u(y^*)/c(y^*) = 9/\ln 10$. For the version with $F = F^e_\alpha$, for which the fraction of people with zero cost of becoming monitored is α (and for which the rest have a prohibitive cost B), we compute optima for a selection of (α, β). For this specification,

$$\beta^* = \frac{1}{1 + \frac{(9/\ln 10)-1}{3}} \approx 0.5077 \text{ and } \bar{\beta} = \frac{1}{1 + \frac{(9/\ln 10)-1}{6}} \approx 0.6735. \tag{30}$$

We report results for these two magnitudes of β and for $\beta = (\beta^* + \bar{\beta})/2$.

For the version with an endogenous set of m people, we choose F's so as to facilitate a comparison between the implied optima and those for a *comparable* F^e_α, an F that implies that α is the exogenous fraction who are monitored. Let $D(\alpha)$ denote the optimal D (see (21)) implied by F^e_α and the other parameters. Then, for given $\mu \geq 0$, let $F_{(\alpha,\mu)}(x)$ be given by $F_{(\alpha,\mu)}(B) = 1$ and

$$F_{(\alpha,\mu)}(x) = \begin{cases} 1 \text{ if } \alpha + \mu[x - D(\alpha)] > 1 \\ 0 \text{ if } \alpha + \mu[x - D(\alpha)] < 0 \\ \alpha + \mu[x - D(\alpha)] \text{ otherwise} \end{cases}. \tag{31}$$

for $x \in [0, B)$. This piecewise linear specification facilitates a comparsion between it and F_α^e in the following sense. For any μ, the optimal allocation for F_α^e is implementable for $F = F_{(\alpha,\mu)}$. Therefore, we can discuss the optimum for $F_{(\alpha,\mu)}$ in terms of how it deviates from the optimum for F_α^e. In particular, we focus on whether the former has more or fewer m-people than the latter. and whether it has a higher or lower rate of return on money holdings of n people.

6 Results

We start by describing a lower-bound benchmark in which everyone is treated as an n person. This is always a feasible choice for the planner (and is the optimum for an economy with prohibitive costs of becoming monitored, one with $F(x) = 0$ for all $x < B$). Table 2 describes this benchmark, which, obviously, does not depend on F.

Table 2. A benchmark lower bound: no m people

β	$\theta^{(n,1)}$	y/y^*	λ	R_0	W_0
β^*	0.38	0.55	1.00	0.18	0.09
$\frac{\beta^*+\bar{\beta}}{2}$	0.45	0.76	1.00	0.21	0.13
$\bar{\beta}$	0.51	1.00	1.00	0.26	0.17

These results are in accord with our qualitative discussion. In particular, both the fraction with money and output are increasing in β. Notice that the return on money is also increasing in β. That may be special to this example.

Below, we use this benchmark in two ways. We report magnitudes for welfare relative to W_0 and we report R relative to R_0.

6.1 Outside money

We begin with the model with $F = F_\alpha^e$, the model with an exogenous fraction who are monitored. The rate of return of money holdings of n people relative to R_0 is given in the following table.

Table 3. R/R_0 when $F = F_\alpha^e$

$\beta \setminus \alpha$	0	1/4	1/2	3/4
β^*	1	0.84	0.81	undefined
$\frac{\beta^*+\bar{\beta}}{2}$	1	0.91	0.88	undefined
$\bar{\beta}$	1	0.95	0.95	1.04

In only one of the cells do we find that the presence of m people is used to raise the return on money holding of m people. The optima when $\alpha = 3/4$ in the first two rows are quite special in a way that is not germane to the subject of this paper (see the appendix for a discussion of those optima). Thus, aside from that one cell, the interpretation is that tax revenue raised from taxes on m people has better uses than raising the return on money for n people.

To better understand that outcome, we describe the optima in detail for $(\alpha, \beta) = (1/4, \frac{\beta^* + \bar{\beta}}{2})$. We do this in two parts. Table 4 describes the aggregate features and Table 5 describes the trades in meetings.

Table 4. Aggregate features: $(\alpha, \beta) = (\frac{1}{4}, \frac{\beta^* + \bar{\beta}}{2})$

W/W_0	Ev^m/W_0	$v^{(n0)}/W_0$	$v^{(n1)}/W_0$	$\theta^{(m1)}$	$\theta^{(n0)}$	$\theta^{(n1)}$	ξ
1.43	3.20	0.36	2.39	0.25	0.57	0.18	0.16

Here W_0 is that in the second row in Table 2. Ex ante welfare is higher than in the benchmark as we expect it to be. However, those who turn out to be monitored do much better than those who turn out to be nonmonitored. Indeed, the latter do worse than in the allocation in the benchmark in which everyone is treated as an n person. (That is, $(\theta_c^{(n,0)}, \theta_c^{(n,1)})v^n = (.83)W_0$.)

The other entries are best discussed simultaneously with the description of trades in meetings that appears in the next table.

Table 5. Trades in meetings: $(\alpha, \beta) = (\frac{1}{4}, \frac{\beta^* + \bar{\beta}}{2})$

stage-1 meeting	$y^{(s,s')}/y^*$	$\lambda_s^{(s,s')}(1)$	$\lambda_{s'}^{(s,s')}(1)$
(n0)(n1)*†	0.573	1	0
(n0)(m1)*	0.573	1	-
(m1)(n0)	0.113	-	0
(m1)(n1)*†	0.381	-	0
(m1)(m1)*	0.381	-	-

In the table, a dash $(-)$ indicates that the variable is not identified. (In these cases, the planner has surplus instruments—state transitions in the meeting and transfers at stage 2.) In this table, a star $(^*)$ denotes a binding producer participation constraint in the meeting, while a dagger $(^\dagger)$ denotes a binding truth-telling constraint in the meeting. As described in Table 4, all m people start a period with money. (This is accomplished through transfers at stage 2 to those m people who spent money at stage 1 or who lost money through the 16% disintegration rate.) Although it is not correct to discuss constraints one at a time, two constraints seem crucial to understanding the outcomes. In meetings in which an $m1$ person is a producer and meets someone with money (the last two rows of Table 5), that $m1$ producer is on the verge of defecting (producing nothing and beginning the next period as an $n1$ person).

That constraint accounts for why output is so low in such meetings. The other constraint is the steady-state condition that the inflow into money holdings of n people is equal to the outflow. One way to reward m people is to have them spend when they are consumers and meet $n0$ producers. Indeed, in order to make such meetings more probable, there are more $n0$ people than in the benchmark. But that produces a large inflow into holdings of money by n people, which can only be partially offset by collecting money in the less frequent $(m1)(n1)$ meetings. That accounts for the 16% disintegration rate, our stand-in for inflation. The lower output in the third row of table 5 is accompanied by a binding truth-telling constraint for the $(n1)$ consumer in the $(m1)(n1)$ meeting.

Finally, we saw in Table 3 that the optimum in this case has a lower rate of return on money than in the benchmark when everyone is treated as an n person. Two things contribute to that lower return. One is the 16% disintegration rate; the other is that n people produce more (per unit of money earned) in $(n0)(m1)$ meetings than they receive (per unit of money spent) in $(m1)(n1)$ meetings. A partial offset to those determinants of the return on money is the higher fraction of meetings in which money can be spent—which, by itself, lowers the average length of time that money acquired is held.

To summarize, the optimum has all the m people with money—presumably because it is desirable that they be able to spend when they meet n producers. But that implies that any production by m people is bounded above by $\beta(v^{m1} - v^{n1})$—and, perhaps, can be interpreted as a tax. However, the size of that bound limits the scope for paying interest on money of n people through more favorable trades when n people are consumers than when they are producers. Thus, of the three possible ways to enhance the return on money held by n people—$\xi = 0$, more favorable trades when n people are consumers than when they are producers, and an altered distribution of money that lowers the average holding period of money—only the third is used at an optimum. The result is a lower return on money than in the benchmark.

Now we turn to the version with a smooth F function. For $(\alpha, \beta) = (\frac{1}{4}, \frac{\beta^* + \bar{\beta}}{2})$, we compute optima for $F = F_{(\alpha,\mu)}(x)$ as described in (31) for $\mu \in \{.2, .4, .6\}$.

Table 6. Aggregate features: $\beta = \frac{\beta^* + \bar{\beta}}{2}$, $F = F_{(\alpha,\mu)}(x)$

μ	W/W_0	Ev^m/W_0	Ev^n/W_0	$\theta^{(m0)}$	$\theta^{(m1)}$	$\theta^{(n0)}$	$\theta^{(n1)}$	R/R_0	ξ
0	1.43	3.20	0.83	0	.250	.574	0.176	0.909	.159
.2	1.35	3.16	0.85	0	.249	.574	0.178	0.909	.156
.4	1.28	3.12	0.86	0	.244	.575	0.181	0.911	.151
.6	1.21	3.06	0.88	0	.235	.579	0.186	0.915	.143

To facilitate the comparison with the model with an exogenous fraction who are monitored ($\mu = 0$), that result appears in the first row. Over this range, the higher is μ, the smaller fraction who choose to be monitored. That is, the planner is choosing to conserve on total one-time costs of becoming monitored. However, the effects are small. Moreover, it should be emphasized that this is not a general result.

6.2 Inside money

Under inside money, a very general proposition implies that the search for optima can be limited to allocations in which $\theta^{(m,1)} = 0$ (see Wallace [21]). **{Numerical results not yet available.}**

7 Concluding remarks

Our analysis is intended to be illustrative. After all, we study essentially one numerical example. And the model is very special in several respects—people are either perfectly monitored or anonymous, and money holdings are in $\{0, 1\}$. Despite that, we think that the results deserve to be taken more seriously than those in existing quantitative analyses that purport to offer recommendations for policy (see, for example, Dotsey and Ireland [8] and Lucas [13]).

First, to the extent that there is any theory of money in the quantitative work, it is that money is some sort of *intermediate* good: it is necessary for the activity of consuming some goods (cash-in advance constraints) or it reduces transaction costs. Agents in our model face cash-in-advance constraints, but only because there are people who are not monitored. And the presence of nonmonitored people has other implications—in particular, for feasible taxation. It is those other implications that are missing in the models that form the basis for existing quantitative work.

Second, consider again the pervasive observation we noted at the outset: currency, the only significant outside component of money, is intensively used in the underground economy. That observation plays no role in any of the quantitative analyses that are widely cited and meant to be estimates for the U.S. economy. But roughly half of U.S. currency is held abroad and a great deal is used in illegal activity. As a consequence, one goal of U.S. policy is to inhibit and control the use of currency through a variety of measures including the selection of a relatively small size for the largest denomination. Can it be that none of that is germane for choosing a rate of return on currency?

8 Appendix

8.1 Proof of the lemma

First we note that the preservation-of-money constraints for the $((n,0),(n,1))$ meeting imply that $\lambda_p^{s,s'} = (1-\lambda,\lambda)$ and $\lambda_c^{s,s'} = (\lambda, 1-\lambda)$ for some $\lambda \in [0,1]$, where λ is the probability that the producer acquires money. Also, because the objective in the definition (see (15)) is concave in y, output, and (linear in) λ and the constraint set is convex, the following first-order conditions are necessary and sufficient for that problem:

$$-c'(y)\left(\gamma+\sigma^s\right)+\left(1-\gamma+\sigma^{s'}\right)u'(y)\begin{cases} =0 \text{ if } y>0 \\ \leq 0 \text{ if } y=0 \end{cases}, \tag{32}$$

$$[(\gamma+\sigma^s)-\left(1-\gamma+\sigma^{s'}\right)]\beta(\varphi^{(n,1)}-\phi^{(n,0)})v^{(n,\cdot)}\begin{cases} \geq 0 \text{ if } \lambda=1 \\ =0 \text{ if } 0<\lambda<1 \\ \leq 0 \text{ if } \lambda=0 \end{cases}, \tag{33}$$

where $\varphi^{(n,1)} = (\xi, 1-\xi)\Phi^n$ and where $\sigma^s \geq 0$ is the Lagrange multiplier associated with (11) and $\sigma^{s'} \geq 0$ is that associated with (12). Notice that if the second of these holds at equality, then $y = y^*$. And if $\lambda = 1$ and the second holds with strict inequality, then $y < y^*$. We assume that $(\varphi^{(n,1)} - \phi^{(n,0)})v^n > 0$ (valued money) because otherwise only $y = 0$ satisfies (11). And valued money and $u^{'}(0) > c^{'}(0)$ imply that any solution to the pairwise core problem satisfies $y > 0$ and $\lambda > 0$. We provide a proof for the case, $a < \rho < b$, which says that $(y,\lambda) = (y^*, 1)$ is interior with respect to constraints (11) and (12). Cases in which $\rho < a$ or $\rho > b$ are similar.

Necessity. There are three cases: $\gamma \in \left[0,\frac{1}{2}\right)$, $\gamma = \frac{1}{2}$, and $\gamma \in \left(\frac{1}{2}, 1\right]$.

If $\gamma \in \left[0,\frac{1}{2}\right)$, then (33) implies that (11) binds. Therefore, $\pi^p = a$. Moreover, $a < \rho$ implies $y = y^*$ and $\lambda \in (0,1)$. The former implies $\pi^c = \varsigma - a = \psi(a)$, where, as noted above, ς is the sum of the payoffs implied by $y = y^*$.

If $\gamma = \frac{1}{2}$ and (11) is slack, then condition (32) implies that $y = y^*$. This yields,

$$\pi^p = -c(y^*) + \beta(\varphi^{(n,1)} - \phi^{(n,0)})v^{(n,\cdot)}\lambda + \beta\phi^{(n,0)}v^n \tag{34}$$

and

$$\pi^c = u(y^*) - \beta(\varphi^{(n,1)} - \phi^{(n,0)})v^n\lambda + \beta\varphi^{(n,1)}v^n = \varsigma - \pi^p = \psi(\pi^p).$$

If (11) is not slack, then we have $\pi^p = a$ as in the first case.

If $\gamma \in \left(\frac{1}{2}, 1\right]$, then the condition (33) and $\rho < b$ imply that $\lambda = 1$. Therefore,

$$\pi^p = -c(y) + \beta\varphi^{(n,1)}v^n \tag{35}$$

and

$$\pi^c = u(y) + \beta\phi^{(n,0)}v^n = u[c(\beta\varphi^{(n,1)}v^n - \pi^p)] + a = \psi(\pi^p),$$

where the second equality comes from substituting for y from (35).

Sufficiency. The proof proceeds by construction. In particular, for each $x \in [a,b]$, we show that the unique (y,λ) that supports $(x,\psi(x))$ and a proposed $(\sigma^s, \sigma^{s'}, \gamma)$ satisfy (32) and (33), which, as noted above, are sufficient to solve the problem that defines the pairwise core. We deal with two cases: $x \in [a,\rho]$, $x \in (\rho, b]$.

If $x \in [a,\rho]$, then, $y = y^*$ and

$$\lambda = \frac{x + c(y^*) - \beta\phi^{(n,0)}v^n}{\beta\left(\varphi^{(n,1)} - \phi^{(n,0)}\right)v^n} \in (0,1].$$

uniquely supports $(x,\psi(x))$. We propose $(\sigma^s, \sigma^{s'}, \gamma) = (0,0,1/2)$ for all such x. Then, (32) and (33) hold (at equality).

If $x \in (\rho, b]$, then $\lambda = 1$ and

$$y = c(\beta\varphi^{(n,1)}v^n - x) \in (0, y^*)$$

uniquely supports $(x,\psi(x))$. We propose $\sigma^s = \sigma^{s'} = 0$ and γ such that (32) holds. Notice that this $\gamma \in (1/2, 1)$. Then (32) holds at equality by the choice of γ and (33) holds (at strict inequality).

8.2 Results when the rate of return is not defined

{To be written}.

References

[1] Antinolfi G, C. Azariadis, J Bullard, The optimal inflation target in an economy with limited enforcement, Federal Reserve Bank of St. Louis Working Paper 2007-037B, November 2007.

[2] Bewley, T. 1980. The optimum quantity of money. In J. Kareken and N. Wallace, eds., *Models of Monetary Economies*, Federal Reserve Bank of Minneapolis.

[3] Cavalcanti, R. and N. Wallace, Inside and outside money as alternative media of exchange. *J. of Money, Banking, and Credit,* 31 (1999, part 2) 443-57.

[4] Correia, I., J. Nicolini and P. Teles, Optimal fiscal and monetary policy: equivalence results. *J. of Political Economy* 168 (2008) 141-170.

[5] Deviatov, A. Money creation in a random matching model. *Topics in Macroeconomics*, 6 No.3, (2006) Article 5.

[6] Deviatov, A. and N. Wallace, A model in which monetary policy is about money. *J. of Monetary Economics,* 56 (2009) 283-88.

[7] Diamond, D. and P. Dybvig, Bank runs, deposit insurance, and liquidity, *J. of Political Economy*, 91 (1983) 401-19.

[8] Dotsey, Michael, Peter Ireland (1996): "The Welfare Cost of Inflation in General Equilibrium," *Journal of Monetary Economics*, 37, 29-47.

[9] Green, E., and R Zhou. Money as a mechanism in a Bewley economy, *International Economic Review*, 46, 2005, pp. 351-371.

[10] Kehoe, T. and D. Levine, 1993. Debt-constrained asset markets. *Review of Economic Studies* 60: 868-888.

[11] Kocherlakota, N. Money is memory. *J. of Economic Theory* 81 (1998) 232-51.

[12] Levine, D., Asset trading mechanisms and expansionary policy. *J. of Economic Theory* 54 (1990) 148-64.

[13] Robert E. Lucas, Jr., Inflation and Welfare, *Econometrica*, Vol. 68, No. 2 (Mar., 2000), pp. 247-274.

[14] Ostroy, J., The informational efficiency of monetary exchange. *American Economic Review* 63 (1973) 597-610.

[15] Phelps, E. (1973), "Inflation in the Theory of Public Finance," *Swedish Journal of Economics* vol. 75, no. 1, January-March 1973, pp. 37-54.

[16] Sargent, T., and B. Smith, Irrelevance of open market operations in some economies with government currency being dominated in rate of return, *American Economic Review* 77 (1987) 78-92.

[17] Shi, S., Money and prices: a model of search and bargaining. *J. of Economic Theory* 67 (1995) 467-98.

[18] Townsend, R.M., Currency and credit in a private information economy. *J. of Political Economy* 97 (1989) 1323-44.

[19] Trejos, A. and R. Wright, Search, bargaining, money and prices. *J. of Political Economy* **103** (1995), 118-41.

[20] Wallace, N., A Modigliani-Miller theorem for open-market operations. *American Economic Review*, 71 (1981) 267-74.

[21] Wallace, N., The mechanism design approach to monetary theory. Manuscript, 2010.