ANNALS OF ECONOMICS AND FINANCE **19-1**, 137–150 (2018)

# Optimal Inflation in a Model of Inside Money: A Further Result*

Wataru Nozawa and Hoonsik Yang†

We extend the Deviatov and Wallace (2014) model of inside money in which they find some examples where inflation is beneficial. Their model was restrictive in that it could not address policies that provide interests on cash due to the small upper bound on money holdings. With a higher upper bound on money holdings, such policies can be engineered without inflation, and it is uncertain whether inflation is necessary for the optima. We investigate this possibility and confirm their results in a more generalized setting. At the optima, interest on cash is not provided, and positive inflation arises in a similar manner to Deviatov and Wallace (2014).

*Key Words*: Friedman rule; Inside-money; Inflation; Monitoring; Optima.
*JEL Classification Numbers*: E31, E42, E50, E52.

## 1. INTRODUCTION

There are several papers that have shown that inflation is a feature of desirable allocations (Kehoe et al. (1992), Molico (2006), Green and Zhou (2005), and Deviatov (2006) to name a few). The optimality of inflation attained in those pure-currency economies (Lucas (1980); Wallace (2014)) relies on two properties of them: 1) lump-sum transfer is beneficial as they provide risk-sharing, 2) there are no available tax instruments other than inflation tax. Wallace (2014) takes stock of this literature and conjectures that positive inflation caused by some transfers which improve risk-sharing or raise the return on money as the Friedman rule is generically optimal in pure-currency economies.

If we go beyond a pure-currency economy, which is extreme, the necessity of inflation is less clear as government transfers can be financed through some other taxes. We argue that inflation may still be optimal but for

* We are grateful to Neil Wallace for his valuable comments and support.

† Nozawa: Department of Engineering, Kyushu University, Fukuoka 819-0395, Japan. Email: nozawa@doc.kyushu-u.ac.jp; Yang: Corresponding author. Department of Economics and Trade, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea. Email: hoonsik.yang@sejong.ac.kr.



1529-7373/2018
All rights of reproduction in any form reserved.

a different reason. We build on Deviatov and Wallace (2014) where they study a model of inside money where there are two types of people in the model, who are subject to public monitoring and can issue inside money (monitored agents), and who are anonymous and cannot issue inside money (non-monitored agents). They find that optima may have positive inflation as it enables monitored agents can spend more than what they earn from non-monitored agents.[1] However, their model is restrictive in that some potentially beneficial policies that are not inflationary policies are physically ruled out due to small upper bound on money holdings. While some taxes are feasible, the small upper bound severely limits the ways to use tax revenue. In particular, their setting does not give scope to the Friedman rule-like policies (providing interest on cash holdings). Hence, one may wonder if positive inflation remains optimal when such policies were feasible.

We extend their model by adopting the smallest upper bound that makes the Friedman rule-like policies become feasible, which is three. This extension provides a rich setting in which we can study whether the Friedman rule-like policy is desirable and whether inflation is necessary. We define a mechanism design problem and study optima by computing those for a range of parameters on the discount factor, the finite marginal utility of consumption at zero, and the fraction of monitored agents in population. This is an additional contribution compared to Deviatov and Wallace (2014), who concentrate on one example.

In all examples we study, positive inflation arises at optimum. While direct monetary transfer from the government is very rarely used (in contrast to optima in pure-currency economies) inflation arises due to net positive creation of inside money from monitored agents. Hence, giving scope to the Friedman rule do not necessarily invalidate the result of Deviatov and Wallace (2014). Optima in economies without monitored agents (outside money economy) and economies with monitored agents (inside money economy) show sharp differences.

## 2. ENVIRONMENT

The environment is borrowed from Deviatov and Wallace (2014), which uses a random matching model in Shi (1995) and Trejos and Wright (1995) as the background. This model is identical to the model in Cavalcanti and Wallace (1999) except the set of money holdings and the way that inflation is modeled. Time is discrete and the horizon is infinite. There is a nonatomic measure of infinitely-lived agents.

[1]See Antinolfi et al. (2016) for a related study.

In each period, pairwise meetings for production and consumption occur in the following way. An agent becomes a producer and meets a random consumer with probability $1/K$, becomes a consumer and meets a random producer with probability $1/K$, or becomes inactive and enters no meeting with probability $1 - 2/K$. In a meeting, the producer can produce $q$ units of a consumption good for the consumer in the meeting at the cost of disutility $c(q)$, where $c$ is strictly increasing, convex, and differentiable and $c(0) = 0$. The consumer obtains period utility $u(q)$, where $u$ is strictly increasing, strictly concave, differentiable on $\mathbb{R}_+$ and satisfies $u(0) = 0$. Also, $q^* = \arg\max_{q \in \mathbb{R}_+}[u(q) - c(q)]$ exists and is strictly positive. The consumption good is perishable: it must be consumed in a meeting or discarded. Agents maximize the expected sum of discounted period utility with discount factor $\beta \in (0, 1)$.[2]

Each agent and the planner have printing presses that can produce an intrinsically useless and indivisible asset which can be used as a medium of exchange. We call the asset money. As monitoring is imperfect as described soon, money can potentially help the economy to achieve a good allocation. Money produced by any agent is distinguishable from that produced by other agents. Individual money holdings are restricted to be in $\{0, 1, 2, 3\}$. While we can consider an economy where the only medium of exchange is outside money when monitored agents exit, we only focus on inside money economy as it is welfare superior. (Deviatov and Wallace (2014))

Monitoring is imperfect in the language of Cavalcanti and Wallace (1999). A fraction of agents are permanently monitored and called $m$-agents, and the rest are permanently not monitored and called $n$-agents. Histories of $m$-agents, who they meet and what they did, are common knowledge, while those of $n$-agents are private. The fraction of $m$-agents is a parameter of the environment and denoted by $\alpha$. Also, agents cannot commit to any future action.

The sequence of actions is as follows. First, a pairwise meeting takes place and people realize their type as a producer or a consumer or become inactive. Depending on their shock realization, people can produce or consume, and period utility is determined here. All information, regarding their type and money holding, is common knowledge between them. Following the mechanism design approach, we will consider broader trading protocol than using a bargaining solution. Given a trading protocol, we assume that people can individually defect and the pair in any meeting can jointly defect to it. When $n$-agents defect, the government cannot punish them consistent to the definition of $n$-agents. When $m$-agents defect, they can be treated as $n$-agents as a punishment. This is a loss as they

[2]This exposition is a simplified version of Shi (1995); Trejos and Wright (1995).

lose their ability to issue valued money. We also allow that $m$-agents can become $n$-agents whenever they want to.

After the pairwise meeting, the government makes monetary transfer and tax. At this point, people can only defect individually. Money holdings of $n$-agents are not visible, and government can ask them to differ the transfer rate depending on their money holdings. Let $\tau_i \in [0, 1]$ be the probability that an agent with $i$ units of money receives another unit for each $i$. In response, people can report any units less than what they have and that imposes that $\tau_i$ must be weakly increasing in $i$. Also, we must assume $\tau_B = 0$. For $m$-agents, all money holdings of them are returned to government and eliminated at this point. After transfer and tax, each unit of money disintegrates with probability $\delta \in [0, 1]$, independent to their money holdings. This is a device to resemble inflation in this class of model, to make agents' money holdings stay in the given set of possible money holdings (Li (1995)).

All our compuations are for $K = 3$; $c(q) = q$, and $u(q) = 1 - e^{-\kappa q}$, which implies that $u'(0) = \kappa$. We study optima for a subset of

$$(\beta, \kappa) \in \{0.1, 0.2, \ldots, 0.8\} \times \{10, 20, 40\}$$

a subset that satisfies

$$\kappa > 1 + \frac{K(1-\beta)}{\beta}$$

This condition is necessary and sufficient for the production of constant positive output with $\alpha = 1$ (perfect monitoring), and it leaves us 21 pairs of $(\beta, \kappa)$. Regarding $\alpha$, we study $\alpha \in \{0, 0.25, 0.5, 0.75\}$. When $\alpha = 0$, this model becomes a version of Shi (1995) and Trejos and Wright (1995), with a richer money holding.

## 3. IMPLEMENTABLE ALLOCATIONS

We restrict our attention to stationary and symmetric allocations[3]: Money holding distribution is invariant to time and people act identically given that they have same monitoring type, money holdings, and shock realization. Therefore, productions and monetary payments are constant over all meetings in which a producer has $k$ units of money and a consumer has $k'$ units of money, a $(k, k')$ meeting. A stationary and symmetric allocation consists of choices for the variables listed in Table 1:

The planner chooses production and payments in meetings, disintegration and transfer rates to maximize ex-ante expected utility before money

[3]This may not be innocuous. In a similar model, Bertolai et al. (2012) finds that there can be ex-ante Pareto superior nonstationary allocations to the stationary optima.

**TABLE 1.**

Variables that constitute an allocation

| | |
|---|---|
| $\pi_k$ | fraction with $k$ units of money before meetings |
| $q(k,k')$ | production in $(k,k')$ meeting |
| $\lambda_p^{k,k'}(i)$ | probability that producer has $i$ money after $(k,k')$ meeting |
| $\lambda_c^{k,k'}(i)$ | probability that consumer has $i$ money after $(k,k')$ meeting |
| $\tau_k$ | transfer rate for agents with $k$ units of money |
| $\delta$ | probability that money disintegrates after meetings |

are assigned. It can be easily shown that ex-ante expected utility is proportional to

$$\sum_{0\leq k\leq B}\sum_{0\leq k'\leq B}\pi_k\pi_{k'}\left[u(q(k,k'))-q(k,k')\right], \tag{1}$$

the expected gains from trade in meetings. The choice is subject to the following constraints.

We define the first-best allocation as production and consumption equal to $q^*$ in every single-coincidence meeting. Accordingly, the first-best welfare level is $u(q^*)-q^*$.

### 3.1. Physical feasibility and stationarity

First, money holdings that result from meetings must be feasible: in $(k,k')$ meeting, if the consumer has $i$ units, then the producer must have $k+k'-i$ units. Also, money holdings cannot be negative or exceed the total amount brought into the meeting.

$$\lambda_c^{k,k'}(i)=\lambda_p^{k,k'}(k+k'-i) \text{ if } 0\leq i\leq k+k' \tag{2}$$

$$\lambda_c^{k,k'}(i)=\lambda_p^{k,k'}(i)=0 \text{ if } i<0 \text{ or } k+k'<i \tag{3}$$

Let $\Lambda(k,k')$ denote the set of pairs of probabilities $(\lambda_c^{k,k'},\lambda_p^{k,k'})$ that satisfy the above constraints.

Given this, the transition probability that a person in state $k$ transits to state $k'$ during pairwise trade meeting stage is

$$t^{(1)}\left(k,k'\right)=$$
$$\begin{cases}\frac{1}{K}\sum_{k_0\in N\cup\{m\}}\pi_{k_0}\left[\lambda_p^{k,k_0}\left(k'\right)+\lambda_c^{k_0,k}\left(k'\right)\right]+\frac{K-2}{K}1_{k=k'} & \text{if } k,k'\in N\\ 1 & \text{if } k'=k=m\\ 0 & \text{otherwise}\end{cases}$$

where $N=\{0,\ldots,B\}$.

The transition probability that a person in state $k$ transits to state $k'$ caused by transfer is denoted by

$$t^{(2)}(k,k') = \begin{cases} 1-\tau_k & \text{if } k \in \{0,1,2\} \text{ and } k' = k, \\ \tau_k & \text{if } k \in \{0,1,2\} \text{ and } k' = k+1, \\ 1 & \text{if } k' = k = 3 \text{ or } k' = k = m, \\ 0 & \text{otherwise,} \end{cases}$$

and the transition caused by disintegration is expressed by

$$t^{(3)}(k,k') = \begin{cases} \left( \begin{array}{c} k \\ k' \end{array} \right) \delta^{k-k'}(1-\delta)^{k'} & \text{if } k \in N \text{ and } k \geq k', \\ 1 & \text{if } k = k' = m, \\ 0 & \text{otherwise.} \end{cases}$$

Denote $T^{(i)}$ the following matrix

$$T^{(i)} \equiv \begin{bmatrix} t^{(i)}(0,0) & t^{(i)}(0,1) & t^{(i)}(0,2) & t^{(i)}(0,3) & t^{(i)}(0,m) \\ t^{(i)}(1,0) & t^{(i)}(1,1) & t^{(i)}(1,2) & t^{(i)}(1,3) & t^{(i)}(1,m) \\ t^{(i)}(2,0) & t^{(i)}(2,1) & t^{(i)}(2,2) & t^{(i)}(2,3) & t^{(i)}(2,m) \\ t^{(i)}(3,0) & t^{(i)}(3,1) & t^{(i)}(3,2) & t^{(i)}(3,3) & t^{(i)}(3,m) \\ t^{(i)}(m,0) & t^{(i)}(m,1) & t^{(i)}(m,2) & t^{(i)}(m,3) & t^{(i)}(m,m) \end{bmatrix}.$$

The stationarity constraint can be stated as

$$\pi = \pi T^{(1)} T^{(2)} T^{(3)}, \tag{4}$$

where

$$\pi = \left[ \begin{array}{ccc} \pi_0 & \cdots & \pi_B, \pi_m \end{array} \right]$$

and $\pi_m = \alpha$.

### 3.2. Incentive compatibility

Incentive compatibility is defined by underlying information and defection assumptions in two stages for transfer and pairwise meeting. Information assumption is on the history of agents and their money holdings, and defection assumptions are whether they can defect jointly or not. As we mentioned already, the history of agents is determined by their monitoring status. Hence the information specification only determines the visibility of money holdings. We assume that (1) in transfer stage, money holdings are private information and people can only defect individually, (2) in pairwise meeting stage, money holdings are visible within each meeting and people can jointly defect within a pair. Although money holdings are private information in transfer stage, they can only hide not overstate, as it can be verified. Hence, the truth-telling constraint on transfer rates is

$$\tau_{i-1} \leq \tau_i$$

for $i \in \{0, 1, 2\}$.

This is one possible specification out of 16, which is the number of combinations of information and defection specifications in each stage. Some of other specifications can be studied in a similar manner. We think it is desirable to allow joint defection within a meeting and hide money holdings in transfer stage, as it seems natural in this pairwise meeting and anonymous environment. If we want to allow joint defection in transfer stage (with visible money holdings), transfer rate must be restricted to linear in money holdings, and this can be easily incorporated. Other than that, our choice of the specification is not more justified than others.

Two small differences from Deviatov and Wallace (2014) are notable. In Deviatov and Wallace (2014), $n$-agents can hide their money in a meeting, and this potentially matters only when $n$ consumers meet $m$ producers. In ours, allowing $n$-agents to hide their money holdings create more complications. Hence, we choose different specification on that. Next, we allow $m$ agents can issue money to $n$ agents even when $m$ producer meets $n$ consumer.[4] This is potentially useful for the government to engineer the transfer, as information in a pairwise meeting is visible while it is not in the transfer stage. This is not beneficial in Deviatov and Wallace (2014) for the same reason why the transfer is not used at optimum. However, in our settings, it potentially is.

Due to stationarity, we omit all time scripts. Following definitions are made under stationarity. Before defining incentive compatibility, it is convenient to define discounted expected utility before meetings. For each money holdings $k \in N$,

$$\begin{aligned} v(k) = \frac{1}{K} \sum_{k' \in N \cup \{m\}} \pi_{k'} \left[ u(q(k', k)) + \beta \sum_{k_0} \lambda_c^{k',k}(k_0) w(k_0) \right] \\ + \frac{1}{K} \sum_{k' \in N \cup \{m\}} \pi_{k'} \left[ -q(k, k') + \beta \sum_{k_0} \lambda_p^{k,k'}(k_0) w(k_0) \right] + \frac{K-2}{K} \beta w(k), \end{aligned}$$

[4]We do not allow $m$ agents to give money to $n$ agents in inactive meetings to reduce number of variables for the computation. To some degree, money-giving by $m$ agents in both active and inactive meetings can be replicated by money-giving by $m$ producers in active meetings. We doubt that money-giving by $m$ agents outside of such range will be used in the optima.

and, for $k = m$,

$$\begin{aligned} v(m) &= \frac{1}{K} \sum_{k' \in N \cup \{m\}} \pi_{k'} \left[ u(q(k', m)) + \beta w(m) \right] \\ &+ \frac{1}{K} \sum_{k' \in N \cup \{m\}} \pi_{k'} \left[ -q(m, k') + \beta w(m) \right] \\ &+ \frac{K-2}{K} \beta w(m). \end{aligned}$$

where $w(\cdot)$ is discounted expected utility before transfer. Under stationarity,

$$w(k) = \sum_{k' \in N} t^{(2)}(k, k') \sum_{i \in N} t^{(3)}(k', i) v(i)$$

for $k \in N$. And for $k = m$,

$$w(m) = v(m)$$

Note that agents with 3 units of money cannot earn more. We also require that monitored agents prefer to stay monitored, rather than becoming non-monitored agents with money they earned in pairwise meeting.

$$w(m) \geq w\left( \max_{k \in N} \{x(m, k)\} \right) \tag{5}$$

where $x(m, k)$ is the amount that $m$ producer receives from a meeting with $n$ consumer with $k$ units of money.

As discussed in Deviatov and Wallace (2014), the monitored agents start a new period with no money at optimum. The money they earned in previous periods are entirely taxed, and they issue money under their name when they need to pay. It is easy to see why it creates a tighter incentive compatibility than (5) if they don't, as $m$ agents can obtain more money than one can get in one meeting.

We also require that trade results in pairwise core to immune to individual defection and joint defection in a pair. If $n$ agents defect to a given trade, it results in no trade in the current period and expecting $w(k)$ next period. If $m$ agents defect, they lose their status and become $n$ agents. To state the constraint for $(k, k')$ meeting for , let $\vartheta(k, k')$ denote a surplus (over no-trade) for a producer in the meeting. The constraint can be stated as follows: $q(k,k')$, $\lambda_p^{k,k'}$, and $\lambda_c^{k,k'}$ solve

$$\begin{aligned} \max_{q\geq 0,\ (\lambda_p,\lambda_c)\in\Lambda(k,k')} & u(q)+\beta\sum_{0\leq i\leq k+k'}\lambda_c(i)w(i) \\ \text{s.t.}\ & -q+\beta\sum_{0\leq i\leq k+k'}\lambda_p(i)w(i)=\beta w(k)+\vartheta(k,k') \\ & u(q)+\beta\sum_{0\leq i\leq k+k'}\lambda_c(i)w(i)\geq \beta w(k') \end{aligned} \tag{6}$$

or some $\vartheta(k,k') \geq 0$.[5] The Karush-Kuhn-Tucker condition is necessary and sufficient for the optimality, and we can derive a set of equations and inequalities from the condition. If either a consumer or a producer is $m$ agent, individual rationality suffices

$$u\left(q(k',k)\right)+\beta\sum_{k_0}\lambda_c^{k',k}(k_0)w(k_0)\geq\beta w\left(k\right),\ \text{for } k\in N\cup\{m\},\ k'=m \tag{7}$$

$$u\left(q(k',m)\right)+\beta w\left(m\right)\geq\beta w\left(0\right),\ \text{for } k'\in N\cup\{m\} \tag{8}$$

$$-q(k,k')+\beta\sum_{k_0}\lambda_p^{k,k'}(k_0)w\left(k_0\right)\geq\beta w\left(k\right),\ \text{for } k\in N\cup\{m\},\ k'=m \tag{9}$$

$$-q(m,k')+\beta w\left(m\right)\geq\beta w\left(0\right),\ \text{for } k'\in N\cup\{m\} \tag{10}$$

The planner maximizes the ex-ante expected utility (1), subject to the physical feasibility conditions, the stationarity conditions, the pairwise core constraints, and the individual rationality constraints.

## 4. COMPUTATION PROCEDURE

We compute solutions for the planner's problem using two solvers that are compatible with the GAMS interface, KNITRO and BARON. KNITRO

[5]Solving the problem is necessary for trades being in the pairwise core. That is also sufficient if the utility function of the producer and the consumer are strictly monotone in consumption goods and money holdings (see, for example, Mas-Collel (1995)). Here, the utility function may not be strictly increasing in money holdings; Some additional units of money may not be valued in some allocations, and hence the value function $w$, which specifies the preference for money holdings in trade meetings, may be non-strictly increasing in a part of the domain. In effect, we are solving a relaxed problem using this formulation. For example, if we find that a numerical solution has non-strictly increasing $w$, it may not be an optimum as solving above problem is not a sufficient condition in that case. It is verified that numerical solutions have strictly increasing $w$, and thus it is assured that the solutions solve the problem of our interest.

is a local solver for large-scale optimization problems. For a given initial point, it quickly converges to a local solution (or shows that it cannot reach one), but it does not guarantee global optimality. This issue is usually dealt with by using a large number of initial values. The solver automatically feeds in different initial values as we change an option that controls the number of initial values. In contrast, BARON (Branch-And-Reduce Optimization Navigator) is a global solver for nonconvex optimization problems. It continues to update an upper bound and a lower bound on the objective by evaluating the values of variables satisfying the constraints and stops when the difference between the two bounds becomes smaller than a threshold. It guarantees global optimality under mild conditions, but it generally takes much longer time to converge than local solvers. Even before it converges, we can terminate it and see its candidate solution. When Baron did not finish in a reasonable time span, we stopped it and checked the candidate solution with the solution from KNITRO.

Due to the complexity of the problem, BARON did not complete its computation after spending a day, which is the limit of time we can use for a single routine. In all examples that we tried, the candidate solution was not updated after roughly 12 hours. The remaining time was being used to verify that other feasible allocations are not better than the candidate solution. The intermediate result earned after one day run is our first candidate of the solution. We ran KNITRO with 1000 initial points and found that its solution coincides with the intermediate output from BARON, which is the best lower bound. With KNITRO, we used 3000 initial starting points and checked if the results from using 5000 initial points are the same. This becomes our second candidate when they coincide. We compare two candidates, the candidates from two solvers, and determine that results are consistent when maximized values and aggregate variables are same within the tolerance level, which is 0.001. The following results meet this criterion.

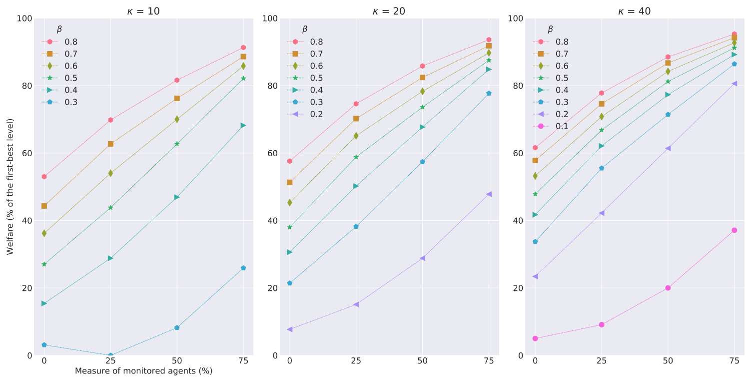
## 5. RESULTS

The first two figures show the welfare (relative to the first-best level) and money supply at the optima for different values of $\beta, \kappa, \alpha$. The money supply is defined as

$$\frac{\pi_1 + 2\pi_2 + 3\pi_3}{1-\alpha}$$

The welfare increases with $\beta$ as patient producers can endure more production, and it increases with $\alpha$ as the economy has the higher monitoring capacity. One exception is $(\beta, \kappa) = (0.3, 10)$, in the first panel where the welfare decreases as $\alpha$ increases from 0 to 0.25. The welfare at $\alpha = 0.25$ is close to zero, suggesting that implementable allocations are very close to an allocation with no production. While Wallace (2010) shows that welfare is weakly increasing in $\alpha$, the proof of this claim is relying on social planner treating some monitored agents as non-monitored agents. For simplicity, we require the expected utility of all monitored agents to be equal, and higher than that of any non-monitored agents'. In this case, welfare is not guaranteed to be weakly increasing in $\alpha$, as we can see in this result. This example shows that social planner may actually want to treat some monitored agents as non-monitored agents, and it might be better to use less than full monitoring capacity.

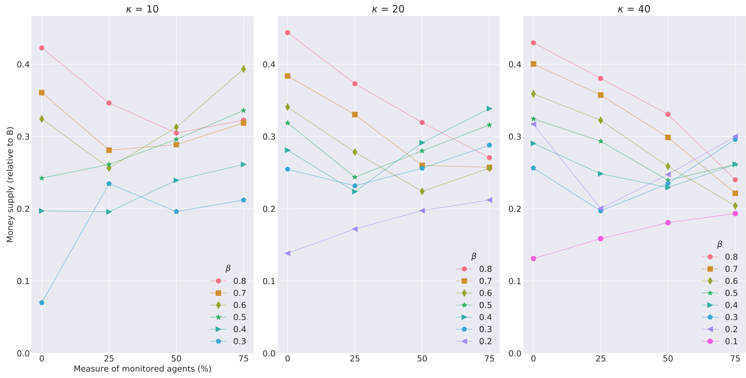
**FIG. 1.** Welfare




Broadly, money supply decreases with $\alpha$ when $\beta$ is high, and it increases $\alpha$ when $\beta$ is low. For intermediate values of $\beta$, money supply decreases with initially and then increases with $\alpha$. In sum, money supply does not change monotonically with $\alpha$, but it is less than $B/2$ in all examples.

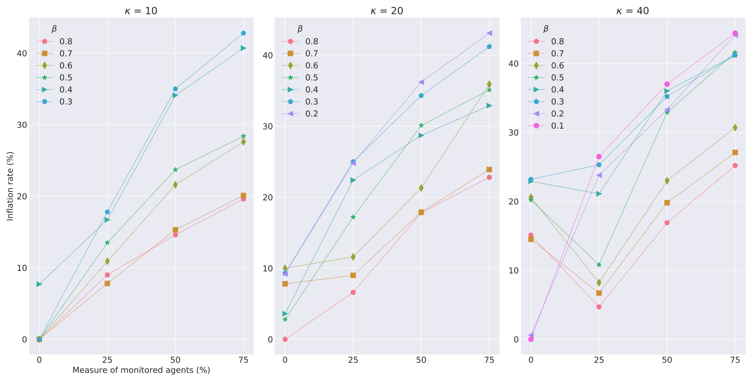
Figure 3 shows the inflation rate at optima. It is strictly positive in all examples with $\alpha > 0$, while transfer rate is positive only in one of them. In this model, both social planner and monitored agents can create money, and that will increase the money supply and lead to inflation. In examples where the inflation rate is positive and transfer rates are zero, inflation is caused by net money creation of monitored agents, which is the amount of money spent by monitored to non-monitored agents net of the amount of money earned by monitored from non-monitored agents. Hence, the result is consistent with Deviatov and Wallace (2014). Even though we extend their model to incorporate policies that are potentially beneficial but not inflationary, it is still not beneficial to tax monitored agents. Even in the example with positive transfer rates, inflation may not be solely caused by

**FIG. 2.** Money supply




government transfer. It can be both government transfer and net money creation of monitored agents that cause inflation.

**FIG. 3.** Inflation Rate




To understand this result, it is helpful to discuss the limit on the role of policies, generated by the one-unit upper bound on money holdings in Deviatov and Wallace (2014). Taxing monitored agentsby making monitored agents spend less money to non-monitored agents than what they earn from non-monitored agentsis feasible even under the one-unit upper bound (given that it is incentive compatible). However, it cannot be beneficial for welfare with the one-unit upper bound, as transfer scheme is too limited. With the one-unit upper bound, who can receive additional units of money through transfers are only people with no money. Consequently, transfer schemes always shift wealth from the rich to the poor, improving risk-sharing. There is no scope for transfer schemes shifting wealth from the poor to the rich and relaxing participation constraints of some producers. With the three-unit upper bounds as in here, taxing monitored agents can be beneficial. The social planner can not only improve risk-sharing through lump-sum transfers, but also relax participation constraints of some producers through policies resembling the Friedman rule. The cost of inflation resulting from such transfers can be reduced by taxing monitored agents. However, it will make monitored agents to consume less and decrease their utility. At the optima, net issue of money from monitored to non-monitored agents is positive, as in the example of Deviatov and Wallace (2014).

Compared with the counterpart pure-currency economies ($\alpha = 0$), transfer is used ($\tau_i > 0$ for any $i$) much infrequently. According to the conjecture in Wallace (2014), there generically exists some government transfer schemes that improve welfare over no transfer in pure-currency economies. We find that government transfer is used in 2/3 of examples with $\alpha = 0$ (See Nozawa and Yang (2016) for a discussion on the upper bound and the conjecture). There are two reasons why the transfer is not used as frequently as in pure-currency economies. First, the transfer is done through monitored agents. We allow not only that monitored agents can produce for non-monitored agents with no money, but also can give money to them. In 1/2 examples with $\alpha > 0$, monitored producers both produce and give money to non-monitored consumers with no money. For the social planner, this is a more efficient way to give money to agents with no money than the direct transfer. While transfer is subject to incentive compatibility constraint ($\tau_0 \leq \tau_1 \leq \tau_2$) , we assume perfect information in pairwise meetings. Hence, risk-sharing can be achieved more efficiently through monitored agents in pairwise meetings, than the direct transfer. Second, as some inflation is already caused by net money creation of monitored agents, the transfer can be too costly as it will further increase the inflation rate.

## 6. CONCLUDING REMARKS

Let us quote Deviatov and Wallace (2014) to highlight the message of this study: this is a counter-example to the view that “inside-money economies should be regulated so as to avoid inflation,” and that “inflation is harmful in the absence of nominal rigidities.” Inflation can be optimal in a model with inside money as it enables those who can issue money spend more than they earn. However, their model does not give scope to policies resembling the Friedman rule because of the small upper bound on money holdings, and the result can potentially be overturned if such policies financed through the tax on monitored (those who can issue money) were possible and optimal. We extend their model by adopting higher upper bound and find that their result is robust. Inflation is optimal for all the examples we considered, and it arises from the net positive money creation by monitored agents, not from the government transfer.

The upper bound of three, which is what we used, is not large but not small either. It is the smallest $B$ that gives potential scope to policies resembling the Friedman rule, and we cannot think of a reason why adopting a higher $B$ will overturn the result.

## REFERENCES

Antinolfi, G.,C. Azariadis, and J. Bullard, 2016. The optimal inflation target in an economy with limited enforcement. *Macroeconomic Dynamics* **20(2)**, 582-600.

Bertolai, J., R. Cavalcanti, and P. Monteiro, 2012. Transitions when money distribution matters. *Mimeo.*

Cavalcanti, R. and N. Wallace, 1999. Inside and outside money as alternative media of exchange. *Journal of Money, Credit and Banking* **31(3)**, 443-457.

Deviatov, A., 2006. Money creation in a random matching model. *Topics in Macroeconomics* **6(3)**, 1-20.

Deviatov, A. and N. Wallace, 2014. Optimal inflation in a model of inside money. *Review of Economic Dynamics* **17(2)**, 287-293.

Green, E. J. and R. Zhou, 2005. Money as a mechanism in a bewley economy. *International Economic Review* **46(2)**, 351-371.

Kehoe, T. J., , D. K.Levineand, and M. Woodford, 1992. The optimum quantity of money revisited. *Economic analysis of markets and games: Essays in honor of Frank Hahn*, 501-526.

Li, V. E., 1995. The optimal taxation of fiat money in search equilibrium. *International Economic Review* **36(4)**, 927-942.

Lucas, R. E., 1980. Equilibrium in a pure currency economy. *Economic inquiry* **18(2)**, 203-220.

Mas-Collel, A., M. D. Whinston, and J. R. Green, 1995. Micreconomic theory.

Molico, M., 2006. The distribution of money and prices in search equilibrium. *International Economic Review* **47(3)**, 701-722.

Nozawa, W. and H. Yang, 2016. Optimal intervention in a random-matching model of money. mimeo.

Shi, S., 1995. Money and prices: A model of search and bargaining. *Journal of Economic Theory* **67(2)**, 467-496.

Trejos, A. and R. Wright, 1995. Search, bargaining, money, and prices. *Journal of Political Economy* **103(1)**, 118-141.

Wallace, N., 2010. The mechanism-design approach to monetary theory. *Handbook of Monetary Economics* **3**, 3-23.

Wallace, N., 2014. Optimal money creation in "pure currency" economies: a conjecture. *The Quarterly Journal of Economics* **129(1)**, 259-274.