The Optimal Inflation Tax

Isabel Correia*

Banco de Portugal
and Universidade Católica Portuguesa

Pedro Teles*

Banco de Portugal
and Universidade Católica Portuguesa

ABSTRACT

We determine the second best rule for the inflation tax in monetary general equilibrium models where money is dominated in rate of return. The results in the literature are ambiguous and inconsistent across different monetary environments. We compare the derived optimal inflation tax solutions across the different environments and find that Friedman's policy recommendation of a zero nominal interest rate is the right one.

Key words: Friedman rule; Inflation tax

JEL classification: E31; E41; E58; E62

1 Introduction

The purpose of this paper is to determine the optimal inflation tax in monetary general equilibrium models where money is dominated in rate of return, i.e. where the issue of the optimal inflation tax is the one of determining a nominal interest rate. This is the issue addressed by Friedman (1969). Friedman (1969) proposes a monetary policy rule that generates a nominal interest rate equal to zero, corresponding to a zero inflation tax, and to a negative rate of inflation. His approach is first best partial equilibrium.

We are interested in the more relevant second best results, i.e. when the government must finance government expenditures without having access to lump-sum taxation. Here, the literature is inconsistent, particularly across different monetary environments. The key inconsistency, that will be our main focus here, is that while in models that explicitly specify transactions technologies, the Friedman rule is a general result, in models with money in the utility function the traditional Phelps (1973) result of an optimal positive inflation tax is still the doctrine. We clarify the issues involved and surprisingly find that the Friedman rule is, in fact, the optimal policy. This result is important in that it translates into a very clean policy recommendation, independent of the parameterization of the economy.

The class of general equilibrium models that incorporate the feature of dominance in rate of return, and where we perform the welfare analysis, are designed in a somehow ad-hoc fashion[1]. Where this is more clearly so is in models where the preferences depend on the real quantity of money, as proposed by Sidrauski (1967) and Brock (1975). The fact that the use of money for transactions is not explicit in these models lead Clower (1967) to propose a cash-in-advance restriction. Lucas (1980) and Lucas and Stokey (1983) used this approach in a general equilibrium framework. A more com-

---

[1]In contrast, models where the purpose is to generate an equilibrium positive price for fiat money are more fundamentally specified. The seminal papers are Samuelson (1958), Grandmont and Younes (1973), Bewley (1980), Townsend (1980) and Kyiotaki and Wright (1989). In these models, the perfect substitutability between money and bonds implies a zero nominal interest rate, and the policy issue is the determination of the real interest rate.

plete transactions technology where it is assumed that time is substitutable for the use of money was addressed by McCallum (1983), Kimbrough (1986) and also McCallum and Goodfriend (1987).

Two major second best taxation sets of rules in the public finance literature have been used to justify the optimal inflation tax results: Diamond and Mirrlees (1971) optimal taxation rules of intermediate goods and Ramsey (1927) taxation rules of final goods, further developed by Atkinson and Stiglitz (1972). The Diamond and Mirrlees (1971) optimal taxation rules, derived for the case of constant returns to scale production functions, are the basis for the results in the literature of monetary models with transactions technologies[2]. In Correia and Teles (1996) we show that the Friedman rule is the optimal solution in these monetary models for all homogeneous transaction costs functions. We also show that the interpretation of this result is not a direct extension of Diamond and Mirrlees theorem but is related to the free good characteristic of money and to the special structure of production and taxation implied in this class of models.

Atkinson and Stiglitz (1972) established that it is optimal not to distort the relative prices between consumption of different goods when the preferences are separable in leisure and homothetic in the consumption goods. These rules were applied to cash-credit goods economies by Lucas and Stokey (1983) and Chari, Christiano and Kehoe (1996). There, the inflation tax translates into discriminated effective taxes on credit and cash goods. The Friedman rule is optimal under the Atkinson and Stiglitz (1972) conditions for uniform taxation.

The Ramsey (1927) rules were also used to explain the results in models with money in the utility function. In this class of models, the general impression is that, as pointed out by Woodford (1990),"...either the Phelps or the anti Phelps result is possible depending upon details of specification". Phelps (1973) uses a money in the utility function model with exogenous factor prices and concludes that "...the optimal inflation tax is positive". Chamley (1985) aims at generalizing his results to a general equilibrium model and concludes that the Friedman rule is the optimum only in the first best case. Siegel (1978) stresses the costless nature of liquidity services but concludes that this characteristic does not affect the result of a strictly

---

[2]See Kimbrough (1986), Guidotti and Végh(1993) and Chari, Christiano and Kehoe (1996)

3

positive tax on those services. Drazen (1979) states that the distinction of costly and costless to produce goods is important in the determination of the second best solution. Nevertheless he concludes that "...it appears difficult to say even whether the optimal inflation rate will be positive or negative". These results are disturbing because they are not consistent with the general optimality result of the Friedman rule in transactions technology models.

It turns out that the generalized use of the Ramsey (1927) rules to justify Phelps result is misleading and explains the ambiguity and the apparent inconsistency in the results in the different monetary environments. The rules on optimal taxation of final goods apply to ad-valorem taxes on costly goods. In general, the optimal ad-valorem consumption taxes are strictly positive. Since the goods are costly, the corresponding unit taxes are also strictly positive. But money is assumed to have a negligible production cost. If this is the case, then the only tax that can generate positive revenue is a unit tax and in any case the nominal interest rate is by construction a unit tax. The general result that the ad-valorem tax rate on real balances is strictly positive, can translate in the limit, when the costs of producing money are made arbitrarily small, into an optimal zero nominal interest rate. This is the basic intuition for why the Friedman rule is a general result.

Another reason for the apparent inconsistency in the optimal inflation tax results is that the approach to the second best problem is not uniform. Some authors impose conditions of stationarity on the second best problem. In this third best solution the Friedman rule is never the optimal solution.

The clarification of these puzzles and the specification of the conditions under which the Friedman rule is optimal is the aim of this paper. We take into account mainly (i) the free good characteristic of real balances, (ii) the fact that models with money in the utility function are reduced forms of more explicit monetary models and (iii) the Ramsey problem is unrestricted. We extend the results for money in the utility function models obtained by Chari, Christiano and Kehoe (1996) and establish the links between the results obtained in the different types of monetary models. In particular, we relate the results in a money in the utility function model to the ones for transactions technology models obtained by Correia and Teles (1996).

The optimal rules in the money in the utility function model are derived in section 2. In section 3, we show that the Friedman rule is a general result by establishing an equivalence between the money in utility function models and the underlying transactions technologies. In section 4 we provide the

4

main intuition for the results. In section 5 we compare the results to the ones in models with credit goods. In section 6 we discuss the robustness of results to alternative specifications of the available taxes and alternative timing structures. Section 7 contains the conclusions.

2 Money in the utility function

We use the general equilibrium model of a monetary economy developed by Sidrauski (1967) and later used by Brock (1975), Woodford (1990) and Chari, Christiano and Kehoe (1996) to discuss the optimality of the Friedman rule in a first best and a second best environments. The economy is populated by a large number of identical infinitely lived households, with preferences given by

$$\sum_{t=0}^{\infty} \beta^t V\left(c_t, \frac{M_t}{P_t}, h_t\right) \tag{1}$$

where c_t, M_t, P_t and h_t represent respectively consumption in period t, money balances held from period t to period $t+1$, the price of the consumption good in units of money in period t, and leisure in period t. The utility function shares the usual assumptions of concavity and differentiability. It is increasing in $\frac{M_t}{P_t}$, as long as $\frac{M_t}{P_t} < m^*(c_t, h_t)$ and non increasing for $\frac{M_t}{P_t} \geq m^*(c_t, h_t)$. $m^*(c_t, h_t)$ is the satiation function that represents the "satiation level", that is the point where "cash balances ... are held to satiety, so that the real return from an extra dollar is zero" (Friedman (1969)). The characterization of this function $m^*(c_t, h_t)$ is crucial for the determination of "the optimum quantity of money", as will be shown in the next section. Although Friedman (1969) does not fully characterize this point, the examples he presents imply that the point of satiation is finite. Phelps (1973) also assumes that there is "full liquidity" or "liquidity satiation" when the nominal interest rate is not strictly positive and the demand for real balances is finite. The discussion in Brock (1975), on the optimum quantity of money, is also for a finite satiation level. The technology of production of the private good and the public consumption good is linear with unitary coefficients.

The representative household (that implicitly solves the problem of the firm) chooses, a sequence $\{c_t, h_t, M_t, B_t\}_{t=0}^{\infty}$, given a sequence of prices and income taxes, $\{P_t, i_t, \tau_t\}_{t=0}^{\infty}$ and initial conditions for $W_0 \equiv M_{-1} + (1+i_{-1})B_{-1}$,

5

to satisfy a sequence of budget constraints:

$$P_t c_t + M_{t+1} + B_{t+1} \leq (1 - \tau_t) P_t (1 - h_t) + M_t + (1 + i_t) B_t \ , \ t \geq 0$$
$$ \tag{2}$$
$$M_0 + B_0 \leq W_0$$

together with a no- Ponzi games condition. B_t is the number of bonds held from period t to period $t + 1$, and i_t is the nominal return on these bonds. $1 - h_t$ is the labor supply.

The set of budget constraints can be written as a unique intertemporal budget constraint:

$$\sum_{t=0}^{\infty} Q_t P_t c_t + \sum_{t=0}^{\infty} Q_t i_t M_t \leq \sum_{t=0}^{\infty} Q_t P_t (1 - \tau_t)(1 - h_t) + W_0 \tag{3}$$

where $W_0 = M_{-1} + (1 + i_{-1}) B_{-1}$ and $Q_t = \frac{1}{(1+i_0)...(1+i_t)}$.

The competitive equilibrium is such that:

$$\frac{V_c(t)}{\beta V_c(t+1)} = (1 + i_{t+1}) \frac{P_t}{P_{t+1}} \ , \ t \geq 0 \tag{4}$$

$$(1 - \tau_t) V_{c_t} = V_{h_t} \ , \ t \geq 0 \tag{5}$$

$$V_{m_t} = i_t V_{c_t} \ , \ t \geq 0 \tag{6}$$

and the resources constraint is

$$c_t + g_t = 1 - h_t \ , \ t \geq 0 \tag{7}$$

where g_t is the level of public spending in period t. We assume that g_t is constant, $g_t = g$.

The first best policy in this economy is given by the maximization of (1) subject to the resources constraint (7). In this solution $V_{c_t} = V_{h_t}$ and $V_{m_t} = 0$, $t \geq 0$. If the government could collect lump-sum taxes, it would be possible to decentralize this solution by setting a constant nominal interest rate equal to zero. It is clear that the Friedman rule is optimal simply because real balances are a free good, in the sense that they do not require resources to be produced. Since the social marginal costs of money are equal to zero, then the level of money balances that characterizes the first best is m^*, i. e. the level for which marginal utility is zero. The solution can be decentralized

by setting the private marginal cost of holding real balances identical to zero, i.e. a zero nominal interest rate.

The optimal policy problem is more interesting when the government cannot levy lump-sum taxes. To determine the second best, Ramsey, solution we construct the implementability constraint, substituting (4), (5) and (6) into the intertemporal budget constraint (3):

$$\sum_{t=0}^{\infty} \beta^t \left(V_{c_t} c_t - V_{h_t} (1 - h_t) \right) + \sum_{t=0}^{\infty} \beta^t V_{m_t} m_t = \left(V_{c_0} + V_{m_0} \right) \frac{W_0}{P_0} \qquad (8)$$

The solution of the maximization of (1) subject to the implementability constraint (8) and the resources constraint (7), and given the initial nominal wealth, W_0, is the following: P_0 is set at an arbitrarily large number and the first order conditions for $\{c_t, h_t, m_t\}_{t=0}^{\infty}$ are as follows:

$$\beta^t V_{c_t} + \beta^t \psi \left[V_{c_t} + V_{c_t c_t} c_t - V_{h_t c_t} (1 - h_t) + V_{m_t c_t} m_t \right] = \lambda_t \ , \ t \geq 0 \qquad (9)$$

$$\beta^t V_{h_t} + \beta^t \psi \left[V_{h_t} + V_{c_t h_t} c_t - V_{h_t h_t} (1 - h_t) + V_{m_t h_t} m_t \right] = \lambda_t \ , \ t \geq 0 \qquad (10)$$

$$\beta^t V_{m_t} + \beta^t \psi \left[V_{m_t} + V_{c_t m_t} c_t - V_{h_t m_t} (1 - h_t) + V_{m_t m_t} m_t \right] = 0 \ , \ t \geq 0 \qquad (11)$$

where ψ is the shadow price of the implementability constraint, i.e. it measures the marginal excess burden of government deficits in this second best world. λ_t measures the shadow price of resources.

This second best allocation can be decentralized using the instruments i_t and τ_t, $t \geq 0$. Given (6), the discussion of whether the Friedman rule is optimal in this environment is equivalent to the discussion of whether V_{m_t} is zero, for $t \geq 0$. Conditions (9)-(11), together with the resources constraint (7) and the implementability constraint (8) define the stationary solution for c_t, h_t, m_t, ψ, and λ_t/β, for $t \geq 0$. From the competitive equilibrium condition $V_{m_t} = i_t V_{c_t}$, the solution for the nominal interest rate is stationary. If the Friedman rule holds, it holds for every period. So the issue is whether, for $t \geq 0$, $V_{m_t} \equiv V_m = 0$, is a solution of the system of equations.

Because money is a free good, as is clear from the resources constraint (7), the multiplier of the resources constraint does not show up in condition (11). In this second best solution the social marginal benefit of using money, for the households, is equal to the marginal "excess burden", i.e. the marginal cost due to the fact that a change in m affects the budget constraint of the government[3]. This can be seen by rewriting (11) as:

$$V_{m_t} = -\psi \left[V_{m_t} + V_{c_t m_t} c_t - V_{h_t m_t}(1 - h_t) + V_{m_t m_t} m_t \right]$$

Given that at the optimum, $\psi > 0$, the relevant issue is the determination of the sign of the term in parenthesis, i.e. the impact on government revenue of an increase in m, holding the quantities of the other goods constant. An increase in m corresponds to a decrease in the nominal interest rate and has a negative impact on the revenue from seigniorage, $V_m m$. So $V_m + V_{mm} m \leq 0$. The sign of the expression depends then on the cross derivatives. Suppose $V_{cm} < 0$ and $V_{hm} > 0$. In this case, the expression would be negative meaning that an increase in m has a negative effect on total government revenue. This means that the marginal "excess burden" would be strictly positive. The implication is that the Friedman rule would not be optimal. This is an example of the general properties that are typically discussed in the literature.

However, it is possible that the marginal excess burden of real balances is zero at the optimum. As we will argue in the next section this is indeed the more reasonable scenario, contrary to the assumptions typically made in the literature. To that end, we interpret models with money in the utility function as reduced forms of environments with a transactions role for money.

Since leisure is not a cash good, we will assume that the satiation point in real balances is a function of consumption only, $m^*(c)$. This assumption will be completely justified in the next section. The following proposition states the main result in the paper:

Proposition 1:

> In models with money in the utility function the Friedman rule is the optimal policy when the satiation point in real balances is such that $m^* = \infty$ or $m^* = \overline{k}c$, where $\overline{k}$ is a positive constant.

---

[3]The implementability constraint was constructed using the budget constraint of the households, but an equivalent constraint can be obtained using the government budget constraint. The marginal effect on the implementability condition corresponds to a symmetric marginal effect on the condition expressed in terms of the government budget constraint.

Proof. We verify whether equation (11) is satisfied when $V_m = 0$. m^* was assumed to be a function of c only, therefore at the satiation point $V_{hm} = 0$.

When $m^* = \infty$, must have $V_{mm}(c, m^*) = 0$ and $V_{mc}(c, m^*) = 0$. We assume that, when $V_m = 0$, the inflation tax revenue, $V_m m$, is zero. Therefore $\lim_{m\to\infty} \frac{\partial(V_m m)}{\partial m} = 0$, or $\lim_{m\to\infty}(V_m + V_{mm}m) = 0$. Then $m^* = \infty$ verifies the equation.

When the satiation point in real balances is finite, we find from $V_m(m^*, c) = 0$, that

$$\frac{V_{cm}(c, m^*)}{V_{mm}(c, m^*)} = -\frac{dm^*}{dc}$$

So, expression (11) evaluated at the satiation point in real balances can be written as

$$V_m(c, m^*)\left[1 + \psi\right] = -\psi V_{mm}(c, m^*)m^*\left[1 - \frac{dm^*}{dc}\frac{c}{m^*}\right]$$

where $V_m(c, m^*) = 0$.

When $m^* = \overline{k}c$, we have that $\frac{dm^*}{dc}\frac{c}{m^*} = 1$ and so, $V_{cm}(c, m^*)c + V_{mm}(c, m^*)m^* = 0$. Therefore $m = \overline{k}c$ also satisfies the Ramsey first order condition, (11) ∎

The elasticity of the satiation function, $m^*(c)$, can also be either strictly bigger or smaller than one. In either case the Friedman rule is not optimal, unless it is a corner solution, which can be the case when the elasticity is higher than one. In the next section we use a theoretical argument to show that these two cases of preference specifications are not the relevant ones. In any case we now provide a numerical argument in favor of the Friedman rule as a very close approximation to the optimal inflation tax.

When the elasticity is higher than one, $V_{cm}(c, m^*)c + V_{mm}(c, m^*)m^* > 0$. If we assume that the marginal benefit of real balances is always non-negative, then for all the preference specifications that we have used, the optimal allocation is $V_m = 0$, corresponding to the Friedman rule.

We have performed a numerical analysis of a particular preferences specification when the elasticity of the satiation function is strictly lower than one, specifically for the limit case, when the elasticity of m^* is equal to zero. The results are depicted in Figures 1, 2, 3 and 4. The calibration is done using as border line cases, the example in Calvo and Guidotti (1993) and in Lucas (1994). The instantaneous utility functions are additively separable $U = c_t + H(h_t) + v^i(m_t)$, where $H(h_t) = h_t - \frac{E}{2}(h_t)^2$ and $i = 1, 2$.

9

We consider two possible v^i functions: $v^1(m_t) = m_t(B - D\ln(m_t)$ and $v^2(m_t) = -A\left[\frac{1}{m_t} + \frac{m_t}{\overline{k}^2}\right]$. A, B, D, E and $\overline{k}$ are parameters, and $\overline{k}$ represents the constant satiation level in real balances. Government expenditures are set $g = .15$. The first utility function, v^1, is initially calibrated with the numbers provided by Calvo and Guidotti (1993), $B = -.65$, $D = .5$, $E = 1$. Figure 1 shows the resulting welfare cost of the inflation tax, in units of consumption, as well as the corresponding equilibrium schedule for the ratio of real balances to income, as a function of the nominal interest rate. The US line is the log-linear $M1$ to NNP schedule estimated by Lucas (1994) to US data (elasticity is .5). The optimal nominal interest rate is large, around 10%, but notice that the calibrated money-income ratio and interest rate schedule is not consistent with the US data. Figure 2 shows the same curves for a different calibration, that fits better the US money demand schedule: $B = -.046$, $D = .1429$. The semi-elasticity is now 7. The optimal nominal interest rate is considerably smaller. Figures 3 and 4 represent still the same two curves, for the second preferences specification, v^2. With this utility function, in the limit, for an arbitrarily large $\overline{k}$, the real balances to consumption ratio, as a function of the nominal interest rate, is a log-linear schedule. For $\overline{k} = 1$, the optimal nominal interest rate is smaller than .1% (Figure 3). Figure 4 shows the optimal nominal interest rate when $\overline{k} = .4$, generating levels of velocity at the satiation point that have been observed for considerably higher nominal interest rates. The optimal nominal interest rate is less than 1%. The conclusion is that for reasonable levels of $\overline{k}$, the Friedman rule is a very good approximation to the optimum. So, we conclude with the following result:

Result 1:

> When Proposition 1 does not apply, the optimal inflation tax is very close to the Friedman rule.

In summary we have shown in this section that the first best and the second best optimal inflation tax rules coincide, when the satiation point in real balances is infinite or when it is characterized by an unitary elasticity with respect to consumption. The reason being that in those cases the increase in real balances has a zero effect on government revenues at the satiation point and consequently the marginal costs are equal to zero in both problems. In the next section we go beyond the reduced form of the money in the utility

function models to inquire, first, how these local properties can be justified, second, whether it can be argued that the case of the unitary elasticity of real balances at the satiation point is the relevant case.

3   Monetary models with transactions technologies

In this section we show that if the money in the utility function models are seen as reduced forms of transactions technologies monetary models, then the preference specifications must be restricted and the Friedman rule is optimal under general conditions.

Feenstra (1986) shows that it is possible to establish an equivalence between a monetary model with a transactions technology, where the preferences depend only on consumption net of transactions costs, and a model with money in the utility function. He establishes a correspondence between the set of assumptions characterizing the transactions technology and the assumptions on the utility function expressed as a function of real balances. Here we extend Feenstra's (1986) results to a world where the original preferences are defined over consumption and leisure, $\sum_{t=0}^{\infty} \beta^t U\left(c_t, h_t^u\right)$, and the transaction costs are measured in units of time. The transactions costs function is represented by $s = l(m, c)$, where s is the time spent in transactions. This is the type of shopping time specification of McCallum (1983)[4].

The maximization problem is as follows:

Problem 1: Choose $\{c_t,\ M_t,\ B_t,\ h_t^u,\ s_t\}_{t=0}^{\infty}$ to maximize

$$\sum_{t=0}^{\infty} \beta^t U\left(c_t, h_t^u\right)$$

subject to

$$P_t c_t + M_{t+1} + B_{t+1} \leq (1 - \tau_t) P_t n_t + M_t + (1 + i_t) B_t \ , \ t \geq 0$$

---

[4]McCallum (1990) constructs (footnote 7) the indirect utility function associated with his type of shopping time costs. However he doesn't derive the properties of this utility function.

$$s_t = l(c_t, m_t)$$

$$1 = h_t^u + n_t + s_t$$

and $M_0 + B_0 \leq W_0$.

The transactions technology is characterized by the following assumption:

Assumption 1:

The transactions costs function $s = l(m, c)$ has the following properties:
a) $s \geq 0$, $l(m, 0) = 0$,
b) $l_c \geq 0$.
c) $l_{cc} \geq 0$, $l_{mm} \geq 0$
d) $l_m(m, c) = 0$ defines $m = \overline{m}(c)$. $l_m < 0$ when $m < \overline{m}$.
e) l is restricted to assure that Problem 1 is a concave problem.

The following are two first order conditions of Problem 1:

$$\frac{U_c(t)}{U_{h^u}(t)} = \frac{1}{1 - \tau_t} + l_c(t), t \geq 0$$

$$-l_m(t) = \frac{1}{1 - \tau_t} i_t , \ t \geq 0$$

The optimal choice of m is such that the private agents choose the point where the private value of using money, $-l_m(1-\tau)$, is equal to its opportunity cost, i. The implementation of the Friedman rule, $i = 0$, implies that $l_m = 0$.

We call the private problem defined in the last section, Problem 2. The assumptions for the utility function $V(c_t, m_t, h_t)$ are:

Assumption 2:

The utility function $V(c_t, m_t, h_t)$ satisfies the following conditions:
a)$\lim_{m \to \infty} V(c_t, m_t, h_t) < \infty$
b)V is concave and so, $V_{cc} \leq 0$, $V_{mm} \leq 0$
c) $V_m(c, m, h) = 0$ defines $m = m^*(c)$, $V_m > 0$ when $m < m^*$.

The following proposition states the equivalence between the two models.

Proposition 2:

Given a monetary model with an explicit transactions costs function (Problem 1), it is possible to construct an equivalent model with money in the utility function (Problem 2), where $h_t \equiv h_t^u + l(c_t, m_t)$ and $U(c_t, h_t - l(c_t, m_t)) \equiv V(c_t, m_t, h_t)$. If Assumption 1 is satisfied in Problem 1, then Assumption 2 is satisfied in Problem 2 and the functions $\overline{m} = m^*$.

Proof.

The equivalence is for a given pair of functions (U, l). The solution of Problem 1 is the vector $\left(\widehat{c}, \widehat{m}, \widehat{h^u}\right)$. Then there is a function V such that $\left(\widehat{c}, \widehat{m}, \widehat{h} = \widehat{h^u} - l(\widehat{c}, \widehat{m})\right)$ solves Problem 2.

The equivalence between the assumptions characterizing the two functions is granted by:

$V_m = -U_{h^u} l_m > 0$ when $m < m^*$ (since $l_m < 0$ when $m < m^*$) and

$V_m = 0 \iff l_m = 0 \iff \overline{m} = m^*$

$V_{cc} = U_{cc} - 2l_c U_{ch^u} + l_c^2 U_{h^u h^u} - l_{cc} U_{h^u} \leq 0$ (from concavity of Problem 1)

$V_{mm} = -U_{h^u} l_{mm} + l_m^2 U_{h^u h^u} \leq 0$

Therefore Assumption 1 implies Assumption 2. ∎

As we discussed in the last section the determination of whether the Friedman rule is optimal or not in a money in the utility function model depends crucially on the functional form of the function $m^*(c, h)$, the function defined by setting the marginal utility of money equal to zero. By the equivalence result we verify that the properties of this function are identical to the properties of the function defined by $l_m = 0$. From this, it results that:

1) We can justify the hypothesis made in the last section that leisure is not an argument of the function m^*.

2) When $l_m \leq 0$, then $V_m \geq 0$. The marginal utility of money is non-negative. This is relevant for the numerical exercises when the elasticity of the satiation function is higher than one.

3) When the transactions costs function is homogeneous of degree q, l_m is homogeneous of degree $q - 1$ and can be represented by

$$l_m = L(m/c)c^{q-1} \tag{12}$$

13

and, at the point $l_m = 0$, $L(m/c) = 0$ defines $m = m^* = \overline{k}c$.

In monetary models with explicit transactions technologies, the first order conditions of the second best problem imply that $l_m = 0$[5]. This solution can be decentralized through a zero nominal interest rate and so the Friedman rule is optimal. These are the main conclusions of Correia and Teles (1996), where it is shown that the Friedman rule is the Ramsey solution in a transactions costs monetary model with homogeneous transaction costs function.

Homogeneous transactions costs function correspond to the case described in section 2, where the function m^* has unitary elasticity and the Friedman rule is always optimal.

4) When the transactions costs function, $s = l(c, m)$, is associated with a transactions production function $c = f(s, m)$ which verifies Inada conditions, $l_m = 0$ is equivalent to $m^* = \infty$. In this case at the point $l_m = 0$, we have $l_{mm} = l_{mc} = 0$. Again the marginal condition of the second best problem is satisfied for $l_m = 0$, that is the Friedman rule is optimal. This corresponds to case of $m^* = \infty$, described in section 2.

5) The case of elasticity of m^* lower or greater than one corresponds to the case of non-homogeneous transactions costs functions. We do not know of any work where it is argued on theoretical or empirical grounds that the transactions costs function ought to be non homogeneous. At the theoretical level the micro-foundations of this function are obtained from Baumol (1952) and Tobin (1956), from the generalization of Barro (1976), from Guidotti (1989) or from Jovanovic (1982). All these forms are homogeneous of degree zero. In Marshall (1992) the proposed and estimated transactions costs function is homogeneous of degree one. Braun (1994) estimates the degree of homogeneity of the transaction cost function to be .98.

In summary we can conclude that once the equivalence between money in the utility function models and explicit transactions technologies models is established, the local properties used in section 2 to derive the optimality of the Friedman rule are associated with global properties of the transactions

---

[5] Correia and Teles (1995) show that this is a global maximum.

costs functions. Besides, for reasonable functional forms of transactions costs functions, the Friedman rule is optimal[6].

4   Money is a free good

In section 2, we show that the Friedman rule is optimal when there is no effect on government revenues of changing real balances, from the full liquidity level. The arguments for the second best taxation rules of final goods in the public finance literature are different from these. There, the optimal taxes on different goods depend on the comparison of the respective marginal effects on government revenues. In particular, for it to be optimal not to tax final goods, when the alternative choice is an income tax, the marginal effect on government net revenues of a change of one unit of labor used to produce any of the goods should be equal. Atkinson and Stiglitz (1972) derived the conditions under which this is the optimal rule.

Our results show that the conditions on preferences to obtain the optimality of the Friedman rule are more general than the ones derived in Atkinson and Stiglitz (1972) and therefore extend the result in Chari, Christiano and Kehoe (1996), that identified those conditions as sufficient conditions for the optimality of the Friedman rule[7]. The homotheticity and separability conditions of Atkinson and Stiglitz (1972) correspond to utility functions where the marginal rate of substitution between consumption and real balances depends only on the ratio of these two variables. This is one example of the conditions in Proposition 1. The conditions in Proposition 1 are much less restrictive though, since they must hold only in the neighborhood of $i = 0$.

We now show how the very appealing argument of the distribution of distortions among different goods in the economy can be reconciled with the

---

[6]The introduction of capital will not alter that result. Suppose that the transactions costs function is defined as before. The introduction of capital as an input in the production of the consumption good has no consequence in the Ramsey solution that defines the optimal choice of m. The main difference is that now the marginal utility of labor is not constant and depends on the level of the stock of capital. The Ramsey solution will have a transitional period but $l_m = 0$, and consequently $V_m = 0$, characterizes also that transition. If the transactions costs function is modified in a way that capital is also an input in the production of transactions the results are maintained once we impose that the transaction costs function is homogeneous in consumption, real balances and capital.

[7]As they point out, the condition is not necessary.

zero inflation tax result derived for monetary economies.

What distinguishes money from any other consumption good is the fact that an additional unit of real money does not require relevant marginal resources. We think that this is the right way of describing fiat money. In the following exercise we analyze how the taxation rules are affected when the cost of producing a good m becomes arbitrarily small. Consider a stationary real economy corresponding to the monetary economy we have studied, but for the fact that m is now a consumption good that is produced with α units of time, which implies that the price of m in units of the other consumption good c, is α. The ad-valorem tax on m is τ^m. The budget constraint of the households is written as:

$$c + \alpha(1 + \tau^m)m = (1 - \tau)(1 - h)$$

Notice that the equivalent unit tax would be $T^m = \alpha\tau^m$ (The corresponding equation in the monetary economy is $c + im = (1 - \tau)(1 - h)$).

According to the optimal taxation rules derived for the monetary economy (so, for the case where α goes to zero), when the impact of a marginal increase in m on government revenues is not zero, then the optimal unit tax can be positive. This corresponds to an optimal ad-valorem tax that becomes arbitrarily large as the cost of producing m is made arbitrarily small.

In the real economy with ad-valorem taxation, the optimal taxation rules of Ramsey (1927) and Atkinson and Stiglitz (1972) apply. Atkinson and Stiglitz (1972) derive necessary and sufficient conditions for optimality of uniform taxation of consumption goods. When the preferences are homothetic in c and m and separable in leisure, then a tax on labor income and a zero ad-valorem tax on both c and m decentralizes the second best, Ramsey, solution. This corresponds to $\tau^m = 0$. The unit tax is always equal to zero. Now suppose the alternative tax is a tax on c, τ^c. The budget constraint is written as

$$(1 + \tau^c)c + \alpha(1 + \tau^m)m = 1 - h$$

The second best solution is $\tau^m = \tau^c$. The corresponding limiting unit tax on m, T^m, as α converges to zero, is zero. This same result applies as long as the optimal ad-valorem tax converges to a finite number. This holds whenever, according to the rules derived for the monetary economy, the marginal impact of m on the government revenue, at the satiation point in real balances, is equal to zero.

There is a sense in which the application of the conditions of Atkinson and Stiglitz (1972) to this problem is misleading. Chari, Christiano and Kehoe (1996) compute the optimal mix between the inflation tax and the income tax, rather than the consumption tax. If the fiscal instrument is an income tax, under the conditions of uniform taxation, it is not optimal to use consumption taxes to distort the consumption of the different goods. That way, a zero inflation tax could apparently mean that consumption and real balances were not being distorted. But that is clearly not the case, consumption is taxed and real balances are not taxed at all.

5   Credit goods

In this section we compare the results in the money in the utility function model with the results in models with credit goods. It is well known that it is possible to establish an equivalence between the two models, by replacing in the money in the utility function model, total consumption for the sum of the consumptions of the two goods and the real balances for the consumption of the cash good. A condition that ensures that real balances are bigger than total consumption guarantees non-negativity of consumption of the credit good.

Chari, Christiano and Kehoe (1996) identify as the necessary and sufficient condition for optimality of the Friedman rule in the cash-credit goods model, the conditions of homotheticity in the two goods and separability in leisure. The explanation for this result is simple. Suppose that real money was costly to produce and consider the typical structure in a cash-credit goods model, where consumption of the cash good requires time and real money in fixed proportions. The production functions of the two goods and of real money are linear. Then, a positive tax on money would not distort the production of the consumption good, due to the Leontieff structure, but would distort the relative consumption of the two goods. When the utility function is homothetic in the consumption of the two goods and separable in leisure it is clearly optimal to set the tax on real balances (ad-valorem or unit) to zero. If the utility function is not homothetic in the two goods, then the inflation tax could be used as a means of achieving the optimal distortion in the two goods[8].

---

[8]If the transactions technology allowed for substitutability between real balances and

17

Separability in leisure and homotheticity in the cash and credit goods imply separability in leisure and homotheticity in real balances and total consumption in the equivalent money in the utility function model. These conditions imply that the conditions in Proposition 1 are met. Separability in leisure is guaranteed by the assumption that the satiation level of real balances does not depend on leisure. Given the constraint that total consumption must be higher than real balances, the real balances-consumption elasticity at the satiation point must be one.

6 Alternative Taxes and welfare criteria

In sections 2 and 3 we have constructed a second best environment assuming that there were two alternative taxes, an inflation tax and a tax on labor income. For the purpose of checking the robustness of the results, in this section we discuss the implications of considering a consumption tax instead of the labor income tax. In addition we will assess the implications of considering alternative timing conventions and welfare criteria in the specification of the second best problem.

6.1 Consumption taxes

When the level of transactions is measured by consumption net of taxes, the tax on consumption, τ_c, does not affect the transactions costs function, i. e. $s = l(c, m)$. In this case the indirect utility function $V(c, m, h)$, associated with the pair (U, l) is the same as described in section 3. The second best formulation in terms of the primal approach and the optimal allocation are independent of the type of instruments available for its decentralization, provided the number of instruments is enough to decentralize the Ramsey solution.

When the alternative tax is a tax on labor income, τ_t, the private first order conditions are (5) and (6). When the alternative tax is a consumption tax, τ_{ct}, these conditions are replaced by

---

time, then the distortion of the consumption of the two goods, would also imply a distortion in production. Clearly, in this case, it would be preferable to discriminate between the consumption taxes on the two goods.

$$V_{ct} = (1 + \tau_{ct})V_{ht} \ , \ t \geq 0$$

$$V_{mt} = \frac{i_t}{(1 + \tau_{ct})}V_{ct} \ , \ t \geq 0$$

The optimal allocation $V_{mt} = 0$, $t \geq 1$, can be decentralized in both cases with $i_t = 0$, $t \geq 0$. So the conditions under which the Friedman rule applies are the same whether the alternative tax is a tax on labor income or a tax on consumption. The Friedman rule is still the optimal inflation tax rule, when the alternative tax is a consumption tax. To conclude, the irrelevance result of the alternative tax in money in the utility function models is extended to models of explicit transaction costs functions, given the equivalence established in section 3.

A number of authors claim that the introduction of a consumption tax should modify the transactions costs function, in the sense that the amount of transactions ought to be measured by consumption gross of taxes: $s = l(c(1+\tau_c), m)$. This introduces some changes. Adopting the same procedure as in section 3, it is clear that the pair (U, l) is equivalent to a utility function V such that:

$$U\left(c, h - l\left(c(1 + \tau_c), m\right)\right) \equiv V\left(c, m, h, (1 + \tau_c)\right)$$

Now preferences depend on the tax parameter τ_c. Under this formulation the Friedman rule is optimal when $m^* = \infty$. It is also optimal when the elasticity of m^* $(c(1 + \tau_c))$ is unitary, if in addition we impose that, at the satiation point, the underlying technology is characterized by $l_{cm} = 0$.

6.2 Alternative timing conventions and welfare criteria

It is a standard view (See Woodford (1990), p. 1092) that alternative timing conventions in the decisions of the private agents affect the result of optimality of the Friedman rule. In the previous sections, the private agents are assumed to choose financial assets, in each time period, so that the resulting

money balances can be used for transactions that same period. This is the timing assumed in Lucas (1982) and Lucas and Stokey (1983)[9].

Alternatively, Woodford (1990) assumes, in line with Svensson (1985)[10], that the money balances that can be used in any one period are decided the period before. The implication of this timing is that there are real effects of unanticipated monetary shocks. In the beginning of time, period zero, the agents cannot adjust the portfolios and therefore it is no longer optimal, for the benevolent government to completely deplete the real value of outstanding monetary balances. The allocation is stationary from period one on, but in period zero the levels of the variables are, in general, different from the corresponding stationary levels. The stationary optimal allocation from period one on corresponds to the Friedman rule. Woodford (1990), for the sake of tractability, and Braun (1994) and Lucas (1995), propose a third best solution concept: the maximization of welfare restricted to the solution being stationary, i. e. the problem is restricted so that the allocation in period zero is the same as from period one on.

When the stationarity restriction is imposed the government faces a trade-off between the low level of initial real balances and the high steady state level. It is intuitive that from this trade off results a stationary level of real balances higher than the initial optimal level of the Ramsey solution and lower than the high steady state level of the same solution. The solution is characterized by less than full liquidity and a strictly positive nominal rate of interest.

Lucas (1994) has computed numerically the optimal policy in a transactions technology model, using this criterium. He concludes that the optimal nominal interest rate, although strictly positive, is very close to zero.

---

[9]Asset markets open in the beginning of period, so that money balances used for transactions in any period are beginning of period balances. Kimbrough (1986) and others assume that end of period money balances are used for transactions that same period.

[10]A discussion of the positive implications of the two timing conventions in a cash-in-advance model is presented in Giovannini and Labadie (1991). Nicolini (1993) discusses time inconsistency in a cash-in-advance model with the two alternative timing conventions.

7 Concluding remarks

In this paper we compute the second best inflation tax rule in models where real balances are an argument in the utility function. We identify local conditions that extend the global conditions of separability and homotheticity in Chari, Christiano and Kehoe (1996), as sufficient conditions for the optimality of the Friedman rule. Furthermore we establish an equivalence between the models with money in the utility function and more fundamental models of transactions technologies, and show that the wide class of transactions technologies where the Friedman rule is optimal satisfy the local conditions for the optimality of the Friedman rule in the money in the utility function specification.

The characteristic of real balances that is determinant for the general optimality of the Friedman rule is the fact that money is a free good meaning that the production cost of money is zero, i.e. the production possibilities in these economies are not affected by a change in the quantity of money. For this reason, the usual intuition of the comparison of the marginal excess burdens of alternative taxes that give the same revenue, no longer applies. The optimal decision here consists in the following comparison: An increase in the quantity of money generates a benefit for the households in terms of utility and a cost equal to the value of the marginal effect on government net revenues. At the point of satiation in real balances the marginal utility benefit is by definition equal to zero. We show that under reasonable preferences specifications the marginal impact on government net revenues is also equal to zero, at that point of full liquidity. Therefore the Friedman rule is optimal. In less adequate specifications for preferences, in terms of its micro-foundations, where the Friedman rule is not optimal, it is nevertheless very close to the optimum.

The main conclusion of this paper is that the optimal taxation results in monetary models are much more robust than the public finance results derived in other economic environments. In particular, the Friedman rule, i.e. a zero inflation tax, is a general result for monetary economic structures with reasonable micro-foundations. This normative result has no counterpart in the public finance literature where the optimal policies depend on the structure of preferences and technologies. The neat and successful practical recommendation of Friedman is reinforced now that it is shown that its optimality extends to a second best environment.

References

[1] Atkinson, A. B. and J. E. Stiglitz, 1972, The Structure of Indirect Taxation and Economic Efficiency, Journal of Public Economics 1, 97-11.

[2] Barro, Robert, J., 1976, Integral Constraints and Aggregation in an Inventory Model of Money Demand, Journal of Finance 31, 77-87.

[3] Baumol, William J., 1952, The Transactions Demand for Cash: An Inventory Theoretic Approach, Quaterly Journal of Economics 66, 545-556.

[4] Bewley, Truman, 1980, The Optimum Quantity of Money, in Models of Monetary Economies, John Kareken and Neil Wallace, eds, Federal Reserve Bank of Minneapolis, 169-210.

[5] Braun, Anton, 1994, Another Attempt to Quantify the Benefits of Reducing Inflation, Quarterly Review, Federal Reserve Bank of Minneapolis, Fall, 17-25.

[6] Brock, William, 1975, A Simple Perfect Forsight Monetary Model, Journal of Monetary Economics 1, 133-50.

[7] Calvo, G. and P. Guidotti, 1993, On the Flexibility of Monetary Policy: The Case of the Optimal Inflation Tax, The Review of Economic Studies 60, 667-687.

[8] Chamley, Christophe, 1985, On a Simple Rule for the Optimal Inflation Rate in Second Best Taxation, Journal of Public Economics 26, 35-50.

[9] Chari, V. V., Lawrence J. Christiano and Patrick J. Kehoe, 1996, Optimality of the Friedman Rule in Economies with Distorting Taxes, Journal of Monetary Economics 37, 203-223.

[10] Clower, Robert W., 1967, A Reconsideration of the Microfoundations of Monetary Theory, Western Economic Journal 6, 1-8.

[11] Correia, Isabel and Pedro Teles, 1995, Is the Friedman Rule Optimal when Money is an Intermediate Good?, WP 4-95, Banco de Portugal.

[12] Correia, Isabel and Pedro Teles, 1996, Is the Friedman Rule Optimal when Money is an Intermediate Good?, Journal of Monetary Economics 38, 223-244.

[13] Diamond, Peter A. and James A. Mirrlees, 1971, Optimal Taxation and Public Production, American Economic Review 63, 8-27, 261-268.

[14] Drazen, Allen 1979, The Optimal Rate of Inflation Revisited, Journal of Monetary Economics 5, 231-248.

[15] Feenstra, Robert C., 1986, Functional Equivalence between Liquidity Costs and the Utility of Money, Journal of Monetary Economics 17, 271-291.

[16] Friedman, Milton, 1969, The Optimum Quantity of Money, in M. Friedman, ed., The Optimum Quantity of Money and other Essays (Aldine, Chicago, Il.).

[17] Giovannini and Labadie, 1991, Assets Prices and Interest Rates in Cash-in-Advance Models, Journal of Political Economy 99, 1215-1251.

[18] Grandmont, Jean-Michel and Yves Younes, 1973, On the Efficiency of a Monetary Equilibrium, Review of Economic Studies XL(2), 149-165.

[19] Guidotti, Pablo, 1989, Exchange Rate determination, Interest rates and an Integrative Approach to the Demand for Money, Journal of International Money and Finance 8, 29-45.

[20] Guidotti, Pablo E. and Carlos A. Végh, 1993, The Optimal Inflation Tax when Money Reduces Transactions Costs, Journal of Monetary Economics 31, 189-205.

[21] Kimbrough, Kent P., 1986, The Optimum Quantity of Money Rule in the Theory of Public Finance, Journal of Monetary Economics 18, 277-284.

[22] Kiyotaki, Nobuhiro, and Randall Wright, 1989, On Money as a Medium of Exchange, Journal of Political Economy 97, 927-954.

[23] Jovanovic, Boyan, 1982, Inflation and Welfare in the Steady State, Journal of Political Economy 90, 561-577.

[24] Lucas, Robert E., Jr, 1980, Equilibrium in a Pure Currency Economy, in Models of Monetary Economies, ed. John Kareken and Neil Wallace, Federal Reserve Bank of Minneapolis, 131-45.

[25] Lucas, Robert E., Jr., 1982, Interet Rates and Currency prices in a Two-Country World, Journal of Monetary Economics 10, 335-59.

[26] Lucas, Robert E., 1994, The Welfare Cost of Inflation, mimeo, The University of Chicago.

[27] Lucas, Robert E., Jr. and Nancy L. Stokey, 1983, Optimal Fiscal and Monetary Theory in an Economy without Capital, Journal of Monetary Economics 12, 55-93.

[28] Marshall, David A., 1992, Inflation and Asset Returns in a Monetary Economy, The Journal of Finance 47, 1315-1342.

[29] McCallum, Bennet T., 1983, The Role of Overlapping Generations Models in Monetary Economics, Carnegie-Rochester Conference Series on Public Policy18.

[30] McCallum, Bennet, 1990, Inflation: Theory and Evidence, in Handbook of Monetary Economics, B. Friedman and F. Hahn eds, North Holland, 963-1012.

[31] McCallum, Bennet T. and Marvin Goodfriend, 1987, Demand for Money: Theorethical Studies, in The New Palgrave: A Dictionary of Economics, ed J. Eatwell, M. Milgate and P. Newman, London, MacMillan, 775-781.

[32] Nicolini, Juan Pablo, 1993, More on the Time Inconsistency of Optimal Monetary Policy, forthcoming Journal of Monetary Economics.

[33] Phelps, Edmund S., 1973, Inflation in the Theory of Public Finance, Swedish Journal of Economics 75, 37-54.

[34] Ramsey, Frank P., 1927, A Contribution to the Theory of Taxation, Economic Journal 37, 47-61.

[35] Samuelson, Paul A., 1958, An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money, Journal of Political Economy 6, 467-82.

[36] Sidrauski, Miguel, 1967, Inflation and Economic Growth, Journal of Political Economy 75, 796-810.

[37] Siegel, Jeremy,1978, Notes on Optimal Taxation and the Optimal Rate of Inflation, Journal of Monetary Economics 4, 297-305.

[38] Svensson, Lars E. O., Money and Asset Prices in a Cash-in-Advance Economy, Journal of Political Economy 93, 919-944.

[39] Tobin, James, 1956, The Interest Elasticity of the Transactions Demand for Cash, Review of Economics and Statistics 38, 241-247.

[40] Townsend, Robert M., 1980, Models of Money with Spatially Separated Agents, in Models of Monetary Economies, John Kareken and Neil Wallace, eds, Federal Reserve Bank of Minneapolis, 265-303.

[41] Woodford, Michael , 1990, The Optimum Quantity of Money, in Handbook of monetary Economics, B. Friedman and F. Hahn, eds, North-Holland, 1067-1152.