Annals of Finance
DOI 10.1007/s10436-008-0103-9

RESEARCH ARTICLE

On the general equilibrium costs of perfectly anticipated inflation

Paulo Barelli · Samuel de Abreu Pessôa

Received: 20 May 2008 / Accepted: 23 June 2008

Abstract In monetary models where M0 has no social costs and a positive demand for cash and deposits is taken as a primitive, we show that the compensating variation in endowment is the exact general equilibrium measure of welfare costs of perfectly anticipated inflation. As a consequence, we show that a good approximation to the welfare costs of inflation is given by the area under the compensated demand for M0, a result that brings us back to Bailey (J Polit Econ 64:93–110, 1956). The estimated welfare costs of inflation are bounded at less than a quarter of a percent of the GDP for the U.S. economy.

Keywords Money · Inflation · Welfare · Financial services

JEL Classification E31 · O42

1 Introduction

The question of measurement of welfare costs of perfectly anticipated inflation has a long tradition in the literature. The early intuition of Bailey (1956) was to use the area below the money demand curve as an estimate of the welfare costs of inflation. Fischer (1981) used Bailey's idea and estimated the welfare cost of inflation to be 0.3% of

P. Barelli (✉)
Department of Economics, University of Rochester, Rochester, NY 14627, USA
e-mail: pbarelli@troi.cc.rochester.edu

P. Barelli
Ibmec São Paulo, São Paulo, Brazil

S. de Abreu Pessôa
Graduate School of Economics, Fundação Getúlio Vargas, Rio de Janeiro, Brazil
e-mail: samuel.pessoa@fgv.br

GDP in the US. Several authors noticed, however, that Bailey's measure was a partial equilibrium measure and that a general equilibrium measure was needed. Dotsey and Ireland (1996) argued that a general equilibrium measure would be three times as large as Bailey's measure. Lucas (2000) used a general equilibrium measure and found that the benefit of reducing a 10% inflation rate to a 0% inflation rate is of the order of 1% of the GDP of the U.S. economy.

In this paper we show that in a class of monetary models the general equilibrium measure of welfare costs of inflation is the compensating variation in endowment. This measure differs from the measures used in the literature in two ways: first it is the compensating variation in *endowment* rather than *consumption* (as in Lucas 2000), and second it uses M0 rather than M1 as the relevant concept of money for welfare comparisons. Endowment is derived as the relevant variable instead of consumption because consumption is an endogenous variable, while endowment is exogenous. Agents will change their optimal choices of consumption and leisure in response to changes in inflation, so the compensating variation in consumption is likely to overestimate the welfare costs of inflation. And M0 is derived as the relevant concept of money because it is produced at no social cost. Since M0 does not include the effects of the money multiplier, our measure of welfare costs of inflation is smaller than the ones in the literature (Lucas 2000). Building on the results of Aiyagari et al. (1998), we estimate that at most a quarter of a percent of the US's GDP is lost due to inflation, a figure that is in line with Fischer's estimates.

That is, if on the one hand Bailey's measure underestimates the true costs of inflation because it is a partial equilibrium measure, on the other hand the proposed general equilibrium measures overestimate the true costs of inflation because of the concept of money used. The interesting twist is that estimates using a general equilibrium measure and M0 are much closer to estimates that use Bailey's measure. Bailey's measure provides a good approximation to a general equilibrium measure. In fact, a corollary of our main result is that, with no capital accumulation effects, Bailey's measure is an exact general equilibrium measure of welfare costs of inflation.

The main result is as follows: for a class of monetary models where welfare is measured as the intertemporal utility of the representative household and M0 has no social cost,[1] the general equilibrium effect of long-run inflation on welfare is given by the discounted sum of the effects of inflation on demand for M0, minus the effect on the initial endowment. The effects of inflation on all other variables, for instance, effects of inflation on capital accumulation, do not matter: except for money, the choice mechanism ensures that costs and benefits of all other choice variables are equalized.

The assumption that M0 is produced at no social cost is a common assumption in the macroeconomics literature that builds on the tradition of Bailey (1956) and Friedman (1969). The production of M1 requires a banking sector that uses up productive resources (capital and labor) that could be employed in other productive activities, so it generates social costs. The social cost of producing M0 is the cost of printing money, which is taken as negligible. On the other hand, one can argue that M0 carries the social cost of the resources used up in protecting from theft and counterfeit,

[1] Unlike M1 or M2, which include costly financial services, M0 is produced at virtually zero social cost. See the discussion in the next paragraph.

and in storage and handling of large amounts of cash. We follow the macroeconomic literature and abstract away such costs.[2]

The implications of the main result are as follows. First, Bailey's proposition holds true when the dynamic adjustment in capital is instantaneous. Moreover, even when the adjustment process is not instantaneous, the main result shows that the effects of the adjustment process are of second order of importance. Hence Bailey's partial equilibrium measure is a good approximation of a general equilibrium measure. Second, the competitive equilibrium in the class of monetary models considered here is Pareto optimal restricted to a given path of real money balances. That is, since all choice variables except money are chosen optimally, we do not have the ambiguous result of the second-best theorem: any policy that induces households to increase money holdings is welfare improving. And finally, in line with this last comment we have an overbanking result: any policy that increases money holdings and, consequently, reduces the banking sector, is welfare improving.

The rest of the paper is organized as follows: Sect. 2 reviews the literature. Section 3 presents the main result using a simple extension of Sidrauski's model, and the corollaries are in Sect. 4. In Sect. 5 contains an estimate of welfare costs of inflation and some concluding comments. In the Appendix we derive our main result in a more general cash-in-advance model, showing that the main result holds true in the class of monetary models where a positive demand for cash and deposits is taken as a primitive.

2 Related literature

Two recent contributions to the topic of welfare costs of perfectly anticipated inflation are Aiyagari et al. (1998) and Lucas (2000).[3] Due to its encyclopedic character, we view Lucas' paper as the representative of the literature that uses a general equilibrium measure of the welfare cost of inflation, instead of Bailey's measure. As mentioned above, our main difference is the concept of money used: we use monetary base, Lucas used M1. Our choice is justified by the main result in this paper. Lucas does not consider the existence of a banking sector, and hence cannot differentiate between M1 and M0.

Aiyagari et al. (1998), henceforth ABE, is the closest paper to ours: to the best of our knowledge, ABE were the first authors to notice the connection between Bailey's measure and a general equilibrium measure of welfare costs of inflation. In a cash-in-advance economy with credit goods, ABE show that the income share of the financial sector is equal to the area below the money demand curve. Since the former is, in ABE's set-up, the general equilibrium measure of costs of inflation, ABE show that Bailey's measure is more general than previously thought. On the other hand, ABE fail to realize the full generality of this result: There are costs of inflation that are not captured by the income share of the financial sector, and these are captured by the

[2] It is important to notice that, under the assumption that M0 is produced at no social cost, M0 is the relevant measure of money for welfare purposes regardless of whether banks are or are not allowed to offer interest-bearing demand deposit (contrary to the views of Lucas 1981, p 45).

[3] Another recent contribution is English (1999), although his main interest is the relation between inflation and the size of the banking sector.

money demand. For instance, time spent queuing in stores to buy goods and oversized financial departments in firms are nonnegligible costs of inflation not captured by ABE's measure. Moreover, the financial sector does provide a welfare improving service of intermediating saving and investment decisions, so it is not necessarily true that the income share of the financial sector represents the welfare costs of inflation. In ABE's model, where the role of the financial sector is solely the provision of substitutes for money, the exact measure of welfare costs of inflation is the income share of the financial sector. But this would not be true in a more general model. Our main result complements and generalizes ABE's findings in that we show that Bailey's result is true regardless of which costs of inflation one includes explicitly in the model.[4]

3 Result

The model economy is an extension of Sidrauski (1967) "money in utility function" model, with an additional banking sector provider of substitutes of money. We make no effort in terms of having the most general model. In the Appendix we derive an analogous of Lemma 1 in a cash-in-advance type of model with endogenous consumption-leisure choices and different technologies in the two sectors, showing that the logic behind Lemma 1 holds true in more general monetary models. The model used here is one of the simplest models with enough structure (namely, with substitutes for money) to convey our main point.

With an infinitely lived representative agent, welfare is given by the intertemporal discounted utility. The impact of inflation on welfare can be computed as the derivative of welfare with respect to the growth rate of money balances. Since all choice variables other than money are chosen optimally, this derivative turns out to be composed of two terms only: the effect of σ on initial capital stock, and the effect of σ on demand for monetary base (where σ is the growth rate of money).

Put differently, the general equilibrium adding up constraints allow one to rewrite the effect of σ on welfare in a simple way, that brings us back to Bailey's early intuition. This point, to the best of our knowledge, has not been considered in the literature.

3.1 Households

There is an infinitely lived representative household that derives positive instantaneous utility from the consumption of goods and negative instantaneous utility from costs involved in transactions, and rents labor and capital to firms. Labor is supplied inelastically. Households can hold two kinds of "money": cash, denoted by m_1, and demand deposit (supplied by banks), denoted by m_2. Both kinds of money reduce transaction costs, so m_1 and m_2 enter as arguments of the instantaneous utility function. In short,

[4] Jones et al. (2004) follow the approach of ABE, and show that welfare cost of inflation is substantially lower in the model with interest-bearing deposits than in models where all monetary assets are assumed to be non-interest bearing.

households choose c_t, m_{1t}, m_{2t} and $k_{h,t}$ that solve the following program:

$$\max \left\{ \int_0^\infty e^{-\rho t} u(c_t, m_{1t}, m_{2t}) \mathrm{d}t \right\}$$

$$\text{s.t. } \dot{a}_t = r_t a_t + \mathrm{w}_t + \chi_t - c_t - p_t m_{2t} - (\pi_t + r_t)(m_{1t} + m_{2t}), \qquad (1)$$

where, omitting time subscripts, $a \equiv k_h + m_1 + m_2$ is households' assets, k_h is households' physical capital stock, r and w are rental and wage rates, χ is government's transfers to households, c is consumption of goods, p is demand deposit's relative price and π is the inflation rate. All terms in the budget constraint (1) are standard, apart from the terms χ and pm_2 which are explained below. All variables are in units of goods per capita, and nominal assets depreciate at rate π.

The instantaneous utility function satisfies the standard assumptions. In addition, we assume that $u_{23} < 0$ to capture substitutability of m_1 and m_2 (subscripts represent partial derivatives with respect to the relevant arguments, for instance $u_{23} \equiv \frac{\partial^2 u}{\partial m_1 \partial m_2}$).

Let λ be the co-state variable associated with the budget constraint (1).[5] The first-order conditions are given by

$$\begin{aligned} u_1 &= \lambda \\ u_2 &= \lambda(\pi + r) \\ u_3 &= \lambda(\pi + r + p) \\ \frac{\dot{\lambda}}{\lambda} &= \rho - r, \end{aligned}$$

together with the transversality condition $\lim_{t\to\infty} e^{-\rho t} \lambda_t a_t = 0$.

The supply side is composed of two productive sectors; sector 1 (firms) produces goods and sector 2 (banks) produces liquidity services. The two sectors operate under the same technology, represented by a neoclassical production function $F(K_i, L_i)$, where K_i and L_i are sector i's capital stock and labor.[6] Let $K = K_1 + K_2$ and $L = L_1 + L_2$ be the economy's capital stock and labor force. In per capita terms, we have $l_1 k_1 + l_2 k_2 = k$, $l_1 + l_2 = 1$, and $\frac{F(K_i, L_i)}{L} = l_i f(k_i)$, where $k_i = \frac{K_i}{L_i}$, $l_i = \frac{L_i}{L}$, $k = \frac{K}{L}$ and $f(k_i) \equiv F(k_i, 1)$, $i = 1, 2$.

3.2 Firms

There is a large number of firms that produce goods, renting capital and labor taking as given the rental and wage rates. There are no dynamic adjustment costs, so firms

[5] The reader should keep in mind that λ is a jump variable.

[6] Recall that in the Appendix we show that Lemma 1 below remains valid with different technologies and consumption-leisure choice.

choose k_1 and l_1 that solve the following program

$$\max\{l_1 f(k_1) - (rk_1 + \mathrm{w})l_1\}$$

with the usual first order conditions

$$r = f'(k_1)$$
$$\mathrm{w} = f(k_1) - k_1 f'(k_1).$$

3.3 Banks

As with firms, there is a large number of banks that employ k_2 and l_2 to produce liquidity services. Differently from firms, banks own capital because households must deposit goods in bank accounts in order to acquire liquidity services. That is, for households to use m_2 units of liquidity services, households must deposit m_2 units of goods in bank accounts. These goods are owned by banks, and can be rented to productive units (firms and banks) as capital. When households transfer m_2 units of goods to a bank, the bank opens a checking account denominated in nominal units, $M_2 = Pm_2$, where P is the price level.

There is a fixed reserves requirement ratio, denoted by ζ, so that banks (as capital owners) can rent $(1-\zeta)\,m_2$ units of capital to firms and banks (as productive units). Let $k_b = (1-\zeta)\,m_2$ denote banks' capital, and note that k_b need not be equal to $k_2 l_2$, which is the amount of capital used by banks in the production of liquidity services.[7] Banks collect revenue rk_b from this renting activity. In addition, banks charge a fee, p, to provide liquidity services, and banks collect inflationary tax because demand deposit is a nominal asset. In short, banks' revenue is $(p + (\pi + r)\,(1-\zeta))m_2$.[8]

The fee, p, is charged because the inflation rate may be too low, making inflationary revenue too low as well.[9] For sufficiently high inflation rates, competition among banks will drive p down, even to negative levels. That is, we do not rule out $p < 0$. Until 1986, Regulation Q prevented such an outcome. But, since Regulation Q was eliminated in 1986 (Greenbaum and Thakor 2007, pp 456 and Feldstein 1995, pp 146), we think it is realistic to allow for $p < 0$.

There is no dynamic adjustment costs to banks either, so banks choose k_2 and l_2 that solve the following program:

---

[7] Note also that $k = k_h + k_b$, and, since k_2 will turn out to be equal to k, we have $k_b = k_2 - k_h$, which need not be equal to $l_2 k_2$.

[8] In other words, demand deposit plays three roles in this economy, and these three roles are captured by the relative price of liquidity services in terms of goods, $p + (\pi + r)\,(1-\zeta)$: (i) demand deposit is a nominal asset in the household's portfolio; (ii) demand deposit is a part of the economy's physical capital, and it is owned by the banks; and (iii) demand deposit is a service purchased by households. The corresponding three components of the relative price are: (i) the inflationary tax $\pi\,(1-\zeta)$; (ii) the rental revenue $r\,(1-\zeta)$ of banks renting their capital stock to firms; and (iii) the fee, p, charged by banks to provide the service.

[9] The assumption is that households demand deposit because of the benefits from the services provided by the bank to the deposit holder. As usual in production theory, the flow of services is assumed proportional to its stock. Examples of deposit services are check redemption, payment of bills, and ATMs.

$$\max\{(p + (\pi + r)(1 - \zeta))m_2 - (rk_2 + \mathrm{w})l_2\}$$
$$\text{s.t. } m_2 = l_2 f(k_2)$$

The constraint above is the technological constraint, stating that the amount of service provided must be equal to the stock of goods deposited. The first order conditions are:

$$r = (p + (\pi + r)(1 - \zeta)) f'(k_2)$$
$$\mathrm{w} = (p + (\pi + r)(1 - \zeta))(f(k_2) - k_2 f'(k_2)).$$

From the first order conditions of firms and banks we have $\frac{\mathrm{w}}{r} = \frac{f(k_1)-k_1 f'(k_1)}{f'(k_1)} = \frac{f(k_2)-k_2 f'(k_2)}{f'(k_2)}$, implying that $k_1 = k_2 (= k)$. That is, since isoquants coincide, the optimal capital-labor ratios must be equal. In turn, $r = f'(k) = (p+(\pi+r)(1-\zeta)) f'(k)$ implies that $p + (\pi + r)(1 - \zeta) = 1$, and the supply side behavior is characterized by

$$r = f'(k) \tag{2}$$
$$\mathrm{w} = f(k) - kf'(k). \tag{3}$$

3.4 Government

The only role of government is to collect seigniorage on the monetary base $b = m_1 + \zeta m_2$ and to redistribute it to households, so that

$$\chi = \dot{m}_1 + \zeta \dot{m}_2 + \pi(m_1 + \zeta m_2). \tag{4}$$

3.5 General equilibrium

Substituting (2) and (3) in firms' profits we see that profits must be zero, so $f(k) = rk + \mathrm{w}$, or

$$rk_h = f(k) - rk_b - \mathrm{w}. \tag{5}$$

Substituting (4) in (1) and using (5), we have

$$\begin{aligned}
\dot{k}_h + \dot{m}_1 + \dot{m}_2 &= r(k_h + m_1 + m_2) + \mathrm{w} + \dot{m}_1 + \zeta \dot{m}_2 + \pi(m_1 + \zeta m_2) \\
&\quad -c - pm_2 - (\pi + r)(m_1 + m_2) \\
\dot{k}_h + (1 - \zeta)\dot{m}_2 &= f(k) - rk_b - \mathrm{w} + \mathrm{w} - \pi(1 - \zeta)m_2 - c - pm_2 \\
\dot{k}_h + \dot{k}_b &= f(k) - (p + (r + \pi)(1 - \zeta))m_2 - c \\
\dot{k} &= f(k) - m_2 - c \\
\dot{k} &= l_1 f(k) - c,
\end{aligned}$$

where we recall that $k_b = (1-\zeta)m_2$, $p+(r+\pi)(1-\zeta) = 1$, and $m_2 = (1-l_1)f(k)$. In summary, the general equilibrium dynamic system is given by

$$\begin{aligned}\dot{k} &= l_1 f(k) - c,\\ \dot{\lambda} &= \lambda(\rho - f'(k))\\ \dot{m}_1 &= m_1(\sigma - \pi).\end{aligned}$$

Note that the long run capital is determined by $f'(k^*) = \rho$, and, hence, does not depend on σ. This is a due to the assumption of equal technologies in both sectors, and is not needed for Lemma 1 below.

The main result is the following:

Lemma 1 *The effect of* σ *on* W *is given by*

$$\frac{dW}{d\sigma} = \int_0^\infty e^{-\rho t}\lambda_t(\pi_t + r_t)\frac{db_t}{d\sigma}dt - \lambda_0\frac{dk_0}{d\sigma}, \tag{6}$$

where $b = m_1 + \zeta m_2$ *is the monetary base.*

Proof We have

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}\sigma} &= \int_0^\infty e^{-\rho t}\left(u_1\frac{\mathrm{d}c}{\mathrm{d}\sigma} + u_2\frac{\mathrm{d}m_1}{\mathrm{d}\sigma} + u_3\frac{\mathrm{d}m_2}{\mathrm{d}\sigma}\right)\mathrm{d}t\\ &= \int_0^\infty e^{-\rho t}\lambda\left(\frac{\mathrm{d}c}{\mathrm{d}\sigma} + (\pi + r)\left(\frac{\mathrm{d}m_1}{\mathrm{d}\sigma} + \frac{\mathrm{d}m_2}{\mathrm{d}\sigma}\right) + p\frac{\mathrm{d}m_2}{\mathrm{d}\sigma}\right)\mathrm{d}t,\end{aligned}$$

where the households' first order conditions were used in the second equality.

Since $l_1 f(k) - c - \dot{k} = 0$ and $(1 - l_1)f(k) - m_2 = 0$, we have

$$\begin{aligned}0 &= \int_0^\infty e^{-\rho t}\lambda\left(f\frac{\mathrm{d}l_1}{\mathrm{d}\sigma} + l_1 f'\frac{\mathrm{d}k}{\mathrm{d}\sigma} - \frac{\mathrm{d}c}{\mathrm{d}\sigma} - \frac{\mathrm{d}\dot{k}}{\mathrm{d}\sigma}\right)\mathrm{d}t\\ 0 &= \int_0^\infty e^{-\rho t}\lambda\left(-f\frac{\mathrm{d}l_1}{\mathrm{d}\sigma} + (1 - l_1)f'\frac{\mathrm{d}k}{\mathrm{d}\sigma} - \frac{\mathrm{d}m_2}{\mathrm{d}\sigma}\right)\mathrm{d}t\end{aligned}$$

Integrating by parts, and using the transversality condition and the boundedness of k,

$$\int_0^\infty e^{-\rho t}\lambda\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}k}{\mathrm{d}\sigma}\mathrm{d}t = \lambda_0\frac{\mathrm{d}k_0}{\mathrm{d}\sigma} - \int_0^\infty e^{-\rho t}\lambda\left(\frac{\dot{\lambda}}{\lambda} - \rho\right)\frac{\mathrm{d}k}{\mathrm{d}\sigma}\mathrm{d}t.$$

Substituting,

$$\frac{dW}{d\sigma} = \int_0^\infty e^{-\rho t}\lambda\left(\frac{dc}{d\sigma}+(\pi+r)\left(\frac{dm_1}{d\sigma}+\frac{dm_2}{d\sigma}\right)+p\frac{dm_2}{d\sigma}+f\frac{dl_1}{d\sigma}+l_1 f'\frac{dk}{d\sigma}-\frac{dc}{d\sigma}\right.$$
$$\left.+\left(\frac{\dot{\lambda}}{\lambda}-\rho\right)\frac{dk}{d\sigma}-f\frac{dl_1}{d\sigma}+(1-l_1)f'\frac{dk}{d\sigma}-\frac{dm_2}{d\sigma}\right)dt-\lambda_0\frac{dk_0}{d\sigma}$$
$$= \int_0^\infty e^{-\rho t}\lambda\left((\pi+r)\left(\frac{dm_1}{d\sigma}+\zeta\frac{dm_2}{d\sigma}\right)\right)dt-\lambda_0\frac{dk_0}{d\sigma},$$

where $p+(\pi+r)(1-\zeta)=1$ and $\rho-\frac{\dot{\lambda}}{\lambda}=r=f'$ were used. □

Notice that the arguments of the instantaneous utility function are m_1 and m_2, and not the monetary base b; that is, b shows up in (6) due to general equilibrium effects.[10] Also, since no assumption on the specific value of σ was made in the argument above, (6) is a global result. It is true for every value for σ.[11] Since the steady state capital stock does not depend on σ, (6) states that the impact of σ on welfare is the present value of the impact of σ on the demand for monetary base. The specific adjustment taking place after a change in σ does not matter: it is incorporated in the demand for monetary base.

Using (6), we have the following definition:

Definition 1 The **welfare cost of inflation** is given by the change in the initial capital stock, $\frac{dk_0}{d\sigma}$, satisfying

$$\int_0^\infty e^{-\rho t}\lambda_t(\pi_t+r_t)\frac{d\bar{b}_t}{d\sigma}dt=\lambda_0\frac{dk_0}{d\sigma}, \tag{7}$$

[10] It is interesting to note that Bali (2000) provides econometric evidence of Lemma 1 without deriving it: using either Bailey's partial equilibrium measure or the compensating variation in consumption measure, Bali shows that the welfare costs of inflation using money as currency-deposit is quite similar to the one obtained using M0, and three times smaller than the one obtained using M1. Bali's result is not exact because he uses the compensating variation in consumption, and in his model he does not consider that the banking sector uses up productive factors.

[11] In the derivation of (6), σ could be any exogenous variable. As an example, say that there is a purchase tax τ for some good. We would get:

$$\left.\frac{dW}{d\tau}\right|_{\tau=0}=\int_0^\infty e^{-\rho t}\lambda_t(\pi_t+r_t)\left.\frac{db_t}{d\tau}\right|_{\tau=0}dt-\lambda_0\frac{dk_0}{d\tau}.$$

The important distinction is that this result would apply only in a neighborhood of zero; in contrast, due to the particular role played by the parameter σ in monetary models - there is no first order condition associated with σ - (6) is a global result.

where $\bar{b}_t$ is the compensated demand for monetary base.

Equation (7) is the dynamic counterpart of the static compensating variation measure, which is defined as the compensating variation of income, the exogenous "endowment". It is important to note that (7) is a result: it is *the measure* of welfare cost of inflation compatible with the model presented here. In the literature (c.f. Lucas 2000) the general equilibrium measure used is the compensating variation in consumption, rather than endowment. Since consumption is endogenous, such a measure tends to overestimate the welfare costs of inflation, as agents will change their optimal choices of consumption and leisure in response to changes in σ.[12]

It is also important to compare (6) and (7) with the analogous expressions in Fischer (1981). Fischer derived an expression for $\frac{dW}{d\sigma}$ (expression (15), p 9) that includes the effect of σ on the inflation rate between today and tomorrow (Fischer used a two period model). Such effect is captured in the derivation of (6): consumers take the paths of prices and inflation as given; a change in σ triggers an once-and-for-all change in π and consumers re-optimize accordingly.

Fischer also derived an expression for the welfare costs of inflation, given by the compensating variation in initial wealth (expression (19), p 10). Such expression is readily comparable to (7) (in fact, it is Bailey's result that we derive in Proposition 1 below). But Fischer worked in a partial equilibrium set up, and assumed that current wealth, real return on bonds, future real income and relative prices were not affected by inflation. In contrast, we obtain a similar result by exploiting the general equilibrium adding up constraints.

4 Some implications

The following three corollaries of (6) are quite interesting.

Proposition 1 (Bailey) *The area below the inverse compensated money demand is the general equilibrium measure of the effects of inflation on welfare.*

Proof Let $R^* = r^* + \pi^*$ be the long-run nominal interest rate. Since λ is a jump variable, from (7) we have:

$$\lambda^* \frac{dk}{d\sigma} = \lambda^* R^* \frac{d\bar{b}}{d\sigma} \int_0^\infty e^{-\rho t} dt = \lambda^* R^* \frac{d\bar{b}}{d\sigma} \frac{1}{\rho}.$$

Integrating:

$$\rho \Delta k_{\text{Compensation for } \sigma} \equiv \rho \int_{-\rho}^{\sigma} \frac{dk}{d\sigma'} d\sigma' = \int_{-\rho}^{\sigma} R^* \frac{d\bar{b}}{d\sigma'} d\sigma' = \int_{\bar{b}(-\rho)}^{\bar{b}(\sigma)} R^* \left(\bar{b}\right) d\bar{b}.$$

[12] It is straightforward to incorporate the consumption-leisure choice in the simple model presented here. In fact, in the Appendix we do so in a cash-in-advance model.

Note that Δk is the compensation in wealth and $\rho \Delta k$ is the flow measure of the welfare cost of inflation. □

Or, in Bailey's words:

"This conclusion, that the area under the observed demand curve for real cash balances during an inflation measures the welfare costs of the reduction of these balances, applies **regardless** of the particular manner in which these costs affect real income and leisure." (Bailey 1956, p. 102, emphasis added.)

It is remarkable that Bailey realized that the area under the observed demand had a general equilibrium interpretation in the 1950s, and the literature failed to fully understand his contribution for over four decades. It is here that the assumption of equal technologies plays a role: with equal technologies, there is no transitional dynamics in k after a change in σ, since k^* is fixed by $f^{'}(k^*) = \rho$. The co-state variable jumps immediately and the economy moves to the new steady state. With different technologies we have to add the condition "if the adjustment process is instantaneous" to Proposition 1 above.

Note, however, that when technologies are different and the adjustment process is not instantaneous, Lemma 1 implies that the effects of the adjustment process are of second order of importance. That is, *Bailey's intuition remains true at first approximation*.[13]

Proposition 2 *The general equilibrium is Pareto-Optimum restricted to a given path of the real quantity of monetary base.*

Proof From (6), for a fixed k_0, we have $\frac{\mathrm{d}W}{\mathrm{d}\sigma} = \int_0^{\infty} e^{-\rho t} \lambda_t (\pi_t + r_t) \frac{\mathrm{d}b_t}{\mathrm{d}\sigma} \mathrm{d}t$. Hence, with the exception of monetary base, social costs and benefits of all other choice variables cancel out. □

In others words, this is a welfare maximizing economy restricted to the fact that the household is consuming less monetary services than the social optimum. That is, a central planner who is neither able to avoid inflation nor to induce the households to increase their money holdings, will do no better than the market. Thus, any policy that increases the present value of money holdings is welfare improving. For instance, increasing the tax rate on any good, from a initial situation with no tax, is welfare improving if and only if it increases the present value of the money holdings (regardless of any other impact on the economy).

The second best theorem states that, in general, we cannot give policy prescriptions for outcomes that are not Pareto optimal. Proposition 2 is a step forward from this dismal position: policies that increase the present value of money holdings are always to be recommended. As a consequence, we have:

Proposition 3 (Overbanking) *Starting from a situation with no tax on banking activities and with no reserve requirement ratio* ($\zeta = 0$)*, an increase in such a tax is welfare improving.*

[13] See also the argument in Sect. 5 below.

Proof Let τ be a tax on banking activities, and assume that the proceeds are transferred back to households in a lump-sum fashion. Repeating the steps of Lemma 1 for a fixed level of k_0, we get:

$$\left.\frac{\mathrm{d}W}{\mathrm{d}\tau}\right|_{\tau=0} = \int_0^\infty e^{-\rho t}\lambda_t(\pi_t + r_t)\left.\frac{\mathrm{d}m_{1t}}{\mathrm{d}\tau}\right|_{\tau=0}\mathrm{d}t,$$

Now, since $u_{23} < 0$, and $\left.\frac{\mathrm{d}m_{2t}}{\mathrm{d}\tau}\right|_{\tau=0} < 0$, we have $\left.\frac{\mathrm{d}m_{1t}}{\mathrm{d}\tau}\right|_{\tau=0} > 0$ and the result follows. □

Johnson (1968) provides the intuition behind the overbanking result:

> "The substantive point is that, because the private cost of holding currency (the interest forgone) substantially exceeds the social cost (raw material, value added, and policing), free competition in banking, by making the private and social cost of deposit holding coincide, would tend to produce a social non-optimum overallocation of resources to the provision of deposit money and underallocation of resources to the provision of currency holding. (...) On the assumption that currency cannot be issued other than as a non-interest-bearing asset, achievement of the "second best" welfare optimum would require a tax on the holding of deposit money at a rate somewhere between zero and the competitive interest rate in deposits (...)" (p 974).

The overbanking result goes against the view that banking regulations necessarily increase welfare costs of inflation. For instance, Fischer (1981, pp 18 and 19) argues that interest rate controls on nominal assets other than M0 increase the welfare costs of inflation. But Fischer overlooked one important general equilibrium effect (he used a partial equilibrium model): as he noted, with interest controls on, say, demand deposit, increases in inflation increase the private cost of holding demand deposit; but the resulting reduced demand for the service frees up productive factors from the banking sector to other sectors of the economy. Lemma 1 shows that these two effects cancel out whenever the interest rate controls are marginally active. And, since interest rate controls increase demand for M0, welfare costs of inflation decrease rather than increase. That is, if inflation is not very high, it may be the case that Regulation Q is a welfare improving regulation. For very high inflations, the impact of Regulation Q on welfare is ambiguous in general, and one would need a specific model to address this particular issue.

The condition $\zeta = 0$ is needed for the result above, since Lemma 1 is in terms of monetary base, $b = m_1 + \zeta m_2$, and for $\zeta > 0$ the net effect of τ on b is ambiguous. But in this model there is no risk of bank runs, so ζ is an additional restriction on banks. So the overbanking result should be read as follows: in the absence of any restriction to banking activities, the introduction of a restriction is welfare improving. Notice that this provides a rationale for $\zeta > 0$ that is not based on risk of bank runs.

A more important point is that it is believed that interest rate controls on nominal assets other than M0 (like Regulation Q) justify the use of M1 instead of M0 to measure welfare costs of perfectly anticipated inflation (Lucas 1981, p 44). The argument above shows that this belief is based on a partial equilibrium view. For welfare purposes, and in general equilibrium, money is the service that does not use up productive resources in its provision, and hence “money” must be M0, and not M1.

5 An estimate of welfare costs of inflation

We mentioned above that this paper complements ABE in showing the generality of Bailey’s result. ABE’s use of the income share of the financial sector can be criticized in two grounds: first, there are other costs of inflation that are not necessarily captured by this measure, like changes in consumption patterns to avoid holding money; second, the financial sector does provide the welfare improving service of financial intermediation. It is true that in ABE’s model the income share of the financial sector is equal to the area under demand for M0, but this need not be the case in other models. The result in this paper is that the area under demand for M0 measures accurately the costs of inflation and is not liable to the two critiques mentioned above. When one estimates the area under the demand for M0, one gets a measure that captures, as costs of inflation, changes in consumption patterns, time spent queuing in stores to buy goods and oversized financial departments, to name a few, with no underlying assumption that the financial sector is always a “bad” for the economy.[14]

In terms of numbers, we also have a complementary result. The estimate of the welfare cost of inflation in ABE is that it is bounded at less than 0.5% of US’s GDP, while our estimate, using ABE’s numbers, is that it is bounded at less than 0.2% of US’s GDP. The curiosity here is that our estimate is the area below the demand for M0, whereas ABE add a component to that area, failing to realize that this component is negligible when transition effects are considered.[15]

> “The second result is that at low to moderate inflation rates, the inflation distortion tax component, which is the difference between the total welfare cost and the misallocation component, is roughly from two to three times the misallocation component” (p 1298).

The misallocation component, that captures the welfare cost as resources are diverted from the goods sector toward the banking sector, is, in ABE’s model, the income share of the financial sector, which is computed as the area under the demand for M0, and hence is our estimate. The distortion tax component is due to the cash-in-advance restriction applied to investment goods, which in effect turns inflation into an increase in the shadow price of capital. ABE report the sum of these two components as total costs of inflation.

---

14 In our model, when the nominal interest rate is zero, the value added in the banking sector is not zero, because banking services and money are not perfect substitutes, as both are arguments of the household’s utility function. This is not the case in ABE’s model (and in the model presented in the Appendix), where a positive demand for money comes from a cash-in-advance constraint which implies that banking services and money are perfect substitutes. In those models, banks add no value to the economy.

15 ABE also fail to realize that an analogue of Lemma 1 is valid in their model as well. In fact, the model presented in the Appendix below includes ABE’s model as a special case.

But Lemma 1 above suggests that the distortion tax component is negligible at first order. And taking numbers from ABE's Figure 9, page 1298, and from ABE's Figure 5, page 1295, we see that total welfare cost, considering transitional effects, and misallocation costs are virtually identical. Once dynamic effects are considered, the distortion tax component is practically eliminated. In other words, total welfare cost is given by the misallocation cost, which is what we expect from Lemma 1.

Note that on top of being in line with Lemma 1, ABE's numbers are also in line with Bailey's intuition: although Bailey's proposition is true only with no transitional effects, ABE's numbers show that it remains approximately true when transitional effects are included.

In summary, the main result in this paper is a derivation of a measure of welfare costs of inflation. In comparison with the literature, the difference is the use of M0 instead of M1. In terms of numbers, ABE's figures together with our result indicate that the costs of perfectly anticipated inflation are bounded at less of 0.2% of GDP, when inflation is low as it is for the US.

Acknowledgments A previous version of this work circulated with the title "Bailey's Rule for The Welfare Cost of Inflation: a Theoretical Foundation." This work benefited from conversations with Marco Antonio Bonomo and Marcos de Barros Lisboa. Ricardo Cavalcanti and Rubens Penha Cysne read a preliminary version. Their comments and suggestions were essential to the completion of the paper. We also thank the feedback of an anonymous referee. Pessôa thanks the Brazilian Government's research assistance agencies - CAPES and CNPq -, and the special program - PRONEX - for financial support.

Appendix

Here we illustrate the generality of Lemma 1 above. A similar result is derived in a cash in advance model. There is a continuum of goods index by $z \in [0, 1]$. Goods are identical from the supply point of view, which means that the producer price P_t is the same, regardless of the type. There are two sectors in the economy: the goods sector, that produces the goods, and the banking sector, that produces a banking service. Each good can be either a cash good or a credit good. Households pay $P_t(z) = P_t$ acquire a cash good and since they must hold cash in order to buy it, then this means they face a cost $P_t(1+R_t)$ from holding the cash. When buying a good as a credit good, households pay P_t to the good's producer plus the intermediation services cost, of $\mathfrak{R}(z)$ units of banking services. Thus, the effective cost of a credit good is $P_t(z) = P_t(1 + p\mathfrak{R}(z))$. Under the assumption that both sectors operate under the same technology, units of banking services correspond to units of goods. The total *per capita* production of goods and services is $f(k_t, n_t)$, where n_t is the *per capita* supply of labor services. Moreover, the transaction services cost function is increasing in the index z and $\mathfrak{R}(0) = 0$. At any moment there is a cut-off index, $\mathfrak{z}_t$, such that any good whose index is lower than the cut-off is bought as credit good, and the others are bought as cash.

The representative household solves

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t) \tag{8}$$

where

$$c_t = \left(\int_0^1 c_t^{\frac{\theta-1}{\theta}}(z)\,\mathrm{d}z \right)^{\frac{\theta}{\theta-1}}$$

is an aggregator function that defines the unit of consumption.

The household faces two sorts of restrictions. One is the cash-in-advance and the other is the budget constraint. Before going to the goods market, the household can go to the credit market and get some cash at no cost. Let M_t, B_t, and X_t be, respectively, the nominal quantity of money and bonds in the household portfolio, and the nominal value of the government's transfer. The cash-in-advance restriction is

$$\frac{M_t + X_t}{P_t} + \frac{B_t}{P_t} - \frac{B_{t+1}}{P_t(1+R_t)} \geq \frac{1}{P_t} \int_{\mathfrak{z}_t}^1 P_t(z)(c_t(z) + i_t(z))\mathrm{d}z. \tag{9}$$

The left hand side of (9) is the amount of cash carried for consumption before going to the goods market, and the right hand side is the nominal cost of cash goods for consumption and investment. The budget constraint is

$$\frac{M_t + X_t}{P_t} + \frac{B_t}{P_t} + \mathrm{w}_t n_t + r_t k_t \geq \frac{1}{P_t} \int_0^1 P_t(z)(c_t(z) + i_t(z))\mathrm{d}z + \frac{M_{t+1}}{P_t} + \frac{B_{t+1}}{P_t(1+R_t)}. \tag{10}$$

Capital accumulates whenever investment exceeds depreciation:

$$k_{t+1} = i_t + (1-\delta)k_t, \tag{11}$$

where i_t is an aggregator function that defines the investment good

$$i_t = \left(\int_0^1 i_t^{\frac{\theta-1}{\theta}}(z)\,\mathrm{d}z \right)^{\frac{\theta}{\theta-1}}.$$

Taking the limit $\theta \to 0$, we have ABE's model; the limit $\theta \to 1$ reproduces Gillman (1993) model if an economy without capital is considered. If the cut-off index, $\mathfrak{z}_t$, is fixed and if there are neither banking services nor transaction services, the model reproduces Lucas and Stokey's (1983) economy under certainty, and if there are no credit goods, the model generates Stockman's (1981) model. Additionally, if capital is a credit good without transaction cost, Lucas (1981) model under certainty is obtained.

First order conditions:

$$\left(\frac{c_t(z)}{c_t}\right)^{-\frac{1}{\theta}} = \left(\frac{i_t(z)}{i_t}\right)^{-\frac{1}{\theta}} = (1+\Re(z))\,\frac{P_t}{Q_t} \text{ if } z \le \mathfrak{z}_t \tag{12}$$

$$\text{and } \left(\frac{c_t(z)}{c_t}\right)^{-\frac{1}{\theta}} = \left(\frac{i_t(z)}{i_t}\right)^{-\frac{1}{\theta}} = (1+R_t)\,\frac{P_t}{Q_t} \text{ if } z > \mathfrak{z}_t, \tag{13}$$

if the household faces the price $(1+\Re(z))\,P_t$ when $z \le \mathfrak{z}_t$, and faces the effective price $(1+R_t)\,P_t$ when $z > \mathfrak{z}_t$, where

$$Q_t \equiv P_t(1+\tau_t) \equiv P_t\left[\int_0^{\mathfrak{z}_t}(1+\Re(z))^{1-\theta}\mathrm{d}z + (1-\mathfrak{z}_t)(1+R_t)^{1-\theta}\right]^{\frac{1}{1-\theta}} \tag{14}$$

is the effective price index faced by the household.

Let $\beta^t\lambda_t\mu_t$, $\beta^t\lambda_t$, and $\beta^t\lambda_t q_t$ be respectively the Lagrange multipliers of (9), (10), and (11). Recalling that $P_t(z) = (1+\Re(z))\frac{P_t(z)}{P_t}$ if $z \le \mathfrak{z}_t$ and that $P_t(z) = P_t$ if $z > \mathfrak{z}_t$, it follows that the first-order conditions for the flow variables, consumption and investment, are

$$\begin{aligned} u_1(c_t, 1-n_t)c_t^{\frac{1}{\theta}}c_t^{-\frac{1}{\theta}}(z) &= \lambda_t(1+\mu_t) \\ \text{and } q_t i_t^{\frac{1}{\theta}} i_t^{-\frac{1}{\theta}}(z) &= 1+\mu_t \text{ if } z > \mathfrak{z}_t; \end{aligned} \tag{15}$$

$$\begin{aligned} u_1(c_t, 1-n_t)c_t^{\frac{1}{\theta}}c_t^{-\frac{1}{\theta}}(z) &= \lambda_t(1+\Re(z)) \\ \text{and } q_t i_t^{\frac{1}{\theta}} i_t^{-\frac{1}{\theta}}(z) &= 1+\Re(z) \text{ if } z \le \mathfrak{z}_t. \end{aligned} \tag{16}$$

The first-order conditions for the labor supply and the cut-off index are

$$\begin{aligned} u_2(c_t, 1-n_t) &= \lambda_t \mathrm{w}_t \\ 1+\mu_t &= 1+\Re(\mathfrak{z}_t). \end{aligned}$$

This last condition states that the relative price of money in units of bonds is equal to the credit cost of the cut-off good. This relative price should be equal to the nominal interest rate in order to keep the budget restriction bounded; otherwise it would be possible to gain money selling (or buying) cash the $\mathfrak{z}_t$ good, and buying (or selling) it as credit good. At each instant the cut-off good is determined with the aim of meeting this non-arbitrage condition. That is

$$\mu_t = R_t = \Re(\mathfrak{z}_t). \tag{17}$$

As Gillman (1993) stressed, (17) is a Baumol-type condition which equates the marginal cost of holding money with the marginal transaction cost.

Substituting (12) and (13) into (15) and (16), and using (14) and (17), we have

$$u_1(c_t, 1-n_t) = \lambda_t(1+\tau_t) \text{ and } q_t = 1+\tau_t. \tag{18}$$

The Euler equations for capital and bonds are respectively

$$\lambda_t(1+\tau_t) = \beta\lambda_{t+1}(1+\tau_{t+1})(1-\delta+\frac{r_{t+1}}{1+\tau_{t+1}}) \tag{19}$$

and,

$$\lambda_t = \beta\lambda_{t+1}(1+R_{t+1})\frac{P_t}{P_{t+1}}.$$

It is clear from (19), after substituting (18), that the cash-in-advance restriction on investment acts as a distortion taxation on capital.

Lemma 1

Proposition 4 *The effect on welfare of an increase in the growth rate of the nominal quantity of money is*

$$\sum_{t=0}^{\infty}\beta^t\lambda_t R_t\frac{dm_t}{d\sigma} + \lambda_0\left[r_0+(1+\tau_0)(1-\delta)\right]\frac{dk_0}{d\sigma} \tag{20}$$

Proof From (8), and (15), (16), (12), and (13), we have

$$\frac{\mathrm{d}W}{\mathrm{d}\sigma} = \sum_{t=0}^{\infty}\beta^t\lambda_t\left[\int_0^{\mathfrak{z}_t}(1+\Re(z))\frac{\mathrm{d}c_t(z)}{\mathrm{d}\sigma}\mathrm{d}z + \int_{\mathfrak{z}_t}^{1}(1+R_t)\frac{\mathrm{d}c_t(z)}{\mathrm{d}\sigma}\mathrm{d}z - \mathrm{w}_t\frac{\mathrm{d}n_t}{\mathrm{d}\sigma}\right]. \tag{21}$$

The GDP equation for this economy is

$$f(k_t,n_t) - \int_0^{\mathfrak{z}_t}(1+\Re(z))(c_t(z)+i_t(z))\mathrm{d}z - \int_{\mathfrak{z}_t}^{1}(c_t(z)+i_t(z))\mathrm{d}z = 0,$$

which means that

$$0 = \sum_{t=0}^{\infty}\beta^t\lambda_t\left[r_t\frac{\mathrm{d}k_t}{\mathrm{d}\sigma} + \mathrm{w}_t\frac{\mathrm{d}n_t}{\mathrm{d}\sigma} - \int_0^{\mathfrak{z}_t}(1+\Re(z))(\frac{\mathrm{d}c_t(z)}{\mathrm{d}\sigma}+\frac{\mathrm{d}i_t(z)}{\mathrm{d}\sigma})\mathrm{d}z\right]$$
$$-\sum_{t=0}^{\infty}\beta^t\lambda_t\left[\Re(\mathfrak{z}_t)(c_t(\mathfrak{z}_t)+i_t(\mathfrak{z}_t))\frac{\mathrm{d}\mathfrak{z}_t}{\mathrm{d}\sigma} + \int_{\mathfrak{z}_t^1}(\frac{\mathrm{d}c_t(z)}{\mathrm{d}\sigma}+\frac{\mathrm{d}i_t(z)}{\mathrm{d}\sigma})\mathrm{d}z\right]. \tag{22}$$

Adding (22) to (21),

$$\frac{\mathrm{d}W}{\mathrm{d}\sigma} = \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\int_{\mathfrak{z}_t}^{1} R_t \frac{\mathrm{d}c_t\,(z)}{\mathrm{d}\sigma} \mathrm{d}z - \mathfrak{R}(\mathfrak{z}_t) c_t(\mathfrak{z}_t) \frac{\mathrm{d}\mathfrak{z}_t}{\mathrm{d}\sigma} \right]$$
$$+ \sum_{t=0}^{\infty} \beta^t \lambda_t \left[r_t \frac{\mathrm{d}k_t}{\mathrm{d}\sigma} - \int_{0}^{\mathfrak{z}_t} (1 + \mathfrak{R}(z)) \frac{\mathrm{d}i_t\,(z)}{\mathrm{d}\sigma} \mathrm{d}z \right.$$
$$\left. - \int_{\mathfrak{z}_t}^{1} \frac{\mathrm{d}i_t\,(z)}{\mathrm{d}\sigma} \mathrm{d}z - \mathfrak{R}(\mathfrak{z}_t) i_t(\mathfrak{z}_t) \frac{\mathrm{d}\mathfrak{z}_t}{\mathrm{d}\sigma} \right]. \tag{23}$$

From the first-order condition for the investment, we have

$$(1 + \tau_t) i_t = \int_{0}^{\mathfrak{z}_t} (1 + \mathfrak{R}(z)) i_t(z) \mathrm{d}z + (1 + R_t) \int_{\mathfrak{z}_t}^{1} i_t(z) \mathrm{d}z,$$

which means that

$$0 = \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\int_{0}^{\mathfrak{z}_t} (1 + \mathfrak{R}(z)) \frac{\mathrm{d}i_t\,(z)}{\mathrm{d}\sigma} \mathrm{d}z + (1 + R_t) \int_{\mathfrak{z}_t}^{1} \frac{\mathrm{d}i_t\,(z)}{\mathrm{d}\sigma} \mathrm{d}z \right]$$
$$+ \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\frac{\mathrm{d}R_t}{\mathrm{d}\sigma} \int_{\mathfrak{z}_t}^{1} i_t(z) \mathrm{d}z - i_t \frac{\mathrm{d}(1 + \tau_t)}{\mathrm{d}\sigma} - (1 + \tau_t) \frac{\mathrm{d}i_t}{\mathrm{d}\sigma} \right]. \tag{24}$$

Since

$$\frac{\mathrm{d}R_t}{\mathrm{d}\sigma} \int_{\mathfrak{z}_t}^{1} i_t(z) \mathrm{d}z - i_t \frac{\mathrm{d}(1 + \tau_t)}{\mathrm{d}\sigma} = 0,$$

adding (24) to (23) yields

$$\frac{\mathrm{d}W}{\mathrm{d}\sigma} = \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\int_{\mathfrak{z}_t}^{1} R_t \frac{\mathrm{d}}{\mathrm{d}\sigma} (c_t(z) + i_t(z)) \mathrm{d}z - \mathfrak{R}(\mathfrak{z}_t)(c_t(\mathfrak{z}_t) + i_t(\mathfrak{z}_t)) \frac{\mathrm{d}\mathfrak{z}_t}{\mathrm{d}\sigma} \right]$$
$$+ \sum_{t=0}^{\infty} \beta^t \lambda_t \left[r_t \frac{\mathrm{d}k_t}{\mathrm{d}\sigma} - (1 + \tau_t) \frac{\mathrm{d}i_t}{\mathrm{d}\sigma} \right]. \tag{25}$$

The second line in (25) can be written as

$$\begin{aligned}
&\sum_{t=0}^{\infty}\beta^t\lambda_t\left[r_t\frac{\mathrm{d}k_t}{\mathrm{d}\sigma}-(1+\tau_t)\left(\frac{\mathrm{d}k_{t+1}}{\mathrm{d}\sigma}-(1-\delta)\frac{\mathrm{d}k_t}{\mathrm{d}\sigma}\right)\right]\\
&\quad=\sum_{t=1}^{\infty}\left[\beta^t\lambda_t r_t+\beta^t\lambda_t(1+\tau_t)(1-\delta)-\beta^{t-1}\lambda_{t-1}(1+\tau_{t-1})\right]\frac{\mathrm{d}k_t}{\mathrm{d}\sigma}\\
&\qquad+\lambda_0\left[r_0+(1+\tau_0)(1-\delta)\right]\frac{\mathrm{d}k_0}{\mathrm{d}\sigma}\\
&\quad=\lambda_0\left[r_0+(1+\tau_0)(1-\delta)\right]\frac{\mathrm{d}k_0}{\mathrm{d}\sigma},
\end{aligned}\tag{26}$$

in which the second equality follows from the first-order condition for capital accumulation, equation (19), the transversality condition, and because capital is bounded. Substituting (26) into (25), we have

$$\begin{aligned}
\frac{\mathrm{d}W}{\mathrm{d}\sigma}&=\sum_{t=0}^{\infty}\beta^t\lambda_t R_t\frac{\mathrm{d}}{\mathrm{d}\sigma}\left(\int_{\mathfrak{z}_t}^{1}(c_t(z)+i_t(z))\mathrm{d}z\right)+\lambda_0\left[r_0+(1+\tau_0)(1-\delta)\right]\frac{\mathrm{d}k_0}{\mathrm{d}\sigma}\\
&=\sum_{t=0}^{\infty}\beta^t\lambda_t R_t\frac{\mathrm{d}m_t}{\mathrm{d}\sigma}+\lambda_0\left[r_0+(1+\tau_0)(1-\delta)\right]\frac{\mathrm{d}k_0}{\mathrm{d}\sigma}.
\end{aligned}$$

□

Notice that the absence of inside money in the model above, which means that we cannot distinguish M0 from M1. But also notice that the inclusion of inside money can be done in the same lines that we did in the main text.

References

Aiyagari, S.R., Braun, R.A., Eckstein, Z.: Transaction services, inflation, and welfare. J Polit Econ **106**, 1274–1301 (1998)

Bali, T.: U.S. money demand and the welfare cost of inflation in a currency-deposit model. J Econ Bus **52**, 233–258 (2000)

Bailey, M.J.: The welfare cost of inflationary finance. J Polit Econ **64**, 93–110 (1956)

Dotsey, M., Ireland, P.: The welfare cost of inflation in general equilibrium. J Monet Econ **37**, 29–47 (1996)

English, W.B.: Inflation and financial sector size. J Monet Econ **44**, 379–400 (1999)

Feenstra, R.C.: Functional equivalence between liquidity cost and the utility of money. J Monet Econ **17**, 271–291 (1986)

Feldstein, M.: The cost and benefit of going from low inflation to price stability. In: Romer, C.D., Romer, D.H. (eds.) Reducing Inflation: Motivation and Strategy, NBER Studies in Business Cycles, vol. 30. Chicago, Chicago University Press (1997)

Fischer, S.: Towards an understanding of the cost of inflation II. In: Cornegie-Rochester Conference Series in Public Policy, vol. 15, pp.5–42 (1981)

Friedman, M.: The optimum quantity of money. In: The Optimum Quantity of Money and Others Essays, pp. 1–50 (1969)

Gillman, M.: The welfare cost of inflation in a cash-in-advance economy with costly credit. J Monet Econ **31**, 97–115 (1993)

Greenbaum, S.I., Thakor, A.V.: Contemporary Financial Intermediation, 2nd edn. Academic Press Advanced Finance, London (2007)

Johnson, H.G.: Problems of efficiency in monetary management. J Polit Econ **76**, 971–990 (1968)

Jones, B., Asaftei, G., Wang, L.: Welfare costs of inflation in a general equilibrium model with currency and interest-bearing deposits. Macroecon Dyn **8**, 493–517 (2004)

Lucas, R.E.: Discussion of Stanley Fischer, towards an understanding of the cost of inflation II. In: Cornegie-Rochester Conference Series in Public Policy, vol. 15, pp. 43–52 (1981)

Lucas, R.E.: Inflation and welfare. Econometrica **68**, 247–274 (2000)

Lucas, R.E., Jr., Nancy, L.S.: Optimal fiscal policy in a economy without capital. J. Monet. Econ. **12**, 55–93 (1983)

Sidrauski, M.: Rational choice and patterns of growth in a monetary economy. Am Econ Rev (Papers and Proceeding) **57**, 534–544 (1967)

Stockman, A.C.: Anticipated inflation and the capital stock in a cash-in-advance economy. J Monet Econ **8**, 387–393 (1981)