Endogenous Money, Inflation and Welfare

Espen Henriksen Finn Kydland

January 2005

What are the welfare gains from adopting monetary policies that reduce the inflation rate? This is among the classical questions in monetary economics and a matter of great interest for researchers and policymakers alike. The traditional way to answer this question and provide an estimate of the welfare cost has been to quantify the opportunity cost of holding non-interest-bearing money and the indirect cost of holdings of real balances that are too low.

In this paper we ask the same question, but we answer it within a model economy with endogenous money supply, i.e. where the ratio of $M1$ (the sum of currency holdings and demand deposits) to $M0$ (the monetary base) is the outcome of households' optimization. Using our model economy, in which currency, reserves, and deposits play distinct roles, this paper makes five contributions.

First, it provides a welfare measure of inflation that has some new features. The welfare cost of inflation reflects three distinct sources. The first is the opportunity cost of holding currency and demand deposits, the second is the cost incurred in order to avoid the inflation tax on currency and deposits, and the third stems from the fact that inflation also distorts labor supply and capital accumulation and, therefore, total output. Among our findings is that for most inflation rates, the opportunity cost of holding currency and demand deposits accounts for less than half of the estimated welfare costs of inflation, and, rather, that most of the costs can be accounted for by actions taken to avoid the inflation tax.

The second contribution is to endogenize the size of the banking sector as a function of the inflation rate. Proxies for its size are holdings of real deposits and use of real deposits for purchases. These measures of the banking sector are also closely related to the enodogenized velocity of the monetary base.

The third contribution is to point out the existence of a lower bound of inflation where inside money is held. This inflation rate is weakly greater than that associated with the Bailey-Friedman rule. At this lower bound of inflation, outside money ceases to exist, and $M1 = M0$.

The fourth contribution is to provide a framework within which we can also analyze the welfare gains of potential improvements in transaction technology. As the costs of transactions are reduced, so will also the cost of inflation.

And fifth, without considering alternative tax schemes, this paper makes a strong case for the inefficiency of inflation as a means of government revenue. In most other models that estimate the cost of inflation, there is a close link between holdings of real balances, the welfare cost of inflation, and the object of government taxation. Within the model economy where, in which the agents can take other actions and incur other costs to avoid inflation, this close link is broken.

Prices and output are fully flexible. Building on Freeman and Kydland (2000) with inside money in the spirit of Freeman and Huffman (1991), households can purchase consumption goods using either currency or interest-bearing deposits. Two transaction costs affect these decisions. One is the time cost of replenishing money balances, and the other is the fixed cost of using deposits for purchases. Faced with these two costs and factors that vary over time in equilibrium, such as over the business cycle, households make decisions that in the aggregate determine velocity of money and the money multiplier.

Closest in spirit to our paper are Aiyagari, Braun and Eckstein (1998) and Gillman (1993), who develop cash-in-advance models where credit services cost resources. Other related papers include Dotsey and Ireland (1996), who answer the question with a cash-in-advance model with shopping time, and Jones, Asaftei and Wang (2004) and Simonsen and Cysne (2001), who generalize money-in-the-utility-function models to include interest-bearing assets. Natural benchmarks for estimated welfare cost of inflations are Cooley and Hansen (1989) and Lucas (2000), who report the gain from reducing inflation from 10% to 0% to be consumption equivalent of 0.38% and around 1%, respectively.

1 Model economy

1.1 Production

Output is given by a constant-returns-to-scale production function with two inputs: capital (k_t) and labor (l_t);

$$y_t = z_t f\left(k_t, l_t\right),$$

The production function is Cobb-Douglas with share parameter α.

The law of motion for the technology level z_t is

$$z_t = \rho z_{t-1} + \varepsilon_t, \qquad z_t \sim \mathcal{N}\left(\mu, \sigma^2\right),\ \mu > 0.$$

The depreciation rate is denoted δ, so the law of motion for the capital stock is

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

where i_t is gross investment.

1.2 Government

The government controls the supply of intrinsically worthless fiat money. The law of motion for the money stock is

$$M_t = \xi M_{t-1}.$$

Net revenues from printing money are transferred to households in a lump sum fashion in the aggregate amount of

$$X_t = \left(\xi - 1\right) M_{t-1}.$$

1.3 Financial intermediation

Banks accept deposits (h), hold required reserves fraction θ as cash, and invest the proceeds in capital. Free entry ensures zero profit, and the rate of return on deposits ($\tilde{r}_{t+1}$) is therefore a linear combination of the real return on capital (r_{t+1}) and the return on holding currency (p_t/p_{t+1}):

$$\tilde{r}_{t+1} = \left(1-\theta\right) r_{t+1} + \theta \frac{p_t}{p_{t+1}}.$$

Total capital stock per household equals the sum of nonintermediated capital (a) and the share of deposits which banks are not required to hold as reserves:

$$k_{t+1} = a_t + \left(1-\theta\right) \frac{h_t}{p_t}. \tag{1}$$

The total stock of fiat money (the monetary base) is by definition equal to the combined stocks of currency (m) and reserves:

$$M_t = m_t + \theta h_t, \tag{2}$$

whereas the total money stock ($M1$) is the sum of nominal deposits and currency. It can be rewritten as the product of the monetary base and the money multiplier:

$$M1_t = m_t + h_t = M_t \left[1 + \frac{h_t (1-\theta)}{m_t + \theta h_t} \right].$$

1.4 Household's problem

There is a continuum of good types of measure c_t^*, ordered by size and indexed by j over $[0, 1]$. The utility of the representative household is given by the following function

$$\mathrm{E} \sum_{t=0}^{\infty} \beta^t u \left[\min \left(\frac{c_t (j)}{(1-\omega) j^{-\omega}} \right), d_t \right]. \tag{3}$$

For a given level c_t^* of total consumption the household will distribute consumption over the continuum of goods according to the following optimizing rule for $c_t (j)$ over $[0, 1]$

$$c_t(j) = (1-\omega) j^{-\omega} c_t^*. \tag{4}$$

There are three vehicles of savings available to the household: nonintermediated capital (a_t), nominal bank deposits (h_t), and currency (m_t). Both bank deposits (h_t) and currency (m_t) can be used to purchase consumption goods, but use of deposits incurs a fixed cost, denoted γ.

At the beginning of each period, households choose their real money holdings and ratio of deposits to currency, which both are maintained throughout the period. In order to purchase a given amount of consumption goods, each household must replenish their money balances n times such that n multiplied by the money balances held equals the worth of the consumption goods. Each time a household replenishes its money balances it spends φ units of time. Total time spent on those transactions in a period then equals φn_t.

Because of this fixed cost of using deposits for purchases, the deposit rate of return net of transaction costs goes to negative infinity as the purchase size j goes to zero. Therefore, some j^* exists below which currency is a preferred means of payment and above which deposits are preferred.

For the representative household, the per-period demand for real deposits is

$$n_t \frac{h_t}{p_t} = \int_{j^*}^{1} c_t(j)\mathrm{d}j = \int_{j^*}^{1} (1-\omega)\, j^{-\omega} c_t^* \mathrm{d}j = \left(1 - (j^*)^{1-\omega}\right) c_t^* \tag{5}$$

and the per period demand for real fiat-money balances is:

$$n_t \frac{m_t}{p_t} = \int_{0}^{j^*} c_t(j)\mathrm{d}j = \int_{0}^{j^*} (1-\omega)\, j^{-\omega} c_t^* \mathrm{d}j = (j^*)^{1-\omega} c_t^*. \tag{6}$$

Substituting from the optimal rule (4), the household's instantaneous utility can be written as

$$u\left(c_t^*, d_t\right) = \frac{1}{1-\nu}\left[\left(c_t^*\right)^{\zeta}\left(d_t\right)^{1-\zeta}\right]^{1-\nu}, \tag{7}$$

and the budget constraint is given by

$$c_t^* + a_t + \frac{h_t}{p_t} + \frac{m_t}{p_t} + \gamma\left(1 - j_t^*\right) = w_t l_t + r_t a_{t-1} + \tilde{r}_t \frac{h_{t-1}}{p_{t-1}} + \frac{m_{t-1}}{p_t} + \frac{X_t}{p_t}. \tag{8}$$

The time available is spent on leisure (d_t), labor (l_t), and replenishment of money balances (φn_t). Normalizing available time to 1, the time constraint is

$$1 = d_t + l_t + n_t \varphi. \tag{9}$$

1.5 Equilibrium

At any point in time, the economy is characterized by the technology level (z), the growth of money stock (ξ), lagged price level (p_{-1}), lagged holdings of non-intermediated capital (a_{-1}), lagged deposits (h_{-1}) and lagged currency holdings (m_{-1}). The vector $s \in S \subset \mathbb{R}^6$ is the state of the economy.

An equilibrium is a sequence of allocations $\left\{l(s), a(s), h(s), m(s)\right\}$, a sequence of prices $\left\{r(s), w(s), p(s)\right\}$ such that

1. Each household solves its optimization problem (3) subject to its liquidity constraints (5) and (6), budget constraint (8) and time constraint (9).

2. The goods market, the asset market for capital (1) and the market for fiat money (2) clear.

2 Calibrating the model

In steady state, investment is one-quarter of output and the annual capital-output ratio 2.5. The depreciation rate is then calibrated to 0.025. The parameter α in the production function is calibrated such that labor share of national income is 0.64.

Setting the average allocation of households' time (net of sleep and personal care) to market activity equal to 0.30 restricts the value of the utility parameter ζ. The risk aversion parameter ν equals 2, and the reserve requirement ratio θ is set equal to 0.10.

In order to calibrate the parameter ω of the utility function, the model business cycle properties resulting from different parameter values are compared. Consistent with Prescott (1986), the autocorrelation coefficient ρ in the technology process is set equal to 0.95, and the shocks have a standard deviation of 0.0076.

2.1 Utility function

In Figure 1, the decision rule $c_t(j)$ in Eq. (4) is plotted for three different values of ω for the case of $c_t^* = 1$. For $\omega > -1$, the amount consumed of different goods is a concave function of the size of the goods, whereas for $\omega < -1$, this amount is a convex function of the size of the goods.

Combining Equations (5) and (6) gives us the cut-off size for purchases, above which deposits are preferred to currency:

$$j^* = \left(1 + \frac{h_t}{m_t}\right)^{\frac{1}{\omega - 1}}. \tag{10}$$

The derivative of j^* is negative, implying that, loosely speaking, the more convex $c_t(j)$ is, the higher is j^*, or conversely, the more concave $c_t(j)$ is, the lower is j^*.

Note that Equations (5) and (6) combined with (10) imply

$$\int_{j^*}^{1} c_t(j)\mathrm{d}j = \left(1 + \frac{h_t}{m_t}\right)^{-1} c_t^*$$

and

$$\int_{0}^{j^*} c_t(j)\mathrm{d}j = \left(1 + \frac{m_t}{h_t}\right)^{-1} c_t^*$$

In words, whereas the cut-off size of purchases for which deposits are preferred to currency is a function of ω, the share of total consumption c_t^* for which deposits are preferred to currency (and vice versa) only depends on the deposit-to-currency ratio.

2.2 Business cycle properties and the parameter ω

In order to quantify the parameter ω, we examine the model's business cycle behavior under different values of ω and for three different policy regimes. Under the first, Policy A, the growth rate of fiat money is fixed at 3 percent in every period. Under the second, Policy B, serially uncorrelated shocks have been added to the supply of fiat money, with a standard deviation of 0.5 percent. And under the third, Policy C, the growth rate of the monetary base is serially correlated with an autoregressive parameter of 0.7, and the shocks have a standard deviation of 0.2.

For these three policies, we examine the business cycle properties for $\omega = \{-0.75, -1.0, -1.5\}$. Table 1 presents the contemporaneous correlations with output.

Table 1: Contemporaneous Correlations with Output

		$M1$	p	R_{nom}	c	i	l
"Policy A" :	$\omega = -0.75$	1	-0.38	-0.73	0.96	0.99	0.99
	$\omega = -1.00$	1	-0.54	-0.29	0.96	0.99	0.99
	$\omega = -1.50$	1	-0.76	0.12	0.96	0.99	0.99
"Policy B" :	$\omega = -0.75$	0.89	-0.09	-0.73	0.96	0.99	0.99
	$\omega = -1.00$	0.85	-0.15	-0.29	0.96	0.99	0.99
	$\omega = -1.50$	0.78	-0.27	0.12	0.96	0.99	0.99
"Policy C" :	$\omega = -0.75$	0.82	-0.07	-0.36	0.96	0.99	0.99
	$\omega = -1.00$	0.78	-0.11	-0.09	0.96	0.99	0.99
	$\omega = -1.50$	0.72	-0.21	0.02	0.96	0.99	0.99

Notice first that the real variables c, i and l hardly are affected by changes in monetary policy or the curvature of the utility function. We also see that $M1$ is strongly correlated with real output. Under Policy A, in which there is no randomness, the correlation is 1. Under the other two policy regimes, $M1$ is slightly less tightly correlated, but still highly correlated.

An interesting pattern is the countercyclical behavior of price level (Figure 2). We see that, for all policies, the price level is more countercyclical the more negative ω is. Further, from the figures we see that, for $\omega = -1.5$, prices are neither leading nor lagging the cycle. Of the three values of ω, for $\omega = -1.5$ the price movements are closest to those reported for U.S. data by Gavin and Kydland (1999).

We notice also that for $\omega = -1.5$ the cyclical behavior of the nominal interest rate is closer to what is observed in the data for $\omega = -1.5$ (Figure 3). Consistent with reported business cycle statistics, for $\omega = -1.5$ the nominal interest rate is weakly procyclical. In contrast, for the other two values of ω, the nominal rate of return (R_{nom}) is countercyclical.

Until we have data from which we can map to ω, we choose $\omega = -1.5$ as our benchmark value, as this value gives business cycle statistics closest to those observed along well-defined dimensions.

3 Quantitative findings

The benchmark economy is calibrated such that for gross annual inflation rate equal to 1.03 the currency-to-deposit ratio is equal to 9 and the non-reserve portion of $M1$ divided by the capital stock is 0.05. This gives us calibrated values for $\gamma = 0.0059$ and $\varphi = 0.00076$, which implies that at this inflation rate the cost associated with deposit purchases, $\gamma(1 - j^*)$, is 0.36 percent of GDP and φ corresponds to approximately 55 minutes per quarter.

Bailey (1956) and Friedman (1969) show that, under given assumptions, the optimal rate of inflation is the inverse of the real rate of return capital and such that the net nominal interest rate is 0. The intuition for this result is that the opportunity cost of holding currency is the nominal interest rate. As long as the net nominal interest rate differs from 0, people will spend time and resources to economize on their nominal currency holdings.

This result is not fully generalizable to our model economy with the presence of inside money. Here there will be a uniquely defined lower bound of inflation, weakly greater than the inverse of the real rate of return on capital. Below this lower bound, no one will hold deposits and the total money stock is equal to the monetary base ($M1 = M$), or, in other words, inside money will cease to exist and all capital held will be non-intermediated. If we also consider inflation rates below the lower bound of the inside money economy, we find that the Bailey-Friedman level of inflation is optimal also in our model economy.

3.1 Steady state

Figure 4 plots the benchmark welfare cost function λ, defined such that

$$u(\lambda c(\pi), d(\pi)) = u(c(\tilde{\pi}), d(\tilde{\pi})),$$

where $\tilde{\pi}$ is set equal to the average inflation rate since 1980 (about 3% annually) and to zero, alternatively.

When comparing steady states, we use the parameter values of our model economy (α, β, δ, φ, γ, ν, ζ, θ, ω, and $\bar{z}$) as they are calibrated for an annual inflation rate equal to 3%, then vary the inflation rate, and solve for combinations of a, c, d, h, l, m, n, p, and w.

Figure 4 plots the steady-state welfare cost for inflation rates ranging from the inverse of the steady state net real return on non-intermediated capital to 0.20. We see that as the inflation rate approaches an annual rate of 0.20, the welfare cost is slightly less than 0.25% of consumption compared with the benchmark.

The most striking feature of the graph is the predicted welfare gain from reducing inflation below 3% annually. The welfare gain from reducing annual inflation towards the lower bound of the inside-money economy, -1.64%, is about the same magnitude as the welfare cost of increasing inflation from 3% to 33%. Further reducing the inflation rate beyond the level at which inside money has ceased to exist also results in a sizeable welfare gain. Attaining the Bailey-Friedman rule will give a consumption equivalent welfare gain of about 0.7% compared with the benchmark.

As Figure 5 shows, as the steady state inflation rate increases further, the associated welfare gain flattens out. If the inflation rate is increased to 100%, the cost of inflation in terms of consumption compensation is still less than 0.6% – and still of a lesser magnitude than the welfare gain associated with a nominal interest rate of 0%.

The welfare gain of inflation can be decomposed into reductions of direct and of indirect costs. The latter are costs associated with changes in labor and capital input, and hence total output. Figure 6 plots the *direct* transaction and opportunity costs of inflation: "Use of deposits for purchases" is equal to the fixed cost of purchasing goods with deposits, $\gamma(1-j^*) = \gamma\left(1-\left(1+\frac{h}{m}\right)^{\frac{1}{\omega-1}}\right)$; "Replenish money balances" is equal to the wage rate times time spent, $wn\varphi$; "Opportunity cost, currency" is equal to the difference in return between non-intermediated capital and currency times real currency holdings, $\left(r-\frac{1}{\xi}\right)\frac{m}{p}$; and "Opportunity cost, deposits" is equal to the difference in return between non-intermediated capital and real deposit holdings, $(r-\tilde{r})\frac{h}{p}=\left(r-\frac{1}{\xi}\right)\frac{\theta h}{p}$. Household optimization implies that the cost of replenishing money balances is equal to the discounted opportunity cost of holding currency and deposits:

$$wn\varphi=\frac{1}{r}\left(r-\frac{1}{\xi}\right)\left(\frac{m+\theta h}{p}\right). \tag{11}$$

For positive inflation rates below about 100% annually, "use of deposits for purchases" is the largest direct cost. As inflation increases, and so also the opportunity cost of holding deposits, households will replenish their money balances more frequently. Almost paradoxically, as inflation increases, the total opportunity cost of holding currency decreases. This is because the households in the economy in part substitute their real money balances from currency to deposits, and in part because households optimally hold less real balances and rather replenish their money balances more frequently. We also notice that, for all inflation rates, the sum of resources spent on using deposits for purchases and on replenishing money balances is greater than the sum of opportunity costs of holding deposits and currency.

Below the lower inflation bound of the inside money economy, we have $h=0$ and $m=M$. Between the Bailey-Friedman rule and this level of inflation, the price level is governed by the liquidity constraint

$$n\left(\frac{m+h}{p}\right)=n\frac{M}{p}=c^{*},$$

and the marginal condition from Eq. (11)

$$wn\varphi=\frac{1}{r}\left(r-\frac{1}{\xi}\right)\left(\frac{M}{p}\right).$$

As the rate of inflation decreases, households choose to replenish money balances less frequently and rather hold higher real money balances. Hence, as the return difference between non-intermediated capital and currency goes to zero, the price level converges towards zero and real balances, $\frac{\bar{M}}{p}$, to infinity.

Figure 7 plots the holdings of real currency (m/p) and real deposits (h/p) as functions of the inflation rate. Real-currency holdings is a monotonously decreasing function for all inflation rates. In contrast, real-deposit holdings is an increasing function of the inflation rate for very low inflation rates, and then converges towards zero as inflation goes to infinity, but at a slower rate than real-currency holdings.

As a corollary to Figures 6 and 7, Figure 8 shows how total net worth (the sum of real currency, real deposits, and non-intermediated capital holdings) is a monotonously decreasing function of the inflation rate. Except for low levels of inflation, where households choose to hold relatively large real deposits, non-intermediated capital as share of total productive capital increases with inflation.

In Figure 8 we notice that total productive capital (K) attains its minimum value at the lower bound of inflation at which the economy still has a positive quantity of inside money. This result is also reflected in total output plotted in Figure 9. Total output reaches it highest level at net annual inflation rate equal to 6.4%. At this inflation rate output is about 0.44% larger than at the lower bound.

Figures 6 and 7 also graph our two proxies for the financial sector of our model economy. As we saw from Figure 6, the cost spent on facilitating transactions using deposits is an increasing, concave function of the inflation rate. On the other hand, and as we have discussed, real deposit holdings reaches its global maximum at an inflation rate slightly above 3%.

3.2 Fiscal considerations

Government steady-state real seignorage income is equal to

$$x = (\xi - 1)\frac{M}{p}. \tag{12}$$

The basis of government revenue is $M = (m + \theta h)/p$.

In contrast to cash-in-advance models, households in our model face more realistic trade-offs and have other margins to avoid inflation tax than holding less real balances. As we described in previous sections, as the rate of inflation increases, households optimally allocate more resources to use of deposits for purchases and to replenishing their money balances more frequently. The flip side of this observation is that they optimally hold less real balances and by that reduce the basis of government revenue.

Figure 10 plots the government seignorage income (x), household cost of holding currency, $\left(\frac{1}{r}\left(r - \frac{1}{\xi}\right)\left(\frac{m+\theta h}{p}\right)\right)$, and total direct cost of inflation, $\left(\gamma\left(1 - j^*\right) + wn\varphi + \frac{1}{r}\left(r - \frac{1}{\xi}\right)\left(\frac{m+\theta h}{p}\right)\right)$, for inflation rates between 0% and 100% annually. As expected, government seignorage income and household cost of holding currency are closely related, and are equal for $\xi = 1+\sqrt{1 - \frac{1}{r}}$. Even at inflation rate equal to 100%, the direct cost of inflation, ignoring general equilibrium effects, are more than 2.7 times larger than government revenue. Compared with a cash-in-advance model, this model emphasizes the inefficiency of inflation as a source of government revenue.

3.3 Uncertainty

Lucas (2000, p. 258) rephrases the belief he claims "[m]any economists [have] that a deterministic framework like Bailey's or [his] misses important costs of inflation that are thought to arise from price or inflation rate *variability*." Lucas continues by stating that "[he is] very confident that the effects of such a modification on the welfare costs estimated [in his paper] would be negligible."

Our results support Lucas' conjecture. For inflation rates above the lower bound of the inside money economy, if we keep the exogenous technology process constant and vary the standard deviation of the monetary process, the first two moments of the consumption and leisure sequences remain almost constant. The economic intuition for this result must be that only a very small fraction of household net worth is held as nominal assets and that therefore the effect of a marginal increase in nominal uncertainty is insignificant.

3.4 Lower bound of inflation

At the lower bound of inflation of the economy with endogenous money supply, the real rate of return net of transaction costs on deposits used on n purchases of the largest consumption goods must equal the rate of return on currency, that is,

$$\tilde{r}_{t+1} - \frac{\gamma n}{(1-\omega)\, c_t^*} r_{t+1} = \frac{p_t}{p_{t+1}},$$

where the second term is the ratio of the cost of n purchases of the largest consumption goods over the size of these purchases multiplied by the alternative rate of return.

The rate of return on deposits ($\tilde{r}_{t+1}$) is a linear combination of the return of non-intermediated capital and currency, so

$$\frac{1}{\xi} = \frac{1}{\beta}\left[1 - \frac{1}{(1-\theta)} \frac{\gamma n}{(1-\omega)\, c^*}\right],$$

or in words, at the margin, the rate-of-return difference between currency and non-intermediated capital is determined by the ratio of transaction cost to purchases of the largest consumption goods and adjusted for the reserve requirement ratio. Substituting for n and c^* at the lower bound as functions of the model's parameters gives us the following expression for the lower bound of inflation

$$\frac{1}{\xi} = \frac{1}{\beta}\left(1 - \frac{\frac{1}{\bar{z}\Phi^\alpha - \delta\Phi} + \frac{1-\zeta}{\zeta}\frac{1}{w}}{\varphi\frac{(1-\omega)(1-\theta)}{\gamma}\left(w\frac{(1-\omega)(1-\theta)}{\gamma} - \frac{1-\zeta}{\zeta} - 1\right)}\right),$$

where

$$\Phi = \left(\frac{\frac{1}{\beta} + \delta - 1}{\alpha\bar{z}}\right)^{\frac{1}{\alpha-1}} \quad \text{and} \quad w = (1-\alpha)\,\bar{z}\Phi^\alpha.$$

Three important technology and policy parameters determine the rate of return difference between non-intermediated capital and currency: the time cost of replenishing money balances (φ), the cost of using deposits for purchases (γ), and the reserve requirement ratio (θ). The rate of return difference at the lower bound is smaller the higher is φ, the lower is γ, and the lower is θ.

As φ approaches infinity, it becomes prohibitively expensive (time consuming) to replenish money balances. Households will then be willing to hold deposits even though the rate-of-return difference is infinitesimal and they face a fixed cost when using deposits for purchases.

As γ approaches zero, there are no costs associated with using deposits for purchases. Given that φ is strictly positive, households will then hold positive deposits for all rate of return differences, and the lower bound of inflation will be equal to the inverse of the rate of return on capital.

Since the rate of return on deposits is simply a linear combination of the rate of return on capital and on currency, the lower the reserve requirement ratio (θ) is, the closer the rate of return on deposits will be to the return on capital. The smaller the interest rate differential between deposits and non-intermediated capital is, the more attractive deposits will be, and the lower the rate of return difference at the lower bound of the economy with endogenous money supply.

3.5 Sensitivity

The benchmark economy is calibrated such that for gross annual inflation rate equal to 1.03 the currency-to-deposit ratio is equal to 9 and non-reserve portion of $M1$ divided by the capital stock is 0.05. This gives us calibrated values for $\gamma = 0.0059$ and $\varphi = 0.00076$, which implies that at this inflation rate the cost associated with deposit purchases, $\gamma(1-j^*)$, is 0.36 percent of GDP and φ corresponds to approximately 55 minutes per quarter.

Both the deposit-to-currency ratio and non-reserve portion of $M1$ divided by the capital stock are hard to measure. Our empirical deposit-to-currency ratio ranges from 12 early in the sample to 7 late in the sample. We have encountered estimates for the non-reserve portion of $M1$ divided by the capital stock as low as 0.03 and as high as 0.20.

Figures 14 and 15 show how the calibration of γ and φ, respectively, are sensitive to the estimates for the deposit-to-currency ratio and non-reserve portion of $M1$ divided by the capital stock. As we see from Figure 14 and 15, the calibrated values of both γ and φ increase in the estimate for $M1$ divided by the capital stock and is insensitive or slightly decreasing in the estimate of the deposit-to-currency ratio.

For the range of $M1/K$ and h/m, the max and min values of γ and φ are (0.00248, 0.01399) and (0.000161, 0.00362). At an inflation rate equal to 0.03, these estimates imply that the cost associated with deposit purchases is in the range between 0.15 and 0.85 percent of GDP and time spent per quarter to replenish money balances is in the range between 11 and 260 minutes.

Figure 16 plots the consumption equivalent welfare gain from reducing inflation from 3% annually to zero for the range of values of γ and φ spanned by Figures 14 and 15. For the boundary cases, $(\gamma, \varphi) = (0.00248, 0.000161)$ and $(\gamma, \varphi) = (0.01399, 0.00362)$, the estimated welfare gain would have been 0.05 and 0.33, respectively, compared with 0.13 for the benchmark case.

Perhaps the most striking feature of Figure 17, is how the effects of changes in γ and φ might offset each other. We see that the lower bound is -0.126 both when (γ, φ) is equal to $(0.00350, 0.000225)$ and when (γ, φ) is equal to $(0.01399, 0.00362)$.

3.6 Improved transaction technologies

One may reasonably assume that, as time goes on, the economy-wide technology level increases, and, as a result, transaction costs fall. The question is then how welfare gains due to potential improvements in transaction technology quantitatively compares to reductions in the inflation rate for a given level of transaction technologies.

Here we do an experiment in which transaction costs are reduced by three-quarters. Figures 11.1 and 11.2 show the welfare gain as φ and γ are reduced, respectively. As we see, the welfare gain if each of the transaction cost parameters are divided by four and the other one kept fixed is about 0.3%. Figure 11.3 shows that if both transaction-cost parameters are simultaneously reduced by three-quarters, the consumption equivalent welfare gain is almost 0.5%.

As implied by Figure 4, the welfare gain associated with this substantially reduction in transaction cost is significantly less than the gain of about 0.7% from reducing inflation to the Bailey-Friedman rule for given transaction costs. At the same time it is larger than the welfare gain associated with reducing inflation to the lower bound of the inside money economy.

Figures 13 and 12 plot the *direct* transaction and opportunity costs of inflation ("Use of deposits for purchases", "Replenish money balances',' "Opportunity costs, currency", and "Opportunity costs, deposits") as the γ and φ are reduced, respectively.

If the time cost of each replenishment of money balances decreases and the cost of using deposits for purchases is kept constant (Figure 12), households will substitute away from deposit holdings over to holding more currency and replenish their money balances more frequently. The results is that both the opportunity cost of holding currency and the total cost of replenishing money balances increase, but the decrease in both the opportunity cost of deposits and the cost of using deposits for purchases more than compensates for that increased cost.

As we see from Figure 13, as the cost of using deposits for purchases (γ) decreases while the time-cost of replenishing money balances is kept constant, households will hold more real deposits and the opportunity cost of holding deposits therefore increases. Nevertheless, the substitution towards holding more real deposits is not larger than that the income effect from improvement of the transaction technology dominates and the total cost of using deposits for purchases decrease.

4 Concluding Remarks

Compared to the existing literature on the welfare cost of inflation, the model in this paper contains several novel features. For one, people make purchases using both inside and outside money. The proportions of purchases made by either are determined by economic decisions in which the liquid assets' relative returns play an important role. The model allows for two transactions costs, one associated with using deposits when making purchases, and one incurred when liquid balances are replenished during the period. The additional margin through which households may to avoid the inflation tax on currency lets us distinguish between costs households incur due to lower return on currency and costs due to actions taken to avoid the inflation tax.

Our welfare-cost estimates are somewhat lower than what Cooley and Hansen (1989) and Lucas (2000) report. An interesting finding is that the welfare-cost curve is quite concave, meaning that the cost goes up steeply with steady-state inflation for low inflation rates (especially around 5 percent and lower) before flattening out considerably.

Another interesting finding is that only a small fraction of the total welfare cost of inflation can be accounted for by the opportunity cost of holding currency, and that most of the cost can rather be accounted for by the direct transaction costs and by the distortions to labor supply and capital accumulation.

The model's two proxies for the banking sector, real deposits and resources spent on transaction services, give somewhat ambiguous answers for how the size varies with inflation. Whereas resources spent on transaction costs increase monotonically in the inflation rate, holdings of real deposits reaches its maximum level at an inflation rate just above 3%.

For strictly positive transaction cost parameters, the lower bound of inflation of the inside money economy is strictly larger than the inverse of the real rate of return. Below this level of inflation, households do not hold any deposits, i.e. outside money cease to exist. This lower bound will be closer to the inverse of the real rate of return the lower are the reserve requirement ratio and the cost of using deposits for purchases and the higher is the time cost of replenishing money balances, and the lower the reserve requirement ratio is.

We have discussed fiscal considerations only briefly and simply pointed to the fact that in an economy such as ours in which the households have another margin to avoid the inflation tax other than by holding less real balances, seignorage is an even less efficient source of government revenue.

We consider the sensitivity of the welfare-cost estimate to several features. In particular, it is quite sensitive to the magnitudes of the two quantities used to calibrate the transaction-cost parameters. This is an interesting discovery as these parameters most likely have changed over the last decades and may be expected to decrease further. As our numerical experiments show, the welfare gain from potential technological progress leading to a decrease of both transaction parameters by 75% is larger than the welfare gain from adopting monetary policies such that the inflation rate reaches the lower bound, but less than the gain from reducing inflation to the inverse of the real rate of return.

References

Aiyagari, S. R., R. A. Braun and Z. Eckstein (1998). "Transaction services, inflation, and welfare." *Journal of Political Economy*, **106**(6), 1274–1301.

Bailey, M. (1956). "The welfare cost of inflationary finance." *Journal of Political Economy*, **64**, 93–110.

Cooley, T. F. and G. D. Hansen (1989). "The inflation tax in a real business cycle model." *American Economic Review*, **79**(4), 733–748.

Dotsey, M. and P. Ireland (1996). "On the welfare costs of inflation in general equilibrium." *Journal of Monetary Economics*, **37**(29-47).

Freeman, S. and G. W. Huffman (1991). "Inside money, output, and causality." *International Economic Review*, **32**(3), 645–67.

Freeman, S. and F. E. Kydland (2000). "Monetary aggregates and output." *American Economic Review*, **90**(5), 1125–1135.

Friedman, M. (1969). *The Optimum Quantity of Money and Other Essays*. Aldine, Chicago.

Gavin, W. T. and F. E. Kydland (1999). "Endogenous money supply and the business cycle." *Review of Economic Dynamics*, **2**, 347–369.

Gillman, M. (1993). "The welfare cost of inflation in a cash-in-advance economy with costly credit." *Journal of Monetary Economics*, **31**, 97–115.

Jones, B., G. Asaftei and L. Wang (2004). "Welfare cost of inflation in a general equilibrium model with currency and interest-bearing deposits." *Macroeconomic Dynamics*, **8**(4), 493–517.

Lucas, R. E. (2000). "Inflation and welfare." *Econometrica*, **68**(2), 247–274.

Prescott, E. C. (1986). "Theory ahead of business cycle measurement." *Minneapolis Fed. Quarterly Review*, pp. 9–22.

Simonsen, M. H. and R. P. Cysne (2001). "Welfare costs of inflation and interest-bearing money." *Journal of Money, Credit and Banking*, **33**(1), 90–100.

A Figures and tables

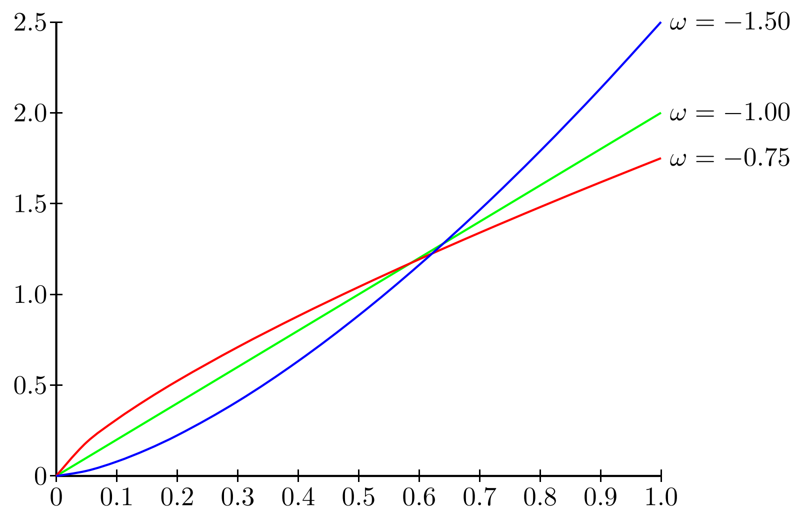
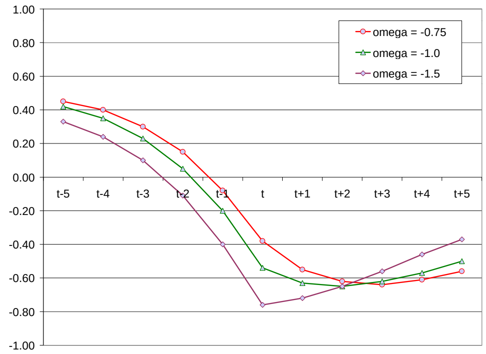


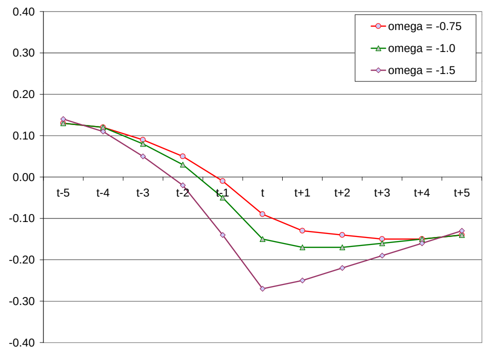
Figure 1: $c_t(j)$ for $0 \leq j \leq 1$

Table 2: Steady state welfare costs, benchmark calibration

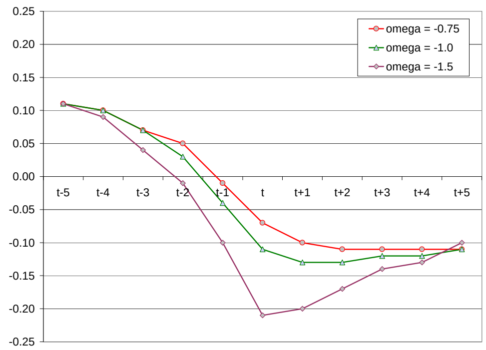
	-0.0164	0.00	0.01	0.03	0.06	0.10	0.25	0.50
n	1.320681	1.192505	1.171063	1.209201	1.309609	1.446476	1.600991	2.311321
l	0.299345	0.299722	0.299879	0.300000	0.300038	0.300018	0.299965	0.299597
a	9.713202	9.566608	9.516885	9.500000	9.521467	9.557323	9.593700	9.700938
h	0.000000	2.331871	3.358631	4.736842	5.927637	6.810877	7.441156	8.736977
m	1.000000	0.766813	0.664137	0.526316	0.407236	0.318912	0.255884	0.126302
h/m	0.000000	3.040992	5.057136	9.000000	14.555766	21.356580	29.080149	69.175101
j^*	1.000000	0.572012	0.486511	0.398107	0.333613	0.288561	0.256264	0.182611
p	3.327290	4.948311	6.309454	8.526316	11.118960	13.827927	16.529927	27.529942
c	0.746847	0.746759	0.746644	0.746420	0.746132	0.745815	0.745490	0.744131
d	0.699650	0.699370	0.699229	0.699079	0.698965	0.698880	0.698816	0.698643
u	−1.398807	−1.399235	−1.399496	−1.399835	−1.400167	−1.400477	−1.400765	−1.401841
k	9.978158	9.990729	9.995970	10.000000	10.001267	10.000614	9.998846	9.986565
l	0.299345	0.299722	0.299879	0.300000	0.300038	0.300018	0.299965	0.299597
y	0.997816	0.999073	0.999597	1.000000	1.000127	1.000061	0.999885	0.998656
$\lambda_{1.03}$	−0.0022	−0.0013	−0.0007	0.0000	0.0007	0.0014	0.0020	0.0043
$\lambda_{1.00}$	−0.0009	0.0000	0.0006	0.0013	0.0020	0.0027	0.0033	0.0057



2.1: "Policy A"

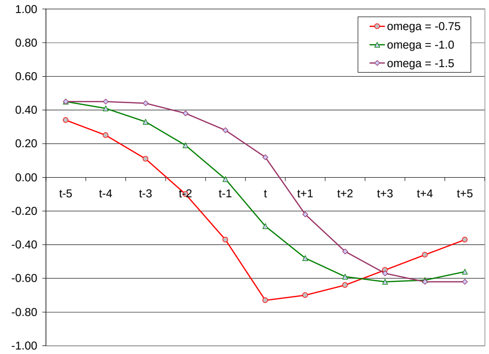


2.2: "Policy B"

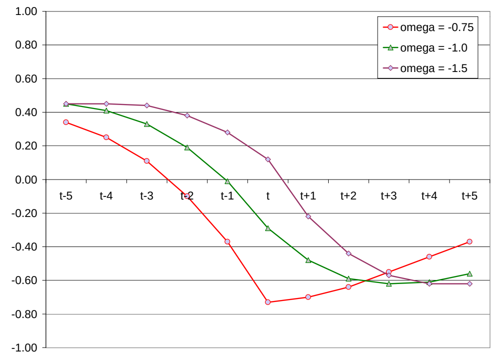


2.3: "Policy C"

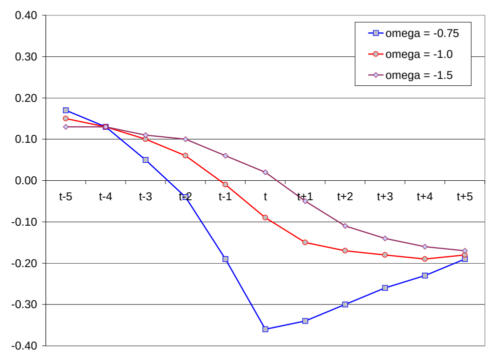
Figure 2: Cross-correlations: Output and Price level



3.1: "Policy A"



3.2: "Policy B"



3.3: "Policy C"

Figure 3: Cross-correlations: Output and Nominal R

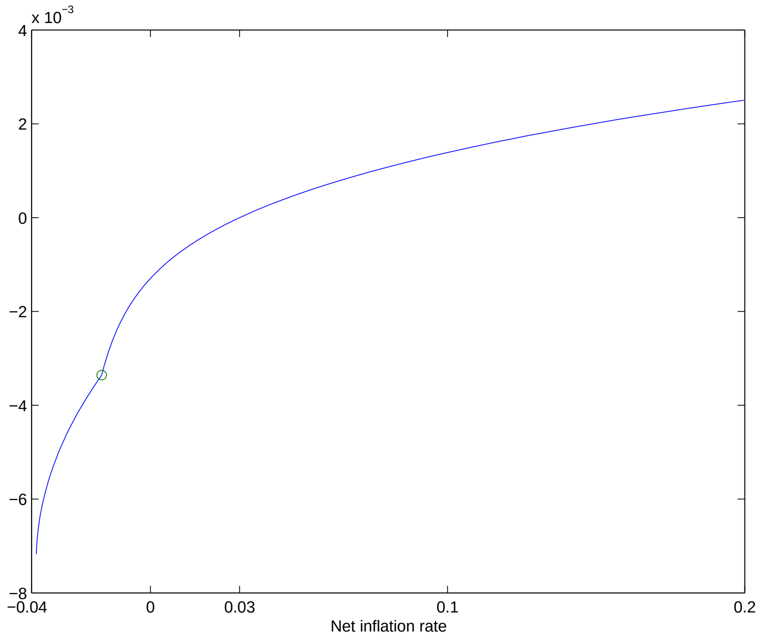


Figure 4: Welfare cost of inflation

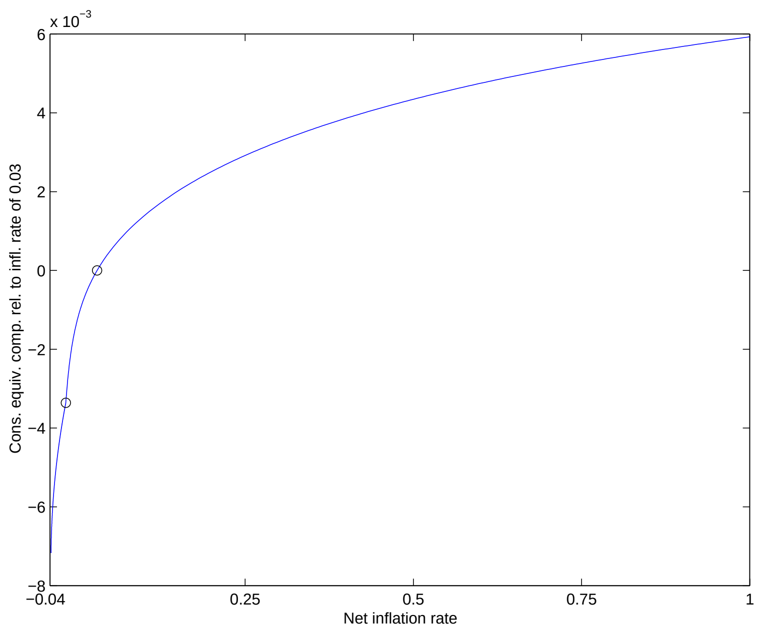


Figure 5: Welfare cost of inflation

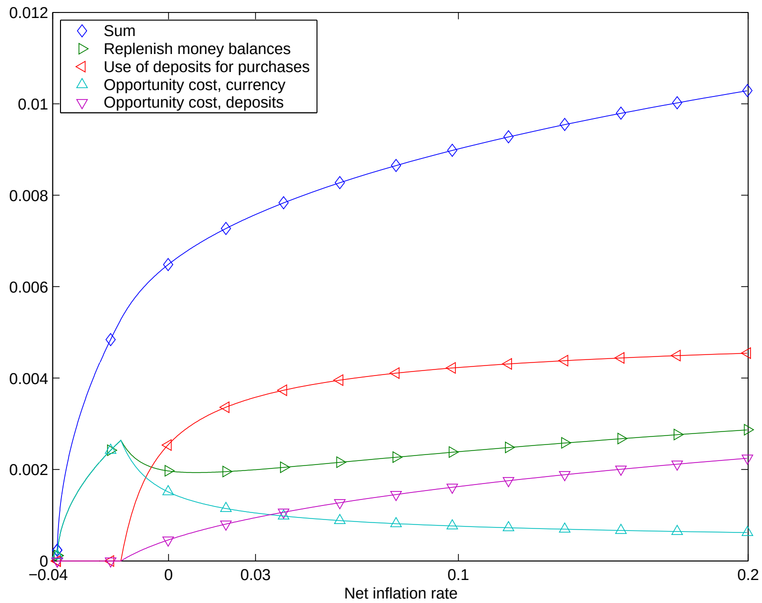


Figure 6: *Direct* transaction and opportunity costs of inflation

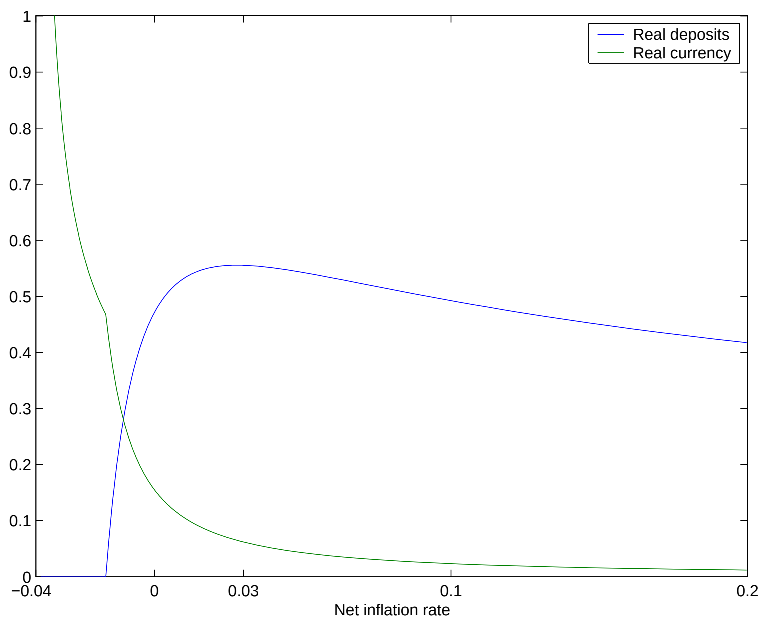


Figure 7: Real deposit and currency holdings

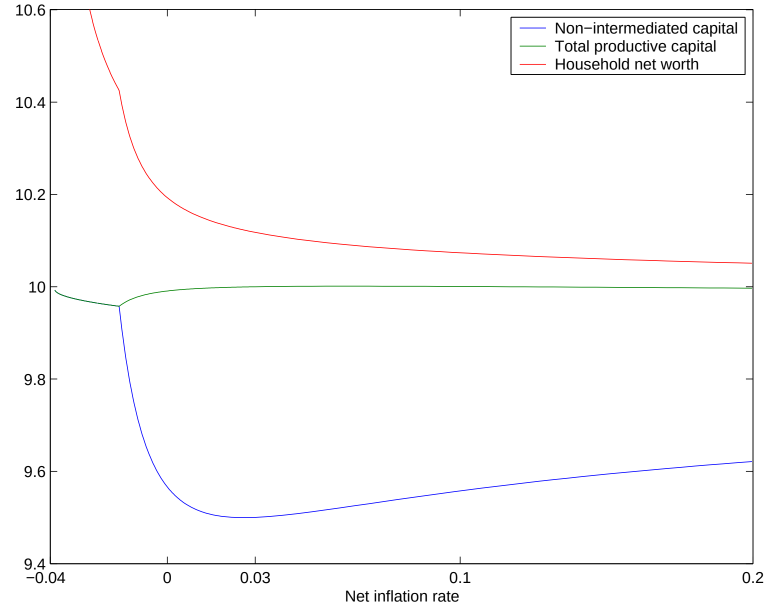


Figure 8: Total productive capital and household net worth

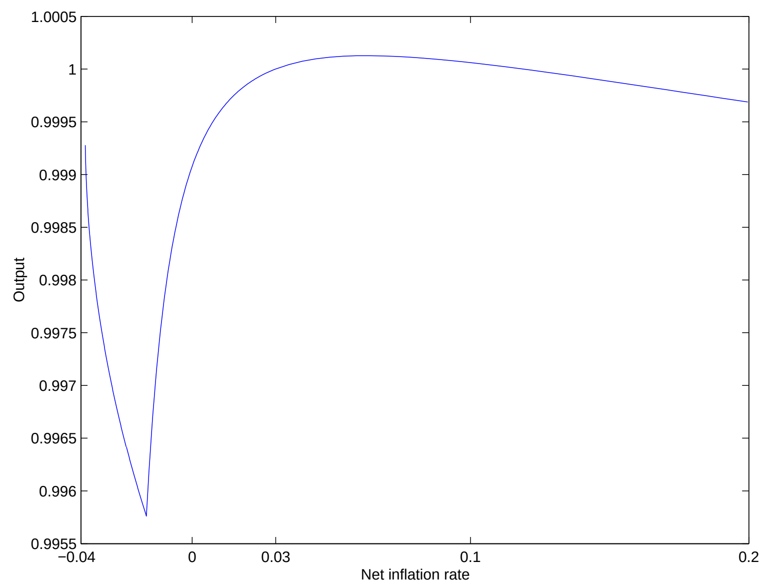


Figure 9: Output

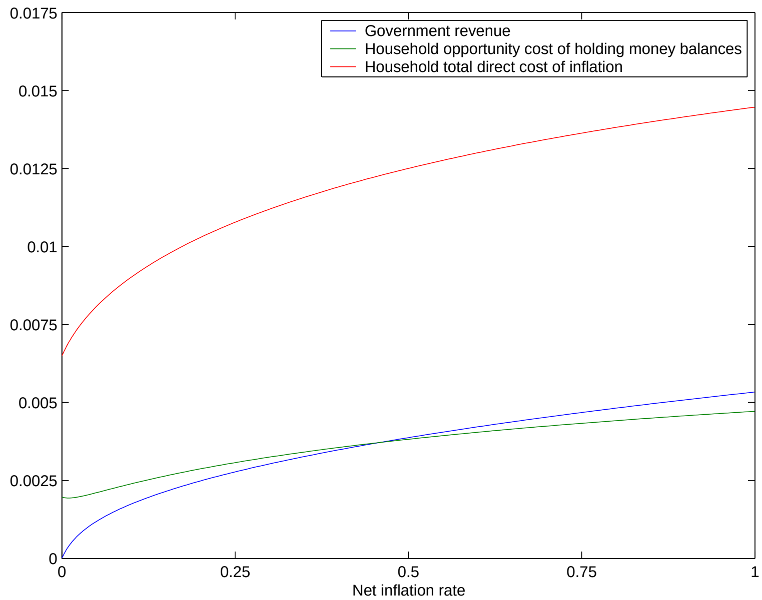
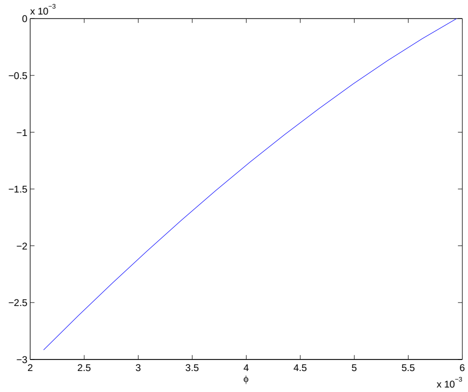
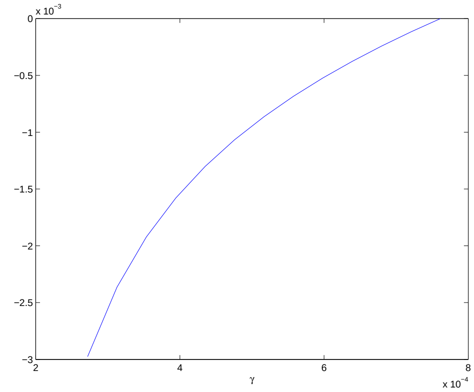


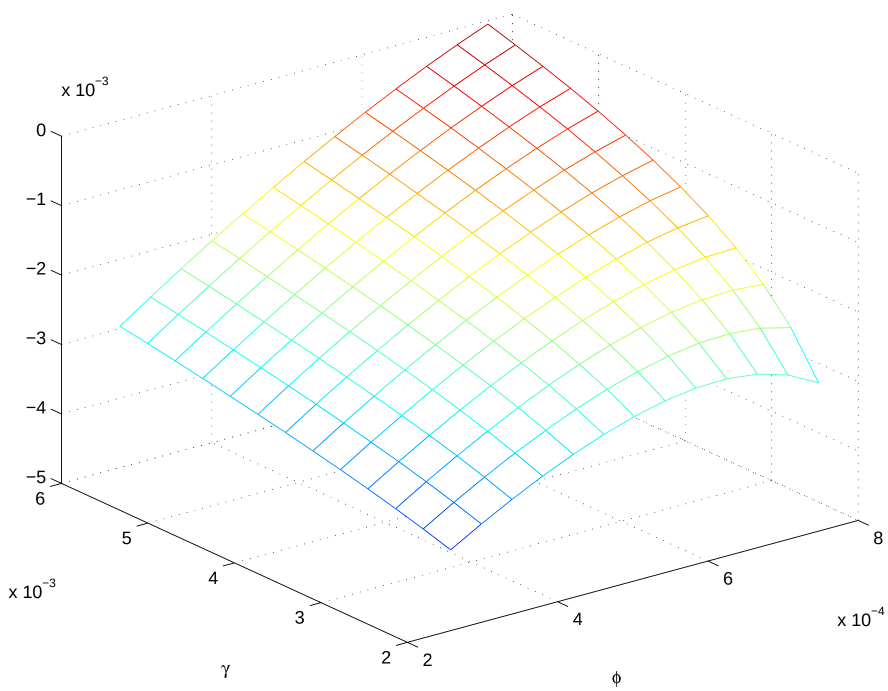
Figure 10: Cost of inflation divided by government inflation tax income



11.1: φ



11.2: γ



11.3: φ and γ

Figure 11: Potential welfare gains from improved transaction technologies

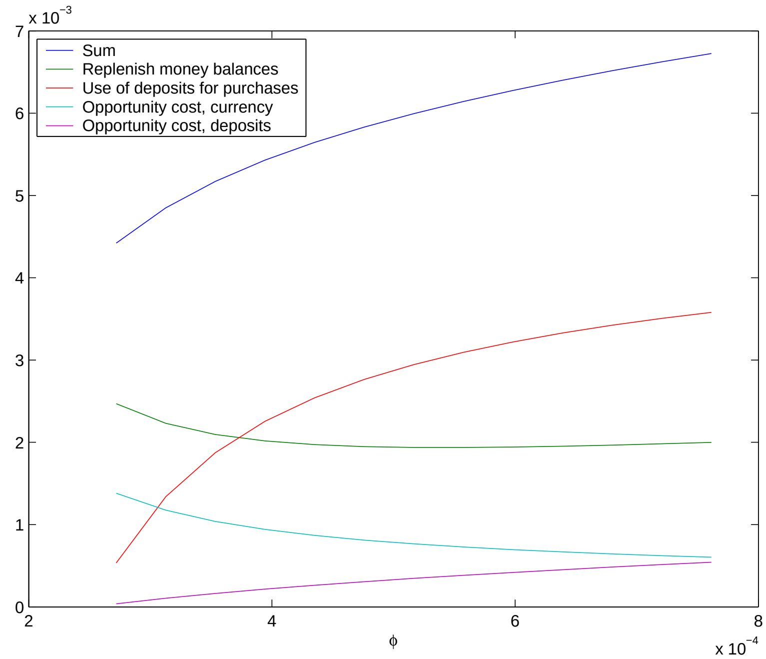


Figure 12: Direct costs of inflation as φ is reduced

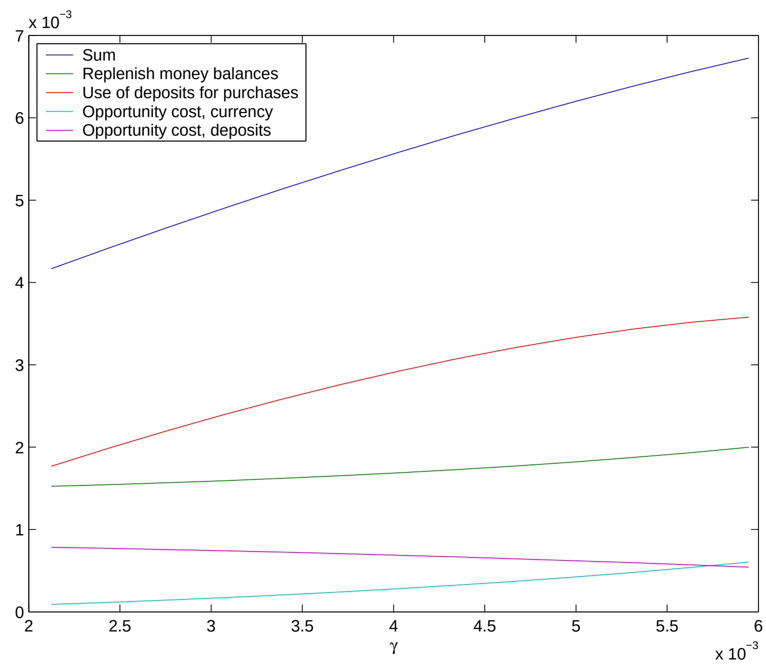


Figure 13: Direct costs of inflation as γ is reduced

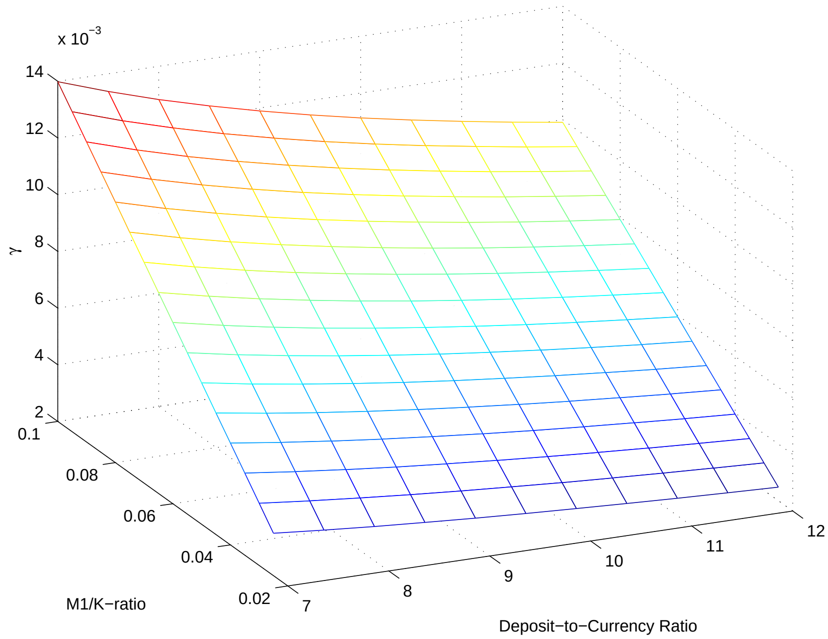


Figure 14: Sensitivity: γ

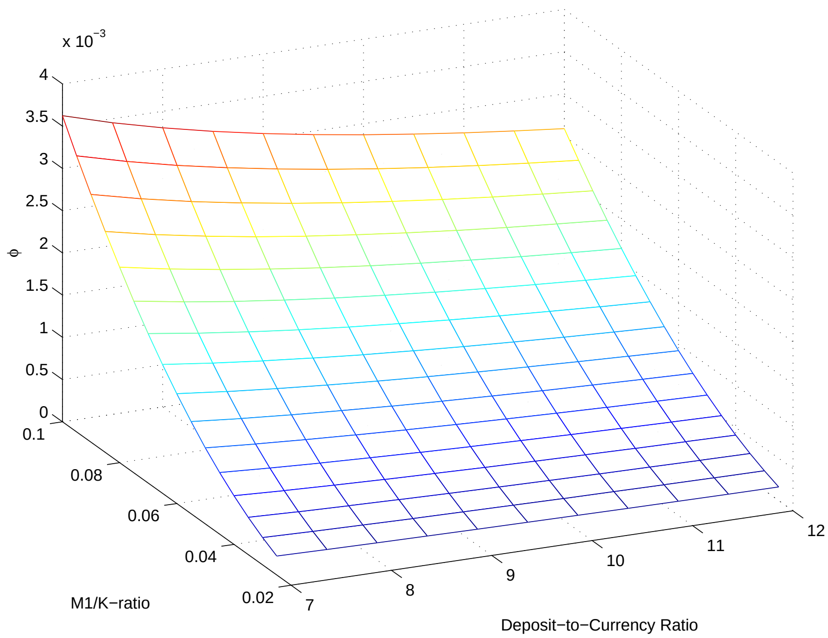


Figure 15: Sensitivity: φ

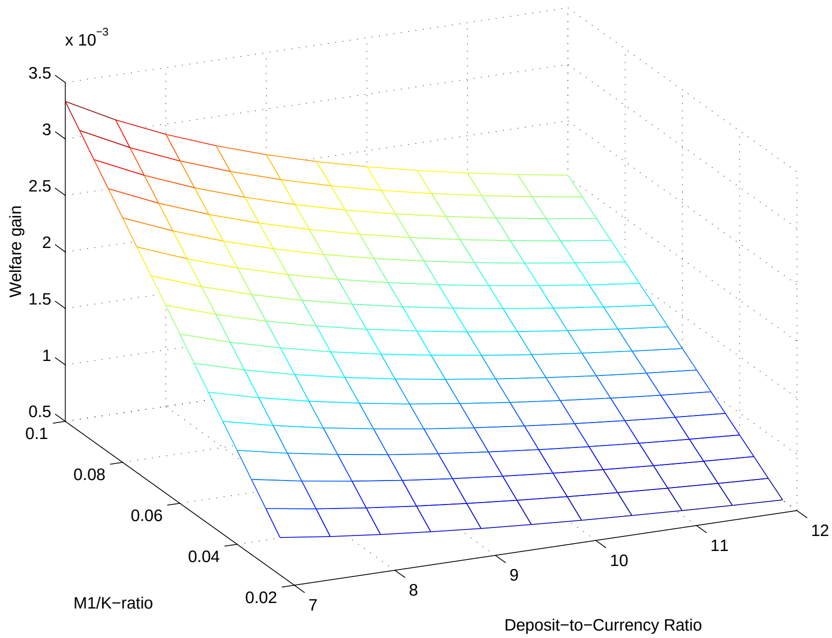


Figure 16: Sensitivity: Welfare

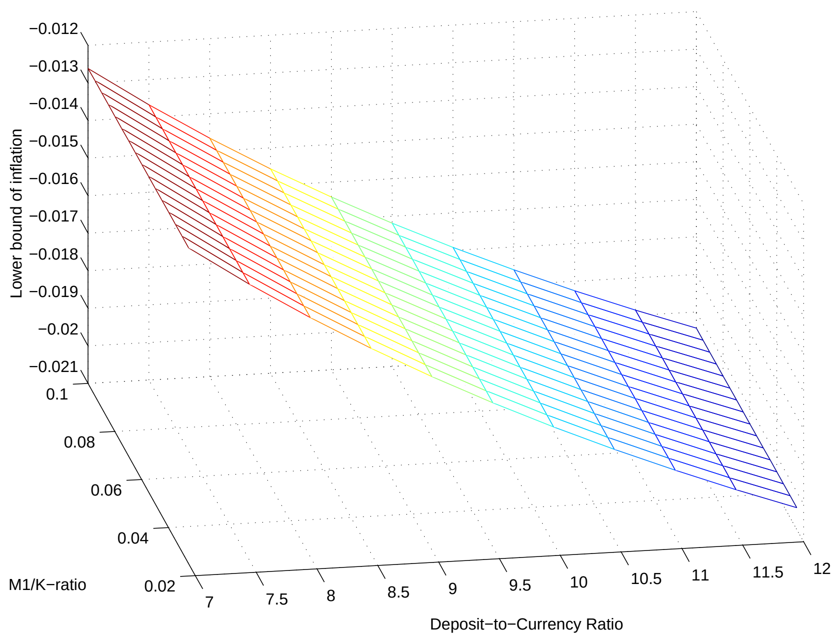


Figure 17: Sensitivity: lower bound