Journal of Monetary Economics 56 (2009) 1004–1013

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jme

The welfare costs of expected and unexpected inflation ☆

Miquel Faig [a,*], Zhe Li [b]

[a] Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada M5S 3G7
[b] Shanghai University of Finance and Economics, 777 Guoding Rd, Yangpu District, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 30 April 2008
Received in revised form
21 August 2009
Accepted 18 September 2009
Available online 26 September 2009

JEL classification:
E40
E52
E32

Keywords:
Monetary search
Imperfect information
Welfare cost monetary cycles
Welfare cost inflation

ABSTRACT

The monetary search model by Lagos and Wright (2005) is extended with imperfect information about nominal shocks as in Lucas (1972). An analytical solution exists with logarithmic preferences. In general, individuals hold precautionary balances. Calibrated to United States postwar data, the welfare cost of the monetary cycle is calculated to be small (below 0.0003% of GDP) compared to the welfare cost of the inflation tax (around 0.25% of GDP). The main reason for the minute welfare cost of the monetary cycle is its low amplitude in 1947–2007. But, monetary crashes, such as those experienced during the Great Depression, can generate important welfare costs.

1. Introduction

Expected inflation represents an implicit tax on money balances. To avoid this tax, individuals reduce their demand for money, thus creating a welfare cost. In the presence of nominal rigidities, unexpected and erratic inflation undermines the information transmitted through the price system, which creates another welfare cost. Most of the literature has studied these two costs separately. However, we discover synergies in the combination of both analyses in the Lagos and Wright (2005) framework extended with imperfect information about nominal shocks as in Lucas (1972).[1]

In our model, the markets where money is essential to purchase goods are subjected to two types of demand shocks: real (stochastic number of buyers) and nominal (stochastic changes in the quantity of money). The sellers in these markets know the distribution of these shocks, but do not know their realized values. As a result, they confuse relative price increases for the product they supply, which are due to the real shocks, with general price increases, which are due to the monetary shocks. This confusion is what makes money non-neutral. With an erratic monetary policy this non-neutrality generates fluctuations on output (the monetary business cycle). In the absence of any other distortion, these output

☆ We are thankful for the comments and encouragement received in the various places where we presented this paper. In particular, we benefitted from participating at the following conferences: Cleveland Fed (August 2006), Bernoulli Center for Economics (May 2007), SED (July 2007), and CMSG (November 2007). We are also grateful to the useful comments from the editor and one referee, the assistance of Juan Ilerbaing, and the financial support of SSHRC of Canada. The usual disclaimer applies.

* Corresponding author. Tel.: +1 416 978 0308; fax: +1 416 978 6713.
E-mail address: miquel.faig@utoronto.ca (M. Faig).

[1] Monetary search markets offer a natural environment where incomplete information can be embedded. For pioneering contributions, but without the tractability of the Lagos and Wright environment, see Wallace (1997) and Katzman et al. (2003).

0304-3932/$ - see front matter
doi:10.1016/j.jmoneco.2009.09.005

fluctuations are inefficient because they result from differences between the marginal valuations of goods by buyers and sellers.

The magnitude of the welfare cost of an erratic monetary policy depends not only on how much variability it generates on real GDP, but also on how concentrated these effects are on a particular segment of the economy, and how tolerant individuals are to output fluctuations in that segment. In the Lagos–Wright model, the transactions affected by monetary shocks coincide with those for which money is essential. As a result, the properties of the demand for money determine not only the welfare cost of expected inflation, but also the welfare cost of the monetary business cycle. In particular, a high elasticity of the money demand to changes in nominal interest rates implies a high welfare cost of inflation, as stressed in a long series of contributions that go back to Bailey (1956). At the same time, this high elasticity implies that individuals are quite willing to accept fluctuations in the transactions they perform with money, thus indicating a high tolerance to the fluctuations of the monetary business cycle.

Surprisingly, with an opportunity cost of holding money, the monetary business cycle may be welfare enhancing because it may partly correct the inefficiencies caused by the inflation tax. Monetary shocks increase the dispersion of prices. As a result, buyers hold more precautionary balances and, in doing so, partly offset the inefficiently low money demand due to the inflation tax. To be clear, this effect does not imply that unexpected monetary shocks are part of an optimal policy. The optimal policy in our model is to follow the Friedman rule, where inflation is both low (negative) and predictable.

Our empirical analysis finds that the welfare gain of eliminating the United States monetary business cycle observed from 1947 to 2007 is a meager 0.0003% of GDP in our preferred simulation, while the welfare gain of reducing the 4% annual rate of inflation, that was experienced on average during those years, to the Friedman rule is several orders of magnitudes larger (around 0.25% of GDP). However, the small welfare cost of the monetary business cycle that we calculate for the United States during the postwar era does not imply that monetary disturbances are of little political concern. A key reason for the low welfare cost that we calculate is the low amplitude of the monetary business cycle from 1947 to 2007. Large monetary crashes such as those experienced in earlier periods can generate important welfare costs. For example, we calculate that the equivalent variation in income of an unexpected 10% drop of the money supply (experienced twice from 1929 to 1946 at the monthly frequency) is 0.5% of GDP.

This paper contributes to the recent literature of models with incomplete information following the work of Morris and Shin (2003a, b). However, the agents in our model are fully Bayesian, so higher order expectations do not matter. In contrast, much of the recent literature deals with the consequences of these higher order expectations. Among these papers, Woodford (2002), Ui (2003), Hellwig (2005), Adam (2007), Angeletos and La'O (2008), and Lorenzoni (2009) deal as we do with monetary policy issues.[2] Some of the points these papers stress, such as finite information processing, endogenous information acquisition, and multiple aggregate shocks, could be interesting additions to the present setup because they generate more realistic dynamics than the one predicted by the present model. They would also allow formulating monetary policy in terms of controlling interest rates, which again is not possible in the present setup.[3]

The rest of the paper is organized as follows. Section 2 describes the model of the paper. Section 3 studies the optimal behavior of individuals and defines a recursive equilibrium. Section 4 characterizes an efficient allocation. Section 5 provides the analytical solution of the model in a special case in which preferences are logarithmic and shocks are log-normally distributed. Section 6 calibrates the model to United States data. This calibration requires using preferences that are not logarithmic, so the model has to be solved numerically. Finally, Section 7 concludes. Lengthy algebraic derivations and the algorithms used for the numerical analysis can be found in a supplementary appendix at the author's web page.

2. The model

Time is discrete and the horizon is infinite. Each period consists of two subperiods to be called day and night. There is a single, non-durable good. During the day, the good is traded in a centralized market with perfect information. During the night, it is traded in $N \geq 2$ decentralized markets, where individuals are anonymous and asset holdings are private information. In the day market, everybody can produce and consume the good. At the night markets, a fraction of the population (the sellers) is able to produce the good, while the remaining fraction (the buyers) gets utility from consuming it. Both the day and night markets are competitive.

There is a continuum of buyers and sellers. Each individual, buyer or seller, has a probability $\pi \in (0,1)$ of reaching a night market and being able to trade. The sellers who are able to trade are evenly distributed across all night markets, and their measure in each one of them is normalized to 1. In contrast, the buyers who are able to trade are randomly distributed across the night markets, so the measure of buyers at market i in period t is a positive stochastic variable n_{it}. The

[2] Earlier work by Lucas (1975) and King (1982) raised some of the issues dealt with in these papers.

[3] See Lucas (1975, Section 13) for a discussion of the difficulties of allowing for centralized asset markets in his theory of cycles based on rational expectations and imperfect information if the only aggregate shock is a monetary shock.

distribution of n_{it}, $i = 1, \ldots, N$ and $t = 0, 1, \ldots$ is identical across markets and time, and independent across time. Moreover, we abstract from aggregate real shocks by assuming $E(n_{it}) = 1$.[4]

Following Lagos and Wright (2005), preferences are assumed to be quasi-linear to generate a tractable distribution of money balances. The instantaneous utility of a buyer who visits night market i at date t is $U^b(x_t^b, y_t^b, q_{it}^b) = v(x_t^b) - y_t^b + u(q_{it}^b)$; where x_t^b and y_t^b are, respectively, quantities consumed and produced during the day, and q_{it}^b is the quantity consumed during the night. Likewise, the instantaneous utility of a seller is $U^s(x_t^s, y_t^s, q_{it}^s) = v(x_t^s) - y_t^s - c(q_{it}^s)$; where x_t^s and y_t^s are, respectively, the quantities consumed and produced during the day, and q_{it}^s is the quantity sold during the night. The lifetime utilities of buyers and sellers are $E\sum_{t=0}^{\infty}\beta^t U^h(x_t^h, y_t^h, q_{it}^h)$, for $h = b$ and s, and $\beta \in (0, 1)$ is the one period discount factor. The functions v, u, and c are all continuously differentiable and increasing. The functions v and u are concave, while c is convex. Moreover, $v(0) = u(0) = c(0) = c'(0) = 0$, and $u'(0) = \infty$. Finally, there exist real numbers x^* and q^* such that $v'(x^*) = 1$ and $u'(q^*) = c'(q^*)$.[5]

In this environment, there is a role for a medium of exchange to facilitate trade, because at night there is a lack of double coincidence of wants and all traders are anonymous. We further assume that the government has the monopoly on issuing money, which is the only counterfeit-proof note in the economy, so money is essential. Money is an intrinsically useless, perfectly divisible, and storable asset. The money supply grows at a random factor $\gamma_t : M_{t+1} = \gamma_t M_t$, where M_t is the quantity of money per buyer. The distribution of γ_t is independent and identically distributed across time, and the unconditional mean of the inverse of γ_t satisfies: $\beta E(\gamma_t^{-1}) < 1$. New money is injected via uniform lump-sum transfers to all buyers at the beginning of the night. This form of introducing money differs from Lucas's classical contributions, which assumed that transfers are proportional to the money balances held by their recipients. While this difference matters little for the effect of monetary shocks on output, it allows for a welfare cost of expected inflation, which is absent in Lucas's contributions.

The same anonymity that makes money essential prevents long-term contracts between buyers and sellers from regulating the transactions that take place at night. As a result, at night liquidity has a direct effect on buyers' demands, and spot prices become the coordinating mechanism. As in the second model of Rocheteau and Wright (2005), these markets are assumed to be competitive. Furthermore, following Lucas (1972, 1973), we focus on recursive equilibria, where prices are functions of the state variables.

A full description of the state of the economy includes the distribution of money and wealth at the beginning of each subperiod, and the distribution of buyers across markets at night. However, as it is well known, quasi-linear preferences imply that consumption and the demand for money during the day are independent of wealth. As a result, the price of the day good for money, P_t, depends only on M_t; and the relative price of a night good for the day good of the same period, p_{it}, depends on the vector of the realized values of the two stochastic variables (γ_t, n_{it}). As it will be proven below, $p_{it}(\gamma_t, n_{it})$ is monotonic in both arguments, and because of recursivity and symmetry, this function is the same in all periods and all night markets.

At the night of period t, the buyers who visit market i know the quantity of money they carry, so they can infer the growth factor of the money supply γ_t. Also, they know the price p_{it}. These two pieces of information allow them to infer n_{it}. In contrast, sellers know neither γ_t nor n_{it}, so in general p_{it} conveys information about γ_t and n_{it}, but does not perfectly reveal either one of these two variables. This is the signal extraction problem emphasized by Lucas (1972, 1973).

As is standard, in a recursive equilibrium the nominal price of the day good P_t is proportional to the quantity of money M_t, because a proportional change in these two variables leaves both individual incentives and opportunities as well as the market clearing conditions unaltered. However, as in Lucas's classical contributions, monetary shocks have real effects in the night markets because sellers observing a high price cannot be certain if there has been a positive monetary shock (γ_t is large) or a positive real shock (n_{it} is large). Hence, monetary shocks affect the incentives that determine the supply of goods.

3. Optimal behavior and equilibrium

Consider a buyer (male) who starts day t with m_{0t}^b units of money and faces a nominal price of goods P_t. Let z_{0t}^b be the real value of m_{0t}^b, that is $z_{0t}^b = m_{0t}^b / P_t$. Consumption, x_t^b, production, y_t^b, and real money balances at the end of the day, z_t^b, are constrained by the following budget:

$$x_t^b + z_t^b = y_t^b + z_{0t}^b, \quad z_t^b \geq 0. \tag{1}$$

[4] If N is finite, n_{it} cannot be independent across markets; otherwise, the aggregating condition $E(n_{it}) = 1$ would not hold. For example, if $N = 2$, n_{2t} must be equal to $1 - n_{1t}$. As in Lucas's seminal contributions, this dependence does not represent a problem for the coherence of the model.

[5] If production is constrained to be non-negative, then our characterization of an equilibrium is valid, as long as the choice of y_t^b and y_t^s is interior for all t.

Denoting V_t^b as the value function at the beginning of the night and Φ as the joint distribution function of p_{it} and γ_t, he chooses $\{x_t^b, y_t^b, z_t^b\}$ to solve the following maximization program:

$$W_t^b(z_{0t}^b) = \max\left\{ v(x_t^b) - y_t^b + \int \left[\pi V_t^b(z_t^b, p_{it}, \gamma_t) + (1-\pi)\beta W_{t+1}^b\left(\frac{z_t^b + \Delta_t}{\gamma_t}\right)\right] d\Phi \right\} \tag{2}$$

subject to (1). The solution to this maximization program defines the value function W_t^b at the beginning of the day.

At night his purchasing power is limited by the money acquired during the day, z_t^b, plus the real value of the lump-sum transfer received from the government, Δ_t. Therefore, he purchases q_{it}^b to solve:

$$V_t^b(z_t^b, p_{it}, \gamma_t) = \max\left[u(q_{it}^b) + \beta W_{t+1}^b\left(\frac{z_t^b + \Delta_t - p_{it}q_{it}^b}{\gamma_t}\right)\right] \tag{3}$$

subject to:

$$p_{it}q_{it}^b \leq z_t^b + \Delta_t. \tag{4}$$

In this expression, it has been used that in a recursive equilibrium $P_{t+1} = \gamma_t P_t$, and that γ_t is known by the buyer at the night of t.

The maximization program of a seller (female) during the day is identical to that of a buyer, so $\{x_t^s, y_t^s, z_t^s\}$ solves:

$$W_t^s(z_{0t}^s) = \max\left\{ v(x_t^s) - y_t^s + \int \left[\pi V_t^s(z_t^s, p_{it}, \gamma_t) + (1-\pi)\beta W_{t+1}^s\left(\frac{z_t^s}{\gamma_t}\right)\right] d\Phi \right\} \tag{5}$$

subject to a budget constraint analogous to (1) with an s replacing the b subscript. In contrast, her maximization program at night has two important differences. First, there is no liquidity constraint binding the optimal supply of goods. Second, the seller does not know γ_t, but she uses p_{it}, the equilibrium relation $p_{it}(\gamma_t, n_{it})$, and the joint distribution $\Phi(p_{it}, \gamma_t)$ to calculate the conditional distribution $F(\gamma_t|p_{it})$. So, the supply of goods q_{it}^s solves:

$$V_t^s(z_t^s, p_{it}) = \max\left[-c(q_{it}^s) + \beta \int W_{t+1}^s\left(\frac{z_t^s + p_{it}q_{it}^s}{\gamma_t}\right) dF(\gamma_t|p_{it})\right]. \tag{6}$$

As in Lagos and Wright (2005), the quasi-linear preferences in (2) and (5) imply that day consumption for buyers and sellers is the efficient quantity that satisfies $v'(x^*) = 1$, $x_t^b = x_t^s = x^*$. Moreover, the day value functions are affine with unit marginal values of real money balances and constant terms which are independent of time: $W_t^b(z_t^b) = w^b + z_t^b$ and $W_t^s(z_t^s) = w^s + z_t^s$ for all t. Combining these functional forms with (3) and (6), the optimal choices at night solve the following programs:

$$\max_{q_{it}^b}\left[u(q_{it}^b) - \frac{\beta p_{it}q_{it}^b}{\gamma_t}\right] \quad \text{subject to } p_{it}q_{it}^b \leq z_t^b + \Delta_t, \tag{7}$$

and

$$\max_{q_{it}^s}\left[-c(q_{it}^s) + \beta \int \frac{p_{it}q_{it}^s}{\gamma_t} dF(\gamma_t|p_{it})\right]. \tag{8}$$

The demand and supply functions that solve (7) and (8) are denoted, respectively, as $\tilde{q}^b(z_t^b, \Delta_t, \gamma_t, p_{it})$ and $\tilde{q}^s(p_{it})$.

The optimal demand for money satisfies the following Euler equation:[6]

$$\int \left\{ \pi \frac{u'[\tilde{q}^b(z_t^b, \gamma_t, p_{it}, \Delta_t)]}{p_{it}} + (1-\pi)\frac{\beta}{\gamma_t}\right\} d\Phi = 1. \tag{9}$$

That is, buyers equate the expected marginal benefit of money at night with the marginal cost of its acquisition during the day. If a buyer has a trading opportunity at night, the marginal value of money is the marginal utility of the goods he can purchase with an extra dollar, regardless of being cash constrained or not. If the buyer does not have a trading opportunity, the marginal value of money is the discounted marginal utility of the day good he can purchase tomorrow. Consequently, the expected marginal benefit of z_t^b in utils is the integral in the left-hand side of (9). To acquire an extra real unit of money, the buyer must supply an extra unit of y_t^b, which costs one util, as stated in the right-hand side of (9).

As a vehicle for saving from period t to period $t+1$, the expected money's return factor is $E(\gamma_{t+1}^{-1})$. Since the utility discount factor is β and, by assumption, $\beta E(\gamma_{t+1}^{-1}) < 1$, no individual should carry money balances that with certainty will not be spent during the night. Therefore, a seller's optimal demand for money is zero: $z_t^s = 0$ for all t. Also, a buyer should face a binding liquidity constraint (4) in at least one of the markets he can potentially visit at night, so in the margin his money balances are not only an instrument for saving, but they also facilitate transactions.

A recursive equilibrium is a set of real numbers z^b and Δ, which, respectively, describe the buyer's money demand z_t^b and the monetary transfer Δ_t for all t, and a set of real functions $p(\gamma, n_i)$, $q^b(\gamma, n_i)$, and $q^s(\gamma, n_i)$, which, respectively, map the

[6] This equation follows from the first order conditions of Eq. (3). The details of its derivation are in the supplementary appendix.

realized values of the shocks γ_t and n_{it} onto the prices p_{it} and the quantities q_{it}^b and q_{it}^s traded at the night markets for all i and t, that satisfy the following conditions[7]:

A. *Optimal behavior*:
$q^b(\gamma, n_i)$ is equal to the buyers' demand function: $\tilde{q}^b[z^b, \Delta, \gamma, p(\gamma, n_i)]$.
$q^s(\gamma, n_i)$ is equal to the sellers' supply function: $\tilde{q}^s[p(\gamma, n_i)]$.
z^b satisfies the Euler equation (9).

B. *Market clearing*:

$$n_i q^b(\gamma, n_i) = q^s(\gamma, n_i). \tag{10}$$

C. *Government budget constraint*:
The lumps-sum transfer satisfies: $\Delta = z^b(\gamma - 1)$.

D. *Rational expectations*:
$F(\gamma|p_i)$ in (8) is the distribution of γ conditional on $p_i = p(\gamma, n_i)$.

In this definition and in what follows, time subscripts are dropped when this does not create ambiguity.

4. Efficient allocation and the Friedman rule

The efficient allocation of output at night $\{q_i^b, q_i^s\}$ maximizes $\int [n_i u(q_i^b) - c(q_i^s)]di$ subject to $n_i q_i^b = q_i^s$ for all i. The solution to this program is: $u'(q_i^b) = c'(q_i^s)$, for all i. Therefore, in a efficient allocation the marginal utility of consumption is equal to the marginal disutility of production. Define the Friedman rule as the monetary policy in which $\sigma_\gamma^2 = 0$ and $\gamma \downarrow \beta$. Under the Friedman rule, the rate of return on money is the subjective discount rate. Therefore, since the day value functions are affine, individuals would be willing to hold money even if it were not the medium of exchange at night. In an equilibrium under the Friedman rule, individuals carry enough money to the night markets for the liquidity constraint (4) not to be binding, so the solutions to the optimization programs (7) and (8) are, respectively, $u'(q_i^b) = p_i$ and $c'(q_i^s) = p_i$. Consequently, under the Friedman rule, the marginal utility of consumption is equal to the marginal disutility of production, which, as shown above, is the condition that characterizes efficiency.

5. The logarithmic case: Lucas (1973)

One of the difficulties of solving for a recursive equilibrium is that, in contrast to Lucas (1972, 1973), buyers do not always spend all the money balances they carry. That is, conditions in the market that a buyer visits may be such that the solution to the optimization problem (7) is interior. This complication can be easily handled with numerical methods but, in general, it prevents an explicit solution to the model. There is an exception to this general rule: when buyers have logarithmic preferences for the night good. In that case, we show in the next paragraph that buyers spend all their money in whatever night market they visit. Moreover, if the sellers's disutility of production and the distribution function of shocks are chosen appropriately, one can obtain an explicit solution that is similar to the solution of the well-known model in Lucas (1973).

Let $u(q_i^b) = \ln(q_i^b)$. The first order condition that results from applying the Kuhn-Tucker theorem to the optimization program (7) is $q_i^b p_i = \gamma \min\{z^b, \beta^{-1}\}$ for all i. If $z^b > \beta^{-1}$, the liquidity constraint (4) is not binding and the optimum is interior. Otherwise, (4) is binding and buyers exhaust their money balances. In any case, buyers' expenditures are the same in all night markets. Intuitively, because a high p_i in market i means that the good for sale is relatively expensive, buyers purchase a low quantity q_i^b; with logarithmic preferences, the demand elasticity of q_i^b with respect to p_i is minus one, so the low q_i^b exactly compensates for the high p_i to make the expenditure $q_i^b p_i$ independent from p_i. Moreover, the optimum cannot be interior in all markets because then the buyer would be carrying money balances in excess of the maximum quantity ever purchased, which would be suboptimal. Consequently, with logarithmic preferences the buyers' demand for goods at night market i obeys

$$q_i^b = \frac{\gamma z^b}{p_i}. \tag{11}$$

Combining (9) with (11) and simplifying, the buyers' demand for money is

$$z^b = \frac{\pi}{[E(\gamma^{-1})]^{-1} - (1 - \pi)\beta}. \tag{12}$$

To obtain an explicit solution, we further assume that $c(q^s) = (q^s)^{1+\alpha}/(1+\alpha)$ and the stochastic shocks are log-normal: $\ln\gamma \sim N(\mu_\gamma, \sigma_\gamma^2)$ and $\ln n_i \sim N(\mu_n, \sigma_n^2)$ for all i. With these functional forms, $E(\gamma^{-1}) = \exp(-\mu_\gamma + \sigma_\gamma^2/2)$ and the optimal supply of

[7] To complete the description of the equilibrium allocation, during the day all individuals consume x^* and produce whatever quantities y^b or y^s are needed for the budget constraints of buyers and sellers to hold.

goods at night implied by (8) is

$$q_i^s = [\beta p_i E(\gamma^{-1}|p_i)]^{1/\alpha}. \tag{13}$$

Therefore, substitution of (11) and (13) into the market clearing condition (10) yields

$$n_i\gamma = \frac{p_i^{1+1/\alpha}[\beta E(\gamma^{-1}|p_i)]^{1/\alpha}}{z^b}. \tag{14}$$

The right-hand side of (14) is a function p_i, while the left-hand side is the product of the realized values of the two stochastic variables unobserved by the sellers. Therefore, observing p_i in a recursive equilibrium carries the same information as observing $n_i\gamma$, so $F(\gamma|p_i) = F(\gamma|\omega_i)$, where $\omega_i = n_i\gamma$. That is, finding $F(\gamma|p_i)$ turns out to be equivalent to the standard signal extraction problem of knowing the unconditional distributions of two log-normally distributed random variables n_i and γ and finding the distribution of one of them conditional on observing the realized value of their product. As explained in Lucas (1973), the distribution of γ conditional on ω_i is log-normal and characterized by the following two moments:

$$E(\ln\gamma|\omega_i) = \theta\mu_\gamma + (1-\theta)(\ln\omega_i - \mu_n), \tag{15}$$

and

$$\mathrm{Var}(\ln\gamma|\omega_i) = \theta\sigma_\gamma^2 \quad \text{where } \theta = \frac{\sigma_n^2}{\sigma_n^2 + \sigma_\gamma^2}. \tag{16}$$

In words, the conditional mean of $\ln\gamma$ is a weighted average between two pieces of information: its unconditional mean μ_γ and the logarithm of the observed product ω_i adjusted by deducting μ_n, so its mean is also μ_γ. If monetary shocks are small or real shocks are large ($\sigma_\gamma^2 \approx 0$ or $\sigma_n^2 \approx \infty$), then θ is close to one, and the unconditional mean μ_γ carries most of the weight in the average. In contrast, if real shocks are small or monetary shocks are large ($\sigma_n^2 \approx 0$ or $\sigma_\gamma^2 \approx \infty$), then θ is close to zero, so current observations of p_i, and so ω_i, carry most of the weight in forming the conditional mean $E(\ln\gamma|\omega_i)$. The conditional variance of $\ln\gamma$ is increasing in both σ_n^2 and σ_γ^2.

Using (15) and (16), the equivalence between $E(\gamma^{-1}|p_i)$ and $E(\gamma^{-1}|\omega_i)$, the assumption $E(n) = 1$, and the fact that γ^{-1} is log-normally distributed, we get $E(\gamma^{-1}|p_i) = \exp[-\theta\mu_\gamma - (1-\theta)\ln\omega_i]$. Substituting this expression into (14) and solving for p_i yields the equilibrium price at the night markets:

$$p_i = A^{-1}(z^b)^{\alpha/(1+\alpha)}(\gamma n_i)^{(1-\lambda)}. \tag{17}$$

The values of z^b and θ are given in (12) and (16), and the values of the remaining constants are $\lambda = \theta/(1+\alpha)$ and $A = [\beta\exp(-\theta\mu_\gamma)]^{1/(1+\alpha)}$. Since $\theta \in [0,1]$ and $\alpha > 0$, λ must belong to the interval (0,1). Therefore, p_i is monotonically increasing in both γ and n_i. Substituting (17) into (10) and (11) yields the equilibrium quantities demanded and supplied at the night markets:

$$q_i^b = A(z^b)^{1/(1+\alpha)}\gamma^\lambda n_i^{-(1-\lambda)} \quad \text{and} \quad q_i^s = A(z^b)^{1/(1+\alpha)}(\gamma n_i)^\lambda. \tag{18}$$

As a result, sellers supply a large quantity of output in response to a large realization of the monetary shock γ. Given the information structure adopted, this is a rational response. The monetary shock increases prices at the night markets, but sellers do not know if the high prices they observe are due to a large realization of γ or to a high realization of n_i. Using whatever information they may have to their best advantage, sellers infer that a high price correlates with a high return on their effort (p_i/γ), so they respond to monetary shocks by supplying more output.

The size of this response depends on the distributions of real and monetary shocks. If monetary shocks are rare and small relative to real shocks, $\sigma_\gamma^2/\sigma_n^2 \to 0$, then $\theta = 1$ and the elasticity of q_i^s with respect to γ is relatively large: $\lambda = 1/(1+\alpha)$. At the other extreme, if the money supply is erratic relative to real shocks, $\sigma_\gamma^2/\sigma_n^2 \to \infty$, then $\theta = 0$ and sellers do not respond to monetary shocks ($\lambda = 0$). In general, θ falls and the elasticity λ decreases with $\sigma_\gamma^2/\sigma_n^2$. In conclusion, suppliers respond most strongly to monetary shocks when these are rare and small.

Since the day market is not affected by monetary shocks and all night markets face the same γ, both aggregate output and aggregate inflation are correlated with γ. Therefore, as emphasized by Lucas (1973), this model generates a short-run upward sloping Phillips curve. However, if monetary authorities were to increase the average rate of inflation by increasing μ_γ, aggregate output would actually fall.[8] This effect of expected inflation on output is not present in Lucas's contributions because with proportional transfers perfectly anticipated inflation has no effect on the demand for money.

6. The general case

The special case studied in the previous section has the double interest of having an explicit solution that is also similar to the classical contribution by Lucas (1973). However, logarithmic preferences rule out potentially interesting forms of

[8] This can be proved combining (18) with (12) and using that γ is log-normally distributed to aggregate the total output produced at the night markets. This derivation can be found in the supplementary appendix.

behavior such as individuals carrying precautionary money balances to be spent only occasionally.[9] Also, with logarithmic preferences the demand for money turns out to be too inelastic to calibrate the model to the United States data. For these two reasons, this section analyzes a version of the model in which the utility of consumption at night has the isoelastic form $u(q_i^b) = (q_i^b)^{1-\eta}/(1-\eta)$. The utility of consumption during the day is still assumed to be logarithmic, $v(x) = B\ln(x)$, since our results have little to do with this functional form.

If the cash constraint (4) is not binding, the new first order condition for optimality of program (7) is $\gamma(q_i^b)^{-\eta} = \beta p_i$. Therefore, once we take into account the cash constraint (4) and use the Kuhn–Tucker theorem, we obtain the following conditional demand for night goods:

$$q_i^b = \min\left\{ \left(\frac{\gamma}{\beta p_i}\right)^{1/\eta}, \frac{z^b \gamma}{p_i} \right\}. \tag{19}$$

This equation together with the Euler condition (9) implicitly determines the optimal demand for money. As with logarithmic preferences, this optimal demand equates the opportunity cost of carrying an extra dollar with the benefit this dollar provides in relaxing the cash constraint in the markets where it is binding. Contrary to logarithmic preferences, the cash constraint is now typically binding in a strict subset of night markets. To see this, notice that (19) implies that the expenditure made by buyers is $q_i^b p_i = (\gamma/\beta)^{1/\eta} p_i^{1(-1/\eta)}$, as long as this amount does not exceed the money balances that they carry ($z^b\gamma$). If $\eta > 1$, this expenditure increases with p_i, so buyers are cash-constrained in markets with high prices. Vice versa, if $\eta < 1$, the expenditure in unconstrained markets is decreasing in p_i and the cash constraint binds in markets with low prices. Therefore, the reason for carrying precautionary balances depends on η. If $\eta > 1$, individuals carry precautionary balances to insure against large reductions in consumption when faced with high prices. In contrast, if $\eta < 1$, individuals carry precautionary balances to make large purchases when they find low prices.

The presence of precautionary balances adds an extra source of elasticity to the demand for money. With precautionary balances, an increase in the opportunity cost of carrying money affects the demand for money along two margins. First, it reduces the demand for night goods in the markets where the cash constraint is binding; second, it increases the fraction of markets where the cash constraint is binding. As a result, as the opportunity cost of holding money increases, buyers economize on money balances at the cost of sacrificing the two advantages of carrying precautionary balances listed above: avoiding reductions in consumption in the presence of high prices (if $\eta > 1$) or taking advantage of price bargains (if $\eta < 1$).

With the demand function (19), the Euler condition (9) does not have an explicit solution for z^b. Consequently, we use numerical methods to explore the predictions of the model. To this end, we calibrate the model using the United States quarterly data on the money supply (M1), output (real GDP), the price level (GDP deflator), and the three-month T-Bill rate from 1947 to 2007. To match the frequency of the data, we assume that one period lasts one quarter. In the baseline calibration, we also assume that buyers and sellers find a trading opportunity every period, so $\pi = 1$. However, to examine the predictions of the model for alternative parameter values, and to facilitate the comparison with the calibrations in Lagos and Wright (2005), we also calibrate the model for $\pi = 0.5$.[10]

Consistent with the quarterly length of one period, the real discount rate is assumed to be 1%, so β is equal to $\exp(-0.01)$. The average growth rate of the money supply μ_γ is set at 0.0086, which is the average quarterly rate of inflation measured using the GDP deflator from 1947 to 2007. The standard deviation σ_γ is set at 0.0036 to match the standard deviation of the innovation in the quarterly growth rate of M1. The remaining four independent parameters[11] (B, η, α, and σ_n) are jointly calibrated to match the average quarterly velocity, its semi-elasticity with respect to the nominal interest rate,[12] and the standard deviations of the cyclical components of real GDP and the GDP deflator induced by monetary shocks (the monetary business cycle).

Although (B, η, α, and σ_n) jointly determine the four moments that we target, each one of these parameters can be associated to the one moment that it affects most directly. The utility weight B determines the output produced during the day, so its value directly affects (positively) the velocity of circulation of money as it is conventionally measured (GDP/M1). The utility parameter η determines the price elasticity of the demand for night goods, so it affects (negatively) the semi-elasticity of velocity with respect to the nominal interest rate. The cost parameter α determines the slope of the supply curve of goods at night, so it controls how responsive prices are to demand shocks. As a result, the standard deviation of the GDP deflator is an increasing function of α. Finally, the standard deviation of real shocks σ_n determines the response of output to monetary disturbances for two reasons. First, it determines how responsive sellers are to nominal prices. If σ_n is large, sellers attribute price increases largely to real shocks, so they respond strongly to the price increases that follow monetary shocks. Second, it affects the incentive for carrying precautionary balances, which determines the fraction of markets that are cash constrained at night. In this case, if σ_n is large, buyers carry abundant precautionary balances. Hence, the fraction of markets that are cash constrained, and responsive to monetary shocks, is small. Consequently, the overall effect of σ_n on the amplitude of the monetary cycle is ambiguous. Numerically, we find that the

[9] See Faig and Jerez (2006) and Telyukova and Wright (2008) for monetary search models that emphasize the role of precautionary balances.

[10] In the calibrations in Lagos and Wright (2005), the probability of trade is even lower than 0.5, e.g. at the monthly frequency they use $\pi = 0.052$. With such low probabilities, we could not generate realistic monetary cycles.

[11] Note that $E(n) = 1$ implies that $\mu_n = -\sigma_n^2/2$.

[12] This is the interest rate on a nominal bond that can only be traded during the day. If it could be traded at night, there would be no need for cash.

effect of σ_n on the standard deviation of GDP is typically humped-shaped: for low values of σ_n the standard deviation of GDP increases with σ_n, but for high values of σ_n the reverse effect takes place.

The four targeted moments in the calibration are calculated as follows. The average quarterly velocity (1.45) is the average ratio of nominal GDP over M1 over our sample period. The interest semi-elasticity of velocity (10.3) is determined regressing the log of velocity on a constant, a time trend, and the three-month Treasury Bill rate. Finally, to take into account the existence of aggregate real shocks abstracted from our model, the cyclical components of real GDP and the GDP deflator are constructed as the projections of the innovations of the growth rates of real GDP or the GDP deflator on a constant and the innovation of the growth rate of M1.[13] The resulting standard deviations of these cyclical components are, respectively, 0.002 and 0.00035 for our sample period.

Our calibration procedure assumes no reverse causation between output and money. With reverse causation, we would be exaggerating the importance and welfare cost of erratic monetary policy, which would only reinforce our conclusion that these costs are small. On the other hand, we are abstracting from long-term effects of monetary shocks (inconsistent with this model), which biases our results in the opposite direction.

Table 1 reports our calibrations of the model. For $\pi = 1$, we found one set of parameter values that fits our targeted moments (column 3). For $\pi = 0.5$, however, we found two sets of parameter values that do so (columns 1 and 2). The reason for this multiplicity is the already mentioned non-monotonic relationship between the standard deviations of real shocks σ_n and real GDP. Comparing columns 1 and 2, σ_n is relatively high in column 2. As a result of this higher value, in this simulation sellers are more responsive to price increases, which they infer to be mainly driven by real shocks. However, the standard deviation of real GDP is the same in both columns. This is in part because other parameter values differ, but mainly because in column 2 the dispersion of prices across markets is larger, so buyers carry more precautionary balances, reducing the fraction of markets that are cash-constrained and subjected to monetary shocks. Precautionary balances are also behind the different values of η in columns 1 and 2. The extra abundance of precautionary balances in column 2 tends to make the demand for money more interest elastic, so the value of η can be much larger in column 2 than in column 1 and still generate the same interest semi-elasticity of velocity in both columns.

In all three calibrations, buyers spend on average a moderate fraction of their money balances, but for different reasons. With a relatively low probability of trade ($\pi = 0.5$), the parameters that fit the data are such that buyers end up spending almost all of their balances (95% and 92%) when they have a trading opportunity. In contrast, if the probability of trade is one, the parameters are such that buyers spend on average a little over half of their money (63%). That is, in the simulation with $\pi = 1$, buyers carry a larger amount of precautionary balances because of a wider dispersion of prices across night markets due to larger real shocks (σ_n is larger) and a steeper supply curve (α is larger).

Once calibrated, we use the model to evaluate the welfare cost of the United States suboptimal monetary policy during the sample period, measured as the equivalent variation of income as a percentage of GDP. That is, we calculate how much residents of the United States would have been willing to pay to face the optimal Friedman rule instead of the prevalent monetary policy. In all three simulations, we find that this payment is approximately 0.25% of GDP. We, also, decompose how much of this welfare cost is due to the monetary business cycle generated by an erratic monetary policy, and how much it is due to a positive opportunity cost of holding money (inflation tax). We find that eliminating the monetary business cycle contributes very little to the welfare gain of implementing the Friedman rule. The welfare gain of reducing σ_γ to zero and adjusting μ_γ to keep the expected return on money ($E\gamma^{-1}$) unchanged[14] is very small in all three simulations ($<0.0003\%$ of GDP). Surprisingly, eliminating the monetary business cycle in the simulations with $\pi = 0.5$ is actually detrimental to welfare. This counter-intuitive sign is due to the interaction between the monetary business cycle and the welfare cost of expected inflation. With monetary shocks, the standard deviation of night prices increases. As a result, buyers decide to hold more precautionary balances, which partly offsets the inefficiently low money demand due to the inflation tax. If the opportunity cost of holding money is close to zero (not in Table 1), then eliminating the monetary cycles does not reduce welfare for any set of parameters. Comparing across simulations, we can observe that the welfare cost of the monetary cycle is inversely related to the fraction of GDP produced in cash-constrained markets. Intuitively, since we control for the amplitude of the cycle, if monetary shocks affect a smaller sector of the economy, the adjustments in this sector must be larger. Hence, they are more detrimental to welfare because welfare cost increases more than proportionally with the deviations from efficiency.

The small gains from reducing the monetary business cycle is just another example of the small welfare gains achieved by further reducing the post-war business cycles in the United States, as emphasized by Lucas (1987, 2003). As Lucas pointed out, the standard deviation of aggregate consumption around its trend has been small during the post-war era. The amplitude of the monetary cycle is even smaller. Therefore, for the welfare gains of further stabilization to be important, one would need that either stabilization raised average consumption, or that the business cycle affected segments of the economy where individuals had a high intolerance to fluctuations. In the present setup, this is not the case. Average consumption is barely changed by σ_γ, and the goods affected by the monetary business cycle are those money-transacted

[13] These innovations were constructed with regressions on a constant and sixteen lags of the three-month Treasury Bill rate and the first differences of the logs of M1, real GDP, and the GDP deflator, allowing for a structural break at the beginning of the Great Moderation (1984-I) in both the intercept and the slope coefficients. The projections used to calculate the monetary cycles of output and prices also allow for this structural break, which statistically is strongly significant.

[14] Without this adjustment, the welfare cost of monetary cycles is even smaller.

Table 1
Except for the trade probabilities (π), the parameters of the three simulations were calibrated to match properties of the demand for money and the monetary business cycle in the United States from 1947 to 2007.

Parameters			
π	0.5	0.5	1
α	0.185	0.193	0.242
B	0.413	0.446	0.422
η	0.043	0.103	0.417
σ_n	0.342	0.508	2.165
Simulation			
Fraction of GDP at night	0.329	0.316	0.433
Fraction of GDP cash-constrained	0.239	0.194	0.096
Average fraction of money used at night	0.477	0.459	0.628
Average fraction of money used in a purchase	0.954	0.918	0.628
Standard deviation p_i	0.048	0.076	0.755
Welfare losses monetary policy			
Total dead-weight-loss (% of GDP)	0.2510	0.2479	0.2505
Monetary business cycle (% of GDP)	−0.0003	−0.0001	0.0003
Expected inflation (% of GDP)	0.2513	0.2480	0.2502
Monetary crashes			
Welfare cost average crash (% of GDP)	0.5383	0.4449	0.4042
Welfare cost monetary business cycle (% of GDP)	0.0054	0.0049	0.0018

The dead-weight-losses reported are percentages of GDP that residents in the United States would be willing to pay to face the optimal policy. These dead-weight-losses are the sum of the losses due to unexpected monetary shocks ($\sigma_\gamma > 0$), and the losses due to a positive opportunity cost of holding money because inflation is above the Friedman rule ($\beta E(\gamma^{-1}) < 1$).

goods for which buyers' tolerance to fluctuations in consumption is high (as implied by the high elasticity of the demand for money). In our simulations there are two additional reasons why the welfare costs of the monetary cycle are small. First, the monetary cycle reduces the inefficiencies caused by the inflation tax, as explained in the previous paragraph. Second, the supply curves of night goods are fairly flat, so sellers are quite ready to accept fluctuations in production. (If one were to increase α without changing the other parameters, the welfare cost of the monetary cycle would increase.)

Even if our simulations find a very small welfare gain of eliminating the monetary cycle experienced in the United States from 1947 to 2007, one should not conclude that monetary disturbances are not important for welfare. To show this, for each set of calibrated parameters, we calculate the stochastic equilibrium with an altered distribution of monetary shocks that includes a small probability (once every 50 years) of a monetary crash, defined as a one period drop of μ_γ to -0.1. (Drops of 10% of M1 at the monthly frequency occurred twice from 1929 to 1946). The second last row of Table 1 shows that the welfare cost of experiencing such a monetary crash in a given period is equivalent to a non-trivial drop of GDP of around 0.5%. However, since monetary crashes are rare in these simulations, the overall welfare cost of monetary cycles continues to be small (last row of the table).

7. Conclusion

Lucas's non-neutrality theory can be successfully incorporated into the framework of Lagos and Wright (2005), where both money is essential and individuals optimize. In general, the solution of the model is more complicated than in Lucas's (1972, 1973) classical contributions because, depending on market conditions, money holders do not spend all the money they carry. However, an exception to this general rule occurs with logarithmic utility for goods purchased with money. In this special case, buyers always spend all the money they carry, and the model can be explicitly solved to obtain a reduced form solution similar to that of Lucas (1973).

Without logarithmic preferences, buyers hold precautionary balances to either insure against high prices, if their demand for goods is inelastic, or take advantage of low prices, if their demand for goods is elastic. These precautionary balances have some remarkable implications. First, they endogenously reduce the velocity of circulation of money and provide an extra source of interest elasticity in the demand for money. As a result, we are able to calibrate the model to the United States data without relying on infrequently trade opportunities or very high elasticities of the demand for cash goods. Second, the presence of precautionary balances implies that eliminating the monetary cycle may be detrimental to welfare. The reason for this counter-intuitive result is that the monetary cycle increases the standard deviation of prices, so it tends to reduce the inefficiencies of the inflation tax by increasing the demand for money.

When the model is calibrated to United States postwar data, we find that the welfare cost of the monetary business cycle is very small ($< 0.0003\%$ of GDP), while the welfare cost of the inflation tax is several orders of magnitude larger (around 0.25% of GDP). However, it would be erroneous to conclude from these calculations that monetary disturbances cannot generate significant welfare costs. The main reason why the welfare cost of the monetary cycle from 1947 to 2007 has been so small in our calculations is because monetary shocks during this period have been small. When we consider the possibility of large drops in the money supply (of the order of magnitude of those experienced during the Great Depression), the welfare costs become substantial.

References

Adam, K., 2007. Optimal monetary policy with imperfect common knowledge. Journal of Monetary Economics 54, 267–301.
Angeletos, G.M., La'O, J., 2008. Dispersed information over the business cycle: optimal fiscal and monetary policy. Manuscript.
Bailey, M.J., 1956. The welfare cost of inflationary finance. Journal of Political Economy 64, 93–110.
Faig, M., Jerez, B., 2006. Precautionary balances and the velocity of the circulation of money. Journal of Money, Credit, and Banking 39, 843–873.
Hellwig, C., 2005. Heterogeneous information and the welfare effects of public information disclosures. UCLA Economics Online Papers 283.
Katzman, B., Kennan, J., Wallace, N., 2003. Output and the price level effects of monetary uncertainty in a matching model. Journal of Economic Theory 108, 217–255.
King, R.G., 1982. Monetary policy and the information content of prices. Journal of Political Economy 90, 247–279.
Lagos, R., Wright, R., 2005. A unified framework for monetary theory and policy analysis. Journal of Political Economy 113, 463–484.
Lorenzoni, G., 2009. Optimal monetary policy with uncertain fundamentals and dispersed information. Review of Economic Studies, forthcoming.
Lucas, R.E., 1972. Expectations and the neutrality of money. Journal of Economic Theory 4, 103–124.
Lucas, R.E., 1973. Some international evidence on output-inflation tradeoffs. American Economic Review 63, 326–334.
Lucas, R.E., 1975. An equilibrium model of the business cycle. Journal of Political Economy 83, 1113–1144.
Lucas, R.E., 1987. Models of Business Cycles. Basil Blackwell, New York.
Lucas, R.E., 2003. Macroeconomic priorities. American Economic Review 93, 1–14.
Morris, S., Shin, H.S., 2003a. Global games: theory and applications. In: Hansen, L., Dewatripont, M., Turnovsky, S. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Proceedings of the Eighth World Congress, vol. 1. Cambridge University Press, Cambridge, pp. 56–114.
Morris, S., Shin, H.S., 2003b. Social value of public information. American Economic Review 92, 1521–1534.
Rocheteau, G., Wright, R., 2005. Money in search equilibrium in competitive equilibrium and in competitive search equilibrium. Econometrica 73, 175–202.
Telyukova, I.A., Wright, R., 2008. A model of money and credit with application to the credit card debt puzzle. Review of Economic Studies 75, 629–647.
Ui, T., 2003. A note on the Lucas model: iterated expectations and the neutrality of money. Manuscript.
Wallace, N., 1997. Short-run and long-run effects of changes in money in a random-matching model. Journal of Political Economy 105, 1293–1307.
Woodford, M., 2002. Imperfect common knowledge and the effects of monetary policy. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps. Princeton University Press, Princeton, pp. 25–58.