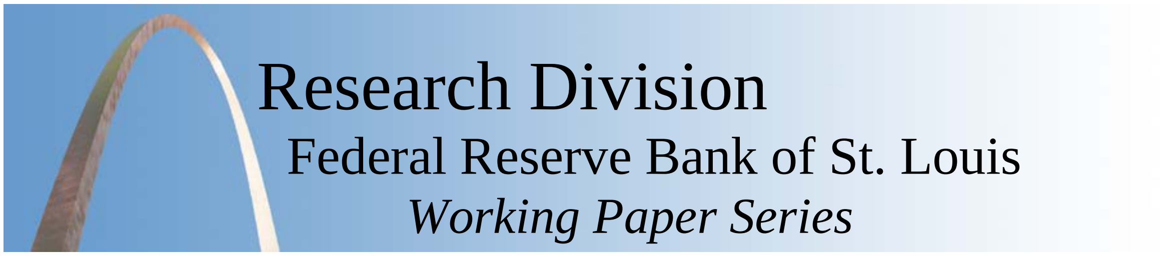

On the Theoretical Efficacy of Quantitative Easing at the Zero Lower Bound

Paola Boel
and
Christopher Waller

Working Paper 2015-027A
http://research.stlouisfed.org/wp/2015/2015-027.pdf

September 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

On the Theoretical Efficacy of Quantitative Easing at the Zero Lower Bound*

Paola Boel†
Sveriges Riksbank

Christopher Waller‡
FRB Saint Louis

September 2015

Abstract

We construct a monetary economy in which agents face aggregate demand shocks and heterogenerous idiosyncratic preference shocks. We show that, even when the Friedman rule is the best interest rate policy, not all agents are satiated at the zero lower bound. Thus, quantitative easing can be welfare improving since it temporarily relaxes the liquidity constraint of some agents, without harming others. Moreover, due to a pricing externality, quantitative easing may also have beneficial general equilibrium effects for the unconstrained agents. Lastly, our model suggests that it can be optimal for the central bank to buy private debt claims instead of government debt.

Keywords: Money, Heterogeneity, Stabilization Policy, Zero Lower Bound, Quantitative Easing

JEL codes: E40, E50

*We thank participants in workshops and seminars at the Chicago Fed, the St. Louis Fed, the Riksbank, Deakin University, University of Sydney and Macquarie University. We especially thank David Andolfatto, Daria Finocchiaro and Per Krusell for helpful conversations. The usual disclaimers apply.

†Research Division, Sveriges Riksbank, SE-103 37 Stockholm, Sweden. Email: paola.boel@riksbank.se.

‡Corresponding author. Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA. Email: cwaller@stls.frb.org.

1 Introduction

In the aftermath of the financial crisis, central banks around the world have pursued a range of unconventional policies to stabilize their economies. The first of these has been driving the policy rate to the zero lower bound (denoted ZLB). Another has been to inject large amounts of liquidity into the economy via large scale asset purchases, also known as quantitative easing (denoted QE). These two policy actions have generated a substantial debate in the economics profession about the role of monetary policy at the ZLB and the efficacy of QE. First, is hitting the ZLB a problem for a central bank? Second, what are the theoretical channels causing QE to have real effects at the ZLB?

Regarding the first question, in several monetary models[1] driving the nominal interest rate to zero, also known as the Friedman rule, is not a constraint. It is instead the optimal policy, and as such it should be implemented all of the time and not just during severe economic downturns.[2] In New Keynesian models, instead, a demand shock of some sort drives the "natural" real interest sufficiently negative. The central bank would like to have a negative nominal interest rate but because of sticky prices and the ZLB the equilibrium real interest rate is too high and economic activity is inefficiently low. In all these models, once the nominal interest rate is zero there is no further role for monetary policy. The Friedman rule is associated with a "liquidity trap" and varying the quantity of money in the economy has no real or nominal effects – households will simply hold onto the excess cash since it is costless to do so.

Regarding the second question, conventional QE policy involves printing money to buy government bonds and/or private assets. Consequently, QE alters the size and composition of a central bank's balance sheet. However, why should the size or the composition of the central bank balance sheet matter for the real economy? In short, does Modigliani-Miller apply to a central bank's balance sheet or not? Wallace (1981) showed that it does and thus QE-type policies will be ineffective. Eggertsson and Woodford (2003) and Cúrdia and Woodford (2011) show a similar result in a New Keynesian model once the ZLB is reached. Moreover, at the ZLB, money and bonds are perfect substitutes. Exchanging one for the other has no real or nominal effects.[3] In a New Monetarist model, Williamson (2014) shows QE can flatten the yield curve, but it also leads to an

[1] For example, MIU, CIA and New Monetarist models.

[2] This is true unless factors such as income shocks, as in Akyol (2004), redistributive issues, as in Bhattacharya, Haslag and Martin (2005) or Ireland (2005), distortionary taxes, as in Phelps (1973) or trading externalities, as in Shi (1997) are taken into account.

[3] This argument has been advanced by John Cochrane in the blogosphere.

increase in real rates and a decrease in inflation at the ZLB, much unlike central bankers' thinking.[4] In summary, showing theoretically that QE can have beneficial real effects has proven difficult to achieve.[5]

Yet, there is substantial empirical evidence showing that quantitative easing has non-trivial effects on the yields of various financial assets.[6] This is why QE is perceived to have had beneficial real effects on the economy even at the ZLB. Consequently, there is a tension between theory and empirical evidence in terms of the effects of QE on the real economy. This is best captured by Ben Bernanke's famous quote "The problem with QE is it works in practice but it doesn't work in theory." In the end, we are still confronted with the question of how QE can be beneficial at the zero lower bound.

Another conundrum with recent experience is that if the ZLB corresponds to the Friedman rule, then this should be associated with deflation. Japan is often cited as an example of this. However, in several of the major economies where the ZLB has been in place for many years, inflation and expected inflation are positive and lie between 1% and 2%.[7] So, why is deflation not occurring in these economies? This puzzle gives rise to yet another one – with nominal interest rates at zero and expected inflation anchored between 1% and 2%, some agents in the economy are willing to accept a negative ex-ante real rate of return on their assets several years out into the future. Why are they willing to do so?

Our objective is to build a New Monetarist model that can be used to address the questions and observations discussed above. Specifically, we amend the New Monetarist model of Berentsen and Waller (2011) by incorporating heterogeneous preferences across agents as in Eggertsson and Krugman (2012). We then study the optimal stabilization response of the central bank to demand shocks that hit the economy. Heterogeneity takes the form of agents receiving iid shocks to their current discount factor. Specifically, in each period, some agents receive shocks such that their discount factor is higher than for the others. We refer to the former agents as patient and the latter ones as impatient. In this setup, patient agents are savers and they are the marginal investors

[4] In other New Monetarist models, Williamson (2012), Berentsen and Waller (2011) and Herrenbrueck (2014) find QE is not welfare improving at the ZLB.

[5] Woodford (2012) and Bhattarai, Eggertsson and Gafarov (2013) have argued QE may have real effects by reinforcing forward guidance – by increasing the size of the central bank balance sheet and exposing it to capital losses if interest rates rise, the central bank commits to keeping interest rates lower than is 'optimal'.

[6] See Krishnamurthy and Vissing-Jorgensen (2011), Hamilton and Wu (2012) and Gagnon and others (2010) for example.

[7] E.g., this was true during the global financial crisis and ensuing recession in the United States, Canada, England and the European Monetary Union among other countries. Sources: Board of Governors of the Federal Reserve System, table H.15, "Selected Interest Rates;" Bureau of Economic Analysis data; Consensus Economics, "Consensus Forecasts;" foreign central bank data.

who price assets. We show that the ZLB may be the best interest rate policy the central bank can implement, although that is not always the case due to a price externality. Nevertheless, even when the Friedman rule is the best policy, it is always second best since only the patient agents are unconstrained in their holdings of real balances. The impatient agents, instead, are still constrained and carry fewer real balances across time. Why? The nominal interest rate is too high at zero for them as the effective nominal rate they face is still positive. The policymaker would like to drive down the nominal interest rate into negative territory to lower the opportunity cost for impatient agents to carry cash, but is constrained in its ability to do so.

We then consider a particular form of QE whereby the central bank purchases private debt via repo arrangements. Since the repos are undone at a later date, QE is a temporary policy, not a permanent increase in the size of the central bank balance sheet. We show that, under this form of QE in response to demand shocks, the central bank is able to temporarily relax the liquidity constraint on impatient agents. This improves their welfare without harming the patient agents. Interestingly, our model also suggests that it can be beneficial for the central bank to buy private debt claims instead of government debt. In this sense, the purchases of mortgage backed securities conducted by the Fed are consistent with our model. Furthermore, we demonstrate that due to the pricing externality mentioned before, QE may also have beneficial general equilibrium effects for the patient agents even if they are unconstrained in their holdings of real balances.[8] Consequently, QE at the ZLB is welfare improving for some if not all agents.

As is common in the New Keynesian literature [e.g., Eggertsson and Woodford (2003) and Bhattarai, Eggertsson and Gafarov (2013)], we then consider a temporary, unanticipated shock to the discount factor. We focus on a shock such that patient agents have a discount factor greater than one. This in turn implies that patient agents are willing to take negative real rates of interest over a short period of time. We show that in this case maintaining the ZLB will generate positive expected inflation in the short run. As before, QE continues to be effective in improving welfare.

The structure of the paper is as follows. Section 2 presents the model. Section 3 discusses the efficient allocation. Section 4 studies stationary monetary equilibrium. Section 5 discusses the optimal inflation rate in an economy without aggregate demand shocks. Section 6 studies optimal stabilization policy at the ZLB in response to aggregate demand shocks. Section 7 concludes. Proofs are in the Appendix.

[8] Rojas Breu (2013) and Berentsen, Huber and Marchesiani (2014) get similar pricing externalities when agents have differing abilities to pay for goods. However, at the ZLB the pricing externality disappears in their models whereas that is not the case in our environment.

2 The model

The model builds on Lagos and Wright (2005), Boel and Camera (2006) and Berentsen and Waller (2011). Time is discrete, the horizon is infinite and there is a large population of infinitely-lived agents who consume perishable goods and discount only across periods. In each period agents may visit two sequential rounds of trade - we will refer to the first as DM and the second as CM.

Rounds of trade differ in terms of economic activities and preferences. In the DM, agents face an idiosyncratic trading risk such that they either consume, produce, or are idle. An agent consumes with probability α_b, produces with probability α_s and is idle with probability $1-\alpha_b-\alpha_s$. We refer to consumers as buyers and producers as sellers. Buyers get utility $\varepsilon u(q)$ from $q > 0$ consumption, where ε is a preference parameter, $u'(q) > 0$, $u''(q) < 0$, $u'(0) = +\infty$ and $u'(\infty) = 0$. Furthermore, we impose that the elasticity of utility $e(q) = qu'(q)/u(q)$ is bounded. Producers incur a utility cost $c(y)$ from supplying $y \geq 0$ labor to produce y goods, with $c'(y) > 0$, $c''(y) \geq 0$ and $c'(0) = 0$. Everyone can consume and produce in the CM. As in Lagos and Wright (2005), agents have quasilinear preferences $U(x) - n$, where the first term is utility from x consumption, and the second is disutility from n labor to produce n goods. We assume $U'(x) > 0$, $U''(x) \leq 0$, $U'(0) = +\infty$ and $U'(\infty) = 0$. Also, let q^* be the solution to $\varepsilon u'(q) = c'[(\alpha_b/\alpha_s)q]$ and let x^* be the solution to $U'(x) = 1$.

2.1 Shocks

The economy is subject to both aggregate and idiosyncratic demand shocks, but agents are heterogeneous only with respect to the latter. Specifically, at the beginning of each CM, agents draw an idiosyncratic time-preference shock $\beta_z \in \{\beta_L, \beta_H\}$ determining their interperiod discount factor. This implies at the beginning of each period an agent can be either patient (type H) with probability ρ or impatient (type L) with probability $1-\rho$. We consider the case $0 < \beta_L < \beta < \beta_H < 1$ with no serial correlation in the draws and β being the average discount factor. Note that time-preference shocks introduce ex-post heterogeneity across households, but the long-run distribution of time preferences is invariant.

We also assume ε is stochastic like in Berentsen and Waller (2011), which allows us to study the optimal response of a central bank to aggregate demand shocks. The random variable ε has support $\Omega = [\underline{\varepsilon}, \bar{\varepsilon}]$, with $0 < \underline{\varepsilon} < \bar{\varepsilon} < \infty$. Shocks are serially uncorrelated and $f(\varepsilon)$ denotes the density function of ε. As shown below, output in the CM is constant so any volatility in total output per period is driven by ε shocks in the DM.

2.2 Information frictions, money and credit

The preference structure we selected generates a single-coincidence problem in the DM since buyers do not have a good desired by sellers. Moreover, two additional frictions characterize the DM. First, agents are anonymous as in Kocherlakota (1998), since trading histories of agents in the goods markets are private information. This in turn rules out trade credit between individual buyers and sellers. Second, there is no public communication of individual trading outcomes, which in turn eliminates the use of social punishments to support gift-giving equilibria. The combination of these two frictions together with the single coincidence problem implies that sellers require immediate compensation from buyers. So, buyers must use money to acquire goods in the DM.

Money is not essential for trade in the CM instead, and indeed agents can finance their consumption by getting credit, working or using money balances acquired earlier. To model credit, we assume agents are allowed to borrow and lend through selling and buying nominal bonds, subject to an exogenous credit constraint A. Specifically, agents lend $-p_{at}a_{t+1}$ (or borrow $p_{at}a_{t+1}$), where p_{at} is the price of a bond that delivers one unit of money in $t+1$, and receive back a_t. We assume that any funds borrowed or lent in the CM are repaid in the following CM. One can show that, even with quasi-linearity of preferences in the CM, there are gains from multi-period contracts due to time-preference shocks. Of course, default is a serious issue in all models with credit. However, to focus on optimal stabilization, we simplify the analysis by assuming a mechanism exists that ensures repayment of loans in the CM.

2.3 Policy tools

We assume a government exists that is in charge of monetary policy and is the only supplier of fiat money, of which an initial stock $M_0 > 0$ exists. Monetary policy has both a long-run and a short-run component. The long-run policy focuses on the trend inflation rate, whereas the short-run one is concerned with the output stabilization response to aggregate shocks. We denote the gross growth rate of money supply by $\pi = M_t/M_{t-1}$, where M_t denotes the money stock in the CM in period t. The central bank implements its long-term inflation goal by providing deterministic lump-sum injections of money $\tau = (\pi - 1)M_{t-1}$, which are given to private agents at the beginning of the CM. If $\pi > 1$, agents receive lump-sum transfers of money, whereas for $\pi < 1$ the central bank must be able to extract money via lump-sum taxes from the economy.

The central bank implements its short-term stabilization policy through state-contingent changes in the stock of money. We let $\tau_1(\varepsilon) = T_1(\varepsilon)M_{t-1}$ and $\tau_2(\varepsilon) = T_2(\varepsilon)M_{t-1}$ denote state-contingent cash injections received by private agents in the DM and CM respectively. We assume injections in the DM are undone in the CM, so that $\tau_1(\varepsilon) + \tau_2(\varepsilon) = 0$. Changes in $\tau_1(\varepsilon)$ thus affect the money stock in the DM without affecting the long-term inflation rate in the CM. This means that the long-term inflation rate is still deterministic since $\tau = (\pi - 1)M_{t-1}$ is not state dependent. Note that the state-contingent injections of cash can be viewed as a type of repurchase agreement - the central bank sells money in the DM under the agreement that it is being repurchased in the CM.

3 Efficient allocation

We start by discussing the allocation selected by a benevolent planner subject to the same physical and informational constraints faced by the agents. We will refer to this allocation as constrained-efficient. The environment's frictions imply the planner can observe neither types nor identities in the DM and therefore has no ability to transfer resources across agents over time in that market. Therefore, the planning problem in the DM corresponds to a sequence of static maximization problems subject to the technological constraints. This implies in the DM the planner must solve:

$$\begin{aligned} &\underset{q,y}{Max} \int_{\Omega} \{\alpha_b \varepsilon u[q(\varepsilon)] - \alpha_s c[y(\varepsilon)]\} f(\varepsilon) d\varepsilon \\ &\text{s.t. } \alpha_b q(\varepsilon) = \alpha_s y(\varepsilon) \end{aligned} \tag{1}$$

In the CM, instead, the planner can transfer resources across agents over time. Therefore, in that market she chooses consumption and labor sequences $\{x_{j0}, x_{j1}, ..\}$ and $\{n_{j0}, n_{j1}, ..\}$ for $j = H, L$ that maximize a weighted sum of individual utility functions subject to feasibility and non-negativity constraints:

$$\begin{aligned} &Max \sum_{j=H,L} \sigma_j \left[U(x_{j0}) - n_{j0} + \sum_{t=1}^{\infty} \beta_j \beta^{t-1} (U(x_{jt}) - n_{jt}) \right] \\ &\text{s.t. } \rho x_{Ht} + (1-\rho) x_{Lt} = \rho n_{Ht} + (1-\rho) n_{Lt} \quad \text{for } t = 0, 1, 2, ... \\ &\text{s.t. } n_{jt} \geq 0 \quad \text{for } j = H, L \text{ and } t = 0, 1, 2, ... \end{aligned} \tag{2}$$

Here σ_H and σ_L are positive utility weights. A solution to this problem is characterized by:

$$U'(x_{j0}) = 1 - \mu_t^j/\sigma_j \qquad \text{for } j = H, L \text{ and } t = 0 \tag{3}$$

$$U'(x_{jt}) = 1 - \mu_t^j/\sigma_j \beta_j \beta^{t-1} \qquad \text{for } j = H, L \text{ and } t \geq 1 \tag{4}$$

where μ_t^j denotes the Kuhn-Tucker multiplier associated with the non-negativity constraint on n_{jt}. Note that the difference between equation (3) and (4) implies a different allocation when $t = 0$ than when $t \geq 1$. The social planner problem is therefore time inconsistent. These differences, however, disappear when $\mu_t^j = 0$. Therefore, the time-consistent solution to the planner's problem consists in $n_{jt} > 0$ and therefore $U'(x_{jt}) = 1$ for $j = H, L$ and $t \geq 0$. The planner wants both types to work and consume a constant and equal amount in every period.

In sum, in the constrained-efficient allocation marginal consumption utility equals marginal production disutility in each market and in every period. Such allocation is therefore stationary and defined by $\varepsilon u'[q(\varepsilon)] = c'[(\alpha_b/\alpha_s)q(\varepsilon)]$ for all ε in the DM and $U'(x) = 1$ in the CM. The constrained-efficient consumption is therefore defined uniquely by $q_H = q_L = q^*$ and $x_H = x_L = x^*$, thus implying equal consumption for type H and type L agents. This allocation is the relevant benchmark in our study, and we will refer to it as efficient.

4 Stationary monetary allocations

In what follows, we want to determine if the constrained-efficient allocation can be decentralized in a monetary economy with competitive markets.[9] Thus, we focus on stationary monetary outcomes such that end-of-period real money and bonds balances are time invariant.

We simplify notation omitting t subscripts and use a prime superscipt and a -1 subscript to denote variables corresponding to the next and previous period respectively. We let p_1 and p_2 denote the nominal price of goods in the DM and the CM respectively of an arbitrary period t. We also let β_j and β_z denote the discount factors drawn in period $t-1$ and t respectively. In addition, we normalize all nominal variables by p_2, so that DM trades occur at the real price $p = p_1/p_2$. In this manner, the timing of events in any period t can be described as follows.

An arbitrary agent of type $j = H, L$ enters the DM in period t with a portfolio $\omega_j = (m_j, a_j)$ listing $m_j = m(\beta_j)$ real money holdings and $a_j = a(\beta_j)$ loans (or savings) from the preceding

[9] Competitive markets in the Lagos and Wright (2005) framework have been studied by Rocheteau and Wright (2005) and Berentsen, Camera and Waller (2007) among others.

period when he experienced a time-preference shock β_j. Trading shocks k and aggregate demand shocks ε are then realized and agents receive a lump-sum transfer $\tau_1(\varepsilon) = T_1(\varepsilon)M_{-1}$. After the DM closes, the agent enters the CM with portfolio $\omega_j^k = (m_j^k, a_j)$, where $m_j^k = m_j^k(\beta_j, \varepsilon)$ denotes money holdings carried over from the DM and $k = s, b, o$ denotes the trading shock experienced in the DM. Here, o identifies an idle agent, while b and s identify a buyer and a seller respectively. Thus, if we let $q_j = q(\beta_j, \varepsilon)$ denote consumption and $y_j = y(\beta_j, \varepsilon)$ production in the DM, individual real money holdings for an agent j evolve as follows:

$$\begin{aligned} m_j^b &= m_j + \tau_1 - pq_j \\ m_j^s &= m_j + \tau_1 + py_j \\ m_j^o &= m_j + \tau_1 \end{aligned} \tag{5}$$

That is, buyers deplete balances by pq_j while sellers increase them by py_j. Idiosyncratic time-preference shocks β_z are realized at the beginning of the CM. Left-over cash is then used to trade and settle bonds positions and x and n are respectively consumption bought and production sold in the CM. Note that bonds positions a_j at the beginning of the CM are not affected by trading shocks in the DM, since they can only be used in the CM. Agents also receive lump-sum transfers $\tau + \tau_2(\varepsilon)$, adjust their money balances $m'_z = m'(\beta_z, \varepsilon)$ and decide whether they want to borrow or lend $a'_z = a'(\beta_z, \varepsilon)$, where m'_z and a'_z denote real values of money holdings and loans (or savings if $a'_z < 0$) at the start of tomorrow's DM. Figure 1 displays the timeline of shocks and decisions within each period:

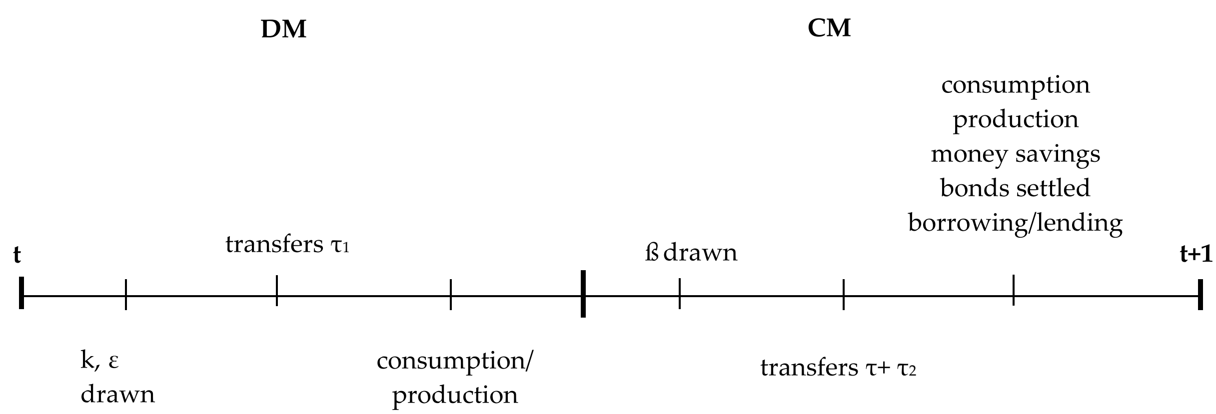

Figure 1: Timing of events within a period

Since we focus on stationary equilibria where end-of-period real money balances are time and state invariant so that $M/p_2 = M'/p_2'$, we have that

$$\frac{p_2'}{p_2} = \frac{M'}{M} = \pi, \tag{6}$$

which implies the inflation rate equals the growth rate of money supply. The government budget constraint therefore is:

$$\tau = (\pi - 1)[\rho m_H + (1-\rho)m_L] \tag{7}$$

Note that the long-run inflation rate is deterministic since the per capita lump-sum transfers τ in the CM are not state dependent.

4.1 The CM problem

Given the recursive nature of the problem, we use dynamic programming to analyze the problem of an agent j at any date, with $j = H, L$. We let $V(\omega_j)$ denote the expected lifetime utility for an agent entering the DM with portfolio ω_j before shocks are realized. We also let $W_z(\omega_j^k)$ denote the expected lifetime utility from entering the CM with portfolio ω_j^k and receiving a discount factor shock β_z at the beginning of the CM.

The agent's problem at the start of the CM then is:

$$\begin{aligned} W_z(\omega_j^k) = \underset{x_{jz}^k, n_{jz}^k, a_z', m_z'}{Max} \quad & U(x_{jz}^k) - n_{jz}^k + \beta_z V(\omega_z') \\ \text{s.t. } & x_{jz}^k + \pi m_z' = n_{jz}^k + m_j^k + p_a \pi a_z' - a_j + \tau + \tau_2 \\ \text{s.t. } & a_z' \leq A \\ \text{s.t. } & m_z' \geq 0 \end{aligned} \tag{8}$$

where $A \geq 0$ is a constant denoting an exogenous borrowing constraint. The resources available to the agent in the CM depend on the realization of the DM trading shock k, as well as the aggregate and idiosyncratic shocks ε, β_j and β_z. Specifically, an agent has m_j^k real balances carried over from the DM and is able to borrow $\pi a_z'$ (or lend if $a_z' < 0$) at a price p_a. Other resources are n_{jz}^k receipts from current sales of goods and lump-sum transfers $\tau + \tau_2$. These resources can be used to finance current consumption x_j, to pay back loans a_j and to carry $\pi m_z'$ real money balances into next period. The factor $\pi = p_2'/p_2$ multiplies a_z' and m_z' because the budget constraint is expressed in real terms.

Rewriting the constraint in terms of n_{jz}^k and substituting into (8) yields:

$$
\begin{aligned}
W_z(\omega_j^k) = \underset{x_{jz}^k, a_z', m_z'}{Max} \; & U(x_{jz}^k) - x_{jz}^k - \pi m_z' + \pi p_a a_z' - a_j + m_j^k + \tau + \tau_2 + \beta_z V(\omega_z') \\
\text{s.t. } & a_z' \leq A \\
\text{s.t. } & m_z' \geq 0
\end{aligned}
$$

Note that here we are focusing on a stationary equilibrium in which all agents provide a positive labor effort. Conditions for $n_{jz}^k > 0$ are in the Appendix, but the intuition is that agents will always choose to work in the CM if the borrowing limit A is tight enough. It follows that in a stationary monetary economy we must have:

$$
1 = \frac{\partial W_z(\omega_j^k)}{\partial m_j^k} = -\frac{\partial W_z(\omega_j^k)}{\partial a_j} \tag{9}
$$

This result depends on the quasi linearity of the CM preferences and the use of competitive pricing. It implies that the marginal valuation of real balances and bonds in the CM are identical and do not depend on the agent's current type z or past type j, wealth ω_j^k or trade shock. This allows us to disentangle the agents' portfolio choices from their trading histories since

$$
W_z(\omega_j^k) = W_z(0) + m_j^k - a_j,
$$

i.e. the agent's expected value from having a portfolio ω_j^k at the start of a CM is the expected value from having no wealth, $W_z(0)$, letting $\omega_j = (0, 0) \equiv 0$, plus the current real value of net wealth $m_j^k - a_j$. Note also that everyone consumes identically in the CM since:

$$
U'(x) = 1 \tag{10}
$$

which also implies $x = x^*$. That is, everyone consumes the same amount x^* independent of current type and past shocks, the reason being that agents in the CM can produce any amount at constant marginal cost. Thus, goods market clearing in the CM requires:

$$
x^* = \alpha_b N^b + \alpha_s N^s + (1 - \alpha_b - \alpha_s) N^o \tag{11}
$$

where $N^k = \rho^2 n^k_{HH} + \rho(1-\rho)(n^k_{LH} + n^k_{HL}) + (1-\rho)^2 n^k_{LL}$ is labor effort for all agents who experienced a trading shock in the DM, with $k = b, s, o$. Let $\mu^m_z \geq 0$ denote the Kuhn-Tucker multipliers associated with the non-negativity constraint for money. Also, let λ^a_z denote the multiplier on the borrowing constraint. The first order conditions for the optimal portfolio choice then are:

$$1 = \frac{\beta_z}{\pi} \frac{\partial V(\omega'_z)}{\partial m'_z} + \mu^m_z/\pi \tag{12}$$

$$-p_a = \frac{\beta_z}{\pi} \frac{\partial V(\omega'_z)}{\partial a'_z} - \lambda^a_z/\pi \tag{13}$$

The left hand sides of the expressions above define the marginal cost of the assets. The right hand sides define the expected marginal benefit from holding the asset, either money or bonds, discounted according to time preferences and inflation. From (12) and (13) we know that money holdings m'_z and bonds a'_z are independent of trading histories and past demand shocks, but instead depend on the current type z and the expected marginal benefit of holding money and bonds in the DM, which may differ across types. We will study this next.

4.2 The DM problem

An agent with a portfolio ω_j at the opening of the DM before aggregate demand and trading shocks are realized has expected lifetime utility:

$$V(\omega_j) = \int_\Omega \left\{ \alpha_b V^b(\omega_j) + \alpha_s V^s(\omega_j) + (1 - \alpha_b - \alpha_s) V^o(\omega_j) \right\} f(\varepsilon) d\varepsilon \tag{14}$$

First, we determine y_j. The seller's problem depends on the current disutility of production and the expected continuation value. Specifically, the seller's problem can be written as:

$$V^s(\omega_j) = \underset{y_j}{Max} \quad -c(y_j) + \rho W_H(\omega^s_j) + (1-\rho) W_L(\omega^s_j), \tag{15}$$

for which the first order conditions, together with (5) and (9), give:

$$c'(y_j) = p \tag{16}$$

Note that (16) implies production is not type dependent, i.e. $y_j = y$ for $j = H, L$.

Now, we determine q_j. A buyer's problem is:

$$V^b(\omega_j) = \underset{q_j}{Max} \quad \varepsilon u(q_j) + \rho W_H(\omega_j^b) + (1-\rho) W_L(\omega_j^b) \tag{17}$$
$$\text{s.t. } pq_j \leq m_j + \tau_1$$

The budget constraint reflects that consumption can be financed with both money holdings and DM transfers. Let λ_j^b denote the multiplier on the buyer's budget constraint. Using (5) and (9), the first order conditions for the buyer's problem imply:

$$\varepsilon u'(q_j) = p(1+\lambda_j^b) \tag{18}$$

From (16) and (18) we know that if the buyer is constrained and $\lambda_j^b > 0$, then $\varepsilon u'[q_j(\varepsilon)] > c'[y(\varepsilon)]$. If instead the buyer is unconstrained and therefore $\lambda_j^b = 0$, then $\varepsilon u'(q_j(\varepsilon)) = c'(y(\varepsilon))$.

Last, an idle agent's problem is simply:

$$V^o(\omega_j) = \rho W_H(\omega_j^n) + (1-\rho) W_L(\omega_j^n)$$

Goods market clearing therefore requires:

$$\alpha_s y(\varepsilon) = \alpha_b[\rho q_H(\varepsilon) + (1-\rho) q_L(\varepsilon)] \qquad \text{for } \varepsilon \in \Omega \tag{19}$$

4.3 Monetary equilibrium

To find optimal savings for an agent j use (8), (14), (15), (16) and (17) to obtain:

$$V(\omega_j) = \int_\Omega \left\{ m_j - a_j + \tau_1 + \alpha_b[\varepsilon u(q_j) - pq_j] - \alpha_s[c(y) - py] + EW(0) \right\} f(\varepsilon) d\varepsilon$$

The expected lifetime utility $V(\omega_j)$ therefore depends on the agent's net wealth and income $m_j - a_j + \tau_1$ and two other elements: the expected continuation payoff $EW(0) = \rho W_L(0) + (1-\rho) W_L(0)$ and the expected surplus from trade in the DM. With probability α_b the agent spends pq_j on consumption deriving utility $\varepsilon u(q_j)$ and with probability α_s, instead, he gets disutility $c(y)$ from production and earns py from his sales. Note that, unlike in the representative-agent case, the expected earnings $p(y - q_j)$ from DM trades might be different from zero since amounts produced and consumed by an agent $j = H, L$ may be mismatched. Hence, we have:

$$\frac{\partial V(\omega_j)}{\partial m_j} = \int_{\Omega} \left\{ 1 + \alpha_b \left[\frac{\varepsilon u'(q_j)}{p} - 1 \right] \right\} f(\varepsilon) d\varepsilon \tag{20}$$

and

$$\frac{\partial V(\omega_j)}{\partial a_j} = -1, \tag{21}$$

which imply money is valued dissimilarly by agents, whereas bonds are valued identically in the economy. Combining (12) with (20) and (13) with (21) one gets that in a monetary equilibrium the following Euler equations must hold:

$$\frac{\pi - \beta_j}{\beta_j} = \int_{\Omega} \left\{ \alpha_b \left[\frac{\varepsilon u'(q_j(\varepsilon))}{c'(y(\varepsilon))} - 1 \right] \right\} f(\varepsilon) d\varepsilon \tag{22}$$

and

$$\pi p_a = \beta_z + \lambda_z^a \tag{23}$$

The expression in (22) tells us that the choice of real balances depends on three components. The first two are standard: the discount factor β_j at the beginning of the period and the real yield on cash $1/\pi$. The third component is $\varepsilon u'(q_j)/c'(y)$. This can be interpreted as the expected liquidity premium from having cash available in the DM and it arises because money is needed to trade in that market. This premium grows with the severity of the cash constraint and the likelihood of a consumption shock α_b. The expression in (23), instead, refers to the choice of bonds, which depends on the discount factor β_z drawn at the beginning of the CM and the real yield $1/\pi p_a$. Note that (23) implies that bonds have no liquidity premium. That is because they are always held until maturity and cannot be used to buy consumption in the DM.

We can now define the equilibrium as the set of values of m_j and a_z for $j, z = H, L$ that solve (22) and (23). The reason is that once the equilibrium stocks of money and bonds are determined, all other endogenous variables can be derived.

Definition 1 *A symmetric stationary monetary equilibrium consists of a m_j satisfying (22) and a_z satisfying (23) for $j, z = H, L$.*

We now want to investigate whether a CM to CM bond a_z for $z = H, L$ would indeed circulate in this economy. We find that the following result holds:

Lemma 1 *A stationary monetary equilibrium exists with impatient agents borrowing and patient* *agents lending at a price $p_a = \beta_H/\pi$. Specifically, $a_L = A$ and $a_H = -(1-\rho)A/\rho$.*

Why are agents interested in trading such a bond in equilibrium? This is somewhat puzzling since we know from (10) they always consume the efficient quantity x^* in the CM. This in turn implies that there is no reason for using bonds for consumption smoothing here due to the quasi linearity of preferences. Bonds, however, allow agents to smooth the labor effort across periods - H agents prefer to work more today and less in the future, whereas L agents would rather do the opposite.[10]

Once we know the price at which these bonds circulate in equilibrium, we can pin down their net nominal yield, which is:

$$i = \frac{1}{p_a} - 1 \quad \Rightarrow \quad i = \frac{\pi}{\beta_H} - 1 \tag{24}$$

We will refer to i as the nominal interest rate in this economy - note it is affected directly by long-term monetary policy through π. We now want to determine the returns on money and bonds that are consistent with equilibrium.

Lemma 2 *Any stationary monetary equilibrium must be such that $\pi \geq \beta_H$, i.e. $i \geq 0$.*

This result derives from a simple no-arbitrage condition - in a monetary equilibrium, the value of money cannot grow too fast with $\pi < \beta_H$ or else type H agents will not spend it.[11] This, together with (24), implies that to run the Friedman rule the monetary authority must let $\pi \to \beta_H$ and cannot target β_L instead.

5 Optimal inflation rate

At this point, we know that given the result in Lemma 2 the monetary authority is constrained in its ability to give a rate of return on money that is attractive for everyone. Given this result, one should expect inefficiencies will arise at $i = 0$[12] and therefore the Friedman rule might not be the optimal policy here. We investigate this next, and in this section we will focus on the optimal inflation rate in an economy without aggregate demand shocks, which we will then reintroduce in Section 6. We find that the following result holds:

[10] See Boel and Finocchiaro (2015) for an economy with borrowing and lending, where such a bond is traded even with permanent heterogeneity in discounting.

[11] Other models with heterogeneous time preferences have analogous results, in that the rate of return on the asset cannot exceed the lowest rate of time preference. See for example Becker (1980) and Boel and Camera (2006). In those models, however, types are fixed.

[12] We know from (24) that it is equivalent to fix i or π.

Proposition 1 *Let $i = 0$. If $c''(y) = 0$, then $q_H = q^*$ and $q_L < q^*$. If $c''(y) > 0$, then $q_H > q^*$ and $q_L < q^*$.*

Proposition 1 implies that, since agents value future consumption differently, the Friedman rule fails to sustain the constrained-efficient allocation in a monetary equilibrium. Indeed, even if the Friedman rule eliminates the opportunity cost of holding money for type H agents, it still fails to provide incentives for everyone to save enough since $\pi > \beta_L$. That is because impatient agents are facing an effective nominal interest rate equal to $\beta_H/\beta_L - 1$, which is positive. We know from Lemma 2 that $\pi \geq \beta_H$ and therefore the central bank is limited in its ability to reduce interest rates even further. Thus, type L agents remain constrained even at the ZLB.

Proposition 1 also implies that the nature of preferences has important consequences for the optimality of the Friedman rule. Specifically, a convex disutility from labor generates a price externality induced by type L agents. This happens because impatient agents consume too little even at $\pi = \beta_H$, and therefore they drive down total output, marginal cost of production and relative price. The low price in turn induces type H agents to consume too much compared to the efficient allocation. This price externality disappears with linear costs, since in that case the marginal cost of production (and hence the relative price in the DM) is constant at any level of output.

In light of the results described in Proposition 1, one must wonder if setting $i = 0$ is still the optimal monetary policy in this economy. The following result clarifies when this is the case.

Proposition 2 *If $c''(y) = 0$, $i = 0$ is always the optimal policy. If $c''(y) > 0$, $i = 0$ is the optimal policy if $c'(y) \geq 1$ and $\beta_L u'(q_L) < \beta_H u'(q_H)$.*

The lesson here is that the Friedman rule is always the optimal policy with linear costs. In that case, even if $i = 0$ fails to sustain the constrained-efficient allocation, such policy delivers a second best allocation that cannot be Pareto improved. With convex costs, however, the Friedman rule is not necessarily optimal. Why not? Because the policy maker needs to take into account the price externality induced by the underconsumption of impatient agents. In this situation, Proposition 2 implies that $i = 0$ is optimal if two conditions hold. First, $p = c'(y) \geq 1$, so that the relative price cannot be too low in order to limit the overconsumption of type H agents. Second, $u'(q_L)/u'(q_H) < \beta_H/\beta_L$, meaning the disparity in consumption between impatient and patient agents must be limited.

Example 1 - optimal inflation with convex costs. In order to derive intuition for the results in Proposition 2, we consider an example with the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = \exp^{y} - 1$$

In this case, the optimal inflation problem to be solved in a monetary equilibrium becomes:

$$\begin{aligned} &\underset{\pi}{Max} \;\; \alpha_b \left[(1-\rho)\varepsilon u(q_L) + \rho\varepsilon u(q_H)\right] - \alpha_s c(y) \\ &\quad \text{s.t.} \;\; \frac{\pi - \beta_H}{\beta_H} = \alpha_b \left[\frac{\varepsilon \exp^{-q_H}}{\exp^{y}} - 1\right] \\ &\quad \text{s.t.} \;\; \frac{\pi - \beta_L}{\beta_L} = \alpha_b \left[\frac{\varepsilon \exp^{-q_L}}{\exp^{y}} - 1\right] \\ &\quad \text{s.t.} \;\; \alpha_s y = \alpha_b[\rho q_H + (1-\rho) q_L] \end{aligned} \tag{25}$$

If we differentiate the objective function in (25), we find that the optimal π must satisfy:

$$\alpha_b \left[(1-\rho)\varepsilon u'(q_L)\frac{dq_L}{d\pi} + \rho\varepsilon u'(q_H)\frac{dq_H}{d\pi}\right] - \alpha_s c'(y)\left[\rho\frac{dq_H}{d\pi} + (1-\rho)\frac{dq_L}{d\pi}\right] \leq 0 \tag{26}$$

In the Appendix, we derive expressions for $dq_L/d\pi$ and $dq_H/d\pi$ from the constraint in (26) and let $\varepsilon = \exp$ so that $\ln(\varepsilon) = 1$. We find that in order for $\pi = \beta_H$ to be optimal in this case it must be that:

$$\frac{\beta_H - \beta_L}{\beta_H} \leq \frac{\alpha_s}{\rho(1-\alpha_b)}$$

Intuitively, the condition above imposes an upper bound on time-preference heterogeneity. This will limit the price externality highlighted in Propositions 1 and 2, and the Friedman rule will be optimal with convex costs. Of course, if the condition above is not satisfied, then $i > 0$ must be optimal. That would be the case, for example, with $\rho = 0.90$, $\beta_H = 0.99$, $\beta_L = 0.70$ and $\alpha_s = \alpha_b = 0.10$. In this case the central bank would choose $i = 1.46\%$.

6 Optimal stabilization policy at the ZLB

We now reintroduce aggregate demand shocks ε. We will first investigate which inefficiencies arise when the central bank does not engage in stabilization policy, i.e. when $\tau_1(\varepsilon) = \tau_2(\varepsilon) = 0$. We will call this passive policy and then compare it to active stabilization policy, i.e. the policy implemented by a central bank whose objective is to maximize the weighted welfare of the agents in the economy.

We will do so for $i = 0$. We find the following result holds with passive policy.

Proposition 3 *Let $i = 0$ and $\tau_1(\varepsilon) = \tau_2(\varepsilon) = 0$. A unique monetary equilibrium exists for $c''(y) \geq 0$ such that: with $c''(y) = 0$, then $q_L(\varepsilon) = q_H(\varepsilon) = q^*(\varepsilon)$ for $\varepsilon \leq \tilde{\varepsilon}$ and $q_L(\varepsilon) < q_H(\varepsilon) = q^*(\varepsilon)$ for $\varepsilon > \tilde{\varepsilon}$, where $\tilde{\varepsilon} \in [0, \bar{\varepsilon}]$; with $c''(y) > 0$, then $q_L(\varepsilon) = q_H(\varepsilon) = q^*(\varepsilon)$ for $\varepsilon \leq \hat{\varepsilon}$ and $q_L(\varepsilon) < q^*(\varepsilon) < q_H(\varepsilon)$ for $\varepsilon > \hat{\varepsilon}$, where $\hat{\varepsilon} \in [0, \bar{\varepsilon}]$.*

Proposition 3 implies that, with passive policy, impatient agents are unconstrained and consume $q_L(\varepsilon) = q^*(\varepsilon)$ in low marginal demand states. They are instead constrained in high demand states, and thus consume $q_L(\varepsilon) < q^*(\varepsilon)$. Why don't type L ever consume $q_L(\varepsilon) > q^*(\varepsilon)$? Because the price externality outlined in Proposition 1 cannot arise when all agents are unconstrained, and therefore $q_L(\varepsilon) = q_H(\varepsilon) = q^*(\varepsilon)$ with $c''(y) \geq 0$ in the low marginal demand states. Agents will use the extra cash to work less in the CM. Type H agents, instead, are never constrained, and they overconsume only in high demand states with $c''(y) > 0$.

We will now move to studying the problem of a central bank engaged in stabilization policy and thus maximizing welfare by choosing the quantities consumed and produced by each type $j = H, L$ in each state subject to the constraint that the chosen quantities satisfy the conditions characterizing a competitive equilibrium. The policy is then implemented by choosing state-contingent injections $\tau_1(\varepsilon)$ and $\tau_2(\varepsilon)$ accordingly. The primal Ramsey problem faced by the central bank is:[13]

$$
\begin{aligned}
&\underset{q_L(\varepsilon), q_H(\varepsilon), y(\varepsilon)}{Max} \int_\Omega \left\{\alpha_b \varepsilon \left[\rho u(q_H(\varepsilon)) + (1-\rho) u\left(q_L(\varepsilon)\right)\right] - \alpha_s c(y(\varepsilon))\right\} f(\varepsilon) d\varepsilon \\
&\text{s.t. } \frac{\pi - \beta_H}{\beta_H} = \int_\Omega \left\{\alpha_b \left[\frac{\varepsilon u'(q_H(\varepsilon))}{c'(y(\varepsilon))} - 1\right]\right\} f(\varepsilon) d\varepsilon \\
&\text{s.t. } \frac{\pi - \beta_L}{\beta_L} = \int_\Omega \left\{\alpha_b \left[\frac{\varepsilon u'(q_L(\varepsilon))}{c'(y(\varepsilon))} - 1\right]\right\} f(\varepsilon) d\varepsilon \\
&\text{s.t. } \alpha_s y(\varepsilon) = \alpha_b [\rho q_H(\varepsilon) + (1-\rho) q_L(\varepsilon)]
\end{aligned}
$$

Note that we are focusing on a monetary equilibrium such that $m_j > 0$ for $j = H, L$. That is why the first two constraints in the Ramsey problem must hold with equality. Moreover, since $i = 0$, then (22) implies that $\varepsilon u'(q_H(\varepsilon)) = c'(y(\varepsilon))$ in every state, and the Ramsey planner simply solves

[13] The objective function of the Ramsey problem faced by the central bank is:

$$\rho(1-\beta_H)V(\omega_H) + (1-\rho)(1-\beta_L)V(\omega_L)$$

where $V(\omega_j)$ is defined in (14). We know that trades are efficient in the CM. Moreover, $\tau_1(\varepsilon) + \tau_2(\varepsilon) = 0$ and $\bar{m} = \rho m_H + (1-\rho) m_L$. Therefore, the central bank has to worry only about maximizing $\int_\Omega \left\{\alpha_b \varepsilon \left[\rho u(q_H(\varepsilon)) + (1-\rho) u\left(q_L(\varepsilon)\right)\right] - \alpha_s c(y(\varepsilon))\right\} f(\varepsilon) d\varepsilon$.

the following problem:

$$\begin{aligned}
&\underset{q_L(\varepsilon),y(\varepsilon)}{Max} \int_\Omega \{\alpha_b\varepsilon\left[\rho u(q_H(\varepsilon)) + (1-\rho)u\left(q_L(\varepsilon)\right)\right] - \alpha_s c(y(\varepsilon))\}\, f(\varepsilon)d\varepsilon \\
&\quad \text{s.t. } \frac{\pi-\beta_L}{\beta_L} = \int_\Omega \left\{\alpha_b\left[\frac{\varepsilon u'(q_L(\varepsilon))}{c'(y(\varepsilon))} - 1\right]\right\} f(\varepsilon)d\varepsilon \\
&\quad \text{s.t. } \alpha_s y(\varepsilon) = \alpha_b[\rho q_H(\varepsilon) + (1-\rho)q_L(\varepsilon)]
\end{aligned} \tag{27}$$

We find the following result holds:

Proposition 4 *If $i = 0$, in a monetary equilibrium with $m_j > 0$ for $j = H, L$ the optimal policy is $q_L(\varepsilon) < q^*(\varepsilon)$ for all states with $c''(y) \geq 0$.*

Why does the Ramsey planner choose $q_L < q^*$ in all states? Because we know from Proposition 3 that without policy intervention agents L would have enough cash to buy q^* in low demand states. In high demand states, however, their cash holdings would constrain their spending to $q_L < q^*$. This would create an inefficiency of consumption across states that can be overcome by stabilization policy.

It is worth emphasizing that our aim here is to determine how stabilization policy can improve welfare compared to the allocation that would be achieved with a passive policy in which agents would only be able to rely on their money balances to finance consumption in the DM. That is why Proposition 4 focuses on an equilibrium where $m_j(\varepsilon) > 0$ for $j = H, L.$ Of course the central bank could also provide L agents with enough liquidity to finance $q^*(\varepsilon)$ in every state, but in that case we would have $m_L(\varepsilon) = 0$ in all states since L agents would have no incentive to bring any money.

Note that Proposition 4 also implies that the central bank is able to temporarily relax liquidity constraints on impatient agents at the ZLB. It can do so simply engaging in repo arrangements that are undone at a later date, i.e. $\tau_1(\varepsilon) + \tau_2(\varepsilon) = 0$.

Example 2 - stabilization policy with linear costs. We now want to get some intuition about optimal stabilization policy at the zero lower bound with different cost functions specifications. We will focus on linear costs first, and specifically we consider the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = y$$

Derivations are in the Appendix, but it is worth pointing out that in this case we find that:

$$q_L(\varepsilon) = \ln(\varepsilon) - \ln\left[\frac{\pi - \beta_L(1-\alpha_b)}{\alpha_b \beta_L}\right]$$

and $q_L(\varepsilon)/q_L(\varepsilon)$ can be expressed as:

$$\frac{\ln(\varepsilon) - \ln\left[(\pi - \beta_L(1-\alpha_b))/\alpha_b\beta_L\right]}{\ln(\underline{\varepsilon}) - \ln\left[(\pi - \beta_L(1-\alpha_b))/\alpha_b\beta_L\right]} = \frac{m_L + \tau_1(\varepsilon)\bar{m}/\pi}{m_L + \tau_1(\underline{\varepsilon})\bar{m}/\pi}$$

Since $\varepsilon > \underline{\varepsilon}$, then it must be that $\tau_1(\varepsilon) > \tau_1(\underline{\varepsilon})$. Thus, the higher the demand for good, the higher the injection $\tau_1(\varepsilon)$ needed to finance the increase in consumption.

Figure 2 illustrates active and passive policies for type L with the specified linear cost function and assuming the following parameter values: $\alpha_s = \alpha_b = 0.3$, $\beta_H = 0.99$, $\beta_H = 0.95$, $\rho = 0.5$. The values for β_H and β_L are consistent with the evidence in Lawrence (1991), Carroll and Samwick (1997) and Samwick (1998) who provide empirical estimates of distributions of discount factors.

The curve labeled "efficient $q_L(\varepsilon)$" represents the constrained-efficient allocation at which $q^*(\varepsilon) = q^*_H(\varepsilon) = q^*_L(\varepsilon)$. The curve "passive $q_L(\varepsilon)$" represents equilibrium consumption for type L under a passive policy, whereas the curve "active $q_L(\varepsilon)$" denotes consumption for the same agents when the central bank behaves optimally.

The important thing to notice here is that the central bank's optimal choice is strictly increasing in ε - the central bank chooses to reduce consumption from the first best in low demand states in order to increase it in higher demand states.

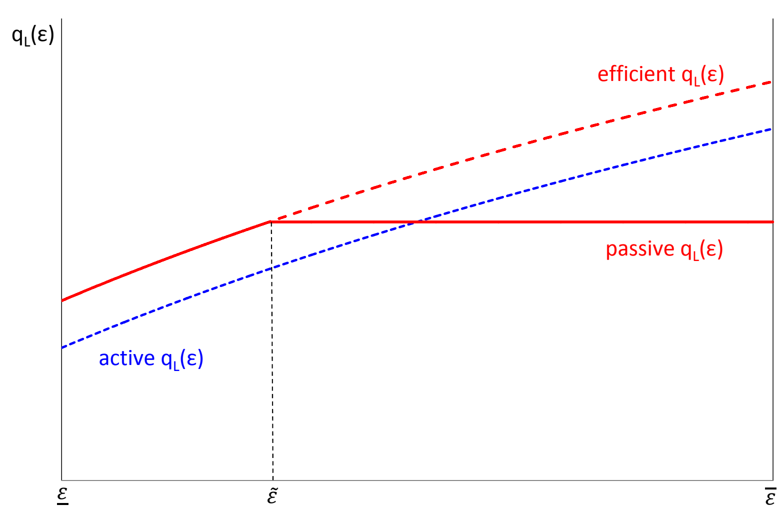

Figure 2: Stabilization versus passive policy for type L agents - linear costs

Example 3 - stabilization policy with convex costs. In this example, we focus on convex costs and consider the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = \exp^{y} - 1$$

Derivations are in the Appendix. Figure 3 illustrates active and passive policies for type L in this example using the same parameter values we used for the linear cost function specification. As in that case, the central bank's optimal choice is strictly increasing in ε - the central bank chooses to reduce consumption from the first best in low demand states to increase it in higher demand states.

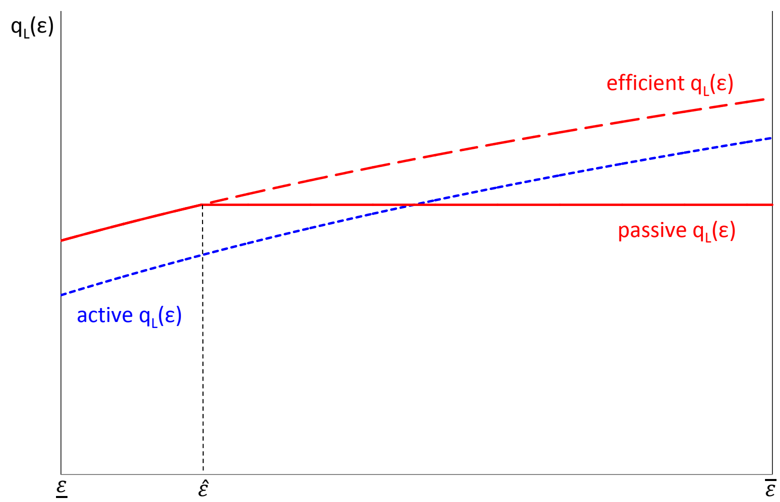

Figure 3: Stabilization versus passive policy for type L agents - convex costs

Note that the central bank is not actively trying to stabilize consumption of patient agents. However, the short-term monetary policy aimed at stabilizing $q_L(\varepsilon)$ generates an externality on $q_H(\varepsilon)$. Figure 4 illustrates. The important thing to notice here is that since $q_L < q^*$ in all states, type H agents always consume more than q^* in light of Proposition 1. We know from Proposition 3 that without policy intervention agents H would buy q^* in some states and $q_H > q^*$ in others. Stabilization policy addresses this discontinuity indirectly and consumption is smoothed so that $q_H > q^*$ in all states.

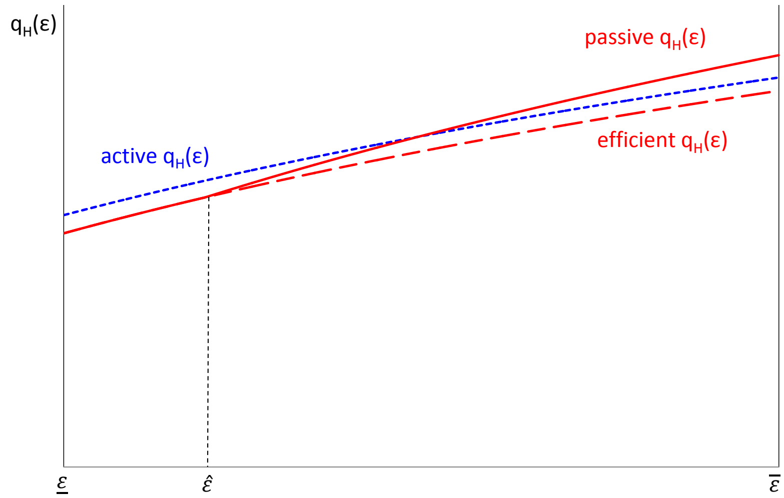

Figure 4: Stabilization versus passive policy for type H agents - convex costs

6.1 Optimal stabilization policy with positive inflation at the ZLB

We now consider the following scenario. We are at the steady state and right before the CM opens the spread between β_H and β_L unexpectedly widens for only one period. Specifically, we consider the case $0 < \beta_L < \beta < 1 < \beta_H$ for one period, but $0 < \beta_L < \beta < \beta_H < 1$ from then on with average β invariant. Note that, given Lemma 2, a discount factor $\beta_H > 1$ implies a positive inflation rate at the zero lower bound, and therefore a negative real interest rate. Note also that, with this particular structure, transversality conditions still hold for H agents in the CM in period t since their expected discount factor is lower than one after that, i.e. $\beta < 1$.[14] The analytical results derived in the previous sections still hold even if the spread between the discount factors widens. Quantitatively, however, results do change since from (22) we have:

$$\frac{dq_j}{d\beta_j} = -\frac{\pi[c'(y)]^2}{\beta_j^2 \alpha_b \varepsilon [u''(q_j)c'(y) - u'(q_j)c''(y)dy/dq_j]}$$

and therefore $dq_j/d\beta_j > 0$. Thus, as β_L decreases, impatient agents will become more constrained at the ZLB. We therefore investigate how the results from the Ramsey problem in this environment compare to the ones where the spread is narrower. For our numerical analysis, we consider $\beta_H = 1.005$ (consistent with an annual 2% inflation) and $\beta_L = 0.935$, so that $\beta = 0.970$ as in our previous exercises. All other parameters and functional forms remain invariant. Figures 5 and 6 illustrate the results for impatient agents for the cases of linear and convex disutility respectively.

[14] See Kocherlakota (1990) for a related discussion in the context of growth economies.

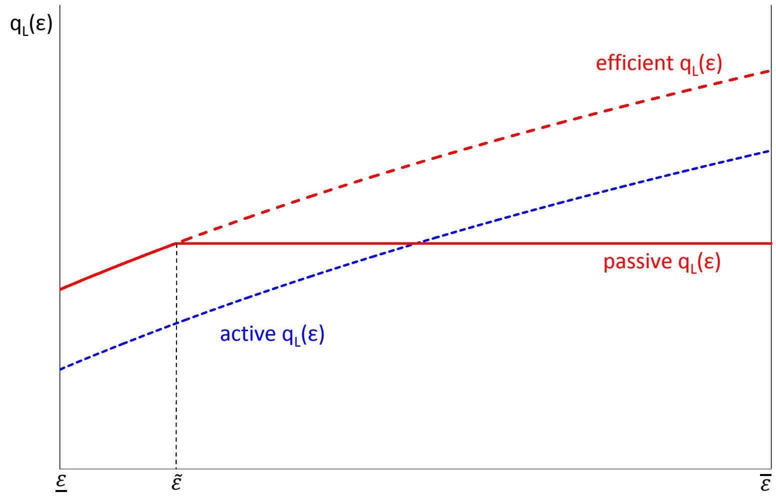

Figure 5: Stabilization versus passive policy for type L agents with positive inflation - linear costs

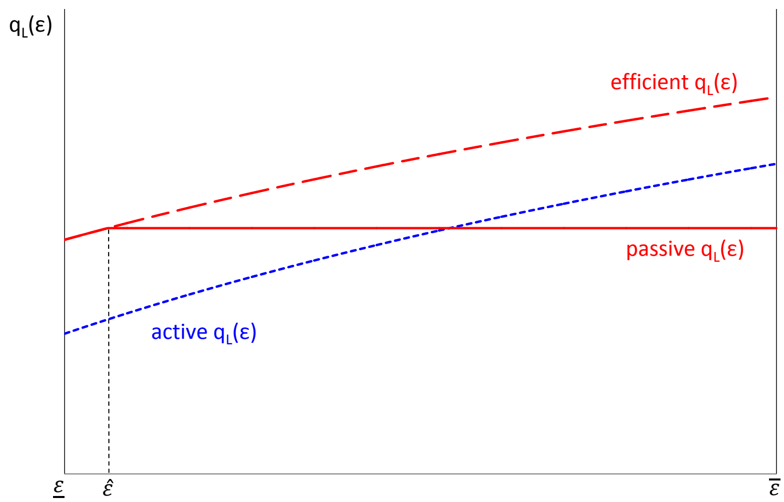

Figure 6: Stabilization versus passive policy for type L agents with positive inflation - convex costs

Qualitatively, the nature of the stabilization policy stays the same even if $\beta_H > 1$ since, with both linear and convex costs, the central bank's optimal choice is strictly increasing in ε. That is, the central bank still chooses to reduce consumption in low demand states to increase it in higher demand states. However, when we compare Figure 2 with 5 and Figure 3 with 6, we notice that as $\beta_H - \beta_L$ increases, the interval of low-demand states for which type L agents are unconstrained decreases. This in turn leads to an increased distance between the "efficient q_L" and the "active q_L."

7 Conclusion

There is substantial empirical evidence showing that QE can have beneficial real effects on the economy even at the ZLB. Yet, at the same time, showing theoretically that QE can be welfare improving has been difficult to achieve. Our study is motivated by the desire to reconcile these observations. For this reason, we construct a New Monetarist model characterized by aggregate and idiosyncratic demand shocks. Agents are heterogeneous with respect to the idiosyncratic shock, so that in every period some agents are more patient than others. This heterogeneity generates a distribution in asset holdings. We then study the optimal stabilization response of the central bank to demand shocks that hit the economy.

We find that several results hold in this environment. First, the ZLB can be the best interest rate policy the central bank can implement although that is not necessarily the case due to a price externality. However, even when the ZLB is the best policy, not all agents are satiated at the Friedman rule and therefore there is scope for central bank policies of liquidity provision. Second, we study a particular form of QE whereby the central bank purchases private debt via repo arrangements in response to demand shocks. We find that such a policy is welfare improving even at the ZLB since it can relax the liquidity constraint of impatient agents without harming the patient ones. We show this is true regardless of whether we have inflation or deflation at the ZLB. Third, due to a pricing externality, QE can be welfare improving for patient agents even if they are unconstrained at the ZLB.

References

Akyol, A., 2004. “Optimal Monetary Policy in an Economy with Incomplete Markets and Idiosyncratic Risk.” *Journal of Monetary Economics* 51 (6): 1245-1269.

Becker, R., 1980. “On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households.” *Quarterly Journal of Economics* 95 (2): 375-382.

Berentsen, A., Camera, G., and C. J. Waller, 2007. “Money, Credit and Banking.” *Journal of Economic Theory* 135 (1): 171-195.

Berentsen, A., and C. J. Waller, 2011. “Price-Level Targeting and Stabilization Policy.” *Journal of Money, Credit and Banking* 43 (7): 559-580.

Berentsen, A., Huber, S., and A. Marchesiani, 2014. “Degreasing the Wheels of Finance.” *International Economic Review*, 55 (3): 735-763.

Bhattacharya, J., Haslag, J., and A. Martin, 2005. “Heterogeneity, Redistribution, and the Friedman Rule.” *International Economic Review* 46 (2): 437-454.

Bhattarai, S., Eggertsson, G.B., and B. Gafarov, 2015. “Time Consistency and the Duration of Government Debt: a Signalling Theory of Quantitative Easing.” NBER Working Paper 21336.

Boel, P., and G. Camera, 2006. “Efficient Monetary Allocations and the Illiquidity of Bonds.” *Journal of Monetary Economics* 53 (7): 1693-1715.

Boel, P., and D. Finocchiaro, 2015. “Money, Credit and Redistribution.” Unpublished Manuscript, Sveriges Riksbank.

Carroll, C. D., and A. A. Samwick, 1997. “The Nature of Precautionary Wealth.” *Journal of Monetary Economics* 40 (1): 41-71.

Cúrdia, V., and M. Woodford, 2011. “The Central-Bank Balance Sheet as an Instrument of Monetary Policy.” *Journal of Monetary Economics* 58 (1): 47-74.

Eggertsson, G. B., and M. Woodford, 2003. “The Zero Bound on Interest Rates and Optimal Monetary Policy.” *Brookings Papers on Economic Activity*, Vol. 1: 139-211.

Eggertsson, G. P., and P. Krugman, 2012. “Debt, Deleveraging, and the Liquidity Trap: a Fisher-Minsky-Koo Approach.” *Quarterly Journal of Economics* 127 (3): 1469-1513.

Gagnon, J., Raskin, M., Remache, J., and B. Sack, 2010. "Large-Scale Asset Purchases by the Federal Reserve: Did They Work?" Federal Reserve Bank of New York Staff Report no. 441 (March).

Hamilton, J. D., and J. C. Wu, 2012. "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment." *Journal of Money, Credit and Banking* 44 (1): 3-46.

Herrenbrueck, L., 2014. "Quantitative Easing and the Liquidity Channel of Monetary Policy." Simon Fraser University Working Paper.

Ireland, P. N., 2005. "Heterogeneity and Redistribution: by Monetary or Fiscal Means?" *International Economic Review* 46 (2): 455-463.

Kocherlakota, N., 1990. "On the Discount Factor in Growth Economies." *Journal of Monetary Economics* 25 (1): 43-47.

Kocherlakota, N., 1998. "Money is Memory." *Journal of Economic Theory* 81 (2): 232-251.

Krishnamurthy, A., and A. Vissing-Jorgensen, 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity*, Vol. 2: 215-287.

Lagos, R., and R. Wright, 2005. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113 (3): 463-484.

Lawrence, E. C., 1991. "Poverty and the Rate of Time Preference: Evidence from Panel Data." *Journal of Political Economy* 99 (1): 54-77.

Phelps, E. S., 1973. "Inflation in the Theory of Public Finance." *Swedish Journal of Economics* 75 (1): 67-82.

Ramsey, F. P., 1927. "A Contribution to the Theory of Taxation." *The Economic Journal* 37: 47-61.

Rocheteau, G., and R. Wright, 2005. "Money in Search Equilibrium, in Competitive Equilibrium and in Competitive Search Equilibrium." *Econometrica* 73 (1): 175-202.

Rojas Breu, M., 2013. "The Welfare Effect of Access to Credit." *Economic Inquiry*, 51 (1): 235-247.

Samwick, A. A., 1998. "Discount Rate Heterogeneity and Social Security Reform." *Journal of Development Economics* 57 (1): 117-46.

Shi, S., 1997. "A Divisible Search Model of Fiat Money." *Econometrica* 65 (1): 75-102.

Wallace, N., 1981. "A Modigliani-Miller Theorem for Open-Market Operations." *American Economic Review* 71 (3): 267-274.

Williamson, S., 2012. "Liquidity, Financial Intermediation, and Monetary Policy in a New Monetarist Model." *American Economic Review* 102 (6): 2570-2605.

Williamson, S., 2014. "Scarce Collateral, the Term Premium, and Quantitative Easing." Federal Reserve Bank of Saint Louis Working Paper 2014-008A.

Woodford, M., 2012. "Methods of Policy Accommodation at the Interest-Rate Lower Bound." In *The Changing Policy Landscape*, Federal Reserve Bank of Kansas City.

8 Appendix

Conditions for $n_{jz}^k > 0$. We now want to provide conditions that guarantee $n_{j,z}^k \geq 0$ in the constrained-efficient equilibrium with $i = 0$. Note that if $n_{HL}^s > 0$, then $n_{jz}^k \geq 0$ in all other cases. We know that $x_{jz}^k = x^*$ for all j, z. This, together with the budget constraint in (8), implies:

$$n_{HL}^s = x^* - m_H^s + \pi m_L - \pi p_a a_L + a_H - \tau - \tau_2$$

From (5), Lemma 1 and (7) the expression above becomes:

$$n_{HL}^s = x^* - m_H - \tau_1 - py + \beta_H m_L - A[\beta_H + (1-\rho)/\rho] - (\beta_H - 1)(\rho m_H + (1-\rho)m_L) - \tau_2$$

Since $\tau_1 + \tau_2 = 0$ and $py = \rho(m_H + \tau_1) + (1-\rho)(m_L + \tau_1)$, rearranging we get:

$$n_{HL}^s = x^* - A[\beta_H + (1-\rho)/\rho] - m_H - \tau_1 - \rho\beta_H[\rho m_H + (1-\rho)m_L]$$

Note that for $\pi = \beta_H$ we have that $m_H - \tau_1 = q^*$. Let $K = \rho\beta_H[\rho m_H + (1-\rho)m_L] + A[\beta_H + (1-\rho)/\rho]$. Then, in order to have $n_{HL}^s > 0$ it must be that:

$$x^* - q^* > K \tag{28}$$

Since $K > 0$, then x^* must be sufficiently bigger than q^* in order to have $n_{HL}^s > 0$. Note that the necessary difference between q^* and x^* will depend on A - a tighter borrowing constraint will generate an incentive to work.

Proof of Lemma 1 From the Euler equation in (23) we have that the following must hold:

$$\beta_L + \lambda_L = \beta_H + \lambda_H$$

Since $\beta_H > \beta_L$, it must be that $\lambda_L > \lambda_H \geq 0$. If $\lambda_L > \lambda_H > 0$, then there is no borrowing or lending. If instead $\lambda_L > \lambda_H = 0$, then $a_L = A$ and given the bonds market clearing condition

$$\rho a_H + (1-\rho)a_L = 0 \tag{29}$$

we have that $a_H = -A(1-\rho)/\rho$. Since $\pi p_a = \beta_H$ from (23), then $p_a = \pi/\beta_H$.■

Proof of Lemma 2 We know from (16) and (18) that $\varepsilon u'(q_j) \geq c'(y)$ for $j = H, L$. This, together with (22), implies that $\pi \geq \beta_H$.■

Proof of Proposition 1 From (22) and (18) we know that if $\pi = \beta_H$ then $\varepsilon u'(q_L) > c'(y)$, thus implying type L agents are constrained and $q_L < q^*$ for $c''(y) \geq 0$. From (22) and (18) we also know that if $\pi = \beta_H$ then $\varepsilon u'(q_H) = c'(y)$, thus implying type H agents cannot be constrained and $q_H \geq q^*$. Assume $q_H = q^*$. Since $q_L < q^*$, then we have that $y < y^*$ where $y^* = (\alpha_b/\alpha_s)q^*$. With $c''(y) = 0$ this would imply $\varepsilon u'(q^*) = c'(y) = c'(y^*)$ since $c'(y)$ is constant and therefore $q_H = q^*$. With $c''(y) > 0$, instead, $y < y^*$ implies $\varepsilon u'(q^*) = c'(y) < c'(y^*)$. Therefore, it cannot be that $q_H = q^*$ and it must be $q_H > q^*$.■

Proof of Proposition 2 The optimal π maximizes

$$\underset{\pi}{Max} \;\; \alpha_b \left[(1-\rho)\varepsilon u(q_L) + \rho\varepsilon u(q_H)\right] - \alpha_s c(y) \tag{30}$$

subject to the constraints

$$\begin{aligned} \text{s.t.} \;\; \frac{\pi - \beta_H}{\beta_H} &= \alpha_b \left[\frac{\varepsilon u'(q_H)}{c'(y)} - 1\right] \\ \text{s.t.} \;\; \frac{\pi - \beta_L}{\beta_L} &= \alpha_b \left[\frac{\varepsilon u'(q_L)}{c'(y)} - 1\right] \end{aligned} \tag{31}$$

and the DM market clearing condition (19). By differentiating the objective function in (30), we know that in order for $\pi = \beta_H$ to be the optimal policy, the following condition must hold:

$$(1-\rho)\left[\varepsilon u'(q_L) - c'(y)\right] \left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} + \rho\left[\varepsilon u'(q_H) - c'(y)\right] \left.\frac{dq_H}{d\pi}\right|_{\pi=\beta_H} \leq 0 \tag{32}$$

From the Euler equations we know that $\varepsilon u'(q_H) - c'(y) = 0$ and $\varepsilon u'(q_L) - c'(y) > 0$ at the Friedman rule. Therefore, (32) becomes

$$(1-\rho)\left[\varepsilon u'(q_L) - c'(y)\right] \left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} \leq 0,$$

which implies the Friedman rule will only be optimal if $\left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} \leq 0$. Now, if we totally differentiate the constraints in (31), we find the following system of equations has to hold:

$$\begin{bmatrix} n\dfrac{\varepsilon u''(q_H)}{c'(y)} - n\dfrac{\varepsilon u'(q_H)c''(y)}{c'(y)^2}\dfrac{n}{s}\rho & -n\dfrac{\varepsilon u'(q_H)c''(y)}{c'(y)^2}\dfrac{n}{s}(1-\rho) \\ \\ -n\dfrac{\varepsilon u'(q_L)c''(y)}{c'(y)^2}\dfrac{n}{s}\rho & n\dfrac{\varepsilon u''(q_L)}{c'(y)} - n\dfrac{\varepsilon u'(q_L)c''(y)}{c'(y)^2}\dfrac{n}{s}(1-\rho) \end{bmatrix} \begin{bmatrix} \dfrac{dq_H}{d\pi} \\ \\ \dfrac{dq_L}{d\pi} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\beta_H} \\ \\ \dfrac{1}{\beta_L} \end{bmatrix}$$

where the determinant D is

$$D = \frac{\alpha_b^2}{c'(y)^2}\left[\varepsilon u''(q_H) - \kappa_H\right]\left[\varepsilon u''(q_L) - \kappa_L\right] - \frac{\alpha_b^2}{c'(y)^2}\kappa_L\kappa_H$$

with $\kappa_H = (\alpha_b/\alpha_s)\rho\varepsilon u'(q_H)c''(y)/c'(y)$ and $\kappa_L = (\alpha_b/\alpha_s)(1-\rho)\varepsilon u'(q_L)c''(y)/c'(y)$. Using Cramer's rule we have:

$$\frac{dq_L}{d\pi} = \frac{\alpha_b}{c'(y)D}\left\{\left[\varepsilon u''(q_H) - \frac{\varepsilon u'(q_H)c''(y)}{c'(y)}\frac{\alpha_b}{\alpha_s}\rho\right]\frac{1}{\beta_L} + \left[\frac{\varepsilon u'(q_L)c''(y)}{c'(y)}\frac{\alpha_b}{\alpha_s}\rho\right]\frac{1}{\beta_H}\right\}$$

and

$$\left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} = \frac{\alpha_b}{c'(y)D}\frac{1}{\beta_L}\left\{\varepsilon u''(q_H) + \left[\frac{u'(q_L)}{u'(q_H)}\frac{\beta_L}{\beta_H} - 1\right]c''(y)\frac{\alpha_b}{\alpha_s}\rho\right\}$$

Since

$$D|_{\pi=\beta_H} = \frac{\alpha_b^2}{c'(y)}[(\varepsilon u''(q_H) - \kappa_H)(\varepsilon u''(q_L) - \kappa_L) - \alpha_b^2\kappa_L\frac{\alpha_b}{\alpha_s}\rho c''(y)/c'(y)], \tag{33}$$

then at $\pi = \beta_H$ it must be that

$$\frac{dq_L}{d\pi} = \frac{\varepsilon u''(q_H) + \left[\frac{u'(q_L)}{u'(q_H)}\frac{\beta_L}{\beta_H} - 1\right]\frac{\alpha_b}{\alpha_s}\rho c''(y)}{\alpha_b\beta_L\left\{\left[\varepsilon u''(q_H) - c''(y)\frac{\alpha_b}{\alpha_s}\rho\right]\left[\varepsilon u''(q_L) - \frac{u'(q_L)c''(y)}{u'(q_H)}\frac{\alpha_b}{\alpha_s}(1-\rho)\right] - \left[\frac{\varepsilon u'(q_L)}{c'(y)^2}\frac{\alpha_b}{\alpha_s}\rho\right]\left[\frac{\alpha_b}{\alpha_s}(1-\rho)\right]c''(y)^2\right\}}$$

Note that if $c''(y) = 0$ then $\left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} < 0$, which implies the Friedman rule is always the optimal policy with linear costs. If instead if $c''(y) > 0$, then the sign of $\left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} < 0$ is indeterminate and therefore the Friedman rule is not necessarily optimal. Note that (33) can be simplified as

$$\begin{aligned} D|_{\pi=\beta_H} &= -\frac{\alpha_b^2}{\alpha_s^2 c'(y)^3}\varepsilon(u'(q_L)\alpha_b^2 c''(y)^2\rho(1-\rho)\left[1 - c'(y)\right] - \alpha_s^2 u''(q_H)c'(y)^2 u''(q_L)\varepsilon) \\ &\quad -\frac{\alpha_b^2}{\alpha_s^2 c'(y)^3}\varepsilon c''(y)\alpha_s c'(y)\varepsilon\left[u'(q_L)u''(q_H)(1-\rho) + \alpha_b u'(q_H)u''(q_L)\rho\right] \end{aligned}$$

Therefore, if $c''(y) > 0$ then $D|_{\pi=\beta_H} > 0$ if and only if $c'(y) \geq 1$. If $c'(y) \geq 1$, then $\left.\frac{dq_L}{d\pi}\right|_{\pi=\beta_H} < 0$ if $u'(q_L)/u'(q_H) < \beta_H/\beta_L$.■

Proof of Proposition 3 From (22) we have:

$$\frac{\pi - \beta_L}{\beta_L} = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left\{ \alpha_b \left[\frac{\varepsilon u'(q_L(\varepsilon))}{c'(y(\varepsilon))} - 1 \right] \right\} f(\varepsilon) d\varepsilon \tag{34}$$

Let $g_L(\varepsilon)$ denote real aggregate spending of type L agents when their trades are efficient, i.e. $g_L(\varepsilon) = \alpha_b(1-\rho)p(\varepsilon)q^*(\varepsilon)$. Now we want to understand how changes in ε affect $g_L(\varepsilon)$:

$$dg_L(\varepsilon) = \alpha_b(1-\rho)\left[q^*(\varepsilon)dp + p(\varepsilon)dq^*\right]$$

The first term denotes the change in the relative price $p(\varepsilon)$ and the second one changes in the efficient quantity $q^*(\varepsilon)$. We can rewrite it as follows:

$$dg_L(\varepsilon) = \alpha_b(1-\rho)pq^* \left[\frac{dp}{p} + \frac{dq^*}{q^*}\right]$$

From (16) we derive that:

$$\frac{dp}{p} = 0$$

The term dq^*/q^*, instead, can be derived from $\varepsilon u'(q^*) = c'[(\alpha_b/\alpha_s)q^*]$:

$$\frac{dq^*}{q^*} = -\frac{\varepsilon u'(q^*)}{\varepsilon u''(q^*) - c''[(\alpha_b/\alpha_s)q^*](\alpha_b/\alpha_s)} \frac{d\varepsilon}{\varepsilon}$$

so that:

$$\frac{dg_L(\varepsilon)}{d\varepsilon} = -\frac{\alpha_b(1-\rho)c'[(\alpha_b/\alpha_s)q^*]q^*u'(q^*)}{\varepsilon u''(q^*) - c''[(\alpha_b/\alpha_s)q^*](\alpha_b/\alpha_s)} > 0 \text{ for } c''(y) \geq 0$$

Let's first consider the case $c''(y) = 0$. If $g_L(\varepsilon) > m_L$, then agents are constrained in all states. If $g_L(\bar{\varepsilon}) < m_L$, then agents are never constrained. If $g_L(\bar{\varepsilon}) \geq m_L \geq g_L(\varepsilon)$, for a given value of m_L there exists a critical value $\tilde{\varepsilon}$ such that $g_L(\tilde{\varepsilon}) = m_L$. This implies that $q_L(\varepsilon) = q^*(\varepsilon) = q_H(\varepsilon)$ for $\varepsilon \leq \tilde{\varepsilon}$ and $q_L(\varepsilon) < q^*(\varepsilon) = q_H(\varepsilon)$ for $\varepsilon > \tilde{\varepsilon}$.

The RHS of (22) is a function of m_L. Note that $\lim_{m_L \to 0} RHS = \infty$ and, for $\bar{m}_L = g(\bar{\varepsilon})$, $RHS|_{\bar{m}_L} = 0 \leq (\pi - \beta_L)/\beta_L$. Since RHS is continuous in m_L then an equilibrium exists.

The RHS of (22) is monotonically decreasing in m_L. To see this use Leibnitz's rule and note that by construction $q_L(\check{\varepsilon}) = q^*(\tilde{\varepsilon})$ to get

$$\frac{\partial RHS}{\partial m_L} = \int_{\check{\varepsilon}}^{\bar{\varepsilon}} \left\{ \alpha_b \left[\frac{\varepsilon\left[u''c' - u'c''(\alpha_b/\alpha_s)(1-\rho)\right]}{(c')^2} \frac{\partial q_L}{\partial m_L} \right] \right\} f(\varepsilon) d\varepsilon < 0$$

Since the right-hand side is strictly decreasing in m_L, we have a unique m_L that solves (34). Consequently, we have $q_L(\varepsilon) = q^*(\varepsilon)$ if $\varepsilon \leq \tilde{\varepsilon}$ and $q_L(\varepsilon) < q^*(\varepsilon)$ otherwise.

The argument is analogous for the case $c''(y) > 0$ and there exists a unique critical value $\hat{\varepsilon}$ such that $g_L(\hat{\varepsilon}) = m_L$. However, we know from Proposition 1 that when type L agents are constrained type H ones consume more than q^*. This implies that $q_L(\varepsilon) = q^*(\varepsilon) = q_H(\varepsilon)$ for $\varepsilon \leq \hat{\varepsilon}$ and $q_L(\varepsilon) < q^*(\varepsilon) < q_H(\varepsilon)$ for $\varepsilon > \hat{\varepsilon}$.■

Proof of Proposition 4 The Lagrangian for (27) is:

$$\mathcal{L} = \underset{q_L(\varepsilon),y(\varepsilon)}{Max} \quad \int_\Omega \{\alpha_b(1-\rho)\varepsilon u(q_L(\varepsilon)) + \alpha_b\rho\varepsilon u(q_H(\varepsilon)) - \alpha_s c\left(y(\varepsilon)\right)\} f(\varepsilon)d\varepsilon$$

$$+\lambda_R\left[\int_\Omega \left\{\alpha_b\left[\frac{\varepsilon u'(q_L(\varepsilon))}{c'\left(y(\varepsilon)\right)} - 1\right]\right\} f(\varepsilon)d\varepsilon - \frac{\pi-\beta_L}{\beta_L}\right] + \mu(\varepsilon)\left[y(\varepsilon) - \frac{\alpha_b}{\alpha_s}\left(\rho q_H(\varepsilon) + (1-\rho)q_L(\varepsilon)\right)\right]$$

Note that μ is a function of ε because the resource constraint varies state by state. The first order condition with respect to $q_L(\varepsilon)$ and $y(\varepsilon)$ are respectively as in (42) and (43). Note that, since $\varepsilon u'(q_H(\varepsilon)) = c'(y(\varepsilon))$ at $i = 0$, from the implicit function theorem we know that

$$\frac{dq_H(\varepsilon)}{dy(\varepsilon)} = \frac{c''\left(y(\varepsilon)\right)}{\varepsilon u''(q_H(\varepsilon))} \tag{35}$$

Therefore, combining (42), (43) and (35) we find that the following expression holds for $c''(y) \geq 0$:

$$\varepsilon u'(q_L(\varepsilon)) - c'\left(y(\varepsilon)\right) = \frac{\lambda_R}{(1-\rho)\,c'\left(y(\varepsilon)\right)^2}\frac{\alpha_b\varepsilon u'(q_L(\varepsilon))c''\left(y(\varepsilon)\right) - \alpha_b\alpha_s u''(q_L(\varepsilon))}{\alpha_s - \alpha_b\rho c''\left(y(\varepsilon)\right)/\varepsilon u''(q_H(\varepsilon))} \tag{36}$$

At this point we need to consider two cases. In the first one, $m_L > 0$ and therefore (27) holds with equality so that $\lambda_R > 0$. Therefore, $\varepsilon u'(q_L(\varepsilon)) > c'\left(y(\varepsilon)\right)$ in (36), which implies $q_L(\varepsilon) < q^*(\varepsilon)$. In the second one, $m_L = 0$ and therefore $\lambda_R = 0$. This in turn implies $\varepsilon u'(q_L(\varepsilon)) = c'\left(y(\varepsilon)\right)$ in (36), so that $q_L(\varepsilon) = q_H(\varepsilon) = q^*(\varepsilon)$. ■

Derivations for Example 1. We consider the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = \exp^{y} - 1$$

In this case, the constraints in the optimal inflation problem become:

$$\begin{aligned}\frac{\pi - \beta_H}{\beta_H} &= \alpha_b \left[\frac{\varepsilon \exp^{-q_H}}{\exp^{y}} - 1\right] \\ \frac{\pi - \beta_L}{\beta_L} &= \alpha_b \left[\frac{\varepsilon \exp^{-q_L}}{\exp^{y}} - 1\right] \\ \alpha_s y &= \alpha_b[\rho q_H + (1-\rho)q_L]\end{aligned} \tag{37}$$

The derivative of the objective function with respect to π yields:

$$\alpha_b \varepsilon \left[(1-\rho)u'(q_L)\frac{dq_L}{d\pi} + \rho u'(q_H)\frac{dq_H}{d\pi}\right] - \alpha_s c'(y)\left[\rho\frac{dq_H}{d\pi} + (1-\rho)\frac{dq_L}{d\pi}\right] \leq 0 \tag{38}$$

We now need to find expressions for $dq_L/d\pi$ and $dq_H/d\pi$. If we simplify the expressions in the constraints in (37) and take logs, we find that:

$$\begin{aligned}q_H &= \frac{\alpha_s Z_H + \alpha_b(1-\rho)\left(Z_H - Z_L\right)}{\alpha_s + \alpha_b} \\ q_L &= \frac{\alpha_s Z_L - \alpha_b \rho \left(Z_H - Z_L\right)}{\alpha_s + \alpha_b}\end{aligned}$$

where $Z_H = \ln(\varepsilon) + \ln\left[\alpha_b \beta_H/(\pi - \beta_H + \alpha_b \beta_H)\right]$ and $Z_L = \ln(\varepsilon) + \ln\left[\alpha_b \beta_L/(\pi - \beta_L + \alpha_b \beta_L)\right]$. We let $\varepsilon = \exp$ so that $\ln(\varepsilon) = 1$. Then:

$$\frac{dq_L}{d\pi} = \frac{\alpha_b \rho}{\left(\alpha_s + \alpha_b\right)\left(\pi - \beta_H + \alpha_b \beta_H\right)} - \frac{\alpha_s + \alpha_b \rho}{\left(\alpha_s + \alpha_b\right)\left(\pi - \beta_L + \alpha_b \beta_L\right)}$$

and

$$\frac{dq_H}{d\pi} = \frac{\alpha_b(1-\rho)}{\left(\alpha_s + \alpha_b\right)\left(\pi - \beta_L + \alpha_b \beta_L\right)} - \frac{\alpha_s + \alpha_b(1-\rho)}{\left(\alpha_s + \alpha_b\right)\left(\pi - \beta_H + \alpha_b \beta_H\right)}$$

By plugging the expressions for $dq_L/d\pi$ and $dq_H/d\pi$ into (38), we find that in order for $\pi = \beta_H$ to be optimal it must be that $(1-\rho)\left[-\left(\alpha_s + \alpha_b \rho\right)\left(\alpha_b \beta_H\right) + \alpha_b \rho\left(\beta_H - \beta_L + \alpha_b \beta_L\right)\right]\left[\exp^{1-Z_L} - 1\right] \leq 0$. Therefore, the following condition must hold:

$$\frac{\beta_H - \beta_L}{\beta_H} \leq \frac{\alpha_s}{\rho(1-\alpha_b)}$$

Derivations for Example 2. In this example, we focus on linear costs and consider the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = y$$

The central bank's problem is as in (27) and the first-order condition with respect to $q_L(\varepsilon)$ is:

$$(1-\rho)\left[\varepsilon u'(q_L(\varepsilon)) - 1\right] = -\varepsilon \lambda_R u'' q_L(\varepsilon),$$

thus implying

$$\lambda_R = -\frac{(1-\rho)\left[\varepsilon u'(q_L(\varepsilon)) - 1\right]}{\varepsilon u''(q_L(\varepsilon))} \tag{39}$$

Note λ_R does not depend on any state ε and hence it must be:

$$\frac{(1-\rho)\left[\varepsilon u'(q_L(\varepsilon)) - 1\right]}{\varepsilon u''(q_L(\varepsilon))} = \frac{(1-\rho)\left[\varepsilon u'(q_L(\underline{\varepsilon})) - 1\right]}{\varepsilon u''(q_L(\underline{\varepsilon}))}$$

This, together with the specific functional forms we chose for this example, implies $\ln(\varepsilon) - \ln(\underline{\varepsilon}) = q_L(\varepsilon) - q_L(\underline{\varepsilon})$. Hence, if $\varepsilon > \underline{\varepsilon}$, we know that $q_L(\varepsilon) > q_L(\underline{\varepsilon})$. Now, assume $\tau_1(\underline{\varepsilon}) = 0$. Note that when agents are constrained it must be that $q_L(\varepsilon) = m_L + \tau_1(\varepsilon)\bar{m}/\pi$ and therefore:

$$q_L(\varepsilon) = \left[\frac{m_L + \tau_1(\varepsilon)\bar{m}/\pi}{m_L + \tau_1(\underline{\varepsilon})\bar{m}/\pi}\right] q_L(\underline{\varepsilon})$$

With the specified functional forms (39) becomes $\lambda_R = (1-\rho)\left[\varepsilon \exp^{-q_L(\varepsilon)} - 1\right]/\varepsilon \exp^{-q_L(\varepsilon)}$ and therefore $\varepsilon \exp^{-q_L} = 1 - \rho/\left[1-\rho-\lambda\right]$. By taking logs of both hand sides, we find that:

$$q_L(\varepsilon) = \ln(\varepsilon) - \ln\left[\frac{1-\rho}{(1-\rho)-\lambda_R}\right] \tag{40}$$

Note that with a linear cost function the constraint in (27) becomes:

$$\frac{\pi - \beta_L}{\beta_L} = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left\{\alpha_b[\varepsilon \exp^{-q_L} - 1]\right\} f(\varepsilon) d\varepsilon$$

If we combine it with (40) we have that $(\pi - \beta_L)/\beta_L = \alpha_b \lambda_R/(1-\rho-\lambda_R)$ and therefore:

$$\lambda_R = \frac{(1-\rho)(\pi-\beta_L)}{\pi - \beta_L + \alpha_b \beta_L} > 0 \tag{41}$$

Thus, combining (40) with (41) we find that:

$$q_L(\varepsilon) = \ln(\varepsilon) - \ln\left[\frac{\pi - \beta_L(1-\alpha_b)}{\alpha_b\beta_L}\right]$$

Since we know that $q_L(\varepsilon) = q_L(\underline{\varepsilon})\left[m_L + \tau_1(\varepsilon)\bar{m}/\pi\right] / \left[m_L + \tau_1(\underline{\varepsilon})\bar{m}/\pi\right]$, then we have that:

$$\frac{\ln(\varepsilon) - \ln\left[(\pi - \beta_L(1-\alpha_b))/\alpha_b\beta_L\right]}{\ln(\underline{\varepsilon}) - \ln\left[(\pi - \beta_L(1-\alpha_b))/\alpha_b\beta_L\right]} = \frac{m_L + \tau_1(\varepsilon)\bar{m}/\pi}{m_L + \tau_1(\underline{\varepsilon})\bar{m}/\pi}$$

Since $\varepsilon > \underline{\varepsilon}$, then it must be that $\tau_1(\varepsilon) > \tau_1(\underline{\varepsilon})$. Thus, the higher the demand for good, the higher the injection $\tau_1(\varepsilon)$ needed to finance the increase in consumption.

Derivations for Example 3. In this example, we focus on convex costs and consider the following functional forms:

$$u(q) = 1 - \exp^{-q} \quad \text{and} \quad c(y) = \exp^{y} - 1$$

As before, the central bank's problem is as in (27). The first-order condition for $q_L(\varepsilon)$ implies:

$$\mu(\varepsilon) = \varepsilon\alpha_s \frac{(1-\rho)u'(q_L(\varepsilon)) + \lambda_R u''\left[q_L(\varepsilon)\right]/c'\left(y(\varepsilon)\right)}{1-\rho} \tag{42}$$

The first order condition with respect to $y(\varepsilon)$ instead yields:

$$\alpha_b\rho\varepsilon u'(q_H(\varepsilon))\frac{dq_H(\varepsilon)}{dy(\varepsilon)} - \alpha_s c'\left(y(\varepsilon)\right) - \lambda_R \frac{\alpha_b\varepsilon u'(q_L(\varepsilon))c''\left(y(\varepsilon)\right)}{c'\left(y(\varepsilon)\right)^2} + \mu(\varepsilon)\left[1 - \frac{\alpha_b}{\alpha_s}\rho\frac{dq_H(\varepsilon)}{dy(\varepsilon)}\right] = 0 \tag{43}$$

Combining (42) with (43) and the fact that with convex costs $dq_H(\varepsilon)/dy(\varepsilon) = c''\left[y(\varepsilon)\right]/\varepsilon u''\left[q_H(\varepsilon)\right]$ and solving for λ_R, we find:

$$\lambda_R = \frac{\left\{\varepsilon u'\left[q_L(\varepsilon)\right] - c'[y(\varepsilon)]\right\}\left[1 - \frac{\alpha_b}{\alpha_s}\rho\frac{c''[y(\varepsilon)]}{\varepsilon u''\left[q_H(\varepsilon)\right]}\right]}{\left\{\frac{\alpha_b\varepsilon u'\left[q_L(\varepsilon)\right]c''[y(\varepsilon)]}{\alpha_s c'[y(\varepsilon)]^2} - \frac{\varepsilon u''\left[q_L(\varepsilon)\right]}{(1-\rho)c'[y(\varepsilon)]}\left[1 - \frac{\alpha_b}{\alpha_s}\rho\frac{c''[y(\varepsilon)]}{\varepsilon u''\left[q_H(\varepsilon)\right]}\right]\right\}} \tag{44}$$

We now consider a uniform distribution with $\underline{\varepsilon} = \exp$ and we proceed as follows. First, we use the first constraint in the Ramsey problem in (27) to solve for $y(\underline{\varepsilon})$:

$$\frac{\pi - \beta_L}{\beta_L} = \int_{\varepsilon}^{\bar{\varepsilon}}\left\{\alpha_b\left[\varepsilon\exp^{\frac{\alpha_b\rho\ln(\varepsilon) - (\alpha_s+\alpha_b)y(\varepsilon)}{\alpha_b(1-\rho)}} - 1\right]\right\}f(\varepsilon)d\varepsilon$$

Second, since λ_R does not depend on any state ε, the following condition must hold for all ε such that $\underline{\varepsilon} \leq \varepsilon \leq \bar{\varepsilon}$ given an arbitrary state ε:

$$\lambda_R = \lambda_R|_{\underline{\varepsilon}} \text{ with } \lambda_R = \frac{\left[\varepsilon \exp^{-q_L(\varepsilon)} - \exp^{y(\varepsilon)}\right]\left[1 - \frac{\alpha_b}{\alpha_s}\rho \frac{\exp^{y(\varepsilon)}}{-\varepsilon \exp^{q_H(\varepsilon)}}\right]}{\left\{\frac{\alpha_b \varepsilon \exp^{-q_L(\varepsilon)} \exp^{y(\varepsilon)}}{\alpha_s [\exp^{y(\varepsilon)}]^2} - \frac{-\varepsilon \exp^{-q_L(\varepsilon)}}{(1-\rho)\exp^{y(\varepsilon)}}\left[1 - \frac{\alpha_b}{\alpha_s}\rho \frac{\exp^{y(\varepsilon)}}{-\varepsilon \exp^{q_H(\varepsilon)}}\right]\right\}}$$

We use this to solve for $y(\varepsilon)$ in terms of $y(\underline{\varepsilon})$ for all ε. Then, we use the resource constraint (19) to solve for $q_L(\varepsilon)$ as a function of $y(\varepsilon)$:

$$q_L(\varepsilon) = \frac{y(\varepsilon)(\alpha_s + \alpha_b \rho) - \alpha_b \rho \ln(\varepsilon)}{\alpha_b (1-\rho)}$$

Last, from $\varepsilon u'[q_H(\varepsilon)] = c'[y(\varepsilon)]$ we find an expression for $q_H(\varepsilon)$ as a function of $y(\varepsilon)$. With the functional forms we chose, the condition is $q_H(\varepsilon) = \ln(\varepsilon) - y(\varepsilon)$.