Robustly Optimal Monetary Policy with Natural Interest Rate Shocks Can Rationalize the Behavior of Central Banks.*

Seth Baylor† Oskar Zorrilla‡

March 27, 2026

Abstract

A central bank commits to a monetary policy in the face of natural interest rate shocks that may result in a nominal interest rate constrained by the zero lower bound (ZLB). However, it is also concerned about the private sector's beliefs of equilibrium outcomes, and faces Knightian uncertainty about these expectations. We define a robustly optimal policy as the best policy under the "worst possible" private sector expectations. We make three contributions. Our main contribution is that the robustly optimal policy is both quantitatively and qualitatively consistent with the observed behavior of central banks, but at odds with the prescriptions of optimal policy under a rational expectations analysis —the robustly optimal solution calls for higher average inflation than that under rational expectations. Our second contribution is that the robustly optimal policy is robust to private sector beliefs and most of the deep parameters of the model. Lastly, the robustly optimal policy is simpler to communicate than the rational expectations policy. A calibration exercise also shows that the optimal policy target is close to three percent, far higher than the target called-for under rational expectations.

*The views expressed here are those of the authors and do not reflect the policy or views of the United States Government, the Department of Defense, or the Department of the Navy.

†Ensign, United States Navy. email: sethwbaylor@gmail.com

‡Department of Economics, United States Naval Academy. email: zorrilla@usna.edu

1 Introduction

Since the beginning of the century, nominal interest rates have been a the zero lower bound for nine years. When the bank's traditional monetary policy instrument prevents it from achieving its desired real interest rate, it must instead rely on managing private sector expectations. One proposal to help manage expectations has been to increase the inflation target.[1] However, estimates of the optimal inflation target for the United States generally range between zero and two percent. Yet central banks have not moved to lower the inflation target. This paper proposes an explanation for this conundrum.

Our key insight is that while rational expectations analyses imply that central bankers can *choose* private sector expectations—or at least have a deterministic mapping between their policy and private sector expectations—central bankers do not believe this to be true in practice. Minutes from the meetings of the Federal Open Market Committee during the Great Recession reveal that while central bankers recognize that they can influence expectations, they are uncertain about the mapping from their policy commitments to the expectations that the private sector forms. For example, consider the following excerpt from the June 2009 meeting

> Mr. Michael Kiley (Assistant Director, Division of Research and Statistics, Board of Governors): *However, an unmooring of expectations is a risk: For example, expectations could shift down markedly in response to the weak economy and decelerating wage gains; alternatively, expectations could ramp up, perhaps in response to our expanding balance sheet.*

While Mr. Kiley highlights possible expectations outcomes, he conspicuously refrains from placing a probability weight on any of them. This is the rule, rather than exception in the minutes of the FOMC. These types of comments suggest that not only does the FOMC not believe there is a deterministic mapping from its policy choice to expectations, it doesn't have a probability distribution over the possible expectations.

We model this concern by relaxing the requirement that expectations of the private sector and the central bank must coincide. We analyze a stylized static model as in Kydland and Prescott (1977) and Barro and Gordon (1983) without rational expectations. Instead, we adopt the robustness framework advocated by Hansen and Sargent (2007). Following the formulation in Woodford (2010), we operationalize policymakers' doubts about the private sector's expectation-formation process by assuming that the central bank has Knightian uncertainty over private sector beliefs. It not only does not know the mapping from its policy announcements and the exogenous variables to private sector beliefs, it is also unable or unwilling to place a probability distribution over the possible mappings. In the face of this uncertainty, the central bank formulates policy that is robust to the "worst case" private sector expectations: those that maximize the central bank's loss function.

We make two contributions. First, we find that a robustly optimal policy is easy to communicate and implement: it consists of setting an inflation target that depends *only* on the natural rate of interest. This result contrasts sharply with that of Eggertsson and Woodford (2003). In a dynamic model under rational expectations, they find that the optimal policy is cumbersome—they consider simpler policies that are close to optimal—and *independent* of the natural rate of interest. Although our result contrasts with results under rational expectations, it is consistent with the behavior from

[1] See, for example, Williams (2009) and Ball (2013).

central banks. A concern for robustness can therefore help explain the widespread adoption of higher inflation targets when rational expectations analysis do not call for them.

Second, we estimate the robustly optimal inflation target for the US and find that it is significantly higher than those called-for under rational expectations analysis. Schmitt-Grohé and Uribe (2011) and Woodford (2003) anchor the lower end of these estimates near zero, driven by the welfare costs of price dispersion under staggered pricing and the logic of Ramsey-optimal tax smoothing. Models including the ZLB, such as Coibion et al. (2012) and Williams (2009), find an optimal rate of 1%–2% for U.S. calibrations. Andrade et al. (2019) extend the analysis to a low natural-rate environment, arguing the optimal target may rise above 2% — possibly to 3%–4% — as the natural rate of interest declines secularly. Dissenting voices push the recommended target higher still: Ball (2013) argues for 4%, and Akerlof et al. (1996) suggest 2%–3% on grounds of downward nominal wage rigidity, while Billi (2008) and Blanco (2021) hold the line near or below 2% once menu costs are properly accounted for.

These numerical estimates are highly sensitive to modeling assumptions. The near-zero results depend on frictionless labor markets, no ZLB constraint, and complete tax systems — conditions under which positive inflation is a pure distortion. Introducing the ZLB, as in Coibion et al. (2012) and Eggertsson and Woodford (2003), converts positive inflation into insurance against potential liquidity traps. The natural rate of interest matters quantitatively: Negro et al. (2017) document a secular decline in the natural rate of interest that mechanically erodes any fixed target's buffer, directly motivating the higher estimates in Andrade et al. (2019). Wage rigidity, as in Akerlof et al. (1996), is a structurally distinct rationale absent from most DSGE frameworks. Finally, Athey et al. (2005) and Woodford (2010) show that credibility and near-rational expectations interact with optimal inflation in ways that generally reinforce the case for low, stable targets — meaning the consensus estimate is also contingent on assumed institutional credibility of the central bank. While this existing literature acknowledges the impact that inflation expectations have on equilibrium, they do not account for the bank's uncertainty about them. Inflation expectations have primarily been treated one of two ways: as a known equilibrium outcome,[2] or a choice variable.[3] In either case, the policymaker is either aware of how expectations are formed —whether model-consistent or otherwise—or how they *will* be formed in response to their policy decision; the central bank has a deterministic mapping from its choice variable to private sector expectations. Our model mirrors the common concerns voiced in the FOMC minutes regarding inflation expectations by relaxing the key assumptions surrounding these processes. To our knowledge, this paper's contribution to the literature is unique in this regard by assuming *as little as possible*: not only is the central bank unable to choose inflation expectations, it also does not place a probability distribution over the possible inflation expectations that may arise, in addition to them being model-inconsistent.

We are not the first to analyze the impact of monetary policy at the ZLB when expectations are not model-consistent. Angeletos and Sastry (2021) consider a model where agents are able to reason through the partial equilibrium effects of monetary policy but not the general equilibrium effects. Like us, they also find a difference between communicating about equilibrium outcomes versus policy instruments. García-Schmidt and Woodford (2019) consider agents who engage in k-level thinking and find that this type of process converges to the rational expectations equilibrium when the nominal interest rate is positive, but will lead to outcomes bounded away from the rational expectations equilibrium at the ZLB. These types of relaxations of the rational expectations hypothesis propose a specific alternative, and they maintain the assumption that the central bank

[2] As in, for example, Athey et al. (2005) and Schmitt-Grohé and Uribe (2011)

[3] As in Eggertsson and Woodford (2003) and Woodford (2003)

knows that alternative. Therefore, as with rational expectations, the bank has a deterministic mapping from its policy commitments to the equilibrium beliefs of the private sector.

Although we are solving for optimal monetary policy, we do not interpret our results here as normative. Our analysis is descriptive rather than prescriptive. The aim of this paper is to show that when the natural rate of interest can be negative—and hence there is a possibility of hitting the ZLB—and central banks are unsure about the beliefs that the private sector holds, then a concern for robustness can *explain* many of the behaviors we observe from central banks which are inconsistent with the prescriptions of a rational expectations analysis.

The rest of the paper is organized as follows: section 2 introduces our model in the case of natural interest rate shocks. section 3 characterizes the solution under rational expectations and robust optimality in subsection 3.2 and subsection 3.3 respectively. section 4 outlines our results and implications, and section 5 concludes.

2 Model

We construct a static commitment model consisting of two agents: the central bank and the private sector. There are two possible states of the world, $j \in \{h, \ell\}$, determined by the value of natural rate of interest, ρ_j. We will refer to $\rho_h > 0$ as the "high state" while we will refer to $\rho_\ell < 0$ as the "low state". Although the state of the world is given by natural rate of interest, we will refer to it as either high or low, and distinguish variables in different states of the world by the appropriate subscript.

Restricting the natural interest rate to only two values offers a simplified approach to analyzing the central bank's problem without loss of generality. Each possible outcome of ρ_j represents a type of regime: one in which the central bank is free to choose an inflation/output gap pair that satisfies the Phillips curve, and one in which it isn't. We are interested in analyzing how the bank formulates robustly optimal monetary policy when it can find itself constrained by the ZLB; allowing ρ_j to take additional values does not affect our analysis. When the bank is unconstrained, it can accommodate any value of the natural rate of interest, and thus the actual value of ρ_h has no impact on the choice of inflation or the output gap. At the zero lower bound, however, the bank can only influence the inflation/output gap pair through expectations. Thus, when unconstrained, the bank can only condition its choice of inflation and output on the *expected* value of the natural rate of interest at the zero lower bound. What matters is the probability of finding itself at the ZLB, and, conditional on that, the expected value of the natural rate of interest.

The central bank and private sector agree that there are two possible states of the world, but potentially disagree about their probabilities:

$$\Pr(\rho_\ell) = \begin{cases} q \in (0,1) & \text{Central Bank} \\ p \in (0,1) & \text{Private Sector} \end{cases}$$

We require that q is in the interior of the simplex; otherwise the problem becomes trivial. If $q = 0$, the central bank believes it is never constrained by the ZLB and we have the standard inflation/output gap trade-off studied by Kydland and Prescott (1977) and Barro and Gordon (1983). If $q = 1$, inflation and the output gap are pinned down by the aggregate demand condition at the ZLB. Additionally, we consider the bank's distribution to be the true distribution, as our analysis is conducted from the perspective of the policymaker. The requirement that p is also

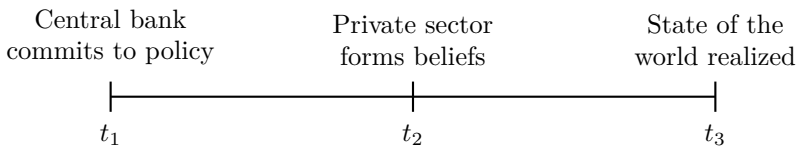


Figure 1: Timing of play

the interior of the simplex is equivalent to assuming absolute continuity. This assumption follows Woodford (2010) and rules the private sector putting positive probability on any additional states of the world besides ρ_ℓ and ρ_h.

The inflation/output gap pair (π_j, x_j) is contingent on the state of the world. In both cases, this pair is subject to the following standard New Keynesian Phillips curve:

$$\pi_j = \kappa x_j + \beta\bar{\pi} \tag{2.1}$$

where $\kappa > 0$ is the slope of the Phillips curve, $\beta \in (0, 1)$ is the representative household's discount factor, and $\bar{\pi}$ are the private sector's inflation expectations. Additionally, because the natural rate of interest can be positive or negative, it is also subject to the inequality constraint given by the aggregate demand curve in both cases:

$$\bar{x} + \bar{\pi} + \rho_j - x_j \geq 0 \tag{2.2}$$

where $\bar{x}$ are the private sector's output gap expectations. Additionally, from absolute continuity of beliefs, we have the following standard constraints on private sector expectations:

$$\bar{\pi} = p\pi_h + (1-p)\pi_\ell \tag{2.3}$$

$$\bar{x} = px_h + (1-p)x_\ell \tag{2.4}$$

While we allow the private sector's beliefs to differ from those of the central bank, we assume that all private agents have the *same* beliefs. Because all agents share the same beliefs, the derivation of the aggregate supply and aggregate demand relations, (2.1) and (2.2), follows in the same way as if agents understand the true distribution of the shocks. Expressions (2.1)–(2.4) define the set of constraints that the central bank faces in choosing state-contingent inflation/output gap pairs.

The timing of play between the central bank and the private sector is sequential and shown in Figure 1. First, the central bank commits to a state-contingent inflation/output pair, (π_j, x_j). Second, the private sector forms unique beliefs, p. Third, the state of the world is realized, ρ_j, along with its corresponding equilibrium outcomes.

The bank chooses state-contingent inflation/output pairs to minimize an augmented loss function. In a standard setting without concern for robustness, the central bank would simply minimize the following quadratic loss function, Q

$$\mathrm{E}_q[Q] = q(\pi_\ell^2 + \lambda(x_\ell - x^*)^2) + (1-q)(\pi_h^2 + \lambda(x_h - x^*)^2) \tag{2.5}$$

where $x^* > 0$ is the efficient rate of output, and $\lambda > 0$ is the relative weight the bank places on the losses associated with inflation and output, subject to the constraints described above. Because the Phillips curve relates the output gap, the inflation rate, and the private sector's inflation expectations, we can rewrite $Q(\cdot)$ as a function of only inflation, π, and private sector beliefs, p.

Equation 2.5 is the *expected* quadratic loss, with the first term denoting the loss when the natural interest rate is positive and the second term when it is negative. This expectation is, of course, taken with respect to the bank's beliefs, q.

To capture the concern for robustness, we follow Hansen and Sargent (2007) and Woodford (2010) and use the Kullback-Leibler divergence as a measure of the distance between the two distributions. [4]

$$\mathrm{E}_p[R] = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right) \tag{2.6}$$

This divergence has been used extensively in both the robustness and ambiguity aversion literature and is closely related to mutual information in the rational inattention literature. It is non-negative, convex in p, and attains its minimum value, 0, when $p = q$. From the bank's perspective, the relative entropy term uses the private sectors' distribution p to highlight that the divergence is set by the beliefs that form in response the bank's policy commitment. More simply, the private sector determines how far distorted their beliefs are compared to the bank; this is relevant for our analysis in subsection 3.3. Thus, this function can be interpreted as a measure of disagreement between the private sector and the central bank. When $p = q$, there is no disagreement. As p moves away from q and approaches the edges of the simplex, this disagreement increases at an increasing rate. We incorporate this distance in the central bank's loss function and assume that the central bank wishes to minimize the following augmented loss:

$$L(\pi, p) = \frac{1}{2}\,\mathrm{E}_q[Q(\pi, p)] - \theta\,\mathrm{E}_p[R(p)] \tag{2.7}$$

where θ parametrizes the relative importance of disagreement in the loss function. The second term places a penalty on agreement. The closer the private sector's beliefs are to the central bank's, the higher the loss. Equation (2.7) nests rational expectations as a special case. When $p = q$, $\mathrm{E}_p[R] = 0$, $\bar{\pi} = \mathrm{E}_q[\pi]$, and equation (2.7) simplifies to (2.5). In that case, the central bank chooses inflation to minimize the quadratic loss with model-consistent expectations.

The concern for robustness arises in this formulation because the central bank does not know what the private sector's beliefs are, and it does not place a probability distribution over the possible beliefs. Without placing a probability distribution on beliefs, it cannot take the expected value of its objective function. Instead, it chooses inflation under the assumption that the private sector's beliefs, for any given choice of inflation, will correspond to the "worst-case." Gilboa and Schmeidler (1989) show that this attitude toward Knightian uncertainty is optimal when agents are ambiguity averse. Thus, the central bank evaluates (2.7) by considering the belief that would maximize that function for any rate of inflation it might choose. In other words, the bank chooses the best among the worst-case outcomes. As Hansen and Sargent (2007) show, this decision problem can be modeled as a Stackelberg zero-sum game between a central bank choosing inflation and an adversary choosing beliefs. The central bank moves first, denoted by t_1 in Figure 1, and chooses a state-contingent inflation rate, π (the output gap is pinned down by the Phillips curve). The adversarial agent moves second, denoted by t_2 in Figure 1; having observed the bank's choice of

[4] Woodford (2010) measures the relative entropy of the *distortion factor*, $m = p/q$, that absolute continuity implies. This formulation is equivalent to ours: $\mathrm{E}_q[m \log m] = \mathrm{E}_p[\log(p/q)]$. In the case of continuous probability distributions, the analysis becomes more tractable if it is done in terms of the change of measure, m, than the distribution, p. In our setting it is straightforward to use p.

state-contingent inflation, the adversarial agent chooses private sector beliefs, p. Because the game is zero-sum, the adversarial agent aims to maximize (2.7), while the central bank aims to minimize it. Backward induction yields the following loss function for the central bank:

$$\mathcal{L}(\pi) = \sup_{p} L(\pi, p)$$

From the perspective of the adversarial agent, θ in (2.7) is the cost of choosing a distribution p that deviates from the bank's distribution. The higher the value θ, the more costly it is for the adversarial agent to disagree with the central bank. Thus, $1/\theta$ parametrizes the central bank's concern for robustness. When θ is small, it is relatively cheap for the adversarial agent to disagree with the central bank. As θ increases, it becomes increasingly costly for the adversarial agent to disagree. In the limit, when $\theta \to \infty$, the adversarial agent will choose $q = p$. In this case, the bank is entirely unconcerned about the possibility that the private sector's beliefs differ from its own, and the problem reverts to the analysis with rational expectations.

3 Solution

In this section, we derive the solution to the central bank's problem in two steps. First, we derive the relationship between central bank policy and private sector expectations as a function of private sector beliefs. The constraints pin down this relationship; therefore, it is independent of the central bank's optimization problem. We then turn to optimality, starting with the special case of rational expectations, followed by the general case of robustly optimal policy. The comparison of these two cases, as well as the discussion of the implications on policy, are left to section 4.

3.1 Private Sector Expectations

Recall that the Phillips curve (2.1), the aggregate demand inequality (2.2), and private sector expectations (2.3) and (2.4) constrain the bank. When the ZLB binds, the weak inequality (2.2) holds as a strict equality. In that case, the aggregate demand and aggregate supply schedules uniquely pin down the inflation/output gap pair (π_ℓ, x_ℓ) for any given expectations $(\bar{\pi}, \bar{x})$. It follows that for any choice of inflation when the natural interest rate is positive, π_h, the constraints pin down all other equilibrium variables, $\{x_h, \pi_\ell, x_\ell, \bar{\pi}, \bar{x}\}$, given private sector beliefs, p. Thus, we can reduce the central bank's problem to choosing only the rate of inflation when the natural rate of interest is positive: π_h. In the remainder of this section we will analyze the central bank's problem as choosing π_h. Unless otherwise noted, whenever we refer to the bank choosing a rate of inflation, or alternatively setting the nominal interest rate, we will mean those associated with a positive natural rate, π_h and i_h.

We can restate the problem of the central bank in terms of π_h and expected inflation $\bar{\pi}$. By increasing π_h, the central bank can, in principle, increase $\bar{\pi}$, and thus achieve a lower real rate of interest when it finds itself at the ZLB. The extent to which this is possible will, of course, depend on p, as the following equation shows:

$$\bar{\pi} = \underbrace{\frac{1}{\gamma}}_{\text{GE effect}} \underbrace{(1-p)}_{\text{PE effect}} \pi_h + \kappa \rho_\ell$$

where $\gamma \equiv \frac{1}{1-p(1+\kappa)}$.[5]

When the bank chooses π_h, the direct marginal effect on expected inflation is given by the probability the private sector places on that state, $1-p$. This is a partial equilibrium effect. There is also a general equilibrium feedback effect. An increase in inflation expectations leads to an increase in the output gap expectations through the Phillips curve, both of which determine π_ℓ.[6] An increase in π_ℓ, in turn, has a general equilibrium feedback on $\bar{\pi}$, with its marginal effect given by p.

The same general equilibrium logic applies to beliefs. For any given inflation rate, π_h, as the probability the private sector places on ρ_h decreases, so does the expected rate of inflation. The general equilibrium logic then follows: a lower $\bar{\pi}$ leads to a lower $\bar{x}$, both of which lead to a lower rate of inflation at the ZLB, π_ℓ. This, in turn, feeds back into the expected rate of inflation. As $p \to 1/(1+\kappa)$, the general equilibrium feedback, given by γ, becomes unbounded. Since the Phillips curve ultimately determines the total effect of expectations on π_ℓ in general equilibrium, the general equilibrium effect depends on the slope of the Phillips curve. The stronger the degree of monetary non-neutrality—the smaller the value of κ—the smaller the GE effect.

The general equilibrium effect arises because when nominal interest rates are at the zero lower bound, inflation is a function of expectations only. This is not the case under cost-push shocks. This distinction makes the problem with natural interest rate shocks fundamentally different from cost-push shocks because the quadratic loss portion of the central bank's objective, $\mathrm{E}_q[Q(\pi_h, p)]$, is a function of beliefs. Unbounded expectations at $p = 1/(1+\kappa)$ will necessarily lead to unbounded losses.

3.2 Rational Expectations

We now turn to the question of optimal monetary policy, starting with the rational expectations benchmark.[7] In this case, the beliefs of the central bank and the private sector coincide ($p = q$). More importantly, this means the bank has a deterministic mapping between its policy choice and private sector expectations, so it can equivalently choose inflation or the private sector's expectations.[8]

As a benchmark, we start with the case where the private sector places a zero probability on a negative natural rate, $p = 0$. This is a useful benchmark because it corresponds to the optimal average rate of inflation under mean-zero cost-push shocks. Setting $p = 0$ simplifies the algebraic solution for this optimal average inflation, denoted ρ^*, where

$$\pi = \rho^* \equiv \left(\frac{\kappa\lambda(1-\beta)}{\lambda(1-\beta)^2 + \kappa^2} \right) x^\star. \tag{3.1}$$

A feature of this solution is that should the private sector place a zero probability on a negative natural rate shock, the bank has no need to distinguish between possible states of the world, as its policy decision when $\rho_j < 0$ will have no impact on expectations. Thus, the bank has a single inflation target: $\pi_h = \pi_\ell = \rho^*$. As the following lemma shows, ρ^* is also an important upper bound.

[5] For the derivation of this result, see Appendix A.

[6] In fact, π_ℓ, is a function only of expectations and the negative natural interest rate shock.

[7] This problem has inequality constraints. See the Appendix B for the algorithm of the solution method.

[8] This is a direct implication of rational expectations and is shared by models of imperfect information as long as there is a unique, bounded equilibrium. Central bankers are nearly unanimous in their opinion that while they can *influence* private sector expectations, they cannot *choose* them.

Lemma 1 (Zero Lower Bound)**.** *The aggregate demand constraint binds at the solution in the low state,* $i_\ell = 0$*, iff* $-\rho_\ell > \rho^*$*.*

Proof. See Appendix A. □

It is important to note that this lemma does not say that it is infeasible for the central bank to choose a rate of inflation for which the aggregate demand constraint does not bind in the low state. Rather, it says that is suboptimal. The lemma also points to the fact that a negative natural rate is not a sufficient condition to hit the ZLB. Therefore, for the rest of the paper we assume that $\Pr(-\rho_\ell > \rho^*) = 1$. It is worth noting that this assumption is consistent with the calibration exercise we conduct in section 4.[9]

Figure 2 shows the central bank's optimal rate of inflation under a positive natural interest rate shock, π_h, as a function of the probability of the natural interest rate being negative, q. Since we have that $-\rho_\ell > \rho^*$, a negative natural rate will always put the bank at the ZLB.

As the figure shows, once hitting the ZLB occurs with positive probability, $q > 0$, the marginal benefit of higher inflation at ρ_h increases with q. At the same time, the marginal cost decreases, since the probability of π_h is decreasing. Therefore, π_h, π_ℓ, and $\bar{\pi}$ also increase. In order to increase the rate of inflation when the bank is free to set the nominal interest rate, however, it must set a lower nominal interest rate. Eventually, when the probability of ρ_ℓ is large enough, the central remains at the ZLB regardless of the state of the world. We denote this probability in the Figure 2 by $\tilde{q}$. We say that the economy finds itself in a *liquidity trap* when the nominal interest rate is zero for all states of the world: $i_h = i_\ell = 0$. As Figure 2 shows, this happens for all $p > \tilde{p}$.

We formalize this argument in the following proposition.

Proposition 1 (Optimal monetary policy under rational expectations)**.** *Under rational expectations* $\exists\ \tilde{p}$ *(equivalently,* $\tilde{q}$*) such that:*

1. *If* $p < \tilde{p}$ *the optimal nominal interest rate is positive,* $i_h^* > 0$*, and the optimal inflation rate is given by the first order condition of the central bank's loss function,* (2.5)*.*

2. *If* $p \geq \tilde{p}$ *the optimal nominal interest rate is zero and the bank finds itself in a liquidity trap* $i_h^* = i_\ell^* = 0$*. In that case the aggregate demand constraint always binds and all variables are pinned down by the set of constraints* (2.1)–(2.4)*.*

Proof. See Appendix A. □

The bank's policy is a piece-wise function with a kink at $\tilde{q}$. As Figure 3 shows, there are three distinct regions of beliefs, which under RE, coincide with the objective probability of being at the ZLB. When $p \in (0,\ \tilde{p})$ the aggregate demand schedule in the high state does not bind. In this region the bank can lower the nominal interest rate when the natural rate is positive to increase the expected rate of inflation.

When $p \in [\tilde{p},\ 1/(1+\kappa)]$ the bank finds itself in a liquidity trap and the Neo-Fisherian effect dominates. The dashed curve represents the rate of inflation, π_h, that the bank would choose if i_h could be negative. Notice that the rate of inflation in this region is *higher* than the bank would otherwise prefer. That is, the bank would achieve a *lower* rate of inflation if were able to set a *lower* nominal interest rate. Therefore, although in this region the Neo-Fisherian effect makes it possible for the bank to increase the nominal interest rate in order to raise inflation, doing so is

[9] See Table 2 for a comparison of ρ^* to the natural rate of interest data used in our estimates.

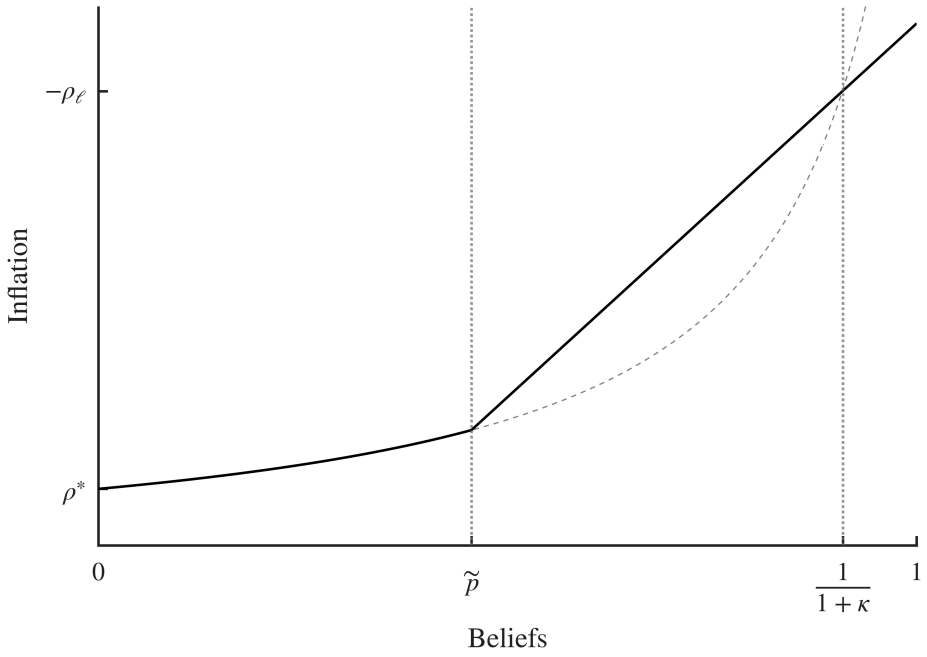


Figure 2: RE: central bank's optimal inflation for varying beliefs. The solid curve represents the chosen optimal rate of inflation, π_h. The dotted curve represents bank's chosen π_h if $\Pr(\rho_j < 0) = 0$.

suboptimal—in this intermediate region the rate of inflation when the natural rate is positive is already too high.

Finally we have the region $p \in (1/(1+\kappa),\ 1)$. In this region the aggregate demand effect once again dominates. While the bank remains in a liquidity trap, the dashed curve is now above the solid line; the rate of inflation is *lower* than the bank would choose if it could set negative nominal rates. If the bank were to raise the nominal interest rate, the rate of inflation would fall even further. We thus have that for intermediate probabilities of experiencing a negative natural rate of interest the Fisher effect dominates but inflation is already too high. When the probability of a negative natural rate of interest is sufficiently high, inflation becomes too low, but the aggregate demand effect dominates.

3.3 Robust Optimality

We now turn to the more general case where the private sector can form different beliefs. We define the central bank's robustly optimal policy, following Woodford (2010), as the best policy under the worst possible unknown beliefs. While these "worst case" beliefs may not reflect those the private sector will hold, faced with ambiguity (rather than uncertainty) the bank guards against potentially catastrophic outcomes by assuming the worst-case scenario. As previously mentioned, we can solve this problem as the Nash equilibrium of a zero-sum game between the central bank choosing inflation and an adversarial agent choosing beliefs.

We solve this problem through backward induction as given by the timing of play in Figure 1.

(t_2) Adversarial agent chooses beliefs as a function of inflation:

$$\hat{q}(\pi_h) \in \arg\max_q L(\pi_h, q) \tag{3.2}$$

(t_1) Central bank chooses inflation as a function of the adversarial agent's best-response function:

$$\hat{\pi}_h \in \arg\min_\pi L(\pi_h, \hat{q}(\pi_h)) \tag{3.3}$$

The solution to this game is given by the following proposition.

Proposition 2 (Robustly optimal monetary policy)**.** *The central bank's robustly optimal inflation rate is independent of* λ, x^*, β, κ, p, q, ρ_h, *and* θ, *and is given by:*

$$\hat{\pi}_h = -\rho_\ell \tag{3.4}$$

Proof. See Appendix A. □

While we leave the details of the proof to the appendix, we will present a sketch here. Because beliefs can yield unbounded losses, the loss function is not concave, so we cannot use standard constrained optimization techniques. We prove our result by ruling out candidate policies.

We first show that the central bank will never choose rate of inflation above $-\rho_\ell$. If $\pi_h > -\rho_\ell$ the aggregate demand constraint is *never* binding for any $q \in (0, 1)$. In that case, by Jensen's inequality, we have that $\pi_h = \pi_\ell$. But since we have that $\rho^* < -\rho_\ell$ we also have that the loss function is increasing in π.

We then show that if the bank chooses any inflation rate below $-\rho_\ell$, it will find itself at the ZLB when the natural rate is negative for any $p \in (0, 1)$. But by equation (3.1) we have that for any rate of inflation that the central bank chooses, the adversarial agent can guarantee itself an arbitrarily large loss for the bank by choosing a belief p arbitrarily close to $1/(1+\kappa)$. We have that losses are bounded if $\pi_h \geq -\rho_\ell$, and unbounded otherwise. Finally, the smallest bounded losses occur at $\pi_h = \rho_\ell$.

This same logic applies when there are more than two states of the world.

Proposition 3 (Robustly optimal monetary policy: general solution)**.** *Suppose there is more than one value of the negative natural interest rate. Assume that the probability distribution has finite support where* $\rho^* = \max \rho_\ell$ *and* $\rho^{**} \equiv \min \rho_\ell$. *The robustly optimal inflation rate is independent of* λ, x^*, β, κ, p, q, ρ_h, *and* θ *and is given by:*

$$\hat{\pi}_h = -\rho^{**} \tag{3.5}$$

Proof. See Appendix A. □

From 2 we have that the central bank faces unbounded losses if it chooses a rate of inflation below ρ_ℓ, but bounded losses above it. When there is more than one possible natural rate that puts the central bank at the ZLB, the "worst-case" beliefs put an arbitrarily high probability on the lowest possible natural rate. In short, Propositions 2 and 3 say that a necessary condition of the robustly optimal policy is that the central bank is never constrained by the ZLB.

To see why this is the case, we can let $1 - p$ continue to be the private sector's belief that the natural rate of interest is less than ρ^*. The general equilibrium feedback that leads to unbounded

expectations arises for *all* values of the negative natural interest rate, so the only relevant belief affecting the feedback is $1-p$. So, we once again end up with a quadratic loss function independent of p, and the global maximum occurs when the private sector's beliefs agree with the central bank's.

The intuition for the robustly optimal policy comes from equation (3.1). When the aggregate demand curve binds, the general equilibrium feedback from beliefs *can* be explosive. Therefore, the robust policy *severs* the effect of beliefs on equilibrium outcomes. It is straightforward to see that when $\pi_h = -\rho_\ell$, γ disappears and $\pi_h = \bar{\pi}$. The robustly optimal policy yields an equilibrium that is independent of p.

4 Implications

In this section we highlight three related implications that arise from our analysis. First, while the central bank is interested in policy that is robust to private sector beliefs, the robustly optimal policy is, additionally, robust to *all* model parameters. Second, the robustly optimal policy is also simpler to communicate than the rational expectations policy. Third, the robustly optimal solution calls for higher average inflation than that under rational expectations.

4.1 General Robustness

The general robustness of the robustly optimal solution implies that is independent of the objective probability, or subjective beliefs, of the state of the world. The rational expectations solution is a function of all parameters; which implies that the central bank's announced policy is contingent on—and thus sensitive to—all parameters. Figure 3 compares the bank's robustly optimal solution to the benchmark rational expectations case. The rational expectations analysis calls for inflation to increase as a function of p and for different rates of inflation depending on the realization of the natural rate of interest. Further, the levels depend on a convolution of model parameters. The robustly optimal solution stands out in the figure for its parsimony. It depends on a single parameter $-\rho^{**}$.

This makes the robustly optimal policy robust in an additional, unexpected, way: the central bank does not need to know the other model parameters in order to implement it. In particular, it does not need to know the slope of the Phillips Curve, κ, household's discount factor, β, the probability of the natural rate becoming negative, p, the efficient rate of output, x^*, or its relative importance, λ. This is notable because in practice central banks do not know these parameters for certain.

4.2 Target Inflation

Unlike the rational expectations solution, which is a function of all model parameters, the robustly optimal solution stands out for its parsimony. Once again, although parsimony is not an explicit objective of the central bank, it is a feature of the solution. And parsimony of the policy implies parsimony in its communication. This result is consistent with the behavior and concerns of central banks. Importantly, the simplicity of our solution is unique. Other models that yield higher inflation rates in response to the positive probability of hitting the ZLB cannot simultaneously explain the bias that most central banks display toward simplicity of communication. In fact, it is generally difficult to reconcile the convoluted solutions under rational expectations analyses with the simplicity with which central banks communicate their policy objectives. In particular,

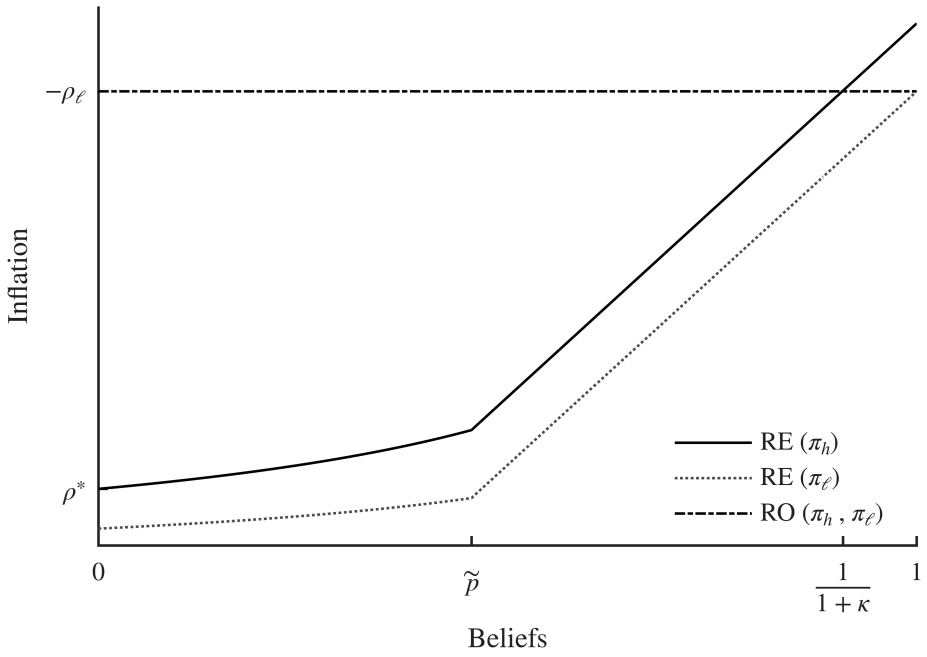


Figure 3: Optimal inflation under Rational Expectations (RE) and Robust Optimality (RO)

inflation targeting is not a direct consequence of rational expectations analysis, as the optimal rate of inflation is highly sensitive to the model parameters. The robustly optimal solution we propose is insensitive to these parameters; its most straightforward characterization is as a strict inflation target.

4.3 The Case for a Higher Inflation Target

As Figure 3 shows, the robustly optimal inflation target is generally higher than the optimal average rate of inflation under rational expectations. In this section we conduct a simple quantitative exercise to estimate the difference. To do this we need to calibrate our model and data on the distribution of the natural rate of interest. Table 1 displays our calibrated values, which come from Woodford (2010).

To estimate the distribution of the natural interest rate we use the dataset from Negro et al. (2017).[10] Negro et al. (2017) provide average natural rate of interest estimates at a quarterly frequency from 1960 to 2021. They also provide the quarterly estimates with confidence intervals of one and two standard deviations from the mean. To calculate p, ρ_h and ρ_ℓ for our model, we take their estimates and reclassify them as either ρ_h or ρ_ℓ. We do so by conditionally averaging every estimate across the 50 year time period that was larger than ρ^*, yielding $\mathrm{E}[\rho_h]$, or less than

[10]The replication package for Negro et al. (2017) containing the data we used can be accessed at https://github.com/FRBNY-DSGE/rstarBrookings2017/blob/master.

Symbol	Description	Value
κ	Slope of Phillips curve	0.025
β	Household's discount factor	0.98
λ	Bank's relative weight of output	.08
x^*	Efficient rate of output	0.2

Table 1: Model Parameters

(a) Mean − 1 S.D

	RE	RO	RO*
π_h	1.33	7.44	3.53
π_ℓ	1.24	7.44	3.53
$\bar{\pi}$	1.30	7.44	3.53

$\mathrm{E}[\rho_\ell] = -3.53, \mathrm{E}[\rho_h] = 2.54, p = 0.33$

(b) Mean

	RE	RO	RO*
π_h	1.27	7.44	2.84
π_ℓ	1.22	7.44	2.84
$\bar{\pi}$	1.26	7.44	2.84

$\mathrm{E}[\rho_\ell] = -2.84, \mathrm{E}[\rho_h] = 3.16, p = 0.23$

(c) Mean + 1 S.D.

	RE	RO	RO*
π_h	1.24	7.44	2.78
π_ℓ	1.18	7.44	2.78
$\bar{\pi}$	1.23	7.44	2.78

$\mathrm{E}[\rho_\ell] = -2.78, \mathrm{E}[\rho_h] = 3.57, p = 0.09$

Table 2: Optimal inflation values for a given probability of being at the ZLB and a negative natural rate shock. Inflation and natural rate values are expressed in percentage terms. RO* denotes a "restricted RO" solution, where the private sector has correct beliefs, conditional on being at the ZLB.

ρ^*, yielding $\mathrm{E}[\rho_\ell]$.[11] To retrieve the probability of the ZLB from these estimates, we take a ratio of the number of quarters that report a negative natural rate to the total number of quarters provided by the data.

The robustly optimal solution is very sharp—it suggests that the central bank should choose the worst-possible outcome of the negative natural rate; in this case, 7.44 percent. This result hinges, however, on the central bank expecting that the private sector will choose to put an arbitrarily high probability on the lowest-possible natural rate, ρ^{**}. This is reported as RO in Table 2 and represents an upper bound on the inflation target. As a lower bound we also consider the strongest restriction on private sector beliefs; conditional on the natural rate being less than ρ^*, the private sector's beliefs agree with those of the central bank. We report this inflation rate as RO*. Unlike the optimal inflation under rational expectations, the robustly optimal inflation targets are significantly higher, and above the current two percent target.

5 Conclusion

We have characterized the central bank's robustly optimal commitment policy in an environment with three key features: natural interest rate shocks, the possibility of finding itself at the zero lower bound, and the absence of a deterministic link between policy and private sector beliefs, which we formalize as Knightian uncertainty. The central bank then chooses monetary policy that is best on under the "worst-case" beliefs that the private sector might hold. This last feature is

[11] We exclude the Q1 2020 estimate from the ρ_ℓ calculation. This estimate was −30.66%; the second most negative estimate was −7.44% from Q4 1974.

consistent with the uncertainty and concern about the impact of monetary policy on private sector expectations captured by the minutes of the FOMC. The robustly optimal policy is an inflation target that depends only on the natural rate of interest. The key takeaway is that, when the central bank is unsure about the mapping between its policy and private sector expectations is to anchor them so deeply that the public has no reason to change them. This logic can explain why the FOMC minutes record so much concern for how expectations might react, which would not be the case if central bankers were sure of how policy would translate into expectations.

Our model is highly stylized; a natural direction for future research would be to consider the robustly optimal policy in a dynamic model with a richer stochastic process of the natural interest rate shocks. It nonetheless explains several discrepancies between the theory and the practice of monetary policy. First, it explains why central bankers prefer simple policies, even though rational expectations solutions are often highly complex. Second, it explains why central bankers have adopted explicit inflation targets even though they are not necessary conditions for optimal monetary policy under rational expectations. And third, it explains why that inflation target has been set at around two percent even though this number is significantly higher than the analysis under rational expectations would suggest.

A Proofs

Proof. Proof of Lemma 1
To show that when $-\rho_\ell \leq \rho^*$ the aggregate demand constraint does not bind we assume it doesn't. By Jensen's inequality, the policymaker will choose $\pi_h = \pi_\ell = \rho^*$. But then the inequality constraint requires that $\rho^* + \rho_\ell \geq 0$, as above. To show that the aggregate constraint binds when $-\rho_\ell > \rho^*$, notice that at $p = 0$, the loss function is flat with respect to p. That means that when the probability of a negative natural rate becomes positive, up to a first order, the central bank's choices are the same. So up to first order the aggregate demand constraint is $\rho^* + \rho_\ell$, which is now negative, so the nominal interest rate binds in the low state. But if it binds at $p = 0$, by continuity it continues to bind as p increases. □

Proof. Proof of Proposition 1

To show that there exist a belief p at which the Bank's choice variables are entirely pinned down, we use the system of constraints from Sequence 2 to algebraically solve the following Lagrangian:

$$\mathcal{L}(\pi, p) = \mathrm{E}_q[Q] + \mu\ G$$

$$G = \begin{cases} \pi_h - \kappa x_h - \beta\bar{\pi} = 0 \\ \pi_\ell - \kappa x_\ell - \beta\bar{\pi} = 0 \\ \bar{x} - p x_\ell - (1-p) x_h = 0 \\ \bar{\pi} - p\pi_\ell - (1-p)\pi_h = 0 \\ \bar{x} + \bar{\pi} + \rho_\ell - x_\ell = 0 \end{cases}$$

Algebra produces optimal inflation output gap pairs, π_h^*/x_h^* and π_ℓ^*/x_ℓ^*, as a function of private sector beliefs. To solve for $\tilde{p}$, we substitute these functions into the aggregate demand constraint, set the nominal interest rate to zero, and solve for p:

$$i_h = \bar{x} + \bar{\pi} - x_h + \rho_h$$

$$i_h = \left(\frac{1-\beta+\kappa}{\kappa}\right)\bar{\pi} - \frac{\pi_h - \beta\bar{\pi}}{\kappa} + \rho_h$$

$$\hat{i_h} = \left(\frac{1-\beta+\kappa}{\kappa}\right)\left(p\hat{\pi_\ell} + (1-p)\hat{\pi_h}\right) - \hat{x_h} + \rho_h = 0$$

Where

$$\hat{\pi_h} = \frac{-\kappa(x^*(-1+\beta)\lambda - p x^*(-1+\beta)(1+\kappa)\lambda + p\rho_\ell((-1+\beta)^2\lambda + \kappa(\kappa+\kappa^2+\lambda)))}{p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda}$$

$$\hat{\pi_\ell} = \frac{-\rho_\ell\kappa^3(-1+p+p\kappa) + (-1+p)x^*(-1+\beta)\kappa(1+\kappa)\lambda - \rho_\ell\kappa((-1+p)(-1+\beta)^2 + p\kappa)\lambda}{p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda}$$

$$\hat{x_h} = \frac{p\rho_\ell\kappa^2(-1+(1-\beta)\kappa) + x^*(-1+\beta)^2\lambda + p(\rho_\ell + x^*)(-1+\beta)(1-\beta+\kappa)\lambda}{p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda}$$

The resulting equation for i_h^* can be rewritten as:

$$\hat{i_h} = \frac{1}{\kappa}\Bigg[\rho_h - \frac{p\kappa^2\big(-1+(-1+\beta)\kappa\big) + x^*(-1+\beta)^2\lambda + p(\rho_\ell + x^*)(-1+\beta)(1-\beta+\kappa)\lambda}{p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda}$$
$$+ (1-\beta+\kappa)\Bigg(\frac{-(1-p)\kappa\Big(x^*(-1+\beta)\lambda - px^*(-1+\beta)(1+\kappa)\lambda + p\rho_\ell\big((-1+\beta)^2\lambda + \kappa(\kappa+\lambda^2)\big)\Big)}{p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda}$$
$$+ p\Bigg(\frac{\rho_\ell\kappa}{p} - \frac{(1-p)\kappa(1+\kappa)\Big(x^*(-1+\beta)\lambda - px^*(-1+\beta)(1+\kappa)\lambda + p\rho_\ell((-1+\beta)^2\lambda + \kappa(\kappa+\lambda^2))\Big)}{(1-p)(1+\kappa)\Big(p\kappa^4 + \kappa^2(1+p(-1+\lambda)) - (-1+p)(-1+\beta)^2\lambda\Big)}\Bigg)\Bigg)\Bigg] \tag{A.1}$$

Figure 4 shows this result as a function of p. As p increases, it is optimal for the bank to lower the nominal interest rate. However, for any beliefs $p > \tilde{p}$, the bank cannot commit to its desired negative nominal interest rate. Therefore, we can solve for the closed form solution of $\tilde{p}$ that is independent of the bank's policy decisions:

$$\tilde{p} = \frac{-\rho_h\kappa^2 - \rho_h\lambda + 2\rho_h\beta\lambda - \rho_h\beta^2\lambda - x^*\kappa\lambda + x^*\beta\kappa\lambda}{(\rho_h - \rho_\ell)\left(-\kappa^2 + \kappa^4 - \lambda + 2\beta\lambda - \beta^2\lambda + \kappa^2\lambda\right)} \tag{A.2}$$

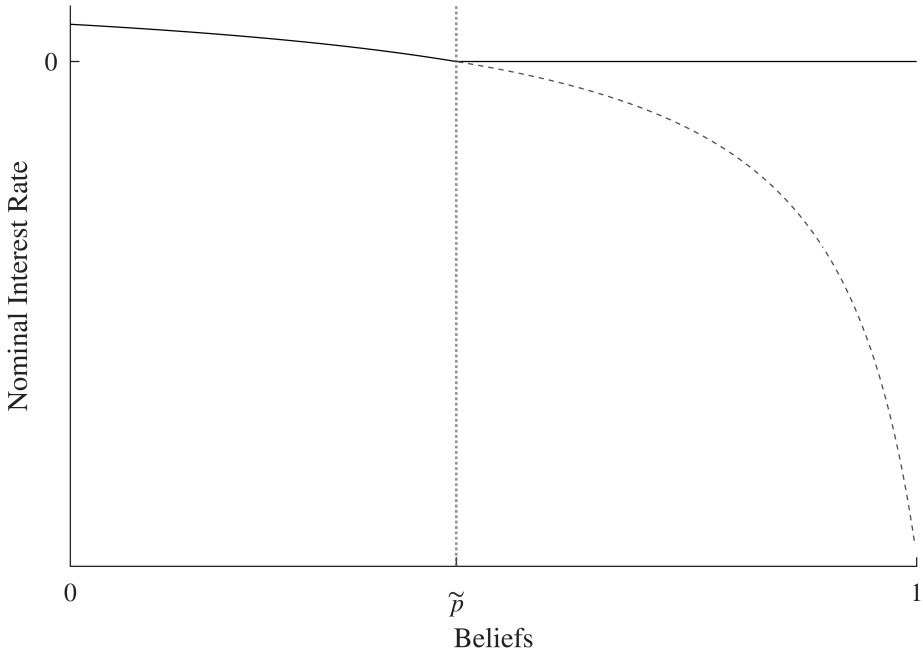


Figure 4: Supplemental Figure for Figure 2 and Proposition 1. The solid curve represents the nominal interest rate chosen by the central bank, i_h^*. The dotted curve represents the nominal interest rate that the central bank would have chosen in the absence of the ZLB when the natural rate is positive.

□

Proof. Proof of Proposition 2

First, we derive the relationship given by Equation 3.1:

$$\gamma\bar{\pi} = (1-p)\pi_h + \kappa\rho_\ell$$

where $\gamma = 1 - p(1+\kappa)$. We start by combining the constraints under a negative natural rate shock with the Phillips curve under a positive natural rate shock. Since π_h is the bank's choice variable, we reorganize the equation to isolate x_h:

$$\begin{aligned} \pi_h &= \kappa x_h + \beta\bar{\pi} \\ x_h &= \frac{\pi_h - \beta\bar{\pi}}{\kappa} \end{aligned} \tag{A.3}$$

Since constraint on the nominal interest rate is binding when the natural rate of interest is negative, we can write the weak aggregate demand constraint as a strict equality constraint:

$$x_\ell = \bar{x} + \bar{\pi} + \rho_\ell$$

We can now derive the relationship between $\bar{x}$, x_ℓ, and $\bar{\pi}$:

$$\begin{aligned} \bar{x} &= (1-p)x_h + px_\ell \\ \bar{x} &= (1-p)\left(\frac{\pi_h - \beta\bar{\pi}}{\kappa}\right) + p\left(\frac{\pi_\ell - \beta\bar{\pi}}{\kappa}\right) \\ \bar{x} &= \frac{1}{\kappa}(1-\beta)\bar{\pi} \\ x_\ell &= \left[\frac{1-\beta+\kappa}{\kappa}\right]\bar{\pi} + \rho_\ell \end{aligned}$$

Now that we have the relationship above, we can derive the relationship between expected inflation and π_ℓ:

$$\begin{aligned} \pi_\ell &= \kappa x_\ell + \beta\bar{\pi} \\ \pi_\ell &= \kappa\left[\frac{1}{\kappa}(1-\beta)\bar{\pi} + \bar{\pi} + \rho_\ell\right] + \beta\bar{\pi} \\ \pi_\ell &= (1+\kappa)\bar{\pi} + \kappa\rho_\ell \end{aligned}$$

With this relationship between π_ℓ and $\bar{\pi}$, we will derive the relationship between expected inflation and inflation under a positive natural rate shock:

$$\begin{aligned} \pi_\ell &= (1+\kappa)\bar{\pi} + \kappa\rho_\ell \\ \bar{\pi} &= (1-p)\pi_h + p\pi_\ell \\ \bar{\pi} &= (1-p)\pi_h + p((1+\kappa)\bar{\pi} + \kappa\rho_\ell) \\ \bar{\pi}(1-p(1+\kappa)) &= (1-p)\pi_h + \kappa\rho_\ell \end{aligned}$$

Thus, yielding the relationship from Equation 3.1:

$$\gamma\bar{\pi} = (1-p)\pi_h + \kappa\rho_\ell$$

First, show that the ZLB binds iff $\pi_h < -\rho_\ell$. To do this we solve the central bank's problem assuming that the ZLB doesn't bind.

$$\mathcal{L}(\pi_\ell, \Omega) = \frac{1}{2} q \left({\pi_\ell}^2 + \lambda (\frac{\pi_\ell - \beta(p\pi_\ell + (1-p)\pi_h)}{\kappa} - x^\star)^2 \right) + \frac{1}{2}(1-q) \left({\pi_h}^2 + \lambda (\frac{\pi_h - \beta(p\pi_\ell + (1-p)\pi_h)}{\kappa} - x^\star)^2 \right) \quad \text{(A.4)}$$

$$\frac{d\mathcal{L}(\pi_\ell, \Omega)}{d\pi_\ell} = 0$$

Solve for π_ℓ:

$$\pi_\ell = \pi_h \cdot \psi_\pi^l + \psi_0^l$$

$$\psi_\pi^l = \frac{-\beta\lambda\Big[(1-q)(1-p)(-1+p\beta) + (1-q)p(-1+(1-p)\beta)\Big]}{(1-q)p^2\beta^2\lambda + q\Big[\kappa^2 + \lambda(\beta p - 1)^2\Big]}$$

$$\psi_0^l = -(1-p(1+\kappa)) \cdot \frac{x^\star\kappa\lambda\Big[(1-q)p\beta + q(-1+p\beta)\Big]}{(1-q)p^2\beta^2\lambda + q\Big[\kappa^2 + \lambda(\beta p - 1)^2\Big]}$$

Substitute the optimal π_ℓ into the aggregate demand constraint:

$$i_\ell = \bar{x} + \bar{\pi} - x_\ell + \rho_\ell$$

$$i_\ell = \left(\frac{1-\beta+\kappa}{\kappa}\right)\bar{\pi} - \frac{\pi_\ell - \beta\bar{\pi}}{\kappa} + \rho_\ell$$

$$i_\ell = \pi_\ell \left[\frac{p(1+\kappa)-1}{\kappa}\right] + \frac{(1-p)(1+\kappa)}{\kappa}\pi_h + \rho_\ell$$

resulting fn for i_ℓ:

$$i_\ell = \psi_0^i + \psi_\pi^i \cdot \pi_h + \rho_\ell$$

Where

$$\psi_\pi^i = \frac{(1-p)(1+\kappa)}{\kappa} - (1-p(1+\kappa)) \cdot C$$

$$C = \frac{\lambda\,\beta\left[(1-q)p\big(1-(1-p)\beta\big) + q(1-p)\big(1-p\beta\big)\right]}{(1-q)\,p^2\,\beta^2\,\lambda + q\Big[\kappa^2 + \lambda(\beta p - 1)^2\Big]}$$

And

$$\psi_0^i = (1-p(1+\kappa)) \cdot D$$

$$D = \frac{x^\star\,\lambda\left[(1-q)p\beta - q(1-p\beta)\right]}{(1-q)\,p^2\,\beta^2\,\lambda + q\Big[\kappa^2 + \lambda(\beta p - 1)^2\Big]}$$

$$\lim_{p \to \frac{1}{1+\kappa}} i_\ell = \pi_h + \rho_\ell \leq 0$$

$$\pi_h \leq -\rho_\ell$$

We thus have that the aggregate demand inequality constraint is violated only when $\pi_h < -\rho_\ell$. Because the loss function is convex, by Jensen's inequality $\mathrm{E}[Q(\pi)] \geq Q(\mathrm{E}[\pi])$, so we have that $\pi_h = \pi_\ell = \bar{\pi}$. Since the min is given by ρ and $-\rho_\ell$ is greater than ρ, $\forall\ \pi > -\rho_\ell$ we have that $Q(\pi) > Q(-\rho_\ell)$. Therefore we have that when $\pi \geq -\rho_\ell$, the loss is minimized at $\pi = -\rho_\ell$.

All that is left is to show that the losses are unbounded when $\pi < -\rho_\ell$. Let

$$p = \frac{1}{1+\kappa} + \varepsilon$$

where $\varepsilon > 0$. Then we have $\gamma = \frac{1}{\varepsilon(1+\kappa)}$. Then $\forall\ \pi_h < -\rho_\ell$ by (3.1), $\forall\ N > 0$, $\exists\ \varepsilon > 0$ such that $Q(\pi_h) > N$. We therefore have that $-\rho_\ell$ is the robustly optimal rate of inflation.

□

Proof of proposition 3 The relationship between inflation and expectations becomes:

$$\gamma\bar{\pi} = (1-p)\pi_h + \kappa \mathrm{E}[\rho_\ell]$$

The worst possible expectations arise when the private sector puts all probability on the lowest possible natural rate.

Proof.

B Rational Expectations Solution Methods

B.1 Model Breakdown

Consider the loss function given by Equation 2.5

$$\mathrm{E}_q[Q] = q(\pi_l^2 + \lambda(x_l - x^*)^2) + (1-q)(\pi_h^2 + \lambda(x_h - x^*)^2) \tag{B.1}$$

We solve this function sequentially via the system of linear equations. First, we solve the bank's problem without enforcing the inequality constraints given by Equation 2.2.

$$\text{Constraints} : \begin{cases} \pi_h - \kappa x_h - \beta\bar{\pi} = 0 \\ \pi_\ell - \kappa x_\ell - \beta\bar{\pi} = 0 \\ \bar{x} - p x_\ell - (1-p)x_h = 0 \\ \bar{\pi} - p\pi_\ell - (1-p)\pi_h = 0 \end{cases}$$

We will refer to this as Sequence 1. We rewrite the constraints with the following linear algebra, and compute the solution using a constrained optimization algorithm.

Sequence 1.

$$I \begin{pmatrix} x_h \\ x_\ell \\ \bar{\pi} \\ \bar{x} \end{pmatrix} = B \begin{pmatrix} x_h \\ x_\ell \\ \bar{\pi} \\ \bar{x} \end{pmatrix} + C \begin{pmatrix} \pi_h \\ \pi_\ell \end{pmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & -\frac{\beta}{\kappa} & 0 \\ 0 & 0 & -\frac{\beta}{\kappa} & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} \frac{1}{\kappa} & 0 \\ 0 & 1 \\ 1-p & 0 \\ p & 0 \end{bmatrix}$$

Without aggregate demand constraints, the bank does not distinguish between states of the world. Thus, the algebraic solution to this system of equations is given by

$$\pi_h = \pi_\ell = \bar{\pi} = \frac{x^\star \kappa \lambda (1-\beta))}{(\lambda(1-\beta)^2 + \kappa^2)} = \rho^\star$$

$$x_h = x_\ell = \bar{x} = \left(\frac{1-\beta}{\kappa}\right)\rho^\star$$

We now check the nominal interest rates in post for each state to determine the possible shocks, if any, that would result in a negative nominal interest rate under the bank's optimal rate of inflation.

$$i_j = (\bar{x} - x_j) + \bar{\pi} + \rho_j$$

$$i_j = \rho^* + \rho_j$$

In the case of a positive shock, i_h is always positive. In the case of a negative shock, i_ℓ is only positive if $\rho_\ell > -\rho^\star$. Therefore, a negative shock will result in the ZLB if $-\rho_\ell > \rho^\star$.

We repeat the same algorithm given by Sequence 1, now imposing (2.2) as an equality constraint in the case of a negative shock.

$$\text{Constraints} : \begin{cases} \pi_h - \kappa x_h - \beta\bar{\pi} = 0 \\ \pi_\ell - \kappa x_\ell - \beta\bar{\pi} = 0 \\ \bar{x} - p x_\ell - (1-p)x_h = 0 \\ \bar{\pi} - p\pi_\ell - (1-p)\pi_h = 0 \\ \bar{x} + \bar{\pi} + \rho_\ell - x_\ell = 0 \end{cases}$$

This is referred to as Sequence 2. Again, we rewrite the constraints with the following linear algebra, and compute the solution using the same constrained optimization algorithm.

Sequence 2.

$$I \begin{pmatrix} x_h \\ x_\ell \\ \pi_\ell \\ \bar{x} \\ \bar{\pi} \end{pmatrix} = B \begin{pmatrix} x_h \\ x_\ell \\ \pi_\ell \\ \bar{x} \\ \bar{\pi} \end{pmatrix} + C \begin{pmatrix} \pi_h \\ \rho_\ell \end{pmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & -\frac{\beta}{\kappa} \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \kappa & 0 & 0 & \beta \\ 1-p & p & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} \frac{1}{\kappa} & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1-p & 0 \end{bmatrix}$$

Sequence 2 reflects the $p < \tilde{p}$ segment of Figure 2. Once again, we check the weak aggregate demand constraint given by Equation 2.2, to find the belief p that results in the ZLB. Once the ZLB is binding in both states of the world, the central bank is entirely pinned down —all constraints are now strict equality constraints.

$$\text{Constraints}: \begin{cases} \pi_h - \kappa x_h - \beta\bar{\pi} = 0 \\ \pi_\ell - \kappa x_\ell - \beta\bar{\pi} = 0 \\ \bar{x} - p x_\ell - (1-p)x_h = 0 \\ \bar{\pi} - p\pi_\ell - (1-p)\pi_h = 0 \\ \bar{x} + \bar{\pi} + \rho_h - x_h = 0 \\ \bar{x} + \bar{\pi} + \rho_\ell - x_\ell = 0 \end{cases}$$

This segment, $p \geq \tilde{p}$, is referred to as Sequence 3.

Sequence 3.

$$I \begin{pmatrix} \pi_h \\ \pi_\ell \\ x_h \\ x_\ell \\ \bar{\pi} \\ \bar{x} \end{pmatrix} = B \begin{pmatrix} \pi_h \\ \pi_\ell \\ x_h \\ x_\ell \\ \bar{\pi} \\ \bar{x} \end{pmatrix} + C \begin{pmatrix} \rho_h \\ \rho_\ell \end{pmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & \kappa & 0 & \beta & 0 \\ 0 & 0 & 0 & \kappa & \beta & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1-p & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-p & p & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

This system can alternatively be solved with the following Lagrangian:

$$\mathcal{L}(\pi, p) = \mathrm{E}_q[Q] + \mu\ G$$

$$G = \begin{cases} \pi_h - \kappa x_h - \beta\bar{\pi} = 0 \\ \pi_\ell - \kappa x_\ell - \beta\bar{\pi} = 0 \\ \bar{x} - p x_\ell - (1-p)x_h = 0 \\ \bar{\pi} - p\pi_\ell - (1-p)\pi_h = 0 \\ \bar{x} + \bar{\pi} + \rho_h - x_h = 0 \\ \bar{x} + \bar{\pi} + \rho_\ell - x_\ell = 0 \end{cases}$$

References

Akerlof, G. A., W. T. Dickens, and G. L. Perry (1996). The macroeconomics of low inflation. *Brookings Papers on Economic Activity 27*(1), 1–76.

Andrade, P., J. Galí, H. L. Bihan, and J. Matheron (2019). The optimal inflation target and the natural rate of interest. Brookings Papers on Economic Activity Fall, Brookings Institution.

Angeletos, G.-M. and K. Sastry (2021). Managing expectations: Instruments versus targets. *The Quarterly Journal of Economics 136*(4), 2467–2532.

Athey, S., A. Atkeson, and P. Kehoe (2005). The optimal degree of discretion in monetary policy. *Econometrica 73*(5), 1431–1475.

Ball, L. (2013). The case for four percent inflation. *Central Bank Review 13*(2), 17–31.

Barro, R. and D. Gordon (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics 12*(1), 101–121.

Billi, R. M. (2008). Optimal inflation for the u.s. economy. Research Working Paper RWP 07-03, Federal Reserve Bank of Kansas City.

Blanco, A. (2021). Optimal inflation target in an economy with menu costs and a zero lower bound. American Economic Journal: Macroeconomics Vol. 13, No. 3, American Economic Association.

Coibion, O., Y. Gorodnichenko, and J. Wieland (2012). The optimal inflation rate in new keynesian models: Should central banks raise their inflation targets in light of the zero lower bound? *The Review of Economic Studies 79*(4), 1371–1406.

Eggertsson, G. B. and M. Woodford (2003). The zero lower bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, 1–48.

FOMC (2009). Fomc statement [press release], january 28, 2009. Federal Reserve Board. https://www.federalreserve.gov/newsevents/pressreleases/monetary20090128a.htm.

FOMC (2012). Comments in the december 11–12, 2012, fomc meeting (transcript). Federal Reserve Board. https://www.federalreserve.gov/monetarypolicy/files/FOMC20121212meeting.pdf.

García-Schmidt, M. and M. Woodford (2019). Are low interest rates deflationary? a paradox of perfect-foresight analysis. *American Economic Review 109*(1), 86–120.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Hansen, L. P. and T. Sargent (2007). *Robustness*. Princeton, NJ: Princeton University Press.

Hazell, J., J. Herreno, E. Nakamura, and J. Steinsson (2022). The slope of the phillips curve: Evidence from u.s. states. *Quarterly Journal of Economics 137*(3), 1299–1344.

Krugman, P. (1998). It's baaack: Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity*, 3–39.

Kydland, F. and E. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy 85*(3), 473–492.

Lepetit, A. (2022). The optimal inflation rate with discount factor heterogeneity. *Journal of Money, Credit and Banking 54*(7), 1971–1996.

Mineyama, T., W. Hirata, and K. Nishizaki (2022, September). Optimal inflation rates in a non-linear new keynesian model: The case of japan and the united states. *International Journal of Central Banking 18*(3), 1–45.

Mishkin, F. S. (2005, May). The inflation targeting debate. Technical report, Columbia University, Graduate School of Business and National Bureau of Economic Research. John Kuszczak Memorial Lecture at the Bank of Canada's Annual Conference, Issues in Inflation Targeting, Ottawa, April 28–29, 2005.

Negro, M. D., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017). Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity 48*(1), 235–316.

Ngo, P. V. (2018). The risk of hitting the zero lower bound and the optimal inflation target. *Macroeconomic Dynamics 22*(2), 402–425.

Rotemberg, J. and M. Woodford (1999). An optimization-based econometric framework for the evaluation of monetary policy. Technical report, National Bureau of Economic Research.

Schmitt-Grohé, S. and M. Uribe (2009). Foreign demand for domestic currency and the optimal rate of inflation. NBER Working Paper 15494, National Bureau of Economic Research.

Schmitt-Grohé, S. and M. Uribe (2011). The optimal rate of inflation. In *Handbook of Monetary Economics*, Volume 3B, pp. 653–722. Benjamin M. Friedman and Michael Woodford (eds.).

Uribe, M. (2022). The neo-fisher effect: Econometric evidence from empirical and optimizing models. *American Economic Journal: Macroeconomics 14*(3), 133–162.

Walsh, C. (2009). Using monetary policy to stabilize economic activity. Technical report, Federal Reserve Bank of Kansas City.

Williams, J. C. (2009). Heeding daedalus: Optimal inflation and the zero lower bound. *Brookings Papers on Economic Activity 40*(2), 1–37.

Woodford, M. (2003). *Interest and Prices*. Princeton, NJ: Princeton University Press.

Woodford, M. (2010). Robustly optimal monetary policy with near-rational expectations. *American Economic Review 100*(1), 274–303.