Managing Expectations: Instruments vs. Targets*

George-Marios Angeletos[†] **Karthik A. Sastry**[‡]

December 5, 2020

Abstract

Should policy communications aim at anchoring expectations of the policy instrument ("keep interest rates at zero until date x") or of the targeted outcome ("do whatever it takes to bring unemployment down to u%")? We study how the optimal approach depends on a departure from rational expectations. People have limited depth of knowledge and rationality, or form otherwise distorted beliefs about the behavior of others and the general equilibrium (GE) effects of policy. The bite of this distortion on implementability and welfare is minimized by target-based guidance if and only if GE feedback is strong enough. This offers a rationale for why central banks should shine the spotlight on unemployment when faced with a prolonged liquidity trap, a steep Keynesian cross, or a large financial accelerator.

Keywords: Zero Lower Bound, forward guidance, Keynesian multipliers, common knowledge, level-k thinking, cognitive discounting, bounded rationality.

JEL Codes: D8, E1, E2, E5.

*An earlier version of this paper was entitled "Managing Expectations without Rational Expectations." We are grateful to the editor, Robert Barro, and four anonymous referees for extensive feedback that greatly improved the paper; to Gaetano Gaballo, Gauti Eggertsson and Marcus Hagedorn for insightful conference discussions; and to Roland Benabou, V.V. Chari, Martin Eichenbaum, John Leahy, Stephen Morris, and various seminar participants for valuable comments. Angeletos acknowledges the financial support of the National Science Foundation (Award #1757198).

†MIT and NBER; angelet@mit.edu.

‡MIT; ksastry@mit.edu.

1 Introduction

Forward guidance is not comprehensive. Even if a central bank can shape expectations about future interest rates, it remains up to the public to predict the consequences for aggregate employment and income. Under what circumstances is it better to do the opposite, anchoring expectations about the targeted economic outcomes and leaving the public to ponder the supporting policy?

The existing literature on instruments and targets for monetary policy (e.g., Poole, 1970; Friedman, 1990; Atkeson, Chari and Kehoe, 2007) emphasizes controllability, accountability, and contingency on shocks. We instead focus on the difficulty people may have in reasoning about the economy, especially during unprecedented times like the Great Recession or the Covid-19 crisis.

We model this difficulty as a structured departure from rational expectations equilibrium (REE). People understand a policy's direct, partial equilibrium (PE) effect but not necessarily its indirect, general equilibrium (GE) effect. For example, although people may understand that lower rates mean cheaper credit for themselves, they may incorrectly perceive how others' spending, and hence aggregate demand, responds to the same policy. Our main result is that, in such circumstances, the optimal communication strategy switches from anchoring expectations of instruments to anchoring expectations of targets like employment or income as Keynesian multipliers, financial accelerators and other GE feedbacks intensify.

This provides a rationale for why experimentations with target-focused communication during the Great Recession, such as the Fed's "unemployment target" in their December 2012 policy announcement or ECB President Mario Draghi's famous "whatever it takes" speech, may have been timely—and why similar strategies may be appropriate in the Covid-19 context as well. These two episodes share the following common threads: intensified GE feedbacks and a lack of comparable prior experiences, which could serve as learning foundations for rational expectations. These are precisely the conditions that, under the lens of our analysis, call for policy commitments that "shine the spotlight on unemployment."

Framework and ZLB application. We use an abstract, minimalistic model to convey the main insights as transparently as possible. But we also show how to nest a New Keynesian economy in a liquidity trap.

In our abstract model, a policymaker interacts with a continuum of small private agents. Each agent's optimal action is an increasing function of their expectations of a policy instrument, τ, and an economic outcome, Y. The latter, in turn, depends on the agents' average action. Together, these relations yield a feedback loop between the targetted outcome and the agents' behavior; this stylizes GE feedback. The policymaker's objective is to minimize the gaps of τ and Y from their first-best counterparts. Finally, the question of interest is whether the best way to achieve this goal is to communicate a commitment to a certain value for τ ("instrument communication") or a certain target for Y ("target communication").

This translates to our main application as follows. The economy is in a liquidity trap. The central bank would like to stimulate aggregate demand but cannot do so by conventional means—i.e., by lowering the current interest rate—because the zero lower bound (ZLB) is binding. Instead, it can only offer a policy commitment about the future. In this context, the agents are consumers deciding how much to spend during the liquidity trap; τ corresponds to the extent that monetary policy will remain lax after the ZLB has ceased to bind; Y is an appropriate measure of aggregate employment or income; and the question of interest is whether the central bank's communication strategy should aim at anchoring the agents' expectations of τ ("keep rates low until 2014") or their expectations of Y ("keep rates low as long as it takes for unemployment to fall below 6.5%").[1]

REE and beyond. Each communication strategy anchors agents' expectations of one object but leaves them to reason the implications for the other object. This is true even when agents are fully rational. But in this fully rational case, agents can flawlessly reason back and forth between τ and Y, or between the extent of monetary loosening and the stimulation of aggregate employment, implying that the policymaker faces no meaningful trade off between the two strategies. Formally, we show that, under rational expectation equilibrium (REE), the implementable combinations of τ and Y are the same under both strategies.[2]

We depart from this benchmark by letting agents have limited depth of knowledge and/or rationality.[3] Such a friction is consistent with the kind of "shallow" higher-order reasoning observed in laboratory experiments (Nagel, 1995; Crawford, Costa-Gomes and Iriberri, 2013). It is also the core common element of a recent theoretical literature upon which we build (Angeletos and Lian, 2018; Farhi and Werning, 2019; Gabaix, 2020; Garcıa-Schmidt and Woodford, 2019). But whereas this literature has restricted the policymaker to instrument-based forward guidance, here we study how the policymaker can regulate the bite of the assumed friction on implementability and welfare by switching to target-based forward guidance.

Main results. Our main results are stated as Theorems 1 and 2. The first revisits implementability away from the REE benchmark. The second characterizes the optimal communication strategy.

Theorem 1 includes three points. First, the sets of τ and Y that can be implemented differ between the two communication strategies. Second, the distance of either set from the rational-expectations counterpart increases with the shallowness of knowledge and rationality. And third, this distance increases with the GE feedback under instrument communication and decreases with it under target communication.

The first two points formalize the idea that the policymaker's choice of whether to anchor the public's expectations of τ or its expectations of Y becomes consequential once we depart from the REE benchmark, and the more so the larger the departure. The last point highlights the differential effect of the GE feedback and holds the key to Theorem 2. To prove these points, we show how each of the policymaker's strategies induces a different game among the public and study the effect of the belief imperfection in each of them.

Under instrument communication, agents play a game of strategic complements: conditional on an path for interest rates, an agent that expects others to spend more also expects higher aggregate income,

[1] What about the alternative strategy of committing to a target for inflation (π), as recommended by Krugman (1998)? Similarly to Eggertsson and Woodford (2003) and the related literature on forward guidance we cite later, our New Keynesian application allows no meaningful distinction between a Y target and a π target, because these two objects are tightly connected to each other via a Phillips curve that lacks any "noise" (i.e., shocks that disentangle inflation from the output gap). But as discussed in Section 6.4, the accommodation of such noise, as well as another empirical consideration, seems to favor an output commitment of the kind we focus on in this paper over the kind of inflation commitment recommended by Krugman (1998).

[2] This irrelevance result is closely related to the equivalence of primal and dual formulations of policy problems in the Ramsey literature (Chari and Kehoe, 1999; Lucas and Stokey, 1983). And it depends not only on rational expectations but also on the uniqueness of the equilibrium in the game(s) played by the agents. We will clarify this point in Sections 3 and 4.2.

[3] "Depth of knowledge" relates to what agents think others believe (and so on, to higher orders). "Depth of rationality" refers to whether agents think others are rational (and so on, to higher orders). Sections 3.3 and 5.3 cover these issues in detail.

so she is willing to spend more herself. The degree of strategic complementarity increases when spending is more sensitive to income, or when the Keynesian cross is steeper.[4]

Under target communication, everything flips. Agents now play a game of strategic substitutability: conditional on a target for aggregate income, an agent that expects others to spend more also expects tighter monetary policy, which reduces the incentive to spend. And because a steeper Keynesian cross maps to a smaller dependance of spending to expectations of interest rates (via which the expectations of others now enter) relative to expectations of income (which are themselves anchored by the policymaker), a steeper Keynesian cross also maps to a lower substitutability in this game.

Why are these game-theoretic observations important? For any finite depth of knowledge and rationality, the deviation of actual behavior from its rational-expectations counterpart increases with the absolute magnitude of the strategic interaction: the more agents care about the behavior of others, the larger the footprint on their own behavior of any mistakes in their reasoning about others. The above observations thus translate as follows: a steeper Keynesian cross increases the deviation from rational expectations under instrument communication and decreases it under target communication.

This sums up the logic behind Theorem 1. And along with the assumption that the REE outcome is efficient, it yields Theorem 2: target communication, or "shining the spotlight on unemployment," is optimal if and only if the Keynesian multiplier or other GE feedback is large enough.

Robustness. The aforementioned assumption that the REE outcome is efficient is conceptually appealing because it isolates bounded rationality as the only source of distortion. But it stretches our ZLB application. In that context, it makes more sense to let the REE outcome be inefficiently low. We explain how this enriches the optimal communication strategy without however upsetting our main lesson.

Our (and the related literature's) preferred departure from REE amounts to having agents systematically under-estimate the responses of others. Assuming the opposite bias, or a form of over-extrapolation, flips the *sign* of the distortion of behavior under both communication strategies. But it does not upset the comparative static of its *magnitude* with respect to the strength of the GE feedback. It follows that Theorem 2 is robust to both kinds of bias. Similarly, a policymaker who suspects that the public "does not fully understand GE" but is not sure of the precise mis-specification thereof could still apply our main lesson.

At the same time, our insights hinge on a departure from full rationality as opposed to pure "noise" or rational inattention. In particular, if we allow agents to observe noisy signals of the policy communications (as, e.g., in Morris and Shin, 2002) but maintain REE, we also maintain the irrelevance of the form of forward guidance for implementability.

Last but not least, our lessons are robust to introducing measurement error, policy trembles, and uncertain fundamentals. These elements, which are the focus of the classics by Poole (1970) and Weitzman (1974), naturally enter the costs and benefits of different policy options. But unlike our approach they do not tie the optimal choice to the relative importance of PE and GE effects.

[4]The discussion here concentrates on the Keynesian cross, but as explained in Section 6 there is an additional GE forces working in the same direction: the feedback between aggregate spending, inflation and real interest rates. Accordingly, the degree of strategic complementarity increases not only with the slope of the Keynesian cross but also with the slope of the Phillips curve.

Discussion and related literature. As mentioned in the beginning of the Introduction, the existing literature on the optimal choice of instruments and targets emphasizes three issues: controllability (or tightness); accountability; and state-contingency. The first refers to the minimization of the "trembles" in the policymaker's hand. The second refers to the alleviation of the type of time-inconsistency problems first highlighted in Kydland and Prescott (1982) and Barro and Gordon (1983). The third refers to the generic necessity of having policy vary with the shocks hitting the economy. See Atkeson, Chari and Kehoe (2007) for a sharp treatment of these issues and Friedman (1990) for an earlier review.

Optimal policy is also generally state-contingent. Insofar as the desirable contingencies can be explicitly articulated (as typically assumed in the Ramsey literature), it suffices for optimality to specify instruments as a functions of exogenous shocks. Otherwise, conditioning instruments on endogenous outcomes may help replicate the missing contingencies. This replication logic, which is the common core of Poole (1970), Weitzman (1974), and the DSGE literature on optimal policy rules, blurs the distinction between instrument- and target-based policies. But it is orthogonal to the logic behind our own results.

A different argument for making instruments contingent on outcomes is to aid equilibrium selection. This relates to the Taylor principle for monetary policy and to the issues discussed in Atkeson, Chari and Kehoe (2010) and Bassetto (2002). None of these considerations are relevant here because, by design, the equilibrium is unique in our setting under both instrument- and target-based policies.

Athey, Atkeson and Kehoe (2005) shift the focus to a trade off between commitment and flexibility. In their setting, tying the monetary authority's hands avoids the familiar time-inconsistency problem at the expense of preventing it from acting on valuable private information about the economy. The optimal policy turns out to be a cap on inflation, which could be read as a target-based policy.[5] Similar trade offs are studied by Amador, Werning and Angeletos (2006), Amador and Bagwell (2013) and Halac and Yared (2018), albeit in different contexts. The core element of all these papers is the interplay of private information and time inconsistency. That of our paper, instead, is the departure from REE.

Similar points distinguish our paper from the literature on policy communication spurred by Morris and Shin (2002).[6] We share this literature's emphasis on higher-order beliefs but drop rational expectations. We also change the meaning of policy communication, from signaling about exogenous fundamentals to regulation of the private agents' strategic interaction via different policy commitments.[7]

Caballero and Simsek (2019) study an economy where, from the policymaker's perspective, private agents have wrong beliefs about a fundamental but correct beliefs about others' behavior. This precludes the kind of flawed GE reasoning and the trade off between instruments- and target-based policies that our paper focuses on. But it shares the theme of finding a policy that persuades the public to do the right thing despite its wrong beliefs. Similar points apply to Hansen and Sargent (2007) and Woodford (2010, 2013).

[5] However, a cap on inflation implements the same outcomes as as a cap on the underlying policy instrument (money growth), so the distinction between instrument- and target-based policy is rather tenuous.

[6] See, e.g., Amador and Weill (2010), Angeletos and Pavan (2007, 2009), Chahrour (2014), and Myatt and Wallace (2012).

[7] In this context, three papers deserve special mention. Angeletos and Pavan (2009) allows a policymaker to regulate the agents' strategic interaction but maintain rational expectations and focus, instead, on the use and the aggregation of information. Cornand and Heinemann (2015) introduces Level-k Thinking but abstract from policy and focus, instead, on the social value of information. Finally, Bassetto (2019) emphasizes the interaction of signaling with commitment.

Last but not least, our paper adds to the literature on the "forward guidance puzzle."[8] Del Negro, Giannoni and Patterson (2015), McKay, Nakamura and Steinsson (2016) and Kaplan, Moll and Violante (2018) have argued that the puzzle is eased by accommodating finite horizons and liquidity constraints. These works maintain rational expectations and the associated irrelevance of instruments versus targets. But the kind of frictions they emphasize map to stronger GE feedbacks, which under the lens of our analysis can favor target-based guidance. Angeletos and Lian (2018), Farhi and Werning (2019), Garcıa-Schmidt and Woodford (2019), Gabaix (2020) and Wiederholt (2016), on the other hand, shift the focus to a belief friction like that captured in our preferred specification. But as previously mentioned, they study only instrument-based guidance. We, instead, highlight that a switch to target-based guidance can ease or even flip the distortion. And we provide a gauge for when such a switch is optimal.

Outline. Section 2 introduces the model. Section 3 describes the relevant REE benchmark, in which the instruments versus targets distinction is irrelevant, and introduces our main specification of bounded rationality. Section 4 presents our main results about implementation and optimal policy. Section 5 explores the robustness of the results and Section 6 translates them to the ZLB context. Section 7 relates our theoretical contribution to the practical experience of the US Federal Reserve in the 2007-2009 financial crisis and immediate aftermath. Section 8 concludes. The Appendix contains proofs omitted from the main text.

2 Model

The economy is populated by a continuum of private agents, indexed by $i \in [0,1]$, and a policymaker. Each private agent chooses an action $k_i \in \mathbb{R}$, the average of which is denoted by $K \equiv \int k_i \, \mathrm{d}i$. The policymaker controls a policy instrument $\tau \in \mathbb{R}$ and is interested in manipulating an aggregate outcome $Y \in \mathbb{R}$.

The workings of the economy are described by two key equations. The first relates the aggregate outcome to the policy instrument and the aggregate behavior of the agents:

$$Y = (1-\alpha)\tau + \alpha K, \tag{1}$$

where $\alpha \in (0,1)$ parameterizes how much of the effect of τ on Y is channeled through the agents' behavior instead of being direct, or mechanical. The second describes the optimal behavior of the typical agent as a function of her expectations of the policy and the outcome:

$$k_i = (1-\gamma)\mathbb{E}_i[\tau] + \gamma \mathbb{E}_i[Y] \tag{2}$$

where $\mathbb{E}_i$ denotes the subjective (possibly non-rational) expectation of agent i and $\gamma < 1$ parameterizes how much private incentives depend on expectations of the aggregate outcome and thereby on the choices of others. In this sense, γ parameterizes the GE feedback.

Interpretation. Our primary application is forward guidance by a central bank during a liquidity trap. In Section 6.1, after presenting our theoretical results, we will spell out the micro-foundations of this application and its mapping to our abstract model. Here, we briefly preview the main ideas to provide context.

[8] This refers to the implausibly large effects that the basic New Keynesian model predicts for forward guidance at the ZLB.

Prices are sticky, a shock has pushed the natural rate of interest into negative territory, and the ZLB is binding. K is aggregate spending during the liquidity trap, Y is aggregate income during and after the trap, and τ is the extent of monetary loosening after the trap. The anticipation of such loosening stimulates spending during the trap through both a partial equilibrium (PE) effect and two general equilibrium (GE) effects. The PE effect captures the impact of lower interest rates on individual spending, holding aggregate income and inflation constant. The two GE effects correspond to the equilibrium response of, respectively, aggregate income and inflation.

Because aggregate income and inflation are tied together via an error-free Phillips curve,[9] there is no need to track inflation explicitly. Instead, the combination of the two GE effects can be represented by a positive dependence of K on Y, as in equation (2). Accordingly, γ is necessarily positive and is also increasing in the following deep parameters: the marginal propensity to consume (MPC), or the slope of the Keynesian cross; and the degree of price flexibility, or the slope of the Phillips curve.

In a second application, described in Online Appendix A, we consider tax policy in a Neoclassical environment. In this context, K is aggregate investment today, Y is aggregate output tomorrow, τ is the negative of future taxation, and γ encapsulates two conflicting GE forces: a real, aggregate demand externality as in Dixit and Stiglitz (1977), and competition for a scarce resource (labor). This example illustrates how a substantially different, flexible-price mechanism could generate the basic structure of equations (1) and (2). It also allows for γ to be either positive or negative, depending on which of the aforementioned GE forces dominates.

Both of the above applications center on consumers and firms instead of financial markets. The latter are certainly more attentive to the fine details of policy communications but could still be subject to bounded rationality. In this context, K could be an aggregate measure of financial trades, or an asset price, which both depends on and feeds into the real economy.[10]

Parametric restrictions. The main analysis restricts to $\gamma > 0$, or positive GE feedback, which is the relevant case for our ZLB application. But as discussed in Section 5.1, our main result extends to $\gamma < 0$. We also require that $\alpha < \frac{1}{2-\gamma}$. This restriction guarantees that the equilibrium is unique in our setting, thus bypassing the equilibrium selection issues that relate to the Taylor principle and are the subject of, inter alia, Atkeson, Chari and Kehoe (2010). More fundamentally, this restriction is necessary and sufficient for behavior not to be unduly sensitive to beliefs of *infinite* order. Without it, the model is ill-behaved: the REE cannot be obtained from iteration of best responses, and tiny relaxations in the agents' depth of knowledge or depth rationality can have arbitrarily large effects on their behavior.

Policy objective. Let $\theta \in \mathbb{R}$ be an exogenous random fundamental that determines the policymaker's ideal, or first-best, values for τ and Y. This maps to aggregate TFP in our ZLB application (Section 6),

[9] Allowing for a random error term in the Phillips curve breaks this tight relation between inflation and output without however upsetting our main lessons. See the discussion in Section 6.4.

[10] Here we have in mind the financial accelerator in Bernanke, Gertler and Gilchrist (1999) and Kiyotaki and Moore (1997), or the positive feedback loop between household wealth and aggregate demand in Caballero and Simsek (2020). But a negative feedback loop as in Caballero and Farhi (2017) could also be possible.

and to the shadow cost of taxation in our Neoclassical variant (Online Appendix A). More generally, θ is a proxy for the kind of state-contingencies that the existing policy literature emphasizes.

The policymaker minimizes the rational expectation of the following quadratic loss function:

$$L(\tau, Y, \theta) \equiv (1-\chi)\left(\tau - \tau^*(\theta)\right)^2 + \chi\left(Y - Y^*(\theta)\right)^2 \tag{3}$$

where $\tau^*(\theta)$ and $Y^*(\theta)$ denote the aforementioned ideal values and $\chi \in (0,1)$ parametrizes the relative importance of the corresponding gaps. In our ZLB application, these gaps maps to two output gaps (one for the ZLB period and another for the period of subsequent monetary loosening).[11]

For our main analysis, the following assumption is also made:

Assumption 1. $Y^*(\theta) = \tau^*(\theta) = \theta$.

As made clear in the next section, the restriction $Y^*(\theta) = \tau^*(\theta)$ amounts to letting the first best be attained under rational expectations. This restriction therefore isolates bounded rationality as the only possible source of inefficiency.[12] Conditional on this, the additional restriction $\tau^*(\theta) = \theta$ is an innocuous normalization. Finally, the assumption that θ does not appear in equations (1) and (2) is largely a simplification: it herein makes sure that there is no scope or need for informing agents about θ per se, but can be relaxed by appropriately redefining K and Y, as we indeed do in our ZLB application (Section 6.1).

Timing. Play occurs in the following three stages, indexed by $t \in \{0,1,2\}$:

0. The policymaker observes θ and, conditional on that, chooses whether to engage in "instrument communication," announcing a commitment to set $\tau = \hat{\tau}$ for the policy instrument, or "target communication," announcing a commitment to achieve $Y = \hat{Y}$ for the outcome.

1. Each agent i hears the policymaker's announcement, forms expectations (in one of the various ways described in the next section), and chooses k_i according to best-response condition (2).

2. K is publicly observed and the pair (τ, Y) is determined as follows. In the case of instrument communication, $\tau = \hat{\tau}$ and Y is given by condition (1). In the case of target communication, $Y = \hat{Y}$ and τ is adjusted so that condition (1) holds with $Y = \hat{Y}$.

This structure embeds three assumptions, which are worth emphasizing.

First, the policymaker always honors in stage 2 any promise made in stage 0. This presumes credibility and equates forward guidance with a policy commitment. The literature has referred to such commitments as "Odyssean" forward guidance (e.g., Campbell et al., 2012). Our paper is about the optimal *form* of such forward guidance.

[11] To be precise, as shown in Section 6.1, the relevant micro-founded gaps map most closely to $(\tau - \tau^*(\theta))^2$ and $(K - K^*(\theta))^2$. But this is inconsequential for our results: as shown in Section 5.2, Theorem 2 readily extends to this case.

[12] In fact, the essential assumption is only that any other distortion is *separable* from that caused by bounded rationality. The logic is similar to that in Correia, Nicolini and Teles (2008) and explain the following subtlety: in the Neoclassical application spelled out in Online Appendix A, $\tau^*(\theta)$ and $Y^*(\theta)$ correspond to a second best à la Barro (1979) and Lucas and Stokey (1983)

Second, the policymaker chooses what to say and do after observing θ. This amounts to letting policy be freely contingent on *any* relevant exogenous shock, as in the textbook Ramsey paradigm.[13] It also allows forward guidance to reveal θ to the public. But because agents do not care to know θ per se, such signaling is irrelevant. There is therefore no room for "Delphic" forward guidance, or for the "information effect" of monetary policy (Nakamura and Steinsson, 2018).

Finally, we allow the policymaker to announce a value for *either* τ or Y, but not on a pair of values for *both* of them. Such joint commitments do not make sense in the model, but for knife-edge cases, owing to the fact that τ and Y have a fixed relationship at $t = 2$ once K is pre-determined. A related rationale, shown in Online Appendix C, rules out a commitment to a value for K. What remains viable, though, is commitment to a flexible *relation* between τ and Y, namely a function f such that $\tau = f(Y;\theta)$. Online Appendix G discusses both why our main insights are robust to such more sophisticated forms of forward guidance and why the simpler forms we focus on in the main text are probably more relevant in practice.

3 Rational Expectations and Beyond

This section shows how rational expectations precludes a meaningful trade off between instrument and target communication, recasts this benchmark in terms of infinite depth of knowledge and rationality, and introduces a specific departure from this benchmark, under which we later derive our main results.

3.1 An Irrelevance Result

Say there is a representative agent, who knows the structure of the economy, observes the policy announcement, and forms rational expectations.[14] In this benchmark, $\mathbb{E}_i[\cdot] = \mathbb{E}[\cdot|\hat{X}]$ for all i, where $\mathbb{E}[\cdot|\hat{X}]$ is the common, rational expectation conditional on announcement $\hat{X}$, with $X \in \{\tau, Y\}$ depending on the form of forward guidance. As a result, $k_i = K$ for all i and condition (2) reduces to the following:

$$K = (1-\gamma)\mathbb{E}[\tau|\hat{X}] + \gamma\mathbb{E}[Y|\hat{X}]. \tag{4}$$

A rational expectations equilibrium (REE) is then defined in the usual fashion. In particular:

Definition 1. *A quadruple $(\hat{X}, \tau, K, Y)$ constitutes a REE if and only if it satisfies conditions (1) and (4) along with $\tau = \hat{X}$ in the case of instrument communication and $Y = \hat{X}$ in the case of target communication.*

What matters to the policymaker is the set of combinations of τ and Y that can be implemented under each form of forward guidance. We will later explain how implementability changes away from REE. For now, we define and characterize implementability within this benchmark.

[13] In particular, if some of the relevant shocks are realized after stage 0 but prior to implementing any announced policy commitment, and if this commitment can itself be contingent on such future shocks, then the analysis goes through. One merely has to think of forward guidance as the description of how $\hat{\tau}$ or $\hat{Y}$ will vary with the future shocks, as opposed to the announcement of a single value for $\hat{\tau}$ or $\hat{Y}$. If, instead, some of the relevant contingencies are ruled out, then Poole (1970) come into the picture. We abstract from this complication in the main analysis but return to it in Section 5.5.

[14] Under the game representations provided in Lemma 1, this maps to imposing complete information and Nash equilibrium.

Definition 2. *A pair* (τ, Y) *is implementable under instrument communication if there is an announcement* $\hat{X}$ *for the policymaker and an action* K *for the representative agent such that* $(\hat{X}, \tau, K, Y)$ *constitutes a REE.*

Denote with $\mathscr{A}_\tau^*$ and $\mathscr{A}_Y^*$ the sets of (τ, Y) that are implementable under, respectively, instrument and target communication. The policymaker's problem can be expressed as follows:

$$\min_{\mathscr{A} \in \{\mathscr{A}_\tau^*, \mathscr{A}_Y^*\}, (\tau, Y) \in \mathscr{A}} \mathbb{E}[L(\tau, Y, \theta)] \tag{5}$$

The choice $\mathscr{A} \in \{\mathscr{A}_\tau^*, \mathscr{A}_Y^*\}$ captures the choice of the optimal form of forward guidance (instrument vs target), whereas the choice $(\tau, Y) \in \mathscr{A}$ captures the optimal pair (τ, Y) implemented under the given form of forward guidance. Both of these choices are conditional on θ.

We now proceed to show that $\mathscr{A}_\tau^* = \mathscr{A}_Y^*$. Using condition (1) to compute $\mathbb{E}[Y]$ and noting that $\mathbb{E}[K] = K$ (the representative agent knows his own action), we can restate condition (4) as

$$K = (1 - \alpha\gamma)\mathbb{E}[\tau|\hat{X}] + \alpha\gamma K$$

Since $\alpha\gamma \neq 1$, this implies that, in any REE,

$$K = \mathbb{E}[\tau|\hat{X}], \qquad Y = (1-\alpha)\tau + \alpha\mathbb{E}[\tau|\hat{X}] \qquad \text{and} \qquad \mathbb{E}[Y|\hat{X}] = \mathbb{E}[\tau|\hat{X}] = K$$

These properties hold regardless of the form of forward guidance. With instrument communication, we also have $\tau = \hat{\tau} = \mathbb{E}[\tau|\hat{X}]$. It follows that, for any $\hat{\tau}$, the REE is unique and satisfies $K = Y = \tau = \hat{\tau}$. With target communication, on the other hand, we have $Y = \hat{Y} = \mathbb{E}[Y|\hat{X}]$. It follows that, for any $\hat{Y}$, the REE is unique and satisfies $K = Y = \tau = \hat{Y}$. Combining these facts, we infer that, regardless of the form of forward guidance, a pair (τ, Y) is implementable if and only if $\tau = Y$. We thus reach the following two results, which serve as benchmarks of comparison for our main analysis.

Proposition 1 (Irrelevance under REE)**.** *With rational expectations, the form of forward guidance is irrelevant for implementability:*

$$\mathscr{A}_\tau^* = \mathscr{A}_Y^* = \mathscr{A}^* \equiv \{(\tau, Y) : \tau = Y\}.$$

Combining this result with Assumption 1, we reach the following property, the first part of which verifies the very meaning of this assumption and the second highlights the relevant policy lesson.

Corollary 1. *The policymaker's ideal, or first-best, combination of* τ *and* Y *is implementable under rational expectations. Furthermore, the policymaker is indifferent between instrument and target communication: he attains* $\tau = Y = \theta$ *(and* $L = 0$*) by announcing* $\hat{\tau} = \theta$*, as well as by announcing* $\hat{Y} = \theta$*.*

That $\mathscr{A}^*$ is a linear locus with slope 1 is a simplifying feature of our environment. The relevant point is that implementability is invariant to the form of forward guidance, or to whether the policymaker commits on a value for τ or a value for Y.

3.2 Unpacking the Assumptions

The kind of "flawless" GE reasoning alluded to above, and our subsequent relaxation of it, can be formalized by recasting our REE benchmark as the combination of two assumptions. The first regards the agents' *own* rationality and awareness, and the second regards the beliefs about *others*.

Assumption 2. *Every agent is rational and attentive in the following sense: she is Bayesian, acts according to condition (2), understands that Y is determined by condition (1) and that the policymaker has full commitment and acts so as to minimize (3), and is aware of any policy communication.*

Assumption 3. *The aforementioned facts are common knowledge.*

Proposition 2. *The REE benchmark studied in the previous subsection is equivalent to the joint of Assumptions 2 and 3.*

This will become evident in Section 4.2 below, when we show how iteration of best responses converges to the REE under the present assumptions on the agents' depth of knowledge and rationality but not once we relax them. With this in mind, we next discuss what Assumptions 2 and 3 mean and how they help structure the forms of "bounded rationality" considered in the rest of the paper.

Assumption 2 imposes that, for any i, agent i's subjective beliefs and behavior satisfy the following three restrictions:

$$\mathbb{E}_i[X] = \hat{X}, \qquad \mathbb{E}_i[Y] = (1-\alpha)\mathbb{E}_i[\tau] + \alpha\mathbb{E}_i[K], \qquad \text{and} \qquad k_i = \left(1-\gamma\right)\mathbb{E}_i[\tau] + \gamma\mathbb{E}_i[Y], \tag{6}$$

where $X \in \{\tau, Y\}$ depending on the form of forward guidance. The first restriction follows from the agent's attentiveness to policy communications and his knowledge of the policymaker's commitment; the second follows from his knowledge of condition (1); the third repeats the assumed best-response condition (2).

Assumption 3, in turn, imposes that agents can reason, with full confidence and no mistakes, that the above restrictions extend from their own behavior and beliefs to the behavior and the beliefs of others, to the beliefs of others about the behavior and the beliefs of others, and so on, ad infinitum. It is such *infinite* depth of knowledge and rationality that our REE benchmark and the textbook policy paradigm alike impose—and that we instead relax by modifying Assumption 3 in the subsequent analysis.

3.3 Higher-Order Doubts

For our main analysis, we replace Assumption 3 with the following:

Assumption 4 ("Doubts about Others' Awareness")**.** *Every agent believes that all other agents are rational but only a fraction $\lambda \in [0,1)$ of them is attentive to or aware of the policy message: every i believes that, for every $j \neq i$, $\mathbb{E}_j[X] = \mathbb{E}_i[X] = \hat{X}$ with probability λ and $\mathbb{E}_j[X] = 0$ with probability $1-\lambda$, where $X \in \{\tau, Y\}$ depending on the form of forward guidance. This fact and the value of λ are common knowledge.*

Relative to Assumption 3 (which can be nested as $\lambda = 1$), this drops common knowledge of the policy communication and introduces a "crisis of confidence" about whether other agents will respond.

The precise form of Assumption 4 draws from a large literature studying lack of common knowledge in macroeconomics and finance. See Abreu and Brunnermeier (2003), Morris and Shin (1998, 2002), and Woodford (2003) for early contributions and Angeletos and Lian (2018) and Wiederholt (2016) for recent applications to the ZLB context. But whereas most of this literature confounds higher-order doubts with noisy information or rational inattention, Assumption 4 isolates the former friction and equates it with a departure from REE.[15] As further explained in Section 5.4, it is the departure from REE, not noisy information or rational inattention per se, that drives our main result.

In so doing, Assumption 4 admits an immediate re-interpretation in terms of "shallow" reasoning. A large literature documents such a phenomenon in the laboratory and accommodates it by replacing REE with Level-k Thinking.[16] The exact mapping for our setting is spelled out in Online Appendix B, but the basic idea is quite simple: doubts about others' rationality (Level-k Thinking) have nearly identical behavioral implications as doubts about others' awareness (Assumption 4). As explained in Online Appendix B, the only difference is that our formulation avoids a certain "bug" that emerges when Level-k Thinking is imported from games of strategic complementarity to games of strategic substitutability.

The experimental literature on Level-k Thinking thus provides indirect empirical support for Assumption 4. Additional support can be found in a large psychology literature that documents how people tend to think they are "better than average" in a variety of contexts (see, e.g., Alicke and Govorun, 2005). The analogue here is that the people think that others are less attentive, or less rational. Assumption 4 is also a close cousin of the form of "cognitive discounting" proposed by Gabaix (2020). Finally, as we will show shortly, Assumption 4 amounts to under-reaction of the average expectations of economic activity (K) to the relevant news (the policy message). There is ample evidence of such under-reaction in surveys of expectations (e.g., Coibion and Gorodnichenko, 2012, 2015), although it is an open question how much of it is due to a departure from full rationality, which is what matters for our purposes, as opposed to noisy information.

4 Main Results

This section contains our main results. We first show how the allowed departure from rational expectations modifies implementability under each of the two forms of forward guidance, and how it breaks the earlier irrelevance result. We then characterize the optimal communication strategy.

4.1 Implementability

With Assumption 4 in place of Assumption 3, we now revisit what pairs of (τ, Y) the policymaker can implement. Under rational expectations, these pairs were given by $\mathscr{A}^* = \{(\tau, Y) : \tau = Y\}$ regardless of the

[15] Similar disentanglements of the role of higher-order beliefs from that of noisy information and first-order beliefs have been employed in Angeletos and La'O (2009) and Angeletos, Collard and Dellas (2018), albeit for different purposes.

[16] For the development of this concept and the related experiments, see Stahl (1993), Nagel (1995) and Crawford, Costa-Gomes and Iriberri (2013). This concept has been recently imported to the New Keynesian model by Garcıa-Schmidt and Woodford (2019) and Farhi and Werning (2019); see also Iovino and Sergeyev (2019) for an application to Quantitative Easing.

communication choice (Proposition 1). With higher-order doubts, we not only break the equivalence, we also observe several economically interesting properties about the deviation from the rational expectations benchmark. These properties are summarized in the following result, which is proved in detail later.

Theorem 1 (Implementability)**.** *Let $\mathscr{A}_\tau$ and $\mathscr{A}_Y$ denote the sets of the pairs (τ, Y) that are implementable under, respectively, instrument and target communication. Then,*

$$\mathscr{A}_\tau = \left\{(\tau, Y) : \tau = \mu_\tau(\lambda,\gamma) Y\right\} \qquad \textit{and} \qquad \mathscr{A}_Y = \left\{(\tau, Y) : \tau = \mu_Y(\lambda,\gamma) Y\right\},$$

where

$$\mu_\tau(\lambda,\gamma) \equiv \left((1-\alpha) + \alpha \frac{1-\alpha\gamma}{1-\lambda\alpha\gamma}\right)^{-1} \geq 1 \quad \textit{and} \quad \mu_Y(\lambda,\gamma) \equiv \mu_Y = \frac{1-2\alpha+\alpha(1-\gamma)\lambda + a^2\gamma}{(1-\alpha)\left(1-\alpha+\alpha(1-\gamma)\lambda\right)} \leq 1$$

Moreover, the following properties hold for (μ_τ, μ_Y):

(i) $\mu_Y(\lambda,\gamma) < 1 < \mu_\tau(\lambda,\gamma)$ for any $\lambda < 1$ and $\gamma \in (0,1)$.

(ii) $\mu_\tau(\lambda,\gamma)$ decreases in λ and $\mu_Y(\lambda,\gamma)$ increases in λ for every $\gamma \in (0,1)$.

(iii) $\left|1-\mu_\tau(\lambda,\gamma)\right|$ increases with γ and $\left|1-\mu_Y(\lambda,\gamma)\right|$ decreases with γ for every $\lambda \in [0,1)$.

The frictionless benchmark is nested by $\lambda = 1$ and results in $\mu_\tau = 1 = \mu_Y$. By contrast, for any $\lambda < 1$ and $\gamma \notin \{0,1\}$, we have $\mu_Y < 1 < \mu_\tau$ and the two implementable sets cease to be the same. An immediate corollary is the following:

Corollary 2. *For any $\lambda < 1$ and $\gamma \in (0,1)$, the policymaker's first best is not implementable under either mode of communication.*

For generic loss functions, the policymaker will therefore experience a trade-off between the two forms of forward guidance, owing to the different sets of implementable outcomes. The specific trade-off that obtains under the assumed objective 3 and its optimal resolution will be characterized in Section 4.3. But first, we expand on the economics behind Theorem 1. Because this result regards only implementability, it applies *regardless* of the policy objective or welfare criterion.

A key lesson, described as points (i) and (ii) of Theorem 1, is that the same friction in beliefs has *opposite* effects on implementability under the two strategies. A larger friction *steepens* the implementability constraint under instrument communication (i.e., it raises μ_τ, the marginal change in τ needed to implement a marginal change in Y) and *flattens* it under target communication (i.e., it lowers μ_Y).

This lesson qualifies the common finding of Angeletos and Lian (2018), Farhi and Werning (2019), Gabaix (2020), Garcıa-Schmidt and Woodford (2019) and Wiederholt (2016). These works have argued that essentially the same friction as that studied here arrests the response of aggregate spending to forward guidance at the ZLB (in our language, it steepens implementability). But whereas these works restrict forward guidance to be *instrument-based*, our result clarifies that this prediction can be reversed with *target-based* forward guidance.[17]

However, what will prove important for our take-home message about optimal communication (the upcoming Theorem 2) is not only the sign of the deviation from rational expectations but also its the comparative statics with respect to the GE feedback, described as point (iii) of Theorem 1. A higher GE feedback *increases* the distortion of the implementability constraint (i.e., the distance of μ_τ or μ_Y from its rational-expectation counterpart) under instrument communication and *decreases* it under target communication.

4.2 Proof of Theorem 1: Under- vs Over-reaction, and the Role of GE Feedback

The proof of Theorem 1 builds upon several intermediate results, the combination of which comprises this paper's main economic argument. The first, and most primitive, such result is that the form of forward guidance affects the agents' strategic interaction and hence the type of reasoning they must engage in.

Consider first the case in which the policymaker announces, and commits on, a value $\hat{\tau}$ for the instrument. This anchors the agents' beliefs of τ but lets them worry what Y will be. In particular, recall that Assumption 2, which imposes individual rationality and attentiveness but allows arbitrary higher-order beliefs, yields the three restrictions given in condition (6). Now that the policymaker has anchored the agents' beliefs of τ, the first restriction becomes $\mathbb{E}_i[\tau] = \hat{\tau}$ and the remaining two reduce to

$$k_i = (1-\gamma)\hat{\tau} + \gamma\mathbb{E}_i[Y] \qquad \text{and} \qquad \mathbb{E}_i[Y] = (1-\alpha)\hat{\tau} + \alpha\mathbb{E}_i[K].$$

To determine their best actions, agents therefore need to predict Y, which is the same as predicting K, or the response of others. Moreover, for any given $\hat{\tau}$, a higher predicted K means higher predicted Y and a higher action k_i. In game-theoretic language, agents' actions are *strategic complements*. And in the language of our liquidity trap application, a consumer who is pessimistic about aggregate spending wants to spend *less*, because she understands that, for fixed nominal interest rates, lower aggregate spending translates to lower income, lower inflation, and higher real rates.

Consider next the case in which the policymaker announces, and commits on, a target $\hat{Y}$ for the outcome. In this case, $\mathbb{E}_i[Y] = \hat{Y}$ and the remaining two restrictions from condition (6) can be rewritten as

$$k_i = (1-\gamma)\mathbb{E}_i[\tau] + \gamma\hat{Y} \qquad \text{and} \qquad \mathbb{E}_i[\tau] = \tfrac{1}{1-\alpha}\hat{Y} - \tfrac{\alpha}{1-\alpha}\mathbb{E}_i[K].$$

Agents now know what the outcome will be but have to figure out the policy that will support it. Predicting τ under target communication, like predicting Y in the previous case, also boils down to predicting K. But the dependence of behavior on the beliefs of K is the opposite: for any given $\hat{Y}$, a higher expectation for K maps to a lower expectation for the value of τ that will be needed to support $\hat{Y}$, and hence to a lower

[17]The kind of higher-order doubts or bounded rationality we have captured via Assumption 4 is the *sole* friction in Farhi and Werning (2019) and Garcıa-Schmidt and Woodford (2019). But this is not the case for Angeletos and Lian (2018), Gabaix (2020) and Wiederholt (2016), which combine the relevant rigidity in higher-order beliefs with a rigidity in first-order beliefs, due the inclusion of noisy information, inattention, or "sparsity." This additional friction, which we allow for in Section 5.4, contributes toward a less effective forward guidance, or a lower response of K to an announcement $\hat{X}$, under *both* communication strategies. But it does not upset our main lessons about their *relative* merits.

action k_i. In game-theoretic language, agents' actions are *strategic substitutes*. And in the language of the liquidity trap, a consumer who is pessimistic about aggregate spending wants to spend *more*, because she understands that the policymaker's commitment to deliver the announced income or employment target will necessitate a more lax monetary policy, or lower interest rates when others spend less.

The following Lemma summarizes the previous points and spells out the precise form of the game played by the agents under the two forms of forward guidance.

Lemma 1 (Game representation)**.** *Say the policymaker announces $X = \hat{X}$ for either $X \in \{\tau, Y\}$. Agents' behavior is given by*

$$k_i = (1-\delta_X)\hat{X} + \delta_X \mathbb{E}_i[K]. \tag{7}$$

where

$$\delta_\tau \equiv \alpha\gamma \in (0,1) \qquad \delta_Y \equiv -(1-\gamma)\frac{\alpha}{1-\alpha} \in (-1,0)$$

The game induced by instrument communication therefore features strategic complementarity, and the game induced by target communication features strategic substitutability.

This insight is true, and Lemma 1 holds, with rational expectations as well. But simple algebra in (7) reveals that the exact value of $\delta_X \in (-1,1)$ is irrelevant for determining the mapping of $\hat{X}$ to K *if and only if* equilibrium expectations are correct or, at least on average, $\mathbb{E}_i[K] = K$.[18]

The important deviation in Assumption 4 is to break this irrelevance in a structured way. The following Lemma demonstrates exactly how this assumption functions:

Lemma 2 (Under-estimating the response of others)**.** *For both modes of communication and for any value $\hat{X}$ of the policy message, $\mathbb{E}_i[K] = \bar{\mathbb{E}}_i[K] = \lambda K$.*

A heuristic argument is the following. If the typical agent believes that only a fraction λ of the population is aware of the policy message like herself, she also expects the same fraction to respond like herself, and the remaining fraction to stay put. That is, $\mathbb{E}_i[K] = \lambda k_i$ for the typical agent and therefore also $\bar{\mathbb{E}}[K] = \lambda K$ in the aggregate. The more precise proof offered in the Appendix demonstrates the formal connection to iteration of higher-order beliefs and makes clearer the intuitive relationship with limited depth of knowledge and rationality.

Combining Lemmas 1 and 2 pins down the behavior of K under instrument and target communication and gets to the heart of the difference between the two methods:

Lemma 3 (Under- vs over-reaction, and the effect of γ)**.** *The realized aggregate action following announcement $\hat{X}$ is given by*

$$K = \kappa_X \hat{X} \qquad \textit{with} \qquad \kappa_X(\lambda,\gamma) \equiv \frac{1-\delta_X}{1-\lambda\delta_X}, \tag{8}$$

[18] The restriction $\delta_X \in (-1,1)$ means that the equilibrium of both games can be obtained via iteration of best responses for any $\lambda \leq 1$, or that beliefs of arbitrarily high order do not have an explosive impact on behavior. Without this restriction, the REE outcome itself is extremely fragile. For instance, Level-k thinking fails to recover it in the limit as $k \to \infty$. This circles back to our discussion of how our framework guarantees, not only a unique equilibrium, but also a vanishing effect of infinite-order beliefs. See also Lemma D.1 in the Online Appendix for the calculation of why $\delta_Y > -1$ maps to $\alpha < \frac{1}{2-\gamma}$.

where $X \in \{\tau, Y\}$ depending on the form of forward guidance and $\kappa_\tau \leq 1 \leq \kappa_Y$. Moreover, the following properties hold for (κ_τ, κ_Y):

(i) $\kappa_\tau < 1 < \kappa_Y$ *for any* $\lambda < 1$ *and* $\gamma \in (0,1)$

(ii) $\kappa_\tau(\lambda,\gamma)$ *increases and* $\kappa_Y(\lambda,\gamma)$ *decreases in* λ *for every* $\gamma \in (0,1)$.

(iii) $\left|1-\kappa_\tau(\lambda,\gamma)\right|$ *increases with* γ *and* $\left|1-\kappa_Y(\lambda,\gamma)\right|$ *decreases with* γ *for every* $\lambda \in [0,1)$.

This result parallels, and basically proves, Theorem 1. Point (i) shows that instrument and target communication result in *opposite* distortions relative to the rational expectations case: the former leads K to *under-react* to the announcement, while the latter leads K to *over-react* to the announcement. This is a direct consequence of the previous discussion of strategic interaction. Point (ii) complements part (i) by showing that a larger friction amplifies the distortion in both cases, increasing under-reaction under instrument communication and increasing over-reaction under target communication. Finally, point (iii) studies how the distortion under each communication method varies with the GE feedback parameter γ. This point, which is crucial for the upcoming characterization of the optimal communication strategy, relates to the mapping from the primitive parameter γ to the strategic interaction parameters $\{\delta_\tau, \delta_Y\}$ and thereby to the role of higher-order beliefs.

Under instrument communication, a higher γ maps to a larger degree of strategic complementarity, or a more positive value for δ_τ. In the ZLB context, for example, a higher γ may correspond to a steeper Keynesian cross, and hence to a larger feedback from aggregate spending to individual spending for given interest rates, or a higher δ_τ. As this happens, any given under-estimation of the response of others' consumption (and hence of aggregate income) results in a larger reduction in individual spending. This maps to a lower κ_τ, or equivalently to a larger deviation of κ_τ from its REE counterpart.

By contrast, with target communication, a higher γ maps to a lower degree of strategic substitutability, or to a less negative value for δ_Y. To understand this, recall that under target communication the role of forecasting K is to forecast the future τthat will support the Y target. As γ increases, expected policy matters less for decisions, and so does the expected response to forward guidance via K. As such, there is less opportunity for the friction to bite. That is, κ_Y gets closer to its REE counterpart as γ increases.

Proving Theorem 1 from this point requires only the following few additional lines of algebra. In the case of instrument communication, replacing $K = \kappa_\tau \hat{\tau}$ in (1) gives $Y = (1-\alpha+\alpha\kappa_\tau)\hat{\tau}$, which together with $\tau = \hat{\tau}$ yields the implementability constraint $\tau = \mu_\tau Y$ with $\mu_\tau = \frac{1}{1-\alpha+\alpha\kappa_\tau}$. In the case of target communication, on the other hand, replacing $K = \kappa_Y \hat{Y}$ in (1) and solving for τ gives $\tau = \frac{1-\alpha\kappa_Y}{1-\alpha}\hat{Y}$, which together with $Y = \hat{Y}$ yields the implementability constraint $\tau = \mu_Y Y$ with $\mu_Y = \frac{1-\alpha\kappa_Y}{1-\alpha}$. The properties of (μ_τ, μ_Y) then follow directly from the properties of (κ_τ, κ_Y).

4.3 Optimal Policy

We now turn to optimal policy. In particular, we show how the optimal choice between instrument and target communication hinges on γ, or the ferocity of GE feedback. We think of this result as a gauge for when, as a function of economic circumstances proxied by γ, a policymaker should prefer one form of forward guidance to the other.

As a prelude to this result, it is useful to consider two extreme cases: $\gamma = 0$ and $\gamma = 1$. When $\gamma = 0$, condition (2) reduces to $k_i = \mathbb{E}_i[\tau]$, which means that agents care to know only τ. When the policymaker commits on a value for τ, she tells agents everything they need to know, eliminates their need to reason about the behavior of others, and neutralizes the bite of the friction on implementability (formally, $\mathscr{A}_\tau(0,\lambda) = \mathscr{A}^*$ for all $\lambda < 1$). By contrast, if the policymaker commits on a target for Y, agents must reason what K will be in order to figure out the value of τ that will support the announced target for Y, and the mistakes in such reasoning distort implementability (formally, $\mathscr{A}_Y(0,\lambda) \neq \mathscr{A}^*$ for all $\lambda < 1$). It follows that instrument communication is strictly optimal when $\gamma = 0$. When $\gamma = 1$, everything flips. In this case, agents only care to know Y and the only way to insulate the economy from belief friction is to offer target-based forward guidance.

These two cases are knife-edge in the sense that, as anticipated in Corollary 2, the first best is unattainable once $\gamma \in (0,1)$. But they illustrate the basic logic behind our main policy lesson: the optimal form of forward guidance aims at minimizing, as much as possible, the agents' need to reason about the economy. Building on the comparative statics of the implementability constraints with respect to γ documented in Theorem 1, we can indeed show that this logic extends to the general case as follows.

Theorem 2 (Optimal Forward Guidance)**.** *For any $\lambda < 1$, there exists a threshold $\hat{\gamma} \in (0,1)$ such that: when $\gamma \in (0,\hat{\gamma})$, instrument-based guidance is strictly optimal for all θ; and when $\gamma \in (\hat{\gamma},1)$, target-based guidance is strictly optimal for all θ.*

4.4 Proof of Theorem 2

Given θ, the policymaker chooses a set $\mathscr{A} \in \{\mathscr{A}_\tau(\lambda), \mathscr{A}_Y(\lambda)\}$ and a pair $(\tau, Y) \in \mathscr{A}$ to minimize her loss:

$$\min_{\mathscr{A} \in \{\mathscr{A}_\tau(\lambda), \mathscr{A}_Y(\lambda)\}, (\tau,Y)\in\mathscr{A}} L(\tau, Y, \theta)$$

where $L(\tau,Y,\theta) = \chi(\tau-\theta)^2 - (1-\chi)(Y-\theta)^2$. We focus on $\lambda < 1$ and $\gamma \in (0,1)$, and let $(\mathscr{A}^{\text{sb}}, \tau^{\text{sb}}, Y^{\text{sb}})$ be the unique "second-best" triplet that attains the minimum.

Given the specification of L and the characterization of the implementability sets in Theorem 1, we can restate the choice of the form of forward guidance as the choice of a slope $\mu \in \{\mu_\tau(\lambda,\gamma), \mu_Y(\lambda,\gamma)\}$ for the equilibrium mapping between τ and Y. Letting $r \equiv \tau/\theta$ and substituting the implementability constraint, we reach the following simpler representation of the policymaker's problem:[19]

$$\min_{\mu\in\{\mu_\tau(\lambda,\gamma),\mu_Y(\lambda,\gamma)\}, r\in\mathbb{R}} \left[(1-\chi)(r-1)^2 + \chi(r\mu^{-1}-1)^2\right] \tag{9}$$

This makes clear that the optimal form of forward guidance is the same for all realizations of θ. It also lets r identify the optimal covariation of τ with θ.

[19]The expression for the optimal r is $r^*(\mu) = \frac{\mu^2(1-\chi)+\mu\chi}{\mu^2(1-\chi)+\chi}$. We can further deduce that, given $\chi \in (0,1)$, $r^*/\mu > 1$ for $\mu \in [0,1]$, $r^*/\mu < 1$ for $\mu > 1$, and $r^*/\mu = 1$ for $\mu = 1$. Further, $r > 0$ as long as $\mu > 0$.

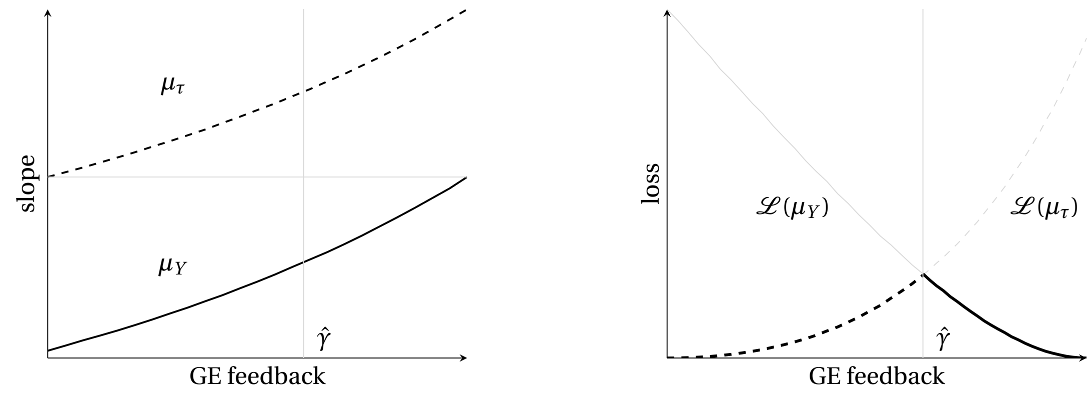

Figure 1: Slope of implementability constraint (left) and welfare loss (right) as function of GE feedback γ.

It is simple to solve for the optimal r in closed form and arrive at the following representation of the policymaker's loss as a function of μ alone:

$$\mathscr{L}(\mu) \equiv \min_{r\in\mathbb{R}} \left[(1-\chi)(r-1)^2 + \chi(r\mu^{-1}-1)^2\right] = \frac{\chi(1-\chi)(1-\mu)^2}{\mu^2(1-\chi)+\chi},$$

which is a U-shaped function of $\mu \in (0,\infty)$, with a minimum equal to 0 and attained at $\mu = 1$ (the frictionless case).[20] The interpretation of this loss function is simple. The closer μ is to 1, the smaller would be the distortion from the frictionless benchmark even if we were to hold r fixed at 1. The fact that the policymaker can adjust r as a function of μ moderates the distortion but does not upset the property that the loss is smaller the closer μ is to 1.

The optimal form of forward guidance can now be found by studying which of the two feasible values of μ yields the smallest value for $\mathscr{L}(\mu)$. Varying γ changes these two values without affecting the loss incurred from any given μ. In particular, raising γ drives μ_τ further way from 1 and brings μ_Y closer to 1 (part (iii) of Theorem 1). It follows that $\mathscr{L}(\mu_\tau)$ is an increasing function of γ, whereas $\mathscr{L}(\mu_Y)$ is a decreasing function of it.[21] Next, note that both $\mathscr{L}(\mu_\tau)$ and $\mathscr{L}(\mu_Y)$ are continuous in γ and recall from our earlier discussion about the extremes $\gamma = 0$ and $\gamma = 1$ that the following properties hold: $\mathscr{L}(\mu_\tau) = 0 < \mathscr{L}(\mu_Y)$ when $\gamma = 0$, and $\mathscr{L}(\mu_\tau) > 0 = \mathscr{L}(\mu_Y)$ when $\gamma = 1$. It follows that there exists a threshold $\hat{\gamma}$ strictly between 0 and 1 such that $\mathscr{L}(\mu_\tau) < \mathscr{L}(\mu_Y)$ for $\gamma < \hat{\gamma}$, $\mathscr{L}(\mu_\tau) = \mathscr{L}(\mu_Y)$ for $\gamma = \hat{\gamma}$, and $\mathscr{L}(\mu_\tau) > \mathscr{L}(\mu_Y)$ for $\gamma > \hat{\gamma}$.

Figure 1 illustrates this argument in a graph, with the slopes (μ_τ, μ_Y) in the left panel and the loss functions $(\mathscr{L}(\mu_\tau), \mathscr{L}(\mu_Y))$ on the right. In a nutshell, because a stronger GE feedback increases the distortion under instrument communication but reduces the distortion under target communication, target communication is optimal if and only if the GE feedback is strong enough.

[20]Note that, from the envelope theorem, $\partial\mathscr{L}/\mu = -2\cdot\chi\cdot r^*\cdot(r^*/\mu-1)/\mu^2$. Combined with the previous footnote's expression for r^*, this suggests that $\partial\mathscr{L}/\mu = 0$ when $\mu = 1$, $\partial\mathscr{L}/\partial\mu > 0$ when $\mu > 1$, and $\partial\mathscr{L}/\partial\mu < 0$ when $\mu \in [0,1]$. Finally, note that we guarantee $\mu_Y > 0$ for any possible $\lambda \leq 1$ given the restriction of $|\delta_Y| < 1$ or $\alpha < \frac{1}{2-\gamma}$.

[21]This is true strictly away from $\gamma \in \{0,1\}$.

5 Extensions

This section explores the robustness of our results to the following extensions: negative GE feedback, alternative policy objectives, alternative departures from rational expectations, the introduction of inattention, and the introduction of confounding shocks as in Poole (1970). Readers eager to see the application of our theory to the Great Recession are invited to jump to Sections 6 and 7.

5.1 Negative GE Feedback

The restriction to positive GE feedback, or $\gamma > 0$, is consistent with our main application. But the opposite scenario is possible in other contexts. For instance, in the Neoclassical example of Online Appendix A, $\gamma < 0$ is obtained if the wage pressure due to competition for labor overcomes the aggregate demand externality. And in the ZLB context, $\gamma < 0$ can obtain from competition for another scare resource, like safe assets as in Caballero and Farhi (2017). Theorem 2 directly extends to such situations. This is readily verified by noting that the proof of Theorem 2 does not actually use $\gamma > 0$.[22] We thus have that target communication is optimal if and only of if the GE feedback is *both* positive in sign and large enough in magnitude.

5.2 Alternative Policy Goals

Our main results focused on implementable pairs of (τ, Y), and their deviations from their first-best counterparts. But what if the policymaker cared also about K per se? The following result shows how to accommodate this possibility:

Proposition 3. *Let the policymaker have the loss function*

$$L = \chi_\tau(\tau-\theta)^2 + \chi_Y(Y-\theta)^2 + \chi_K(K-\theta)^2 \tag{10}$$

for some non-negative weights $\chi_\tau, \chi_Y, \chi_K$. The optimal communication strategy has a threshold form for some $\hat{\gamma} \in (0,1)$, as in Theorem 2, if at least two of the three weights are positive.

Our Neoclassical investment example, as shown in Online Appendix A, maps to $(\chi_\tau, \chi_Y) > 0$ and $\chi_K = 0$ and is hence covered directly by Theorem 2. By contrast, our liquidity trap example, as shown in Section 6.1, maps to $(\chi_\tau, \chi_K) > 0$ and $\chi_Y = 0$, hence it requires Proposition 3. Either way, the basic logic is the same. When expectations are rational, the policymaker can close all gaps at once, and can do so with both forms of forward guidance. Otherwise, a trade-off obtains. For instance, the policymaker can close the gap $\tau - \theta$ by announcing $\hat{\tau} = \theta$, but as long as $\lambda \neq 1$ this leads to *both* $K \neq \theta$ and $Y \neq \theta$. And because this kind of distortion increases with γ, a higher γ favors a switch from instruments to targets.[23]

Things become more tricky only if we relax Assumption 1, that is, if we let the REE benchmark itself deviate from the policymaker's ideal point. In such circumstances, the logic "the optimal policy aims at

[22] The proof requires only the weaker restriction $|\delta_X| < 1$, which means that beliefs of arbitrarily high order have a vanishing effect on behavior. See Online Appendix D for additional details.

[23] This holds, of course, as long as the policymaker cares about at least two of the gaps seen in (10). Otherwise, there is trivially no trade off and the form of forward guidance is indeterminate.

minimizing the bite of bounded rationality on implementability and welfare" does not necessarily hold. Instead, the following version of the generic second-best argument applies: if the distortion induced by bounded rationality happens to go in the opposite direction than another distortion, the policymaker may want to leverage on the former to offset the latter.[24]

However, the mere existence of another distortion does not necessarily upset our result. For instance, in our Neoclassical example, the REE outcome is not first-best efficient because lump-sum taxes are unavailable. Still, our result goes through because the tax distortion is invariant to bounded rationality. This guarantees that welfare can still be expressed as in (10), modulo a reinterpretation of the policymaker's ideal point: instead of representing the unconstrained first best, this point represents the kind of second best characterized in Barro (1979) and Lucas and Stokey (1983). The essence of Assumption 1 is therefore not to rule out all other distortions but rather to abstract from the possibility that the agents' deviation from full rationality is used by the policymaker as a tool for correcting other problems in the economy (as it is the case, for instance, in Gabaix, 2020, and Farhi and Gabaix, 2020).

5.3 Arbitrary Mistakes in Reasoning

Our main specification equated agents' bounded rationality to lack of common knowledge of others' awareness and rationality, which amounted to under-estimation of the responses of others. This captures the common core feature of the theoretical literature on which we build (Angeletos and Lian, 2018; Farhi and Werning, 2019; Gabaix, 2020; Garcıa-Schmidt and Woodford, 2019). But a policymaker themselves could be forgiven for not having complete confidence that these theories are correct.

Consider the following two alternative stories. The first is that individual agents will be startled and over-estimate the response of the economy to the policy news. The second is that agents' behavior will be swayed by animal spirits (extrinsic waves of optimism and pessimism about the behavior of others) or by purely random errors in their reasoning about GE effects (such as, perhaps, those caused by a misspecified belief about the structure of the economy). To capture these possibilities in a structured yet flexible manner, we now consider the following specification of the beliefs about others' behavior:

Assumption 5 (General distorted reasoning)**.** *Average beliefs satisfy* $\bar{\mathbb{E}}[K] = \lambda K + \sigma\varepsilon$ *for some* $\lambda > 0$ *(possibly* $\lambda > 1$*) and* $\sigma \geq 0$*, where* ε *is a unit-variance noise term unknown to the policymaker and independent of the policy announcement.*

Unlike the main analysis, the friction is now introduced directly in the expectations of K as opposed to in the depth of agents' knowledge and rationality. This shortcut lets us focus on how the friction matters for behavior as opposed to how it is micro-founded. But the missing details can easily be filled in.

[24] For instance, suppose that, due to a production externality or some other failure of the first welfare theorem, the rational-expectations equilibrium itself exhibits over-reaction of K to τ relative to the first best, or the policymaker's ideal point. Then, instrument communication may bring the equilibrium closer to the first best by letting the belief friction induce the opposite distortion. Furthermore, if this consideration happens to be more important when γ is large, this could overturn the comparative statics of the optimal strategy with respect to γ.

For instance, letting $\lambda > 1$ is akin to modifying the higher-order beliefs in Assumption 4 in the following way: let agents believe that with positive probability others will be "startled" and over-react to the policy message instead of being "sleepy" and under-reactive.[25] Basically the same applies to Level-k Thinking with a level-0 belief of the form $\eta\hat{X}$, for some $\eta > 1$. And just as $\lambda < 1$ captures the form of "cognitive discounting" assumed in Gabaix (2020), $\lambda > 1$ captures the opposite bias, "cognitive hyperopia" or a form of over-extrapolation (Bordalo, Gennaioli and Shleifer, 2017; Bordalo et al., 2020). As for $\sigma > 0$, this can be micro-founded by introducing either or shocks to higher-order beliefs (Angeletos and La'O, 2013; Angeletos, Collard and Dellas, 2018) or an "erratic" level-0 belief; and it can be interpreted as random errors in equilibrium reasoning, or as animal spirits operating within a unique equilibrium.

The upshot for implementable sets is the following extension of Theorem 1:

Proposition 4. *When Assumption 5 replaces Assumption 3, a pair (τ, Y) is implementable if and only if*

$$\tau = \mu_X(\lambda,\gamma)Y + \psi_X(\sigma,\gamma)\varepsilon$$

where $X \in \{\tau, Y\}$ indexes the form of forward guidance, $\mu_\tau(\lambda,\gamma)$ and $\mu_Y(\lambda,\gamma)$ are as in Theorem 1, and

$$\psi_\tau(\sigma,\gamma) \equiv -\sigma\alpha\frac{\alpha\gamma}{1-\lambda\alpha\gamma+\alpha^2\gamma(\lambda-1)} \qquad \text{and} \qquad \psi_Y(\sigma,\gamma) \equiv -\sigma\frac{\alpha^2(1-\gamma)}{(1-\alpha)((1-\alpha)+\lambda\alpha(1-\gamma))}$$

Compared to the case with under-reactive beliefs ($\lambda < 1$), the case with over-reactive beliefs ($\lambda > 1$) flips the *sign* of distortion: implementability is now flattened under instrument communication ($\mu_\tau < 1$) and steepened under target communication ($\mu_Y > 1$). Nevertheless, the comparative statics of the *size* of distortion with respect to the strength of the GE feedback remain the same: as γ increases, the distortion under instrument communication gets larger and that under target communication gets smaller.

The distortions induced by random perturbations ($\sigma > 0$) share this comparative static, too. The common mechanism is that a higher γ increases the dependence of behavior on any mistakes about Y relative to any mistakes about τ. This is true regardless of whether these mistakes are positively correlated, negatively correlated, or uncorrelated with the announcement itself.

Putting these ideas together, it is easy to show that our main policy result and the intuition about minimizing the distortion also remain for any λ and σ.

Proposition 5. *When $\lambda = 1$ and $\sigma = 0$, the optimal form of forward guidance is indeterminate. When instead $\lambda \neq 1$ and/or $\sigma \neq 0$, Theorem 2 continues to hold: there exists $\hat{\gamma} \in (0,1)$ such that target communication is optimal if and only if $\gamma > \hat{\gamma}$.*

The following corollary is then immediate:

Corollary 3 (Robustness to unknown distortions)**.** *Assume that, from the policymaker's perspective, the parameters $(\lambda,\sigma) \in \mathbb{R}_+^2$ are random and drawn from some non-degenerate prior distribution. Then Theorem 2 continues to hold.*

In this sense, a policymaker who suspects that the public has the wrong model of how policy works in GE but is not sure of the precise model thereof could still apply our main result.

[25] The only twist is that λ is no more the aforementioned probability, but rather a mixture of this probability and the perceived over-reaction of others.

5.4 Inattention

An important simplification in our model is that agents hear forward guidance perfectly clearly. This contrasts with ample evidence of inattention and compatible theories (Sims, 2003; Gabaix, 2014). We now explain why inattention per se, or noisy information, does not upset the irrelevance of the form of forward guidance that served as our starting point in Section 3—it is only the departure from rational expectations that breaks this irrelevance and that opens the door to the trade off .

Consider the simplest example of rational inattention or noisy information, with Gaussian signals. Let the fundamental θ be Gaussian with mean 0 and let the policy announcement X be linear in θ. Next, let each agent receive only a noisy version of X, given by $x_i = X + u_i$, where u_i is idiosyncratic Gaussian noise. Finally, let x_i have a fixed signal-to-noise ratio with respect to X, regardless of the form of forward guidance. This can be justified as the optimal attention choice in a model where the cost of attention is an increasing function of the Shannon mutual information between x_i and X.

In this model, which resembles Morris and Shin (2002), Woodford (2003), and the topical applications of Angeletos and Lian (2018) and Wiederholt (2016), the following properties hold under rational expectations. First, the first-order beliefs satisfy $\bar{\mathbb{E}}[X] = qX$ for some $q < 1$, with q being a positive transformation of the aforementioned signal-to-noise ratio (or of the parameter that regulates the cost of attention). Second, the higher-order beliefs satisfy $\bar{\mathbb{E}}^h[X] = \lambda^{h-1}\bar{\mathbb{E}}[X]$ for all $h \geq 2$, with $\lambda = q$. And third, the implementable combinations of τ and Y are *invariant* to the form of forward guidance.

The first two properties are commonplace in the literature. The third, which is central to our purposes and is proved in Online Appendix F, clarifies that noisy information *alone* does not upset the irrelevance property of the *noiseless* REE benchmark we studied in Section 3. A simple intuition is that, in the new context, which has noisy but still rational expectations, τ and Y are both functions of the same fundamental and these functions are themselves correctly understood by the agents. It follows that a signal of one is just as good as a signal of the other and, as a result, there is still no meaningful trade-off between anchoring the expectations of τ and anchoring the expectations of Y. We summarize this lesson below.

Proposition 6. *Insofar as expectations remain rational, the introduction of noisy information or inattention, as modeled above, preserves the irrelevance of the form of forward guidance.*

By the same token, the crucial feature of Assumption 4 was not the rigidity in higher-order beliefs per se, which is present in the above model, but rather the systematic error in equilibrium reasoning, which is absent in the above model. This squares well with the basic premise that our paper is all about systematic errors in equilibrium reasoning, as opposed to mere inattention or rational confusion.

What if inattention coexists with flawed equilibrium reasoning? We study this case in detail in Online Appendix F. The upshot is the following. If inattention is rational and efficient, in line with the microfoundations put forward in Sims (2003, 2006) and the Welfare Theorems for inattentive economies proved in Angeletos and Sastry (2019), the errors in equilibrium reasoning remain the only source of inefficiency and our main lessons (Theorem 2) go through. Otherwise, the second-best argument "use one distortion to fight another" may once again become relevant.

5.5 A Bridge to Poole (1970): Imperfect Control and Additional Shocks

Much of the contemporary discussion of instrument and targets follows the durable logic of Poole (1970): that the optimal implementation device is the one best hedged against confounding shocks. Such hedging is front-and-center in actual policy design, and while our focus in this paper is transparently different, it is useful to study whether the two justifications conflict with one another.

In Online Appendix E, we enrich our model with two Poole-like elements. The first is "uncertainty about future fundamentals," or the existence of an unobserved shock to Y in (1) on which policy cannot be contingent. Formally, we modify (1) as follows:

$$Y = (1-\alpha)\tau + \alpha K + u, \tag{11}$$

where u is Gaussian, orthogonal to θ, and unobserved by both the policymaker and the private agents. The second element is "imprecise implementation," or noisy measurement of Y or τ. Formally, we let the policymaker announce and commit to a value for $\tilde{\tau}$ or $\tilde{Y}$ (instead of, respectively, τ or Y), where

$$\tilde{\tau} = \tau + u_\tau, \qquad \tilde{Y} = Y + u_Y, \tag{12}$$

and the u's are independent Gaussian shocks, orthogonal to θ, and unpredictable by both the policymaker and the private agents. The shock u_τ may capture the policymaker's imperfect control over mortgage rates (the kind of interest rates that govern consumer spending), whereas the shock u_Y may capture measurement error in macroeconomic statistics, or other source of "noise" in the mapping from such statistics to the true outcomes of interest.[26]

Regardless of interpretation, the key assumption here is the lack of sufficiently flexible contingency of the policy on the disturbances u, u_τ and u_Y. This assumption is at the heart of Poole (1970): if the policymaker could freely condition the policy on these disturbances, the trade-off studied in that paper would disappear and our own analysis could proceed *as if* these disturbances were absent. It is therefore only the absence of such contingency that gives rise to the considerations articulated in Poole (1970).

These considerations may naturally favor one or the other form of forward guidance, regardless of whether expectations are rational or not. For instance, committing on a value for τ helps insure against shocks that cause fluctuations in aggregate output but should not influence future interest rates. But, unlike our approach, such considerations do *not* necessarily induce a dependence of the optimal policy on the relative importance of PE and GE considerations, as captured by the structural parameter γ.

Proposition 7. *Allow for the Poole-like elements described above. When expectations are rational, the optimal choice between instrument and target communication is invariant to γ. And otherwise, Theorem 2 continues to hold.*

What is more, the logic that the instrument-vs-target choice is irrelevant for implementability in the rational-expectations benchmark (but not away from it) generalizes in the following "average" sense:

[26] As an example, the Fed was very concerned in March 2014 that unemployment figures were falling toward the pre-committed 6.5% threshold for the "bad reason" that individuals were leaving the labor force, while primitive labor market conditions were not improving so much (Blinder et al., 2017).

Proposition 8. *Let $\mathbb{E}_p[\tau]$ and $\mathbb{E}_p[Y]$ be the policymaker's expectation of τ and Y at stage 0, where the expectation is taken over the possible realizations of the future shocks or measurement errors. When $\lambda = 1$, $\mathbb{E}_p[\tau] = \mathbb{E}_p[Y]$ regardless of the policymaker's strategy. When instead $\lambda \neq 1$, $\mathbb{E}_p[\tau] = \mu_\tau \mathbb{E}_p[Y]$ under instrument communication and $\mathbb{E}_p[\tau] = \mu_Y \mathbb{E}_p[Y]$ under target communication, with $\mu_\tau \neq \mu_Y$. Furthermore, μ_τ and μ_Y are exactly the same as in our main analysis.*

With rational expectations, and from the policymaker's perspective at the time she has to choose whether to commit on a value for τ or a value for Y, there continues to exist *no* trade-off in terms of how steep or flat the implementability constraint is. What varies between these two choices is only the extent of insurance provided against future shocks or measurement error. By contrast, with bounded rationality, implementability is fundamentally altered: the average relation between τ and Y depends on the form of forward guidance, and on its interaction with γ, essentially in the same way as in our main analysis.[27]

We close this section with another example of the novel considerations our approach brings to light. In our baseline analysis, we assumed that the policymaker already knew all the shocks upon which the optimal value of τ or Y should be conditioned on; and in the present extension, we relaxed this assumption but, as in Poole (1970), prevented the policymaker from making state-contingent commitments for τ or Y. Suppose now that such state-contingencies are allowed but are also hard to decipher by the agents, in the sense that the inclusion of more contingencies makes agents more prone to mistakes in equilibrium reasoning. Then, announcing and committing to a simple, non-contingent plan could be optimal for the policymaker because it offers more "clarity," or a smaller departure from rational expectations.

This reasoning favors simple, sharp, communications such as Mario Draghi's "do whatever it takes" over the kind of more complicated plans, detailed with all kinds of contingencies, found in the typical FOMC announcement. And it provides a novel rationale for curtailing state-contingencies, in addition to previously established results related to time inconsistency (Athey, Atkeson and Kehoe, 2005). The full exploration of these ideas seems an interesting angle for future research.

6 Application: Monetary Policy in a Liquidity Trap

In this section, we apply our insights to the main application of interest. We first show how to nest a micro-founded New Keynesian economy at the ZLB in our abstract framework. We then translate our main policy lesson, encapsulated in Theorem 2, into the following more practical lesson: central banks should switch from talking about interest rates to talking about unemployment when the Keynesian cross is steeper or the deflationary spiral intensifies.

[27] A corollary of Proposition 8 is that, when and only when $\lambda < 1$, the switch from instrument to target communication is associated with a reduction in the expected value of τ needed to achieve the desired target in Y. This anticipates a point we make in Section 6: the 2012 shift in the Fed's communication strategy may have, unintentionally but favorably, help shorten the time the economy had to spend at the ZLB.

6.1 Micro-foundations

Consumers and Firms. There are countably infinite periods, indexed by $t \in \{1,2,\ldots\}$, and a unit measure of households, or consumers, indexed by $i \in [0,1]$. Household i consumes $C_{i,t}$ of the good and works $N_{i,t}$ hours in period t. Let $\beta_t = \exp(-\bar{\rho} - \rho_t)$ be each consumer's subjective discount rate, parametrized by a long-run level $\bar{\rho} > 0$ and a shock ρ_t, and let $\mathbb{E}_{i,t}[\cdot]$ be an individual consumer's subjective expectation operator at t. Preferences are given by the following:

$$\mathcal{U}_{i,t} = \mathbb{E}_{i,t}\left[\left(\log C_{i,t} - \tfrac{1}{2}N_{i,t}^2\right) + \beta_t \mathcal{U}_{i,t+1}\right] \tag{13}$$

Each consumer also faces the following, standard flow budget constraint:

$$C_{i,t} + B_{i,t} = R_{t-1}\frac{P_{t-1}}{P_t}B_{i,t-1} + Y_{i,t},$$

where $Y_{i,t}$ is the consumer's income, $B_{i,t}$ is her savings in a one-period, risk-free bond, P_t is the price level, R_{t-1} is the nominal interest rate between $t-1$ and t, and $R_{t-1}\frac{P_{t-1}}{P_t}$ is the corresponding real rate.

There is also a continuum of intermediate goods firms, indexed by $j \in [0,1]$. Each such firm hires $N_{j,t}$ units of labor and produces quantity $X_{j,t} = N_{j,t}$. Aggregate output is produced by a competitive firm with technology $Y_t = e^{\theta_t}F((X_{j,t})_{j\in[0,1]})$, where θ_t is an aggregate TFP shock and $F(\cdot)$ is a standard constant-elasticity-of-substitution aggregator. The intermediate-goods firms thus face the same demand and operate the same technology. We let this symmetry extend to prices ($P_{j,t} = P_t$ for all j) but add nominal rigidity by imposing the following, *ad hoc*, backward-looking Phillips curve:

$$\Pi_t \equiv \frac{P_t}{P_{t-1}} = \left(\frac{Y_{t-1}}{Y_{t-1}^*}\right)^{\xi} \tag{14}$$

where Π_t is (one plus) the inflation rate, Y_t^* is the natural or first-best rate of output, Y_t/Y_t^* is therefore a measure of the output gap, and $\xi \in [0,1)$ is a slope parameter. Perfectly rigid prices are nested with $\xi = 0$.[28]

Fundamentals, First-best, and Liquidity Trap. We assume the following structure for the shocks (ρ_t, θ_t). For all $t \geq 3$, both the discount factor and aggregate productivity are at their steady-state values: $\rho_t = 0$ and $\theta_t = 0$. At $t = 2$, the discount factor is at its steady state, or $\rho_2 = 0$, but productivity is $\theta_2 = \theta$ for some random θ. At $t = 1$, the discount factor is weakly higher than one, or $\rho_1 = -\bar{\rho} - \Delta$, for some $\Delta \geq 0$, and productivity is the same as in period 2, or $\theta_2 = \theta_1 = \theta$.

Given these assumptions, the first-best level of output and the associated natural rate of interest are

$$\log Y_t^* = \begin{cases} \theta & \text{for } t = 1 \\ \theta & \text{for } t = 2 \\ 0 & \text{for } t \geq 3 \end{cases} \qquad \text{and} \qquad \log R_t^* = \begin{cases} -\Delta & \text{for } t = 1 \\ \bar{\rho} - \theta & \text{for } t = 2 \\ \bar{\rho} & \text{for } t \geq 3 \end{cases} \tag{15}$$

[28] Our version of the Phillips curve is purely backward looking. This contrasts with the textbook version of the New Keynesian model, in which the Phillips curve is purely forward looking. But the kind of hybrid Phillips curves that best fit the data (Galí and Gertler, 1999) and that populate the DSGE literature (Christiano, Eichenbaum and Evans, 2005) often assign a relatively greater importance to backward-looking elements. In any case, the extension of our analysis to more realistic settings, featuring richer dynamics and both backward- and forward-looking elements, is left for future work. Finally, as it will become clear momentarily, the restriction $\xi < 1$ is needed in order to make sure that the inflation-spending spiral is not explosive.

The zero-lower-bound (ZLB) constraint, on the other hand, requires $R_t \geq 1$. Without the constraint, the first best would be implemented with $R_t = R_t^*$ and zero inflation. With the constraint, the policy in (15) can still be implemented for $t \geq 2$,[29] but the ZLB necessarily binds at $t = 1$, weakly if $\Delta = 0$ and strictly if $\Delta > 0$. This situation defines a liquidity trap and motivates the study of the following policy problem.

The Policy Problem. The monetary authority is bound by the ZLB during the trap ($R_t = 1$ at $t = 1$) and is also committed to replicating flexible-price outcomes in the long run ($R_t = R_t^*$ and $Y_t = Y_t^*$ at $t \geq 3$). But it is free to lower R_2 below R_2^*, and can offer forward guidance about any such plan at $t = 1$, in an attempt to stimulate aggregate demand during the trap. Finally, the monetary authority chooses its form of forward guidance—a commitment for R_2 versus a commitment for aggregate employment and output—so as maximize the representative household's welfare.

6.2 Mapping to the Abstract Model

Let all lowercase variables be in log deviations from a steady state in which $\rho_t = \theta_t = 0$, $R_t = \exp(\bar{\rho})$ and $\Pi_t = 1$.[30] Similarly to Angeletos and Lian (2018), optimal consumption can be expressed as follows:

$$c_{it} = \mathbb{E}_{it}\left[(1-\beta)b_{it} + (1-\beta)\sum_{k\geq 0}\beta^k y_{t+k} - \sum_{k\geq 0}\beta^k(r_{t+k} - \pi_{t+k+1} - \rho_t)\right] \tag{16}$$

where $\beta \equiv \exp(-\bar{\rho}) \in (0,1)$ is the steady-state discount factor. This is the Permanent Income Hypothesis, modified to allow for a time-varying real interest rate and a discount-rate shock. Note in particular that $\sum_{k\geq 0}\beta^k y_{t+k}$ captures permanent income and $(1-\beta)$ captures the marginal propensity to consume (MPC). The following Lemma summarizes how we can use this elementary result along with market clearing and the Phillips curve (14) to derive simple expressions for aggregate income (equivalently, aggregate spending) in periods 1 and 2.

Lemma 4. *Aggregate income in periods 1 and 2 satisfy*

$$y_1 = \bar{\mathbb{E}}\left[-\beta^2 r_2 + (1-\beta+\beta\xi)(y_1+\beta y_2) + M_1\right] \qquad \textit{and} \qquad y_2 = -\frac{1}{1-\xi}r_2 + M_2, \tag{17}$$

up to constants (M_1, M_2), *which are functions of* θ, $\bar{\rho}$, *and* Δ *but are invariant to policy.*

The first equation is a modified Keynesian cross: it combines the GE feedback between income and spending with the GE feedback between inflation, real interest rates, and spending. The second equation shows that period 2 income is directly proportional to the policy instrument.

To map to the abstract model, we define τ as a rescaling of $-r_2$, the degree of policy looseness; k_i as a rescaling of c_{i1}, spending during the liquidity trap; and Y as a rescaling of $y_1 + \beta y_2$, the relevant notion of permanent income. We can then derive the following result.

[29] For the ZLB not to bind at $t = 2$, and otherwise not matter as long as $\rho_t \equiv 0$, we assume that $\theta < \bar{\rho}$ always.

[30] This rule holds with the exception of b_{it}, which is defined as the simple, linear deviation from steady state, because the steady state value of assets is zero.

Lemma 5. *The the liquidity-trap context described above maps to conditions* (1) *and* (2) *with parameters*

$$\gamma = 1 - \beta^2(1-\xi)^2 \qquad \textit{and} \qquad \alpha = \frac{1}{1+\beta(1-\xi)}. \tag{18}$$

And by implication, the degrees of strategic complementarity and substitutability in the games following instrument and target communication are given by, respectively,

$$\delta_\tau = 1 - \beta(1-\xi) \qquad \textit{and} \qquad \delta_Y = -\beta \cdot (1-\xi). \tag{19}$$

A higher MPC (lower β) and a steeper Phillips curve (higher ξ) thus both map to a more positive δ_τ and a less negative δ_Y, which via Lemma 3 translates to the following:

Lemma 6. *A higher MPC and/or a steeper Phillips curve raises* $|\kappa_\tau - 1|$*, the distortion in the response of* K *to forward guidance under instrument communication, and reduces* $|\kappa_Y - 1|$*, the corresponding distortion under target communication.*

This offers a first clue about how to translate our earlier, abstract insights to the present context. But to complete the translation, we must verify that the applicable policy objective can be represented in the way assumed in our abstract analysis.

The second-order approximation of welfare is

$$\mathscr{W} = \mathscr{W}^* - (y_1 - y_1^*)^2 - \beta(y_2 - y_2^*)^2,$$

where $\mathscr{W}^*$ is the first-best level and $y_t - y_t^*$ is the output gap in period t. Using the fact that $y_1^* = y_2^* = \theta$ and the applicable transformation of variables, we reach the following result:

Lemma 7. *The welfare losses relative to the first best can be represented as*

$$L = (1-\chi)(K - \theta - (1-\xi)\Delta)^2 + \chi(\tau - \theta)^2, \tag{20}$$

where $\chi \equiv \frac{\beta(1-\xi)^2}{1+\beta(1-\xi)^2}$ *and where, recall,* Δ *measures the distance of the natural rate of interest at* $t = 1$ *from the ZLB.*

This helps clarify the following two points. First, the policy objective obtained in (20) is nested in Proposition 3 if and only if $\Delta = 0$, or the ZLB is "weakly" binding. This is the analogue of Assumption 1 in the present context and its role is explored in the next subsection. Second, in the model considered thus far, the β behind χ is the same as the β that regulates the MPC and enters γ and α. However, if we consider an overlapping generations (OLG) extension along the lines of Del Negro, Giannoni and Patterson (2015), Farhi and Werning (2019) and in particular Angeletos and Huo (2020, Section 7), we can disentangle the two objects and interpret a high MPC, or a low β in (18) and (19) for given χ in (20), as a proxy for liquidity constraints. We adopt this interpretation throughout.

6.3 Optimal Forward Guidance at the ZLB

Building on the above results, we reach the following translation of Theorem 2:

Proposition 9. *Suppose $\lambda \neq 1$ and $\Delta = 0$, and let m measure the marginal propensity to consume. There exists a critical threshold $\hat{m}$ such that target communication is optimal if and only if $m \geq \hat{m}$. Furthermore, $\hat{m}$ decreases in ξ, the slope of the Phillips curve.*

In other words, "talking about unemployment rather than interest rates" becomes more desirable when the Keynesian cross gets steeper, as in the case of worsening credit conditions, or the deflationary spiral gets stronger. And although our model is too stylized for quantitative purposes, the following back-of-the-envelope exercise offers a useful illustration.

Interpret a period in our model as four years, so that the liquidity trap has a realistic length in light of the Great Recession; let the policymaker weigh equally the output gaps during and after the trap; let $\lambda = .75$, which amounts to assuming that 75% of the population are fully rational, level-∞ agents and the remaining are unsophisticated, level-0 agents; and finally let $\xi = 0$, which amounts to assuming completely rigid prices and unresponsive inflation. In this case, target-based forward guidance is optimal whenever the annualized MPC exceeds 0.14.

The typical estimate of the average annualized MPC in the US and other advanced countries is close to 0.30 (e.g., Jappelli and Pistaferri, 2010), suggesting that the condition $m > \hat{m}$ is likely to be satisfied in practice. What is more, previous work has shown how realistic, joint heterogeneity in the MPC and the business cycle exposure maps to a higher slope of the Keynesian cross (Patterson, 2019) and how this in turn amplifies the importance of higher-order doubts (Section 7 in Angeletos and Huo, 2020). In the light of our results, such HANK-like heterogeneity seems likely to favor target-based forward guidance.

The above back-of-the envelope exercise assumed $\xi = 0$, or perfectly rigid prices. If instead we set the slope of the Phillips curve to $\xi = 4 \times 0.1$, which translates a 1% output shortfall to an annual deflation of merely 0.1%, then the relevant threshold reduces to $\hat{m} = 0.09$. This illustrates how target-based forward guidance becomes more desirable as the deflationary spiral kicks in.

Let us now consider the role of Δ. The translation of Theorem 2 offered in Proposition 9 relies on $\Delta = 0$, which means that the ZLB is weakly binding, or equivalently that Assumption 1 holds. In the more realistic case in which the ZLB is strictly binding ($\Delta > 0$ and $R_1^* < 1$), Assumption 1 no more holds. This of course does not affect our results about implementability (Theorem 1 and Lemma 6), but enters the policymaker's calculation as follows:

Proposition 10. *For $\lambda < 1$, increasing Δ marginally from $\Delta = 0$ favors target communication if $\theta > 0$ and instrument communication if $\theta < 0$. And the opposite is true for $\lambda > 1$.*

That is, the influence of $\Delta > 0$ on the ranking between instruments and targets flips sign with the sign of the productivity shock, as well as with the direction of the departure from rational expectations.

Let us explain why. When $\Delta > 0$, the policymaker is combating an inefficient recession at $t = 1$, regardless of the value of θ. But the value of θ influences whether the policymaker would like the public to over- or under-react. When $\theta > 0$, the policymaker wants to engineer a boom at $t = 2$ and therefore has "good news" to share with the economy. And because such good news can alleviate the inefficient recession at $t = 1$, policymaker prefers on the margin that the public *over-react* to it. Target communication, in our main specification of $\lambda < 1$, fulfills this role. When instead $\theta < 0$, the policymaker would prefer for the "bad news" to not overly affect behavior at $t = 1$; and with $\lambda < 1$, this goal is achieved with instrument communication. Finally, with $\lambda > 1$, the logic about how how the sign of θ influences the desirability of over- or under-reaction remains valid, but the means of accomplishing such over- and under-reaction flip.

In short, our take-home lesson remains valid at least "on average" (i.e., across realizations of θ or parameterizations of λ). One way to formalize this idea is the following:

Corollary 4. *Suppose that the policymaker is restricted to using the same form of forward guidance across all θ and to implementing a linear relation between θ and τ (or Y) Then, Proposition 9 extends to any $\Delta > 0$, and the threshold $\hat{m}$ is invariant to Δ.*

The above restriction forces the policymaker to calculate things "on average" (in expectation over realizations of θ), balances the θ-dependent effects of $\Delta > 0$, and recovers the result for $\Delta = 0$. The same applies if we replace this restriction space with appropriate uncertainty about λ, the error in people's reasoning.

6.4 Talking About Inflation

So far we have let the monetary authority communicate a commitment for either interest rates or aggregate income. This was the most direct mapping between the application and our abstract framework. But the application raises the possibility that the policymaker tries to communicate other kinds of commitments.

Consider, in particular, Krugman (1998)'s famous recommendation that an economy should "inflate its way out of a liquidity trap." There are three possible interpretations of this strategy in our context: a commitment for π_3, a commitment for π_2, and a commitment for $\pi_2 + \beta\pi_3$. We first explain that these options do not enrich the set of implementable outcomes in our model, due to the absence of "noise" in the Phillips curve. We then argue that the accommodation of such noise, as well as an additional empirically relevant consideration, may naturally favor output commitments over inflation commitments.

By the Phillips curve (14), a commitment for π_3 is the same as a commitment for y_2; and by the second part of Lemma 4, this is also the same as a commitment for r_2, or "instrument communication." A commitment for π_2, on the other hand, is the same a commitment for y_1, or for K. But as noted earlier (and explained in more detail in Online Appendix C.1), such a commitment is not viable in our model. Finally, a commitment for cumulative inflation, $\pi_2 + \beta\pi_3$, is the same as a commitment for $y_1 + \beta y_2$, or for Y, that is, "target communication."

This verifies the claim that, in our stylized New Keynesian model, the set of implementable outcomes is not expanded by considering inflation commitments on top of output commitments. The same is true in Krugman (1998), Eggertsson and Woodford (2003), and the works on forward guidance cited earlier. And the reason is the same in all cases: the absence of unobserved cost-push shocks or other forms of "noise" in the structural relation between inflation and output gaps.

What if we accommodate such noise? Suppose, in particular, that we modify (14) as follows:

$$\Pi_t \equiv \frac{P_t}{P_{t-1}} = \left(\frac{Y_{t-1}}{Y^*_{t-1}}\right)^{\xi} \upsilon_t \tag{21}$$

where υ_t is a random shock, independently distributed over time and unrelated to θ or any other fundamental in the economy. Suppose further that a policy commitment cannot be contingent on this shock.[31] Because, with the assumed microfoundations, welfare depends only on the output gap and not on inflation per se, it is immediate that an inflation commitment is now strictly inferior to an output commitment: the two strategies are still equally good in regulating the agents' strategic interaction and the bite of bounded rationality, but only the output commitment insulates the output gap for the unwanted noise.

One may of course quibble with our assumption that υ_t is pure noise as opposed to some fundamental shock that the policymaker should care for. This circles back to our discussion of Poole (1970) and begs the question of what exactly υ_t is. Still, we contend that our treatment of υ_t as pure noise is a reasonable approximation of reality, because the bulk of the business-cycle variation in inflation appears to be unrelated to fundamentals such as productivity and labor costs.

An additional rationale for favoring output commitments over inflation commitments is the following. In the theory, an inflation commitment stimulates aggregate demand because it maps in GE to a lower real interest rate and to higher aggregate output. But in the real world, people may fail to understand these GE relations and, instead, may read a commitment for higher inflation merely as bad news about real wages and purchasing power. This possibility, which is both close to the spirit of our theoretical contribution and consistent with the empirical evidence reviewed in Candia, Coibion and Gorodnichenko (2020), reinforces the argument that central banks should talk about unemployment or output instead of inflation—at least when addressing the general public, who care primarily about unemployment and income, as opposed to bond traders and banks, who instead tend to obsess about inflation and the yield curve.

7 The Fed's Experiment with Target Communication

This paper's analysis is normative rather than positive. Still, it is useful to relate our insights to actual experience so as to further contextualize our results. We focus on forward guidance in the United States during the Great Recession and its aftermath.

The Fed's initial approach to forward guidance was based on a timeline for interest rates to remain near zero. This is typified in the August 9, 2011, statement that economic conditions merited "exceptionally low levels for the federal funds rate at least through mid-2013." In December 2012, however, the Fed sharply pivoted to a strategy that highlighted a 6.5% unemployment rate as a specific pre-requisite for increasing interest rates.[32]

[31] Otherwise, it is *as if* the shock is absent, in the sense that the aforementioned equivalence between inflation and output commitments continues to hold.

[32] The strategy was popularly dubbed the "Evans rule" in reference to one of its chief proponents, Chicago Fed President Charles Evans. For additional context and retrospective policy analysis, see reviews by Blinder et al. (2017) and Feroli et al. (2017).

While the actual statement hedged with goals for inflation and other contingencies, it is not unreasonable to call this, in our language, as a switch from instrument communication to target communication. As Alan S. Blinder (2018) wrote in a retrospective analysis, much of the public seemed only to hear the following:

> The Fed would begin to raise rates as soon as the unemployment rate dipped below 6.5 percent. Period.

What is more, this seems to have been largely expected by the Fed, if not fully intended. During the December 2012 FOMC deliberations, San Francisco Fed President John C. Williams argued:

> We should recognize we are shining a very bright spotlight on the unemployment rate. ... When we stated a specific date for lift-off, the spotlight was cast on the calendar, and that's what everyone focused on, for better or for worse. Once we start talking in terms of an unemployment threshold, it will be the unemployment rate that takes center stage, commanding all of the attention of our audience.

Minneapolis Fed President Narayana Kocherlakota similarly commented that, absent "the perfect description of a reaction function," attempting to describe more complicated contingencies would be "letting the perfect be the enemy of the good."

Our theory offers a new way to think about these issues. The Great Recession was an unusual episode during which two conditions were satisfied. First, GE feedbacks were amplified because of both the ZLB and the credit crunch. Second, economic agents were presumably more likely to make mistakes in reasoning about how the economy works, or how other may respond to policy, simply because this time was unprecedented. Under the lens of Theorem 2, such an episode is exactly the right time to "shine the spotlight on unemployment." And for the reasons alluded to at the end of Section 5.5, the attempt to describe more complicated contingencies could have backfired by confusing people and amplifying the mistakes in their equilibrium reasoning.

The proximate cause of the switch may have been uncertainty about the length and severity of the recession. But it is anyone's guess whether this was merely uncertainty about the economy's hard fundamentals, in the sense that may be best captured by a Poole (1970) model, or also uncertainty about when and how public confidence would be restored, which is the focus of our own analysis. To the extent that the latter maps in the theory to higher-order beliefs or mistakes in equilibrium reasoning, the actual reasons behind the switch may have been correlated with the issues considered in this paper, even though policymakers at the time may not have reasoned through our theory.

Did the switch work as predicted by our theory? While the data from this one episode is insufficient to test our theory, there are hints of a trade-off between anchoring instrument and target expectations, in line with our theory.

Consider first the Fed's explicit date-based guidance in late 2011 and early 2012. On the one hand, the Fed was able to move expectations about interest rates to an extent that it could not during the previous, weaker regime of communication (see, e.g., Swanson and Williams, 2014; Williams, 2016). On the other hand, Andrade et al. (2019) show in professional forecast data that the same focusing of interest rate expectations coincided with an *increase* in the dispersion of GDP forecasts. This is consistent with our theory's claim that instrument communication is successful in anchoring expectations of interest rates only at the expense of increasing the uncertainty about how the economy would respond to them.[33]

Finally, the December 2012 shift in the Fed's communication strategy coincided with a reduction in GDP forecast dispersion (see again Andrade et al., 2019) at the expense of an increase in interest-rate forecast dispersion. This is again consistent with our theory's proposition that target communication helps anchor expectations of outcomes. And it raises the possibility that the policy shift may have—perhaps unintentionally but favorably—reduced the period of time that the economy had to be stuck at the ZLB.[34]

8 Conclusion

Should a policymaker offer clarity about policy instruments, for instance by outlining a time-path for interest rates or a dollar amount for fiscal stimulus, or should she instead shine the spotlight on the relevant outcome, for instance by promising to do whatever it takes to hit a target for unemployment?

We first showed that this choice is irrelevant in a frictionless, Ramsey benchmark where the policymaker has full commitment, the relevant state-contingencies can be fully articulated, and the public is unboundedly rational. We then relaxed the last assumption, allowing people to have a shallow or distorted understanding of others' behavior and of the GE implications of policy. We explained why and how this breaks the aforementioned irrelevance and provided a new gauge for how forward guidance should be conducted in such circumstances. This gauge is summarized below:

> *The optimal strategy shifts from instrument-based to target-based forward guidance, or from anchoring expectations of interest rates to anchoring expectations of unemployment and income, as Keynesian multipliers, financial accelerators, and other GE feedbacks get larger.*

Why? Instrument-focused communication leaves the public to reason about the effect of interest rates on aggregate employment and income. Target-focused communication does the opposite, sacrificing clarity about the policy for more anchoring of the expectations of targeted outcome. A larger GE multiplier makes expectations of aggregate outcomes such as employment and income more essential for private decisions, and any mistakes thereof more detrimental for welfare, which tilts the balance toward target-focused communication.

The irrelevance result that served as our point of departure echo related results from the Ramsey literature about the equivalence of different implementations. From this perspective, a high-level contribution of our paper is to illustrate both how such results hinge on *infinite* depth of knowledge and rationality,

[33] To be precise, the kind of anchoring featured in our theory manifests in the cross section of beliefs as soon as we introduce heterogeneity in λ.

[34] This is the point we anticipated at the end of Section 5.5: the reduction in the expected value of τ needed to induce the desired expected value of Y translates, in the present context, to less monetary loosening needed *after* the economy has exited the liquidity trap, or a faster "lift-off" date, while achieving the same stimulating effect during the trap.

and how more "sophisticated" policies can regulate the bite of bounded rationality. Exploring our insights in other contexts, such as Ricardian equivalence (Barro, 1974) or the equivalence of monetary policy and taxation (Correia, Nicolini and Teles, 2008; Correia et al., 2013), is left for future work.

Our analysis focused on "Odyssean" forward guidance: we abstracted from signaling, or the "information effect" of monetary policy, and equated different communication strategies with different commitments. At the same time, our analysis abstracted from commitment problems: forward guidance was fully credible. Commitment problems can exist with rational expectations and have been the topic of a large literature. Perhaps more intriguingly, our analysis has hinted at how the private sector's bounded rationality could itself be the source of time-inconsistency. The exploration of this issue, too, is left for future work.

Another interesting direction for future research is the application of our insights to the theory of optimal policy rules, and in particular optimal Taylor rules for monetary policy. The conventional approach is based on either the idea of replicating certain state-contingencies or the logic of Poole (1970). Our approach instead highlights how different policy rules can influence the strategic interaction among the private agents and, thereby, the bite of higher-order beliefs and of related forms of bounded rationality.[35]

Finally, our paper has provided a new theoretical context for a growing empirical literature on the effectiveness of central bank communication in moving the beliefs of both experts (Campbell et al., 2012; Ehrmann et al., 2019) and the general public (Coibion, Gorodnichenko and Weber, 2019; Coibion et al., 2019). Most of this literature focuses on information revelation—think, in particular, of the communication of the central bank's forecasts about future economic conditions. We instead have focused on the commitments embedded in central bank communications and have argued that such commitments may fruitfully manage private sector expectations especially when the latter are not fully rational. We hope this perspective can inform more empirical or experimental work.

[35] A formal treatment of this idea within our framework is offered in Online Online Appendix G. The application to richer, dynamic, macroeconomic models remains to be done.

Appendix

This Appendix contains the proofs of the results in the main text, except for a few that are less crucial and are relegated to the Online Appendix.

Proposition 1, Corollary 1, Theorem 1, Corollary 2, and Lemma 1

In main text.

Lemma 2

This proof supplements the simpler argument given in the main text and reveals the role of higher-order beliefs. By iterating the best responses in the representation $K = (1-\delta_X)X + \delta_X \bar{\mathbb{E}}[K]$, provided that $|\delta_X| < 1$, we can express the expectation of K as a weighted average of the higher-order beliefs about X:

$$\bar{\mathbb{E}}[K] = \bar{\mathbb{E}}\left[(1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\bar{\mathbb{E}}^h[X]\right]. \tag{22}$$

Because we have let the typical agent believe that only a fraction λ of the other agents is aware of the policy message, her second-order beliefs satisfy

$$\mathbb{E}_i\left[\bar{\mathbb{E}}^1[X]\right] = \mathbb{E}_i\left[\mathbb{E}_j[X]\right] = \lambda\hat{X} + (1-\lambda)0 = \lambda\hat{X}.$$

By aggregation and induction, for any $h \geq 1$,

$$\bar{\mathbb{E}}^h[X] = \lambda^h \hat{X}. \tag{23}$$

Relative to the frictionless benchmark (nested here with $\lambda = 1$), higher-order beliefs are therefore more rigid (i.e., anchored to 0), and the more so the higher their order.

It follows that $\bar{\mathbb{E}}[K]$, a weighted average of higher-order beliefs, is also rigid. By direct calculation, the action K is

$$\begin{aligned} K &= (1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\bar{\mathbb{E}}^h[X] \\ &= (1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\lambda^h X \end{aligned} \tag{24}$$

while the expectation thereof is

$$\begin{aligned} K &= \bar{\mathbb{E}}\left[(1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\bar{\mathbb{E}}^h[X]\right] \\ &= (1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\lambda^{h+1}X \\ &= \lambda K \end{aligned} \tag{25}$$

The last line is the desired result.

Lemma 3

By direct calculation, using the best response $K = (1-\delta_X)X + \delta_X \bar{\mathbb{E}}[K]$ and the result from Lemma 2 that $\bar{\mathbb{E}}[K] = \lambda K$, we compute

$$\kappa_X(\lambda,\gamma) = \frac{1-\delta_X}{1-\lambda\delta_X} \tag{26}$$

for each communication method. Observe that:

1. For $\lambda < 1$, $\kappa_X > 1$ if and only if $\delta_X < 0$; otherwise, if $\delta_X > 0$, $\kappa_X < 1$.

2. The derivative of κ_X in λ is

$$\frac{\partial\kappa}{\partial\lambda} = \delta_X \frac{(1-\delta_x)}{(1-\lambda\delta_X)^2} \tag{27}$$

   which has the same sign as δ_X provided that $\delta_X < 1$, which is always satisfied.

3. The derivative of κ_X in γ is

$$\begin{aligned}\frac{\partial\kappa}{\partial\gamma} &= \frac{\partial\kappa}{\partial\delta_X}\cdot\frac{\partial\delta_X}{\partial\gamma} \\ &= -\frac{1-\lambda}{(1-\lambda\delta_X)^2}\cdot\frac{\partial\delta_X}{\partial\gamma}\end{aligned} \tag{28}$$

   which has the opposite sign as $\frac{\partial\delta_X}{\partial\gamma}$.

We now review how to apply the previous points to derive each of the results for instrument and target communication.

Instrument Communication.

1. $\delta_\tau = \alpha\gamma \in (0,1)$ provided that $\alpha \in (0,1)$ and $\gamma \in (0,1)$. Thus $\kappa_\tau < 1$.

2. By the same argument, $\frac{\partial\kappa}{\partial\lambda} < 0$.

3. $\frac{\partial\delta_X}{\partial\gamma} = \alpha > 0$ so $\frac{\partial\kappa}{\partial\gamma} < 0$. As $\kappa_\tau < 1$, this implies $|1-\kappa_\tau|$ increases in γ.

Target communication.

1. $\delta_Y = -(1-\gamma)\alpha/(1-\alpha) < 0$ for any $\alpha \in (0,1)$ and $\gamma < 1$. Thus $\kappa_Y > 1$.

2. By the same argument, $\frac{\partial\kappa}{\partial\lambda} > 0$.

3. $\frac{\partial\delta_X}{\partial\gamma} = \alpha/(1-\alpha) > 0$ so $\frac{\partial\kappa}{\partial\gamma} < 0$. As $\kappa_Y > 1$, this implies $|1-\kappa_Y|$ decreases in γ.

Proposition 3

To prove this proposition, we will directly establish the monotonicity of the loss functions $\mathscr{L}_\tau(\gamma)$ and $\mathscr{L}_Y(\gamma)$, corresponding to the loss under method τ or Y, evaluated at the optimal announcement, as a function of the underlying parameter γ for fixed values of all other parameters.

Let the implementable outcomes under each communication method take the form

$$\mathscr{B}_X = \{(\tau,K,Y) : \tau = r\theta, K = a_X\tau, Y = b_X\tau\} \tag{29}$$

for some $r \in \mathbb{R}$ and coefficients (a_X, b_X). Observe that, based on previous calculations, $a_\tau = \kappa_\tau < 1$, $b_\tau = \mu_\tau^{-1} < 1$, $a_Y = \kappa_Y \mu_Y^{-1} > 1$, and $b_Y = \mu_Y^{-1} > 1$.

For a given set of coefficients (a_X, b_X), the policy problem simplifies to choice of r:

$$\mathscr{L}_X = \max_r \chi_\tau (r-1)^2 + \chi_Y (r b_X - 1)^2 + \chi_K (r a_X - 1)^2 \tag{30}$$

The optimal choice of r solves

$$r^* = \frac{\chi_\tau + \chi_Y b_X + \chi_K a_X}{\chi_\tau + \chi_Y b_X^2 + \chi_K a_X^2} \tag{31}$$

Observe that $r^* > 1$ for instrument communication, owing to the fact that $a_\tau < 1$ and $b_\tau < 1$, and $r^* \in (0,1)$ for target communication, owing to the fact that $a_Y > 1$ and $b_Y > 1$.

The derivative of $\mathscr{L}_X$ in γ is the following, by the envelope theorem:

$$\frac{\partial \mathscr{L}_X}{\partial \gamma} = 2\chi_Y r (r b_X - 1)\frac{\partial b_X}{\partial \gamma} + 2\chi_K r (r a_X - 1)\frac{\partial a_X}{\partial \gamma} \tag{32}$$

Given that $r > 0$, observe that the sign of the above is the same as the sign of

$$\chi_Y (r b_X - 1)\frac{\partial b_X}{\partial \gamma} + \chi_K (r a_X - 1)\frac{\partial a_X}{\partial \gamma} \tag{33}$$

Instrument communication. Observe first that $\mathscr{L}_\tau(0) = 0$ as $a_\tau = b_\tau = 1$. It remains to establish $\frac{\partial \mathscr{L}_\tau}{\partial \gamma} > 0$ for $\gamma \in (0,1)$. This requires showing

$$\chi_Y (r b_\tau - 1)\frac{\partial b_\tau}{\partial \gamma} > -\chi_K (r a_\tau - 1)\frac{\partial a_\tau}{\partial \gamma} \tag{34}$$

Note that $\frac{\partial a_\tau}{\partial \gamma} = \frac{\partial \kappa_\tau}{\partial \gamma} < 0$, according to Lemma 3, that $b_\tau = (1-\alpha) + \alpha a_\tau > \alpha_\tau$, and $0 > \frac{\partial b_\tau}{\partial \gamma} = \alpha \frac{\partial \kappa_\tau}{\partial \gamma} > \frac{\partial a_\tau}{\partial \gamma}$. Hence the above simplifies to

$$\frac{\chi_K}{\chi_Y}(1 - r a_\tau) > \alpha (r b_\tau - 1) \tag{35}$$

Observe next that, since $a_\tau < b_\tau$, it is guaranteed that

$$r a_\tau = \frac{\chi_\tau + \chi_Y a_\tau b_\tau + \chi_K a_\tau^2}{\chi_\tau + \chi_Y b_\tau^2 + \chi_K a_\tau^2} < \frac{\chi_\tau + \chi_Y b_\tau^2 + \chi_K a_\tau^2}{\chi_\tau + \chi_Y b_\tau^2 + \chi_K a_\tau^2} = 1$$

and hence the right condition is

$$\frac{\chi_K}{\chi_Y} > \alpha \frac{r b_\tau - 1}{1 - r a_\tau} \tag{36}$$

An immediate sufficient condition, which helps reveal the economics, is $r b_\tau < 1$: as long as the relation between the optimal policy only partially offsets the distortion, so that Y remains less responsive to θ than under REE, the result goes through. Otherwise, the result will still be true provided the relative importance of the output gap is small. Direct calculation shows that (36) is equivalent to

$$(1 - \alpha\lambda\gamma)\chi_\tau + (1-\alpha)(1 - \alpha\gamma(\lambda + (1-\lambda)\alpha)) > 0 \tag{37}$$

which is verified in our parameter range.

Target communication. The argument for target communication is symmetric. Observe first that $\mathscr{L}_Y(1) = 0$ as $a_Y = b_Y = 1$. It remains to establish $\frac{\partial \mathscr{L}_Y}{\partial \gamma} < 0$ for $\gamma \in (0,1)$. This requires showing

$$\chi_Y(rb_\tau - 1)\frac{\partial b_\tau}{\partial \gamma} > -\chi_K(ra_\tau - 1)\frac{\partial a_\tau}{\partial \gamma} \tag{38}$$

Note that $\frac{\partial b_Y}{\partial \gamma} = -\frac{1}{\mu_Y^2}\frac{\partial \mu_Y}{\partial \gamma} > 0$, by Theorem 1. Next $a_Y = \kappa_Y \mu_Y^{-1} = \frac{\kappa_Y(1-\alpha)}{1-\alpha\kappa_Y} > 0$. By direct calculation,

$$\frac{\partial b_Y}{\partial \gamma} = \frac{\alpha(1-\alpha)}{(1-\alpha\kappa_Y)^2}\frac{\partial \kappa_Y}{\partial \gamma} \qquad \text{and} \qquad \frac{\partial a_Y}{\partial \gamma} = \frac{1-\alpha}{(1-\alpha\kappa_Y)^2}\frac{\partial \kappa_Y}{\partial \gamma} \tag{39}$$

It follows that the relevant condition is

$$\frac{\chi_K}{\chi_Y}(ra_\tau - 1) > \alpha(1 - rb_\tau) \tag{40}$$

Observe next that, since $a_Y > b_Y$, it is guaranteed that $ra_Y > 1$ and hence the right condition is, in direct analogy to (36),

$$\frac{\chi_K}{\chi_Y} > \alpha\frac{rb_Y - 1}{1 - ra_Y} \tag{41}$$

Direct calculation shows that (41) is true if and only if

$$1 + \alpha(-2 + \alpha\gamma + \lambda(1-\gamma)) > 0 \tag{42}$$

but the above is guaranteed, for all λ, by $\alpha < 1/(2-\gamma)$ which was assumed earlier to ensure $\delta_X \in (-1,1)$.

Proposition 4

First, for each of instrument and target communication, we solve the best-response fixed point

$$K = \kappa_X X + \kappa_X^\varepsilon \varepsilon = (1-\delta_X)X + \delta_X(\lambda(\kappa_X X + \kappa_X^\varepsilon \varepsilon) + \sigma\varepsilon)$$

where κ_X coincides exactly with the values given in Lemma 3 and κ_X^ε is a new loading on the belief shock given by

$$\kappa_X^\varepsilon = \frac{\sigma\delta_X}{1-\lambda\delta_X}$$

Then, to get the implementability constraints, we solve the fixed-point

$$\tau = \mu_X Y + \psi_X \varepsilon = \frac{1}{1-\alpha}(Y - \alpha(\kappa_X X + \kappa_X^\varepsilon \varepsilon))$$

Which has solutions μ_X given in Proposition 1 and ψ_X given by

$$\psi_\tau = -\kappa_\tau^\varepsilon \frac{\alpha}{1-\alpha+\alpha\kappa_\tau}$$
$$\psi_Y = -\kappa_Y^\varepsilon \frac{\alpha}{1-\alpha}$$

The expressions given in the text follow from plugging the primitive parameters into the previous expression.

Proposition 5

This is an extension of the proof of Theorem 2. Let us assume that $\lambda \in [0, 1 + 1/\alpha)$. This plays a similar role as the restriction $\alpha < \frac{2}{1-\gamma}$: it ensures convergence of iterated best-responses and thereby also that $(\mu_\tau, \mu_Y) \in \mathbb{R}^2_+$.

First, note that, using the same expressions from the proof of Proposition 1, μ_τ and μ_Y both *decrease* in γ for $\lambda > 1$. Importantly, the comparative statics in the distortion $|\mu_X - 1|$ are the same.

Second, consider the comparative statics for the variance contributions (ψ_τ, ψ_Y). Note that

$$\frac{\partial \psi_\tau}{\partial \gamma} = -\frac{\sigma \alpha^2}{(\alpha^2 \gamma(\lambda - 1) - \alpha\gamma\lambda + 1)^2} < 0$$

and $\psi_\tau \leq 0$ as long as $\lambda < 1/\alpha + 1$. Finally, $\psi_\tau = 0$ if $\gamma = 0$. Next,

$$\frac{\partial \psi_Y}{\partial \gamma} = \frac{\sigma \alpha^2}{(\alpha(1-\gamma)\lambda + 1 - \alpha)^2} > 0$$

and $\psi_Y \leq 0$ for any value of λ. Finally, $\psi_Y = 0$ if $\gamma = 1$.

Consider now the policy loss function. After substituting in the implementability constraint, the loss function under target communication for a given value of θ is given by

$$\mathscr{L}_\tau \equiv \min_{r_\tau \in \mathbb{R}} \left[\theta^2 (1-\chi)(r_\tau - 1)^2 + \theta^2 \chi (r_\tau / \mu_\tau - 1)^2 + \chi \psi_\tau^2 \sigma^2 \right]$$

where we suppress the dependence on θ for simplicity. The same problem with an unconditional expectation on θ (i.e., an ex ante choice of method) would replace θ^2 with the fundamental's variance σ_θ^2.

Let us now prove that, for fixed $(\theta, \lambda, \sigma)$, optimal policy is characterized by a threshold rule. For target communication, the appropriate translation of the loss function is

$$\mathscr{L}_Y \equiv \min_{r_Y \in \mathbb{R}} \left[\theta^2 (1-\chi)(r_Y - 1)^2 + \theta^2 \chi (r_Y / \mu_Y - 1)^2 + (1-\chi) \psi_Y^2 \sigma^2 \right]$$

where we once again suppress the dependence on θ. Via an identical argument to the one pursued in the proof of Theorem 2, each loss function is monotone in γ holding fixed the value of ψ_x. That is to say, if the loss functions were each re-written in the form $\mathscr{L}_X = \ell_X(\mu_X, \psi_X)$, then, for all θ,

$$\frac{\partial \ell_\tau}{\partial \mu_\tau} \frac{\partial \mu_\tau}{\partial \gamma} > 0 \qquad \frac{\partial \ell_Y}{\partial \mu_Y} \frac{\partial \mu_Y}{\partial \gamma} < 0 \tag{43}$$

Showing the equivalent monotonicity via the second channel is simple using the previously proven comparative static for the ψ_X:

$$\frac{\partial \ell_\tau}{\partial \psi_\tau} \frac{\partial \psi_\tau}{\partial \gamma} = 2\chi\sigma^2 \psi_\tau \frac{\partial \psi_\tau}{\partial \gamma} > 0 \qquad \frac{\partial \ell_Y}{\partial \psi_Y} \frac{\partial \psi_Y}{\partial \gamma} = 2\chi\sigma^2 \psi_Y \frac{\partial \psi_Y}{\partial \gamma} < 0 \tag{44}$$

and combining (43) and (44) is sufficient to prove

$$\frac{\partial \ell_\tau}{\partial \gamma} > 0 \qquad \frac{\partial \ell_Y}{\partial \gamma} < 0$$

Finally, note that previous argument about continuity and extreme values from the proof of Theorem 2 also carries over. In particular, for all θ,

$$\mathscr{L}_\tau|_{\gamma=0} = \mathscr{L}_Y|_{\gamma=1} = 0$$

and

$$\mathscr{L}_\tau|_{\gamma=1} \geq 0 \qquad\qquad \mathscr{L}_Y|_{\gamma=0} \geq 0$$

with the last two inequalities being strict so long as $\theta \neq 0$ or $\sigma > 0$. Hence, generically speaking (for $\theta \neq 0$ or $\sigma > 0$), there exists a $\hat{\gamma} \in (0,1)$ such that target communication is strictly optimal when $\gamma \in (\hat{\gamma}, 1]$, instrument communication is optimal when $\gamma \in [0, \hat{\gamma})$, and the policymaker is indifferent at $\gamma = \hat{\gamma}$.

Corollary 3

Let us consider the case of uncertain distortions. We assume, in this thought experiment, that the government chooses its communication *method* unconditionally and then can make its actual announcement contingenent on the realized value of (λ, σ). Note that, for every (λ, σ) and every value of $\theta \neq 0$, $\gamma \mapsto \mathscr{L}_\tau(\cdot; \lambda, \sigma)$ and $\gamma \mapsto \mathscr{L}_Y(\cdot; \lambda, \sigma)$ (the loss functions conditioned on the optimal message) have all the relevant monotonicity and limit-value properties because of Proposition 5 (see also the proof thereof). It follows that the average $\gamma \mapsto \mathbb{E}_{\pi(\lambda,\sigma)}[\mathscr{L}_X(\cdot; \lambda, \sigma)]$, where the relevant expectation is over possible values of (λ, σ) in accordance with prior π, maintains the same properties. Hence these "expected loss functions" must cross at some $\hat{\gamma} \in (0,1)$. This implies that the optimal choice of communication strategy takes the desired "threshold" form.

Proposition 6

See Online Appendix F.

Propositions 7 and 8

See Online Appendix E.

Lemmas 4, 5, 6, and 7

We prove these results together by detailing the mapping to the abstract model step-by-step.

The average consumption of agents at time t can be expressed as the following function of current and future real interest rates, income, and discount rate shocks:

$$c_t = (1-\beta) \sum_{j=0}^{\infty} \beta^j \bar{\mathbb{E}}_t[y_{t+j}] + \beta \sum_{j=0}^{\infty} \beta^j \bar{\mathbb{E}}_t[-(r_{t+j} - \pi_{t+j+1} - \rho_t)] \tag{45}$$

where $\beta \equiv \exp(-\bar{\rho})$ is the steady-state discount factor. This expression is obtained by substituting the consumer's Euler equation into his lifetime budget constraint and solving for consumption. It represents the optimal consumption in period t, as a function of the expected path of income and interest rates, and identifies $1-\beta$ with the marginal propensity to consume out of income.

Inflation is

$$\pi_t = \xi(y_{t-1} - \theta_{t-1}) \tag{46}$$

Note that $y_t = r_t = \pi_{t+1} = 0$ for $t \geq 3$ as the economy returns to steady state.

Consider first aggregate consumption and income at $t = 2$. We assume that, at this point, all agents have the same (rational) expectations. It is simple to apply forward-looking rational expectations in (45) to get

$$c_2 = -\beta r_2 + (1 - \beta + \beta\xi) y_2 - \beta\xi\theta$$

which solves, after imposing market clearing, to

$$c_2 = y_2 = -\frac{1}{1-\xi}(r_2 + \xi\theta) \tag{47}$$

where the last step uses the fact that $r_s = 0$ for all $s \geq 3$.

Next let us solve for individual consumption in period 1. In this case, (45) reduces to the following after substituting out inflation rates as a function of the relevant output gaps:

$$c_{i,1} = \mathbb{E}_i\left[(1-\beta)b_{i,1} - \beta^2 r_2 + (1-\beta+\beta\xi)(y_1 + \beta y_2)\right] + \beta(\rho + \bar{\rho}) - \beta(1+\beta)\xi\theta \tag{48}$$

Note that the first term disappears because we assume initial asset positions are symmetric, which along with the fact that aggregate assets are zero gives $b_{i1} = 0$ for all i.

We can now derive the mapping to the abstract model. Let

$$\tau \equiv -r_2 \tag{49}$$

be the negative real interest rate at $t = 2$. A higher τ corresponds to looser monetary policy after the trap. Next, let

$$K \equiv (1-\xi)^2 c + b_k$$

be a normalized measure of aggregate spending during the crisis, defined up to constant $b_k = \xi(2-\xi)\theta - (1-\xi)(\rho + \bar{\rho})$. Finally, let

$$Y \equiv \frac{1}{\beta + (1-\xi)^{-1}}\left(y_1 + y_2\right) + \frac{b_k}{\beta(1-\xi)^2 + (1-\xi)} \tag{50}$$

be a normalized measure of total output during and right after the liquidity trap.

We can re-write the definition of Y in the following form:

$$Y = \frac{\beta(1-\xi)}{1+\beta(1-\xi)}\tau + \frac{1}{1+\beta(1-\xi)}K, \tag{51}$$

which matches condition (1) in our abstract framework for

$$\alpha = \frac{1}{1+\beta(1-\xi)}. \tag{52}$$

Condition (48), on the other hand, can be written as:

$$k_i = \beta^2(1-\xi)^2\mathbb{E}_i[\tau] + (1-\beta^2(1-\xi)^2)\mathbb{E}_i[Y], \tag{53}$$

which matches condition (2) in our abstract framework for

$$\gamma = 1 - \beta^2 (1-\xi)^2. \tag{54}$$

The expressions for δ_τ and δ_Y follow from direct calculation. The implications for κ_τ and κ_Y, described in Lemma 6, follow from identical logic as the proof of Lemma 3.

The final step is to calculate the welfare objective as in Lemma 7. We approximate the policymaker's utilitarian objective around the case in which the first-best is implemented by a (possibly infeasible) monetary policy. In log deviations, this yields the following loss function, up to rescaling:

$$L = (y_1 - \theta)^2 + \beta (y_2 - \theta)^2. \tag{55}$$

Each quadratic term incorporates both the concavity of the utility function and the disutility of labor. This is because, for given θ, labor hours and output have a fixed proportion to one another.

Plugging in our definitions of K and τ yields

$$L = (K - \theta + (1-\xi)(\rho + \bar{\rho}))^2 + \beta (1-\xi)^2 (\tau - \theta)^2, \tag{56}$$

This is the same as (20), with

$$\chi \equiv \frac{\beta(1-\xi)^2}{1+\beta(1-\xi)^2} \in (0,1) \qquad \text{and} \qquad \Delta \equiv -(\rho + \bar{\rho}) \geq 0.$$

Proposition 9

Observe that

$$\delta_\tau = \alpha\gamma = 1 - \beta(1-\xi) \qquad\qquad \delta_Y \equiv -\frac{\alpha}{1-\alpha}(1-\gamma) = -\beta(1-\xi) \tag{57}$$

And since $\beta \in (0,1)$ and $\xi \in (0,1)$,

$$-1 < \delta_Y < 0 < \delta_\tau < 1.$$

As noted in the main text, both δ_τ and δ_Y decrease in β and increase in ξ. Identical arguments to those in Lemma 3, and the proof thereof, show that, under instrument communication, $K = \kappa_\tau \tau$ where

$$\kappa_\tau(\beta, \xi) = \frac{\beta(1-\xi)}{1 - \lambda + \lambda\beta(1-\xi)} \tag{58}$$

obeys the following properties for all $\lambda \in (0,1)$: $\kappa_\tau < 1$ and $|1-\kappa_\tau|$ decreases with $-\beta$ and ξ. The implementable sets under instrument communication have the form

$$\mathscr{A}_\tau = \{(\tau, K) : \tau = r\theta, K = a_\tau \tau\} \tag{59}$$

where $a_\tau = \kappa_\tau$.

Next, direct calculation shows that the implementable set under target communication is a locus

$$\mathscr{A}_Y = \{(\tau, K) : \tau = r\theta, K = a_Y \tau\} \tag{60}$$

where

$$a_Y(\beta,\xi) = \kappa_Y(\beta,\xi) \cdot \frac{1-\alpha}{1-\alpha\kappa_Y(\beta,\xi)} = \frac{1}{\lambda} \tag{61}$$

This slope therefore has no dependence on deep parameters (β,ξ).

Let us now establish the two desired properties of the optimal communication strategy. Consider first the dependence of optimal communication on the discount rate β. Let $\ell_\tau(\beta,\xi;\chi)$ denote the policymaker's loss under target communication as a function of (β,ξ), *holding fixed* the value of $\chi = \frac{\bar{\beta}(1-\bar{\xi})^2}{1+\bar{\beta}(1-\bar{\xi})^2}$ for some $\bar{\beta} \in (0,1)$ and $\bar{\xi} \in (0,1)$. This loss function can be re-written as

$$\ell_\tau(\beta,\xi;\chi) = \min_r \left[(1-\chi)(r-1)^2 + \chi(a_\tau(\beta,\xi)\cdot r - 1)^2\right] \tag{62}$$

where a_τ depends on β and ξ. As in the proof of Theorem 2, after establishing that $a_\tau(\beta,\xi) < 1$ and $a_\tau(\beta,\xi)\cdot r^* < 1$, it is straightforward to show that $\partial\ell_\tau/\partial\beta$ has the same sign as $\partial a_\tau/\partial\beta$. The latter is negative so the former is negative: loss decreases as β increases.

The analogous loss function defined for target communication,

$$\ell_Y(\beta,\xi;\chi) = \min_r \left[(1-\chi)(r-1)^2 + \chi(1/\lambda\cdot r - 1)^2\right] \tag{63}$$

is invariant as a function of β. Let $L(\chi)$ denote its level. It follows that a decrease in β strictly increases the difference $\ell_\tau(\beta;\chi) - \ell_Y(\beta;\chi) = \ell_\tau(\beta;\chi) - L(\chi)$ for any value of χ.

Toward showing the desired threshold characterization of the optimal policy, observe there are three possible cases:

1. If $\ell_\tau(1;\chi) > L(\chi)$, then $\ell_\tau(\beta;\chi) > L(\chi)$ for all $\beta \in (0,1)$. Thus target communication is optimal over the entire parameter space, the critical threshold for β is $\hat{\beta} = 1$.

2. If $\ell_\tau(0;\chi) < L(\chi)$, then $\ell_\tau(\beta;\chi) < L(\chi)$ for all $\beta \in (0,1)$. Thus instrument communication is optimal over the entire parameter space, $\hat{\beta} = 0$.

3. If neither of the previous is true, then $\ell_\tau(1;\chi) < L < \ell_\tau(0;\chi)$. Note also that and $\ell_\tau(\beta;\chi)$ is a continuous, increasing function of its first argument. It follows that there exists some threshold $\hat{\beta} \in (0,1)$ below which target communication is optimal.

Next, we establish the comparative static in ξ. Note first that, for a fixed policy objective, higher ξ favors target communication for the exact same reason described above. The log difference between loss with instrument and target communication is

$$\begin{aligned} D &\equiv \log\ell_\tau(\beta,\xi;\chi) - \log\ell_Y(\beta,\xi;\chi) \\ &= 2\log\left(\frac{1-a_\tau^{-1}}{1-a_Y^{-1}}\right) - \log\left(\frac{a_\tau^{-2}(1-\chi)+\chi}{a_Y^{-2}(1-\chi)+\chi}\right) \end{aligned} \tag{64}$$

Only the last term depends on χ, and since $a_\tau < 1$ and $a_Y > 1$, then $\frac{\partial D}{\partial\chi} > 0$. Finally, observe that $\frac{\partial\chi}{\partial\xi} > 0$, so an increase in ξ increases Δ through its effect on the policy objective holding fixed its effect on (a_τ, a_Y).

Next, observe that $\ell_\tau(\beta,\xi;\chi)$ is a strictly increasing function of ξ, holding fixed preferences, according to the same arguments used in the context of β; $\ell_Y(\beta,\xi;\chi)$ is invariant in ξ; and hence $\ell_\tau(\beta,\xi;\chi)-\ell_Y(\beta,\xi;\chi)$ strictly increases in ξ, holding fixed preferences.

The two facts together establish that the preference for target communication increases in ξ. Given the previous characterization of optimal policy as a threshold in $\hat{\beta}$, it must be the case that the aforementioned threshold is non-increasing in ξ. Finally, to complete the proof, note that a threshold in $\hat{\beta}$ can be trivially translated into a threshold in the annualized MPC $m = (1-\beta)/T$, where T is a period length.

Proposition 10

The objective function, up to terms outside the policymaker's control, can be written as

$$L = \min_r \left[\theta^2(1-\chi)(r-1)^2 + \theta^2\chi(a_X r - 1)^2 + 2\chi\theta\cdot\Delta\cdot a_X r\right] \tag{65}$$

Let L_X denote be the loss function under strategy X,

$$L_X = \theta^2(1-\chi)(r-1)^2 + \theta^2\chi(a_X r_X - 1)^2 + 2\chi\theta\cdot\Delta\cdot a_X r_X \tag{66}$$

where r_X solves the optimization in (65).

By the envelope theorem, the change in objective from a marginal change in Δ is

$$\frac{\partial L_X}{\partial \Delta} = 2\chi\theta\cdot a_X r_X \tag{67}$$

Applying the previous,

$$\frac{\partial(L_\tau - L_Y)}{\partial\Delta} = 2\chi\theta(a_\tau r_\tau - a_Y r_Y) \tag{68}$$

As established in the proof of Proposition 9, when $\Delta = 0$ and $\lambda < 1$, we have $a_\tau r_\tau < 1$ and $a_Y r_Y > 1$ for any value of $1-\beta$ or ξ, and hence $\frac{\partial(L_\tau - L_Y)}{\partial\Delta} < 0$.

This implies that when Δ marginally decreases, or $|\Delta|$ marginally increases, then

1. Losses marginally increase for instrument relative to target communication when $\theta > 0$.
2. Losses marginally decrease for instrument relative to target communication when $\theta < 0$.

This implies that the region of parameters for which target communication is optimal increases in the first case and the region of parameters for which target communication is optimal decreases in the second case.

Next, consider the case of $\lambda > 1$. It is straightforward to extend Lemma 6 and Proposition 9 to show (i) $a_\tau > 1$ and $a_\tau r_\tau > 1$ and (ii) $a_Y < 1$ and $a_Y r_Y < 1$ in this context. The results follow from reversing the previous logic in this proof to this new case.

Corollary 4

Taking into account the implementability constraint, the policymaker's problem can be re-expressed as follows:

$$\min_{r,X\in\{\tau,Y\}} \int \left[\theta^2(1-\chi)(r-1)^2 + \theta^2\chi(a_X r-1)^2 + 2\chi\theta\cdot\Delta\cdot a_X r\right] \mathrm{d}F(\theta) \tag{69}$$

where $F(\theta)$ is the distribution from which θ is drawn. Given that θ has mean 0 and variance σ_θ^2, this problem reduces to

$$\min_{r,X\in\{\tau,Y\}} \sigma_\theta^2 \left[\theta^2(1-\chi)(r-1)^2 + \theta^2\chi(a_X r-1)^2\right] \tag{70}$$

This is identical to the policy problem solved in Proposition 9. Hence all properties of the solution identified in Proposition 9 carry over.

References

Abreu, Dilip, and Markus K Brunnermeier. 2003. “Bubbles and crashes.” *Econometrica*, 71(1): 173–204.

Alicke, Mark D, and Olesya Govorun. 2005. “The better-than-average effect.” In *The Self in Social Judgment.* , ed. Joachim I. Krueger, Mark D. Alicke and David Dunning, Chapter 5, 85–106. Taylor & Francis.

Amador, Manuel, and Kyle Bagwell. 2013. “The Theory of Optimal Delegation With an Application to Tariff Caps.” *Econometrica*, 81(4): 1541–1599.

Amador, Manuel, and Pierre-Olivier Weill. 2010. “Learning from Prices: Public Communication and Welfare.” *Journal of Political Economy*, 118(5): 866–907.

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. “Commitment vs. Flexibility.” *Econometrica*, 74(2): 365–396.

Andrade, Philippe, Gaetano Gaballo, Eric Mengus, and Benoit Mojon. 2019. “Forward Guidance and Heterogeneous Beliefs.” *American Economic Journal: Macroeconomics*, 11(3): 1–29.

Angeletos, George-Marios, and Alessandro Pavan. 2007. “Efficient Use of Information and Social Value of Information.” *Econometrica*, 75(4): 1103–1142.

Angeletos, George-Marios, and Alessandro Pavan. 2009. “Policy with Dispersed Information.” *Journal of the European Economic Association*, 7(1): 11–60.

Angeletos, George-Marios, and Chen Lian. 2018. “Forward Guidance without Common Knowledge.” *American Economic Review*, 108(9): 2477–2512.

Angeletos, George-Marios, and Jennifer La’O. 2009. “Incomplete Information, Higher-order Beliefs and Price Inertia.” *Journal of Monetary Economics*, 56: 19–37.

Angeletos, George-Marios, and Jennifer La’O. 2013. “Sentiments.” *Econometrica*, 81(2): 739–779.

Angeletos, George-Marios, and Karthik Sastry. 2019. “Inattentive Economies.” National Bureau of Economic Research Working Paper 26413.

Angeletos, George-Marios, and Zhen Huo. 2020. “Myopia and Anchoring.” *MIT miméo.*

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. 2018. “Quantifying Confidence.” *Econometrica*, 86(5): 1689–1726.

Athey, Susan, Andrew Atkeson, and Patrick J. Kehoe. 2005. “The Optimal Degree of Discretion in Monetary Policy.” *Econometrica*, 73(5): 1431–1475.

Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe. 2007. “On the optimal choice of a monetary policy instrument.” National Bureau of Economic Research.

Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe. 2010. “Sophisticated Monetary Policies.” *The Quarterly Journal of Economics*, 125(1): 47–89.

Barro, Robert J. 1974. “Are Government Bonds Net Wealth?” *Journal of Political Economy*, 82(6): 1095–1117.

Barro, Robert J. 1979. “On the determination of the public debt.” *Journal of Political Economy*, 87(5, Part 1): 940–971.

Barro, Robert J., and David B. Gordon. 1983. “Rules, Discretion and Reputation in a Model of Monetary Policy.” *Journal of Monetary Economics*, 12(1): 101 – 121.

Bassetto, Marco. 2002. "A Game–Theoretic View of the Fiscal Theory of the Price Level." *Econometrica*, 70(6): 2167–2195.

Bassetto, Marco. 2019. "Forward guidance: Communication, commitment, or both?" *Journal of Monetary Economics*, 108: 69 – 86.

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist. 1999. "The financial accelerator in a quantitative business cycle framework." *Handbook of Macroeconomics*, 1: 1341–1393.

Blinder, Alan, Michael Ehrmann, Jakob de Haan, and David-Jan Jansen. 2017. "Necessity as the mother of invention: monetary policy after the crisis." *Economic Policy*, 32(92): 707–755.

Blinder, Alan S. 2018. "Through a Crystal Ball Darkly: The Future of Monetary Policy Communication." *American Economic Review, Papers and Proceedings*, 108: 567–71.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2017. "Diagnostic Expectations and Credit Cycles." *The Journal of Finance*, 73(1): 199–227.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110(9): 2748–82.

Caballero, Ricardo, and Alp Simsek. 2019. "Monetary Policy with Opinionated Markets." MIT miméo.

Caballero, Ricardo, and Emmanuel Farhi. 2017. "The Safety Trap." *Review of Economic Studies*, 85(1): 223–274.

Caballero, Ricardo J, and Alp Simsek. 2020. "A Risk-Centric Model of Demand Recessions and Speculation*." 3.

Campbell, Jeffrey R., Charles L. Evans, Jonas D. M. Fisher, and Alejandro Justiniano. 2012. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brookings Papers on Economic Activity*, 43(1 (Spring)): 1–80.

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko. 2020. "Communication and the Beliefs of Agents." Working Paper.

Chahrour, Ryan. 2014. "Public Communication and Information Acquisition." *American Economic Journal: Macroeconomics*, 6(3): 73–101.

Chari, Varadarajan V., and Patrick J. Kehoe. 1999. "Optimal fiscal and monetary policy." *Handbook of Macroeconomics*, 1: 1671–1745.

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, 113(1): 1–45.

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120(1): 116–159.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review*, 105(8): 2644–78.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Michael Weber. 2019. "Forward Guidance and Household Expectations." *Miméo.*

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2019. "Monetary Policy Communications and Their Effects on Household Inflation Expectations." *Miméo.*

Cornand, Camille, and Frank Heinemann. 2015. "Limited Higher Order Beliefs and the Welfare Effects of

Public Information." *The Economic Journal*, 118(528): 718–742.

Correia, Isabel, Emmanuel Farhi, Juan Pablo Nicolini, and Pedro Teles. 2013. "Unconventional Fiscal Policy at the Zero Bound." *American Economic Review*, 103(4): 1172–1211.

Correia, Isabel, Juan Pablo Nicolini, and Pedro Teles. 2008. "Optimal Fiscal and Monetary Policy: Equivalence Results." *Journal of Political Economy*, 116(1): 141–170.

Crawford, Vincent P, Miguel A Costa-Gomes, and Nagore Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature*, 51(1): 5–62.

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." *FRB of New York miméo.*

Dixit, Avinash K, and Joseph E Stiglitz. 1977. "Monopolistic competition and optimum product diversity." *The American economic review*, 67(3): 297–308.

Eggertsson, Gauti, and Michael Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity*, 34(1): 139–235.

Ehrmann, Michael, Gaetano Gaballo, Peter Hoffmann, and Georg Strasser. 2019. "Can more public information raise uncertainty? The international evidence on forward guidance." *Journal of Monetary Economics*, 108: 93 – 112.

Farhi, Emmanuel, and Iván Werning. 2019. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *American Economic Review*, 109(11): 3887–3928.

Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal Taxation with Behavioral Agents." *American Economic Review*, 110(1): 298–336.

Feroli, Michael, David Greenlaw, Peter Hooper, Frederic S. Mishkin, and Amir Sufi. 2017. "Language after liftoff: Fed communication away from the zero lower bound." *Research in Economics*, 71(3): 452 – 490. Special issue on Macroeconomics.

Friedman, Benjamin M. 1990. "Targets and Instruments of Monetary Policy." In *Handbook of Monetary Economics*. Vol. 2. 1 ed., , ed. B. M. Friedman and F. H. Hahn, Chapter 22, 1185–1230. Elsevier.

Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, 129(4): 1661–1710.

Gabaix, Xavier. 2020. "A Behavioral New Keynesian Model." *American Economic Review*, 110(8): 2271–2327.

Galí, Jordi, and Mark Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics*, 44(2): 195–222.

Garcıa-Schmidt, Mariana, and Michael Woodford. 2019. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review*, 109(1): 86–120.

Halac, Marina, and Pierre Yared. 2018. "Instrument-Based vs. Target-Based Rules." National Bureau of Economic Research Working Paper 24496.

Hansen, Lars Peter, and Thomas J. Sargent. 2007. *Robustness.* Princeton, NJ:Princeton University Press.

Iovino, Luigi, and Dmitriy Sergeyev. 2019. "Central Bank Balance Sheet Policies Without Rational Expectations." *Bocconi University miméo.*

Jappelli, Tullio, and Luigi Pistaferri. 2010. "The consumption response to income changes." *Annual Re-*

view of Economics, 2(1): 479–506.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante. 2018. "Monetary Policy According to HANK." *American Economic Review*, 108(3): 697–743.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *The Journal of Political Economy*, 105(2): 211–248.

Kocherlakota, Narayana. 2012. "Comments in the December 11-12, 2012, FOMC Meeting (Transcript)." Accessible at https://www.federalreserve.gov/monetarypolicy/files/FOMC20121212meeting.pdf.

Krugman, Paul R. 1998. "It's baaack: Japan's slump and the return of the liquidity trap." *Brookings Papers on Economic Activity*, 1998(2): 137–205.

Kydland, Finn E, and Edward C Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica*, 50(6): 1345–1370.

Lucas, Robert E. Jr., and Nancy L. Stokey. 1983. "Optimal Fiscal and Monetary policy in an Economy without Capital." *Journal of Monetary Economics*, 12(1): 55–93.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The Power of Forward Guidance Revisited." *American Economic Review*, 106(10): 3133–3158.

Morris, Stephen, and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-fulfilling Currency Attacks." *American Economic Review*, 587–597.

Morris, Stephen, and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review*, 92(5): 1521–1534.

Myatt, David P., and Chris Wallace. 2012. "Endogenous Information Acquisition in Coordination Games." *Review of Economic Studies*, 79(1): 340–374.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review*, 85(5): 1313–1326.

Nakamura, Emi, and Jon Steinsson. 2018. "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect." *The Quarterly Journal of Economics*, 133(3): 1283–1330.

Patterson, Christina. 2019. "The Matching Multiplier and the Amplification of Recessions." *Northwestern University mimeo.*

Poole, William. 1970. "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model." *The Quarterly Journal of Economics*, 84(2): 197–216.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3): 665–690.

Sims, Christopher A. 2006. "Rational inattention: Beyond the linear-quadratic case." *The American Economic Review*, 96(2): 158–163.

Stahl, Dale. 1993. "Evolution of Smart Players." *Games and Economic Behavior*, 5(4): 604–617.

Swanson, Eric T., and John C. Williams. 2014. "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates." *American Economic Review*, 104(10): 3154–85.

Weitzman, Martin. 1974. "Prices vs. Quantities." *Review of Economic Studies*, 41(4): 477–491.

Wiederholt, Mirko. 2016. "Empirical Properties of Inflation Expectations and the Zero Lower Bound."

Goethe University Frankfurt miméo.

Williams, John C. 2012. “Comments in the December 11-12, 2012, FOMC Meeting (Transcript).” Accessible at https://www.federalreserve.gov/monetarypolicy/files/FOMC20121212meeting.pdf.

Williams, John C. 2016. “Discussion of "Language after Liftoff: Fed Communication Away from the Zero Lower Bound.” Remarks at the 2016 US Monetary Policy Forum.

Woodford, Michael. 2003. “Imperfect Common Knowledge and the Effects of Monetary Policy.” *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.*

Woodford, Michael. 2010. “Robustly Optimal Monetary Policy with Near-Rational Expectations.” *American Economic Review*, 100(1): 274–303.

Woodford, Michael. 2013. “Macroeconomic Analysis Without the Rational Expectations Hypothesis.” *Annual Review of Economics*, 5: 303–346.