Managing Expectations without Rational Expectations

George-Marios Angeletos* **Karthik A. Sastry†**

February 15, 2019

Abstract

Should a policymaker offer forward guidance by committing to a path for the policy instrument or a target for an equilibrium outcome? We study how the optimal approach depends on plausible bounds on agents' depth of knowledge and rationality. Agents make mistakes in predicting, or reasoning about, the behavior of others and the GE effects of policy. The optimal policy minimizes the bite of such mistakes on implementability and welfare. This goal is achieved by fixing and communicating an outcome target if and only if the GE feedback is strong enough. Our results suggest that central banks should stop talking about interest rates and start talking about unemployment when faced with a steep Keynesian cross or a prolonged liquidity trap.

*MIT and NBER; angelet@mit.edu.
†MIT; ksastry@mit.edu.

Monetary policy is 98 percent talk and only two percent action.

Ben Bernanke, 2015

1 Introduction

Forward guidance, the art of managing expectations, is rarely comprehensive. For example, even if a central bank can move the yield curve by announcing an intended path for policy rates, it remains up to the market to predict the consequences for GDP or unemployment. Under what circumstances, we ask, is it better to engage in the opposite type of forward guidance, committing to a sharp target for the outcome of interest and leaving the market to ponder what policy will support this target?

We study how the answer to this question depends on certain kinds of bounded rationality or higher-order uncertainty. We work with an abstract but flexible framework, focusing on the key conceptual issues. The textbook policy paradigm imposes that agents make no mistakes in predicting, or reasoning about, the general equilibrium (GE) consequences of policy shifts. Our framework allows such mistakes to exist and sheds light on how the policymaker can minimize their impact.

Our main result is a sharp dependence of the optimal form of forward guidance on the feedback between aggregate outcomes and individual actions ("GE considerations"). Offering forward guidance in the form of a target for the relevant outcome instead of a value for the policy instrument is optimal if and only if this feedback is sufficiently high, as in situations with a strong aggregate demand externality, a steep Keynesian cross, or a prolonged liquidity trap.

Framework. The following example, which is nested in our abstract framework, helps fix ideas. There are many firms whose investment determines aggregate output (the targeted outcome) and a policymaker who controls a future subsidy (the policy instrument). The promise of a higher subsidy can encourage investment, but its effectiveness depends on how firms expect other firms to respond, due to an aggregate demand externality (the GE feedback). A more topical application recasts the firms as consumers during a liquidity trap, the policy instrument as the interest rate set once the zero lower bound (ZLB) ceases to bind, and the GE feedback as the Keynesian income-spending multiplier.

Communication is necessary because the policymaker's objectives vary with a shock only she observes. This shock has no direct effect on the private agents' payoffs. Its sole role is to make sure that the agents do not *a priori* know what the policymaker will do—they need "forward guidance."

The policymaker chooses between two forms of such guidance. In the first, she announces, and commits to, a value for the policy instrument (the subsidy or the interest rate). In the second, she does the same with a target for the relevant outcome (aggregate output). We refer to the former strategy as *instrument communication* and to the latter one as *target communication*.

This approach equates each form of forward guidance with a commitment to a policy plan. But whereas the literature has been concerned primarily with commitment problems or robustness to fundamental uncertainty (Atkeson, Chari and Kehoe, 2007; Poole, 1970), our analysis shifts the focus to how the choice and communication of a policy plan influences the equilibrium bite of any mistakes agents may make when predicting the behavior of others and the GE effects of the policy plan.

Rational expectations and beyond. Like the textbook policy paradigm, our frictionless benchmark imposes a representative, rational-expectations agent.[1] This rules out the aforementioned kind of mistakes. It also guarantees the implementable combinations of policy and outcome are invariant to the policymaker's choice between the aforementioned two forms of forward guidance.

This irrelevance depends, not only on the assumption that the typical agent is *herself* rational and aware of the policy announcement, but also on the assumption that such rationality and awareness are common knowledge ("I know that you know..."). Our analysis centers on relaxations of the second, stronger assumption: such relaxations operationalize the idea of a friction in the agents' reasoning about the behavior of others and the GE effects of policy.

The existence of such a friction is taken as given. Our contribution is to study how the policymaker can work around it, under a few complementary specifications of this friction.

Anchored beliefs. Our main specification introduces a friction by letting agents doubt about the attentiveness or the awareness of others. This amounts to removing common knowledge of the announced policy plan while preserving every agent's own knowledge of it; it is engineered with the use of heterogeneous priors; and it recasts the kind of anchored higher-order beliefs previously documented in common-prior settings (Abreu and Brunnermeier, 2003; Morris and Shin, 1998, 2002; Woodford, 2003) as a structured departure from rational expectations. A variant specification lets this departure take the form of Level-k Thinking (Nagel, 1995; Stahl, 1993).

Both of these specifications capture essentially the same friction: they anchor the expected responses of others to policy communications. In the first, an agent expects others to respond less than in our frictionless benchmark because he worries that others may be inattentive to policy communications. In the second, the same expectation is justified by the agent's limited depth of reasoning, or by her belief that others are less sophisticated than herself. The available evidence is inconclusive on which interpretation is most relevant, but it generally supports the existence of such a friction.[2]

Main results. Our take-home lesson is that, in the presence of the aforementioned friction, forward guidance in terms of targets rather than instruments is preferable when and only when the GE feedback is sufficiently strong. This lesson builds on two more elementary insights, which we explain next.

Our first insight regards the implementability constraint faced by the policymaker, namely the equilibrium relation between the policy instrument and the targeted outcome. This relation is invariant to the form of forward guidance in our rational-expectations benchmark but not away from it.

With instrument communication, the agents play a game of strategic *complementarity*. In the investment example, this is due to the aggregate demand externality; in the Keynesian context, it is due to the income-spending multiplier. In such a game, the belief friction produces *attenuation*: when an agent expects the others to respond less, she responds less herself. As a result, the implementability

[1] For our purposes, the assumption of a representative agent is equivalent to that of a common belief, or a common expectation operator. Incomplete markets and heterogeneity in payoffs are not the relevant issue here.

[2] For example, see Coibion and Gorodnichenko (2012) and Coibion et al. (2018) for evidence based on surveys of expectations, and Crawford, Costa-Gomes and Iriberri (2013) and Nagel (1995) for experiments.

constraint is steeper than its frictionless counterpart: a larger change in subsidies or interest rates is needed in order to induce the same change in output.

With target communication, everything flips. Conditional on an announced GDP target, a firm that expects a higher aggregate investment also expects a lower required subsidy to that target, which reduces the incentive to invest; and similarly, a household that expects higher aggregate spending also expects a higher interest rate, which reduces the incentive to consume. That is, the agents now play a game of strategic *substitutability*. In such a game, the belief friction produces *amplification*: when an agent expects the others to respond less, she responds *more*, not less. As a result, the implementability constraint is now flatter.

Our second insight relates to the interaction between the form of forward guidance and the underlying GE mechanism. On the one hand, the form of forward guidance regulates which object the agents have to forecast or reason about: fixing a value for the instrument burdens the agents with the task of predicting the outcome, setting a sharp target for the latter lets them ponder what the requisite policy will be. On the other hand, the GE feedback regulates which of these two objects is relatively more important in shaping actual behavior. When this feedback is weak, agents care relatively more about subsidies or interest rates; when it is strong, they care more about aggregate demand.

Combining these observations, we reach the following result: when the GE effect is weak, the bite of the belief friction on implementability is minimized by promising to fix a certain value for the policy instrument; when the GE effect is strong, the same goal is accomplished by promising to meet a certain target for aggregate output. Under the assumption that the first best is attainable with rational expectations,[3] this naturally leads to take-home less stated above: target communication is the best means for managing expectations if and only if the GE effect is sufficiently strong.

Monetary policy. A recent literature has studied the implications of the assumed friction for monetary policy under the restriction that forward guidance takes the form of a commitment on the value of the policy instrument.[4] Our result that the friction attenuates the effectiveness of instrument communication is essentially the same as the results obtained in that literature. The main added value here is to show how the policymaker can bypass or even flip this friction by engaging in a *different* form of forward guidance, and to shed light on *when* she should switch from the one form to the other.[5]

The application of our insights to monetary policy thus suggests the following: as the economy transitions from normal times to unconventional times (a liquidity trap), the central bank should stop talking about interest rates and instead start promising to do what "whatever it takes" to bring unemployment down. This is because a liquidity trap switches on powerful GE feedback chains between income, spending, and inflation, which in turn tilt the balance in favor of target communication.

[3] This assumption is not needed for the two insights discussed above. But it sharpens our policy lesson by letting the bounded rationality, or the wedge between first- and higher-order beliefs, be the *only* source of distortion.

[4] Angeletos and Lian (2018) and Wiederholt (2016) attribute the friction to lack of common knowledge; Garcıa-Schmidt and Woodford (2018) and Farhi and Werning (2017) to Level-k Thinking; Gabaix (2017) to "cognitive discounting." Iovino and Sergeyev (2017) offer a related application to quantitative easing.

[5] A secondary contribution is to put the different approaches considered in that literature under the same umbrella.

Broader scope. Although most of our analysis centers on *anchored* beliefs, the main insights also apply to a variant featuring *erratic* instead of anchored beliefs. This captures situations in which the economy's response to policy is perturbed by random shocks to higher-order beliefs, as in Angeletos and La'O (2013), Benhabib, Wang and Wen (2015) and Huo and Takayama (2015), or where the policymaker is uncertain about the agents' sophistication. The policy recommendation remains the same because the form of forward guidance regulates the distortion caused by higher-order uncertainty or limited depth of reasoning regardless of its volatility and comovement with the policy announcement.

This insight extends to more sophisticated forms of forward guidance, such as when the policymaker promises to link the instrument and the outcome via a rule. Although such strategies may be less practical or harder to communicate in certain contexts, their consideration helps illustrate how our analysis offers a new perspective on policy rules more broadly.

Finally, our conclusions are robust to the introduction of measurement error in the policymaker's observation of the outcome, trembles in her control of the instrument, and other shocks that affect payoffs but not the agents' reasoning. These features embed in our analysis a similar tradeoff as that considered in Poole (1970). They influence the optimal strategy, but not in a way that depends on the GE feedback or the belief friction.

Related literature. Our paper's most direct contributions are to the literature on forward guidance, which has already been discussed, and to the literature on policy targets and instruments that follows the lead of Poole (1970). The latter considers how a policymaker can react to different shocks (e.g., to supply or demand) under different policy regimes (such as fixing the interest rate or the growth rate of money), and studies how the optimal regime depends on the composition of shocks. The same logic underlies the modern literature on optimal Taylor rules, as well a line of work that adds time-inconsistency considerations (Atkeson, Chari and Kehoe, 2007). Our paper, instead, highlights a novel aspect: how different policy commitments can regulate the impact of bounded rationality and higher-order uncertainty. The same basic point also distinguishes our contribution from Weitzman (1974)'s classic on "prices vs quantities."

Our focus on the interaction of policy communication and higher-order beliefs is reminiscent of the literature spurred by Morris and Shin (2002).[6] In this literature, policy communication means varying the precision of a public signal about an exogenous fundamental while holding constant the agents' strategic interaction. In our paper, instead, it means regulation of the strategic interaction, and of the equilibrium impact of higher-order beliefs or reasoning, via commitment to a specific policy rule. An additional difference is that in our setting the communication of the exogenous fundamental, namely the shock to the policymaker's preferences, is both irrelevant and ineffective. Instead, what matters is the communication of the policymaker's plan of action.[7]

[6] See, *inter alia*, Amador and Weill (2010), Angeletos and Pavan (2007), Chahrour (2014), James and Lawler (2011), Morris and Shin (2007), Myatt and Wallace (2012), and Svensson (2006).

[7] Similar points distinguish our paper from the literature on Bayesian persuasion and information design (Bergemann and Morris, 2013, 2018; Kamenica and Gentzkow, 2011; Inostroza and Pavan, 2018).

Last but not least, there is a long tradition of relaxing rational expectations within representative-agent models. See, for example, Hansen and Sargent (2007) and Woodford (2010) on robustness, or Sargent (1993), Evans and Honkapohja (2001) and Marcet and Nicolini (2003) on learning. We depart from this tradition by dropping the representative-agent assumption in the direction of accommodating strategic considerations and a wedge between first- and higher-order beliefs.

Layout. Section 2 introduces our framework. Section 3 studies our rational-expectations benchmark. Section 4 lays down the foundations of the subsequent analysis. Section 5 contains our main specification, anchored higher-order beliefs. Section 6 considers the variant with Level-k Thinking. Section 7 discusses the application to monetary policy. Section 8 considers the variant with erratic beliefs. Section 9 considers additional policy options and richer trade-offs. Section 10 concludes.

2 Framework

Here we introduce the physical environment, the objective of the policymaker, the incentives of the private agents, and the timing of actions. We postpone, however, the specification of how agents form expectations, or how they reason about the behavior of others, until later.

Basic structure. 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 policymaker controls a policy instrument $\tau \in \mathbb{R}$ and is interested in manipulating an aggregate outcome $Y \in \mathbb{R}$.

The aggregate outcome is related to the policy instrument and the behavior of the agents as follows:

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

where $K \equiv \int k_i \, \mathrm{d}i$ is the average action of the private agents and $\alpha \in (0,1)$ is a fixed parameter. This parameter controls how much of the effect of the policy instrument τ on the outcome Y is direct, or mechanical, rather than channeled through the endogenous response of K.

The behavior of the private agents, in turn, is governed by the following best responses:

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

where $\mathbb{E}_i$ denotes the subjective expectation of agent i and $\gamma \in (0,1)$ is a fixed parameter. Depending on assumptions made later on, the operator $\mathbb{E}_i$ may or may not be consistent with Rational Expectations Equilibrium (REE). The parameter γ controls how much private incentives depend on expectations of the aggregate outcome, which in turn depends on the behavior of others.

Key features and interpretation. Our framework stylizes three features likely shared by many applications. First, individual decisions depend on two kinds of expectations: the expectations of a policy instrument, such as a tax or the interest rate set by the central bank, and the expectations of an aggregate outcome, such as aggregate output. Second, the realized aggregate outcome depends on the realized aggregate behavior. And third, the policy instrument has a direct effect on the aggregate outcome even if we hold constant the decisions under consideration.

The first two assumptions capture the interdependence of economic decisions such as firm investment and consumer spending. In macroeconomics, this interdependence typically reflects general-equilibrium (GE) interactions. Accordingly, the parameter γ, which plays a crucial role in the subsequent analysis, may be interpreted as a measure of the strength of the GE interaction. The third assumption and the parameter α, on the other hand, play a more mechanical function. Had $1-\alpha$ been zero, the policymaker could not possibly commit to a specific target for Y "no matter what" (i.e., regardless of K). Letting $\alpha < 1$ makes sure that such a commitment is viable.

To fix ideas, we next discuss how our framework can nest a stylized neoclassical economy, in which τ represents a subsidy, K represents investment, and γ indexes the strength of an aggregate-demand externality. A more topical application, discussed in Section 7, recasts τ as the negative of the interest rate set by the central bank after the economy exits a liquidity trap, K as consumer spending during the trap, and γ as the strength of the Keynesian income-spending multiplier.

A micro-foundation. There are three periods, $t \in \{0,1,2\}$; a continuum of entrepreneurs, $i \in [0,1]$, who choose investment at $t=1$; and a policymaker, who can subsidize production at $t=2$. Two additional agents, a competitive final-good firm and a competitive worker, serve only auxiliary roles.

The first period, $t=0$, identifies the moment when the policymaker announces and commits to a policy plan. No other decision takes place at this moment. At $t=1$, each entrepreneur i is endowed with a unit of a good that can either be consumed or invested in the form of differentiated capital good, x_i. At $t=2$, the entrepreneur sells his capital to the final-good firm at price p_i and consumes the profits. His budget is therefore given by $c_{i,1}+x_i=1$ at $t=1$ and by $c_{i,2}=p_i x_i$ at $t=2$, where $c_{i,t}$ denotes consumption in period t. His lifetime utility is $u_i = c_{i,1}+c_{i,2}$.

The final-good firm operates at $t=2$. Its output is given by $Q = X^{\eta}N^{1-\eta}$ and its revenue by $(1-r)Q - wN - \int p_i x_i \mathrm{d}i$, where r is the rate of taxation, $X \equiv (\int {}_i^{1-\rho} \mathrm{d}i)^{1/1-\rho}$ is a CES (constant elasticity of substitution) aggregator of the differentiated capital goods, N is the labor input supplied by the worker, and $\rho \in [0,1]$ and $\eta \in [0,1]$ parametrize, respectively, the elasticity of substitution of the differentiated inputs and the income share of capital. Finally, the worker lives, works, and consumes only in period $t=2$ and has utility $v = wN - \frac{1}{1+\phi}N^{1+\phi}$, where w is the real wage and $\phi > 0$ parameterizes the Frisch elasticity.

A log-linear approximation of this economy is directly nested to our abstract framework. The detailed derivations can be found in Appendix B.1. Here, we sketch the main argument of how this example offers a micro-foundation of conditions (1) and (2).

Introduce the following transformations:

$$k_i \equiv \tfrac{1+\eta\phi}{1+\phi}(\log x_i - \log \bar{x}), \quad \tau \equiv \tfrac{1+\eta\phi}{\phi(1-\eta)}\log(1-r), \quad \text{and} \quad Y \equiv \log Q - \log \bar{Q}.$$

where $\bar{x}$ and $\bar{Q}$ are constants (essentially the "steady-state" quantities corresponding to $r=0$).

Consider the determination of aggregate output at $t = 2$. Impose optimality for the final-good firm and the worker, plus market clearing. Aggregate output can then be expressed as in (1), namely

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

with $\alpha \equiv \eta \frac{(1+\phi)^2}{(\eta+\phi)(1+\eta\phi)} \in [0, 1]$. The first term captures the effect of the subsidy on labor supply and thereby on output. The second captures the role of capital in production.

Consider next the investment decisions of the entrepreneurs at $t = 1$. Impose optimality and knowledge of the structure of the economy, but not necessarily rational expectations. The optimal investment decision of entrepreneur i can then be expressed as in (2), namely

$$k_i = (1 - \gamma)\,\mathbb{E}_i[\tau] + \gamma \mathbb{E}_i[Y]. \tag{3}$$

with $\gamma \equiv \frac{(1+\eta\phi)(\eta\rho+\phi(\eta+\rho-1))}{\eta\rho(1+\phi)^2} < 1$. The first term captures the direct or PE effect of the subsidy on investment. The second term combines two GE effects. On the one hand, because of the aggregate-demand externality, higher aggregate output raises the individual return to investment for given wages and given τ. On the other hand, higher aggregate output boosts aggregate labor demand, raises wages, and lowers the return on capital. The sign of γ depends on the relative strength of these GE channels. We assume the aggregate-demand externality is sufficiently strong for $\gamma > 0$.[8]

The derivation of condition (3) does not require rational expectations: it holds for possibly arbitrary subjective beliefs about τ and Y. Also, changing the deep parameter ρ, which controls the aggregate-demand externality, varies the reduced-form parameter γ while keeping constant α. This micro-foundation therefore justifies the interpretation of the comparative static in γ for fixed α in our abstract framework as variation in the "underlying GE feedback."

Policy objective. The policymaker minimizes the rational expectation of the following loss function:

$$L = L(\tau, Y, \theta) \equiv (1 - \chi)(\tau - \theta)^2 + \chi(Y - \theta)^2. \tag{4}$$

where $\chi \in (0, 1)$ is a fixed scalar and θ is a zero-mean random variable that represents the policymaker's ideal or first-best combination of the instrument and the outcome.

The micro-foundations of this objective are left outside the analysis. The main insights regarding the bite of bounded rationality on implementability and the regulation of this bite by the form of forward guidance do not depend at all on the specification of the policymaker's objective. The adopted specification only sharpens the normative exercise by letting the policymaker attain her first best (zero loss) in the rational-expectations benchmark studied in the next section.

The realization of θ is observed by the policymaker but not by the private agents. Because we assume full commitment, this does not introduce incentive problems. And because θ does not enter conditions (1) and (2), the agents do not care to know θ *per se*; they only care to know what the policymaker plans to do and how this may affect the behavior of others. As anticipated in the Introduction,

[8]The alternative possibility, in which the combined GE effect turns negative ($\gamma < 0$), is discussed in Appendix D. Although some of the results and intuitions have to be modified in this case, the take-home message remains largely the same.

the sole purpose of letting θ be random and unobserved to the agents is therefore to motivate why the agents do not a priori know what the policymaker will do—they need "forward guidance."

Timing. There are three stages, or periods, which are described below:

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

1. Each agent i chooses k_i.

2. K is observed by the policymaker and (τ, Y) are 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 the assumption of that the policymaker always honors in stage 2 any promise made in stage 0. Different communications are therefore equated to different commitments: instrument communication means forward guidance in the form of a commitment to a value for τ and, similarly, target communication means forward guidance in the form of a commitment to a target for Y. However, the choice between these two strategies has nothing to do with time-inconsistency considerations, because commitment is full. As it will become clear in the sequel, this choice only has to do with the management of the expectations agents form in stage 2 about the behavior of others.

3 A Rational Expectations Benchmark

In this section, we explain why the choice of the form of forward guidance is irrelevant in a benchmark that, like the textbook policy paradigm, imposes a representative agent who is attentive to policy communications and forms rational expectations.[9]

In this benchmark, $\mathbb{E}_i[\cdot] = \mathbb{E}[\cdot|\hat{X}]$ for all i, where $\mathbb{E}[\cdot|\hat{X}]$ is the rational expectation conditional on announcement $\hat{X}$, with $X \in \{\tau, Y\}$ depending on the mode of communication. As a result, $k_i = K$ for all i and condition (2) reduces to the following condition, which describes the optimal behavior of the representative agent:

$$K = (1-\gamma)\mathbb{E}[\tau|\hat{X}] + \alpha\mathbb{E}[Y|\hat{X}]. \tag{5}$$

We can thus define the sets of the combinations of the policy instrument, τ, and the outcome, Y, that can be implemented under each form of forward guidance as follows.

Definition 1. *A pair (τ, Y) is implementable under instrument [respectively, target] communication if there is an announcement $\hat{\tau}$ [respectively, $\hat{Y}$] and an action K for the representative agent such that conditions (1) and (5) are satisfied, expectations are rational, and $\tau = \hat{\tau}$ [respectively, $Y = \hat{Y}$].*

[9] This is the same as imposing complete information and Nash equilibrium.

This definition embeds Rational Expectations Equilibrium (REE). In the subsequent sections, we will revisit implementability under different solution concepts. In the rest of this section, we formulate and solve the policymaker's problem in a manner that parallels the analysis in the subsequent sections.

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

$$\min_{\mathcal{A}\in\{\mathcal{A}_\tau^*,\mathcal{A}_Y^*\},(\tau,Y)\in\mathcal{A}} \mathbb{E}[L(\tau, Y, \theta)] \tag{6}$$

The choice $\mathcal{A} \in \{\mathcal{A}_\tau^*, \mathcal{A}_Y^*\}$ captures the choice of the optimal mode of communication (instrument vs target). The choice $(\tau, Y) \in \mathcal{A}$ captures the optimal choice of the pair (τ, Y) taking as given the mode of communication. Both of these choices are conditional on θ.

We now proceed to show that $\mathcal{A}_\tau^* = \mathcal{A}_Y^*$. 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 (5) 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 mode of communication. 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 mode of communication, a pair (τ, Y) is implementable if and only if $\tau = Y$.

Proposition 1. $\mathcal{A}_\tau^* = \mathcal{A}_Y^* = \mathcal{A}^* \equiv \{(\tau, Y) : \tau = Y\}$.

That $\mathcal{A}^*$ is a linear locus with slope 1 is a simplifying feature of our environment. The relevant point here is that the implementability constraint faced by the planner is invariant to the form of forward guidance,[10] which in turn implies the following.

Proposition 2. *The policymaker attains her first best* $(L = 0)$ *by announcing* $\hat{\tau} = \theta$*, as well as by announcing* $\hat{Y} = \theta$*. The optimal form of forward guidance is therefore indeterminate.*

In fact, the first best is attained even if the policymaker only announces the shock θ itself, as opposed to announcing a policy plan. For, once θ is known, every agent can reason, without the slightest grain of doubt and without any chance of error, that all other agents will play $K = \theta$ and that the policymaker will set $\tau = \theta$, in which case it is optimal for him to play $k_i = \theta$ as well.

These findings suggest where we are heading next. The rest of the paper is devoted on relaxing this kind of flawless reasoning and on understanding how this makes the form of forward guidance an effective tool for managing expectations.

[10]This invariance mirrors the equivalence of the "dual" and "primal" approaches in the Ramsey literature (Chari and Kehoe, 1999): in our setting, $\mathcal{A}_\tau^*$ corresponds to the primal problem, where the planner chooses instruments, and $\mathcal{A}_Y^*$ corresponds to the dual, where she chooses allocations.

4 Preliminaries

Any departure from rational expectations has to be done in a structured way, or else "anything goes." In this section, we explain the common structure underlying the various departures considered in the rest of the paper. We also show how the form of forward guidance regulates the strategic interaction of the agents and thereby the effect of higher-order beliefs. We thus lay down the foundations for the analysis in the subsequent sections.

4.1 Rational Expectations and Common Knowledge

We start by recasting our rational-expectations benchmark as the joint of two assumptions: one regarding the agents' *own* rationality and awareness; and another regarding the beliefs about *others*.

Assumption 1. *Every agent is rational and attentive in the following sense: he is Bayesian (although possibly with a mis-specified prior), acts according to condition (2), understands that the outcome is determined by condition (2) and that the policymaker has full commitment and acts so as to minimize (4), and receives any message sent by the policymaker.*

Assumption 2. *The aforementioned facts are common knowledge.*

Proposition 3. *Provided $\alpha < \frac{1}{2-\gamma}$, the REE benchmark studied in the previous section is equivalent to joint of Assumptions 1 and 2.*

This claim will be verified shortly, as part of the arguments developed in the rest of this section. The basic idea is that, for any policy announcement made at stage 0, the joint of Assumptions 1 and 2 yield a unique rationalizable outcome in stages 1 and 2, which coincides with the REE outcome obtained in the previous section. The restriction $\alpha < \frac{1}{2-\gamma}$ is needed for the uniqueness of the rationalizable outcome, but not for the uniqueness of the REE and can be dispensed with for most of the applied lessons. We next discuss what Assumptions 1 and 2 mean and how they help structure the forms of "bounded rationality" considered in the rest of the paper.

Assumption 1 imposes that, for any i, agent i's subjective beliefs and behavior satisfy the following three restrictions:

$$\mathbb{E}_i[X] = \hat{X}, \qquad \mathbb{E}_i[Y] = (1-\alpha)\mathbb{E}_i[\tau] + \alpha\mathbb{E}_i[K], \qquad \text{and} \qquad k_i = (1-\gamma)\,\mathbb{E}_i[\tau] + \gamma\mathbb{E}_i[Y], \tag{7}$$

where $X \in \{\tau, Y\}$ depending on the mode of communication. 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 condition (2).

Assumption 2, in turn, imposes that agents can reason, with full confidence and no mistake, 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 *boundless* knowledge and rationality that our frictionless benchmark and the textbook policy paradigm alike impose—and that we instead seek to relax.

This explains the approach taken in the rest of the paper: we modify Assumption 2 while maintaining Assumption 1. This aims at isolating the role of any mistakes agents make when trying to predict or reason about the behavior of others and the GE consequences of any policy plan.

With this point in mind, the rest of this section proceeds to develop two insights that hold true for *any* possible relaxation of Assumption 2. The first is that the form of forward guidance controls the agents' strategic interaction and, thereby, the equilibrium impact of the aforementioned kind of mistakes. The second is that such mistakes can be mapped to higher-order beliefs.

4.2 Forward guidance and strategic interaction

Consider first the case in which the policymaker announces, and commits on, a value $\hat{\tau}$ for the instrument. Recall that Assumption 1 yields the three restrictions given in condition (7). Under instrument communication, the first restriction becomes $\mathbb{E}_i[\tau] = \hat{\tau}$ and the remaining two restrictions reduce to

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

The first equation highlights that, under instrument communication, agents only need to predict Y. The second highlights that predicting Y is the same as predicting the behavior of others, or K. Combining them gives the following result.

Lemma 1. *Let $\delta_\tau \equiv \alpha\gamma \in (0,1)$. When the policymaker announces and commits to a value $\hat{\tau}$ for the instrument, agents play a game of strategic complementarity in which best responses are given by*

$$k_i = (1-\delta_\tau)\hat{\tau} + \delta_\tau\mathbb{E}_i[K]. \tag{8}$$

Note that the level of the best responses in this game is controlled by $\hat{\tau}$, the announced value of the policy instrument, while their slope is given by δ_τ. The latter encapsulates how much aggregate behavior depends on the forecasts agents form about one another's behavior relative to the policy instrument—or, equivalently, how much aggregate investment depends on the perceived GE effect of the subsidy relative to its PE effect.[11]

Consider now the case in which the policymaker announces a target $\hat{Y}$ for the outcome. In this case, $\mathbb{E}_i[Y] = \hat{Y}$ and the remaining two restrictions from condition (7) can be rewritten as

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

The first equation highlights that, under target communication, agents need to predict the subsidy that will support the announced target. The second shows that, for given an announced target $\hat{Y}$,

[11] The game obtained above is similar to the beauty-contest games studied in, *inter alia*, Morris and Shin (2002), Woodford (2003), Angeletos and Pavan (2007), and Bergemann and Morris (2013), with $\hat{\tau}$ corresponding to the "fundamental," or the shifter of best responses, in these papers. There are, however, two subtle differences. First, whereas the fundamental in those papers is exogenous, here $\hat{\tau}$ is controlled by the policymaker. Second, whereas these papers let the fundamental be observed with noise, here $\hat{\tau}$ is perfectly observed.

the expected subsidy is a *decreasing* function of the expected K : an agent who is pessimistic about aggregate investment expects the policymaker to use a higher subsidy in order to meet the given output target. Combining these two equations, we reach the following counterpart to Lemma 1.

Lemma 2. *Let $\delta_Y \equiv -\frac{\alpha}{1-\alpha}(1-\gamma) < 0$. When the policymaker announces and commits to a target $\hat{Y}$ for the outcome, agents play a game of strategic substitutability in which best responses are given by*

$$k_i = (1-\delta_Y)\hat{Y} + \delta_Y \mathbb{E}_i[K]. \tag{9}$$

This game is similar to that obtained in Lemma 1 in the following respect: in both cases, the policymaker's announcement controls the intercept of the best responses. The two games are nevertheless different in the following key respect: whereas the game obtained in Lemma 1 displayed strategic complementarity ($\delta_\tau > 0$), the one obtained here displays strategic substitutability ($\delta_Y < 0$). In the first scenario, an agent who expects the others to invest more has a higher incentive to invest, because higher K maps to higher Y and hence to higher returns for fixed τ. In the second scenario, the same agent has a *lower* incentive to invest, because a higher K means that a lower subsidy will be required in order to meet the announced target for Y.

We summarize this elementary, but important, point in the following corollary.

Corollary 1. *Switching from instrument communication to target communication changes the game played by the agents from one of strategic complementarity to one of strategic substitutability.*

4.3 The role of higher-order beliefs

The results developed above follow from Assumption 1 *alone*. They therefore also hold in our REE benchmark. But they turn out to be inconsequential in that benchmark because of its stronger assumption regarding agents' depth of knowledge and rationality, namely Assumption 2. We now explain how this assumption imposes a tight restriction on higher-order beliefs, which in turn drives the irrelevance result. In so doing, we also prove Proposition 3.

With $X \in \{\tau, Y\}$ indexing the mode of communication, the best responses obtained in Lemmas 1 and 2 are nested in the following form:

$$k_i = (1-\delta_X)\mathbb{E}_i[X] + \delta_X \mathbb{E}_i[K]. \tag{10}$$

Necessarily, $\delta_\tau \in (0,1)$ and $\delta_Y < 0$. For the rest of the paper, we restrict $\alpha < \frac{1}{2-\gamma}$, which guarantees that $\delta_Y > -1$.[12] We thus $|\delta_X| < 1$ for both $X \in \{\tau, Y\}$.

[12] This allows the characterization of beliefs and behavior by repeated iteration of the best responses. In this section, this guarantees that the joint of Assumptions 1 and 2 yields the REE outcome. In the next section, it guarantees the equivalence of Assumptions 1 and 3 to a certain PBE outcome. When this restriction is violated, our main lessons continue to apply as long as one focuses directly on the relevant REE and PBE outcomes. What is lost is only the "global stability" of these outcomes, in the sense that the fixed points are no more obtainable as the limit of iterated best responses.

Common knowledge of rationality—which is one half of Assumption 2—implies that every agent can aggregate and iterate condition (10) to obtain her forecast of the aggregate action as follows:

$$\mathbb{E}_i[K] = \mathbb{E}_i\left[(1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\bar{\mathbb{E}}^h[X]\right], \tag{11}$$

where $\bar{\mathbb{E}}^h[\cdot]$ denotes the h-th order average forecast. This is defined recursively by letting $\bar{\mathbb{E}}^1[\cdot] \equiv \int \mathbb{E}_i[\cdot]\mathrm{d}i$ and $\bar{\mathbb{E}}^h[\cdot] \equiv \bar{\mathbb{E}}[\bar{\mathbb{E}}^{h-1}[\cdot]]$ for all $h \geq 2$.

Consider how an agent's expectation of the behavior of others, $\mathbb{E}_i[K]$, varies with $\hat{X}$. This captures the agent's expectation or reasoning about the GE consequences of the announced policy plan. In the investment example introduced earlier on, this corresponds to how much the typical entrepreneur expects other entrepreneurs to invest in response to forward guidance about future subsidies. And in the New Keynesian context studied in Section 7, it corresponds to how much the typical consumer expects other consumers to spend in response to forward guidance about future monetary policy.

Condition (11) allows us to represent this kind of expectation or reasoning as a function of the higher-order beliefs about $\hat{X}$. From this perspective, imposing a structure on higher-order beliefs is synonymous to imposing a structure on how agents form expectations or reason about GE effects.

In our REE benchmark, this structure takes the form of the perfect coincidence of higher-order and first-order beliefs, due to Assumption 2. To see this, note that an agent i who believes that agent $j \neq i$ is rational and attentive has the following second-order-beliefs: $\mathbb{E}_i[\mathbb{E}_j[X]] = \mathbb{E}_i[X] = \hat{X}$. By induction, common knowledge of rationality and attentiveness—which is Assumption 2—gives

$$\bar{\mathbb{E}}^h[X] = \bar{\mathbb{E}}^1[X] = \hat{X} \quad \forall h \geq 1, \tag{12}$$

which verifies the claim that higher-order and first-order beliefs coincide.

Using the above into (11) to get $\mathbb{E}_i[K] = \hat{X}$ and replacing the latter into (10), we then also have

$$k_i = K = \hat{X},$$

which replicates the behavior in our REE benchmark. Not only does this verify Proposition 3, but also sheds light on *how* the predictions of that benchmark depends on "boundless" knowledge and rationality, in the form of Assumption 2. And conversely, the forms of "bounded" knowledge and rationality considered in the rest of the paper represent relaxations of this assumption.

5 Anchored Beliefs

We now turn to the core of our contribution, which is to characterize the optimal strategy for managing expectations in the presence of a particular friction, one that anchors the beliefs about the responses of others to the policy announcement. As noted in the Introduction, the kind of anchored beliefs we are after seems consistent with both survey and experimental evidence. In this paper, we capture this friction by replacing Assumption 2 with the following.

Assumption 3 (Anchored Beliefs)**.** *Every agent believes that all other agents are rational but only a fraction $\lambda \in [0,1]$ of them is attentive. In particular, 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 mode of communication. This fact and the value of λ are common knowledge.*

Relative to Assumption 2, Assumption 3 maintains common knowledge of rationality but drops common knowledge of attentiveness. This amounts to changing the solution concept from REE to Perfect Bayesian Equilibrium with the following heterogeneous, and mis-specified, priors: each agent i receives a private signal s_i of the announcement; believes *correctly* that his signal is a drawn from a Dirac measure at $\hat{X}$; and believes *incorrectly* that, for any $j \neq i$, s_j is drawn from a Dirac measure at $\hat{X}$ with probability λ and from a Dirac measure at 0 with probability $1-\lambda$.

A similar specification has been used in Angeletos and La'O (2009) to introduce belief inertia in the New Keynesian model. As already noted, this specification is grounded on a literature that studies the role of higher-order uncertainty in common-prior settings (e.g., Angeletos and Lian, 2018; Morris and Shin, 2002, 2006; Woodford, 2003). The heterogeneous-prior approach affords, not only a higher level of tractability, but also an extra degree of freedom: the belief anchoring is not necessarily tied to the cross-sectional dispersion in forecasts. This extra degree of freedom represents a departure from rational expectations—which, for the present purposes, is a feature, not a bug.

As explained in Section 6, Level-k Thinking produces similar results as this specification—in effect by relaxing the part of Assumption 2 that pertains to common knowledge of rationality, as opposed to the part that pertains to the common knowledge of attentiveness. We thus invite the reader to interpret the results presented in this section as the product of introducing plausible bounds on *either* the depth of knowledge or the depth of rationality.

5.1 Beliefs or reasoning

Because Assumption 3 maintains common knowledge of best responses, we can once again express an agent's beliefs about K as a function of higher-order beliefs:

$$\mathbb{E}_i[K] = \mathbb{E}_i\left[(1-\delta_X)\sum_{h=1}^{\infty}(\delta_X)^{h-1}\bar{\mathbb{E}}^h[X]\right]. \tag{13}$$

What is different from Section 4 is the structure of higher-order beliefs. Because the typical agent believes that only a fraction λ of the other agents is attentive, 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 induction, for any $h \geq 1$,

$$\mathbb{E}_i\left[\bar{\mathbb{E}}^h[X]\right] = \mathbb{E}_i\mathbb{E}_{j_1}...\mathbb{E}_{j_h}[X] = \lambda^h\hat{X}. \tag{14}$$

Relative to the frictionless benchmark, higher-order beliefs are therefore tilted towards zero, and the more so the higher their order.

As anticipated, this property is similar to that obtained in common-prior settings with the use of dispersed private information (e.g., Morris and Shin, 2002; Woodford, 2003). The only difference is that we have relaxed the common-prior assumption so as to disentangle this property from noisy information and we have recast it as a form of bounded rationality.

By substituting (14) into (13), we infer that the typical agent's expectation of K following announcement $\hat{X}$ is given by

$$\mathbb{E}_i[K] = \frac{\lambda - \lambda\delta_X}{1 - \lambda\delta_X}\hat{X}. \tag{15}$$

Instrument communication corresponds to $\delta_X = \delta_\tau > 0$, whereas target communication corresponds to $\delta_X = \delta_Y < 0$. In either case, however, the responsiveness of $\mathbb{E}_i[K]$ to $\hat{X}$ is bounded between 0 and the frictionless counterpart:

$$0 \leq \frac{\lambda - \lambda\delta_X}{1 - \lambda\delta_X} \leq 1.$$

Furthermore, the above ratio is strictly increasing in λ, which means that the following is true.

Corollary 2. *Under either form of forward guidance, a higher friction (lower λ) reduces the expected response of K to the announced $\hat{X}$ and, in this sense, reduces the perceived GE effect of the policy announcement.*

5.2 Attenuation vs Amplification

Although the *nature* of the belief friction is qualitatively the same between the two forms of forward guidance, its *impact* on actual behavior is qualitatively different. Indeed, replacing (15) in the best-response condition (10) and aggregating across agents, we reach the following result.

Lemma 3. *The realized aggregate investment following announcement $\hat{X}$ is given by*

$$K = \frac{1 - \delta_X}{1 - \lambda\delta_X}\hat{X}, \tag{16}$$

where $X \in \{\tau, Y\}$ depending on the mode of communication.

Recall that the frictionless benchmark had $K = \hat{X}$. When $\delta_X > 0$, the ratio $\frac{1-\delta_X}{1-\lambda\delta_X}$ is strictly lower than 1 for every $\lambda < 1$ and is increasing in λ. When instead $\delta_X < 0$, this ratio is strictly higher than 1 for every $\lambda < 1$ and is decreasing in λ. The following is therefore true.

Corollary 3. *Anchored beliefs attenuate the response of K under instrument communication, and amplify it under target communication. Furthermore, a larger friction (lower λ) translates to larger attenuation in the first case and to larger amplification in the second case.*

This result explains how the mode of communication regulates the impact of the introduced friction on actual outcomes. When agents play a game of strategic complementarity, anchoring the beliefs of the behavior of others causes each agent to respond less than in the frictionless benchmark. When instead agents play a game of strategic substitutability, the same friction causes each agent to respond more than in the frictionless benchmark. The result then follows directly from our earlier observation that the mode of communication changes the nature of the strategic interaction.

5.3 Implementability

We now spell out the implications of the preceding observations for the combinations of τ and Y that are implementable under each mode of communication.

With instrument communication, the value τ of the instrument is pegged at $\hat{\tau}$. Condition (16) then becomes $K = \frac{1-\delta_\tau}{1-\lambda\delta_\tau}\hat{\tau}$ and condition (1) gives the outcome as

$$Y = \left((1-\alpha) + \alpha\left(\frac{1-\delta_\tau}{1-\lambda\delta_\tau}\right)\right)\hat{\tau}$$

With target communication, instead, the outcome Y is itself pegged at $\hat{Y}$. Condition (16) then becomes $K = \frac{1-\delta_y}{1-\lambda\delta_y}\hat{Y}$ and condition (1) gives the value of the instrument needed to hit the target $\hat{Y}$ as

$$\tau = \left(\frac{1}{1-\alpha} - \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-\delta_y}{1-\lambda\delta_y}\right)\right)\hat{Y}$$

Combining these findings, using the formulas for δ_τ and δ_y, and noting that the policymaker is free to choose any $\hat{\tau}$ in the first case and any $\hat{Y}$ in the second case, we reach the following result.

Proposition 4 (Implementation with anchored beliefs). *Let $\mathcal{A}_\tau$ and $\mathcal{A}_Y$ denote the sets of the pairs (τ, Y) that are implementable under, respectively, instrument communication and target communication. Then,*

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

where

$$\mu_\tau(\lambda,\gamma) \equiv \left((1-\alpha) + \alpha\frac{1-\alpha\gamma}{1-\lambda\alpha\gamma}\right)^{-1} \qquad \textit{and} \qquad \mu_Y(\lambda,\gamma) \equiv \left(1 + \frac{\alpha^2(1-\lambda)(1-\gamma)}{1+\alpha(\lambda(1-\gamma)+\alpha\gamma-2)}\right)^{-1}.$$

The frictionless benchmark is nested by $\lambda = 1$ and results in $\mu_\tau = 1 = \mu_Y$. By contrast, for any $\lambda < 1$, we have $\mu_\tau \neq \mu_Y$. That is, the two implementable sets cease to coincide as soon as we move away from the frictionless benchmark. The same is true in the variant with Level-k Thinking studied in Section 6 for any finite depth of thinking.

The next proposition, which is proved in Appendix A, offers a sharper characterization of how $\mu_\tau(\lambda)$ and $\mu_Y(\lambda)$, the slopes of the two implementability constraints, compare to one another, as well as to the frictionless counterpart.

Proposition 5 (Slope of budget lines). *(i) $\mu_\tau(\lambda,\gamma) \geq 1$ with equality only when $\lambda = 1$ or $\gamma = 0$.*

(ii) $\mu_Y(\lambda,\gamma) \leq 1$ with equality only when $\lambda = 1$ or $\gamma = 1$

(iii) $\mu_\tau(\lambda,\gamma)$ increases in λ and $\mu_Y(\lambda,\gamma)$ decreases in λ.

The belief friction under consideration has opposite effects on the slope of the "budget lines" faced by the policymaker. With instrument communication, a higher friction (smaller λ) increases the slope, meaning that a higher variation in τ is needed to attain any given variation in Y. With target communication, the opposite is true. The distortion of the implementability constraint, as measured by the absolute value of $\mu_X(\lambda) - 1$, therefore increases in both cases, but the sign is different.

5.4 Role of the GE feedback

Let us now turn attention to the role played by γ. Recall that γ proxies for the strength of the underlying GE feedback—the aggregate demand externality in the investment example of Section 2, the Keynesian income-spending multiplier in the application to monetary policy discussed in Section 7. The next proposition, whose proof can be found in Appendix A, studies how this interact with the belief friction in shaping the distortion of the implementability constraints.

Proposition 6. *Fix any $\lambda \in (0,1)$ and $\alpha \in (0,1)$. As γ increases, both $\mu_\tau(\lambda,\gamma)$ and $\mu_Y(\lambda,\gamma)$ increase. Furthermore, $\mu_\tau(\lambda,\gamma) \to \mu_{\tau,1} > 1$ and $\mu_Y(\lambda,\gamma) \to \mu_Y^* = 1$ as $\gamma \to 1$, whereas $\mu_\tau(\lambda,\gamma) \to \mu_\tau^* = 1$ and $\mu_Y(\lambda,\gamma) \to \mu_{y,0} < 1$ as $\gamma \to 0$.*

As the GE effects gets stronger (γ increases), the distortion is *exacerbated* under instrument communication, in the sense that μ_τ gets further away from μ_τ^*, whereas it is *alleviated* under target communication, in the sense that μ_Y gets closer to μ_Y^*. The logic is best illustrated by considering the extremes in which $\gamma = 0$ and $\gamma = 1$.

Consider first the case in which the GE effect is absent, or $\gamma = 0$. Behavior is pinned down purely by the direct or PE effect of the policy: $k_i = \mathbb{E}_i \tau$ for all i. As a result, announcing and committing on a value $\hat{\tau}$ for the instrument guarantees that that $K = \hat{\tau}$, regardless of λ. Condition (1) then gives $Y = \hat{\tau}$, which means that $\mathcal{A}_\tau = \mathcal{A}_\tau^*$, for all $\lambda < 1$. That is, there is no distortion with instrument communication—but there is one with target communication. For when $\gamma = 0$, target communication transforms the game played among the agents from one with a null strategic interaction to one with a non-zero strategic substitutability (indeed, $\delta_\tau = 0$ but $\delta_y < 0$ when $\gamma = 0$), thus also allowing the belief friction to influence the implementability constraint.

The converse is true when the GE effect is maximal, or $\gamma = 1$. Behavior is then pinned down exclusively by expectations of the outcome: $k_i = \mathbb{E}_i Y$ for all i. The distortion is then eliminated by, and only by, announcing and committing to a target for Y.

5.5 Optimal communication and the second best

The previous discussion implies that, in the extreme cases of $\gamma \in \{0,1\}$, the first-best outcome remains implementable under one and only one form of forward guidance: instrument communication when $\gamma = 0$, target communication when $\gamma = 1$. Each strategy, in its most favorable case, sidesteps the friction entirely by eliminating agents' need to forecast, or reason about, others' actions.

What about the intermediate cases $\gamma \in (0,1)$? Neither strategy completely eliminates the need to reason about others' behavior. With instrument communication, the agents do so to predict the outcome; with target communication, they do so to predict the value of the instrument that will be required to honor the target. The policymaker can no longer sidestep the friction.

Still, the continuity and monotonicity properties of the implementable sets with respect to γ suggest that target communication is strictly preferred to instrument communication if and only if the GE effect is strong enough. The next theorem verifies this intuition.

Theorem 1 (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 communication is strictly optimal for all θ; and when $\gamma \in (\hat{\gamma},1)$, target communication is strictly optimal for all θ.*

A detailed proof is provided in Appendix A. Below we sketch the main argument and also complete the characterization of the optimal strategy by identifying the optimal pair (τ,θ) that is implemented for any θ.

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

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

Let $(\mathcal{A}^{\text{sb}}, \tau^{\text{sb}}, Y^{\text{sb}})$ be the (unique) triplet that attains the minimum. Then, $\mathcal{A}^{\text{sb}}$ identifies the optimal mode of communication; $(\tau^{\text{sb}}, Y^{\text{sb}})$ identifies the second-best combination of the instrument and the outcome; and the communicated message is given either by $\hat{\tau} = \tau^{\text{sb}}$ or by $\hat{Y} = Y^{\text{sb}}$, depending on whether $\mathcal{A}^{\text{sb}} = \mathcal{A}_\tau$ or $\mathcal{A}^{\text{sb}} = \mathcal{A}_y$.

Given the assumed specification of L and the characterization of the implementability sets in Proposition 4, we can restate the problem as the following choice of a *slope* between τ and Y:

$$\min_{\mu\in\{\mu_\tau(\lambda),\mu_\tau(\lambda)\},(\tau,Y)\in\mathbb{R}^2} \left[(1-\chi)(\tau-\theta)^2 + \chi(Y-\theta)^2\right]$$
$$\text{s.t.} \quad \tau = \mu Y$$

Solving the constraint for Y as τ/μ, substituting this in the objective, and letting $r \equiv \tau/\theta$, we reach the following even simpler representation:

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

This makes clear that the optimal strategy is the same for all realizations of θ and lets r identify the optimal covariation of τ with θ. The policy problem reduces to choosing a value for $r \in \mathbb{R}$ and a value for $\mu \in \{\mu_\tau(\lambda), \mu_Y(\lambda)\}$. That is, if we let $(r^{\text{sb}}, \mu^{\text{sb}})$ be the solution to the above problem, the second-best values of the instrument and the outcome are given by, respectively, $\tau^{\text{sb}} = r^{\text{sb}}\theta$ and $Y^{\text{sb}} = (r^{\text{sb}}/\mu^{\text{sb}})\theta$.

Consider the "inner" problem of choosing r for given μ. The optimal r is given by

$$r(\mu) \equiv \arg\min_r \left[(1-\chi)(r-1)^2 + \chi(r/\mu-1)^2\right] = \frac{\mu^2(1-\chi)+\mu\chi}{\mu^2(1-\chi)+\chi}$$

and the resulting payoff is

$$\mathcal{L}(\mu) \equiv \min_r \left[(1-\chi)(r-1)^2 + \chi(r/\mu-1)^2\right] = \frac{\chi(1-\chi)(1-\mu)^2}{\mu^2(1-\chi)+\chi}$$

We thus have that the optimal r satisfies $r(\mu) < 1$ for $\mu < 1$, $r(\mu) = 1$ for $\mu = 1$, and $r(\mu) > 1$ for $\mu > 1$; and that the resulting payoff is a U-shaped function of $\mu \in (0,\infty)$, with a minimum equal to 0 and attained at $\mu = 1$ (the frictionless case).

How do we explain this shape? Recall that $\mu = 1$ is not feasible away from the frictionless benchmark. Instead, the policymaker has to choose either $\mu = \mu_\tau > 1$ (with instrument communication) or $\mu = \mu_Y < 1$ (with target communication). The policymaker can moderate the incurred loss by adjusting r, the responsiveness of τ to θ, away from $r = 1$, the frictionless value. Conditional on instrument communication, it is indeed optimal to choose $r > 1$, that is, to let the subsidy vary more strongly with the fundamental than in the frictionless benchmark. This offsets the attenuated response of Y to τ, which in turn helps reduces the wedge between $Y$and Y^{fb}; but since this comes at the cost of a large wedge between τ and τ^{fb}, the policymaker chooses an $r > 1$ that only partly offsets the distortion. A similar logic applies with target communication, except that now the effects flip: the policymaker chooses $r < 1$ in order to moderate the amplification effect.

Let us now turn to the optimal choice of μ, which encodes the communication choice. The magnitude of the policymaker's loss increases in the distance between μ and 1. 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.

Varying γ changes the feasible values of μ without affecting the loss incurred from any given μ. In particular, raising γ drives μ_τ further way from 1, brings μ_Y closer to 1, and leaves $\mathcal{L}(\mu)$ unchanged. It follows that $\mathcal{L}(\mu_\tau)$ is an increasing function of γ, whereas $\mathcal{L}(\mu_Y)$ is a decreasing function of it. Next, note that both $\mathcal{L}(\mu_\tau)$ and $\mathcal{L}(\mu_Y)$ are continuous in γ and recall from our earlier discussion that $\mathcal{L}(\mu_\tau) = 0 < \mathcal{L}(\mu_Y)$ when $\gamma = 0$ and $\mathcal{L}(\mu_\tau) > 0 = \mathcal{L}(\mu_Y)$ when $\gamma = 1$. It follows that there exists a threshold $\hat{\gamma}$ strictly between 0 and 1 such that $\mathcal{L}(\mu_\tau) < \mathcal{L}(\mu_Y)$ for $\gamma < \hat{\gamma}$, $\mathcal{L}(\mu_\tau) = \mathcal{L}(\mu_Y)$ for $\gamma = \hat{\gamma}$, and $\mathcal{L}(\mu_\tau) > \mathcal{L}(\mu_Y)$ for $\gamma > \hat{\gamma}$.

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 effect is strong enough.

Finally, from the above results, we have the optimal combination of τ and Y is given by $\tau = \rho(\lambda, \gamma)\theta$ and $Y = \pi(\lambda, \gamma)$, where the functions ρ and π are defined by $\rho(\lambda, \gamma) \equiv r(\mu_X(\lambda, \gamma))$ and $\pi(\lambda, \gamma) \equiv \frac{\rho(\lambda,\gamma)}{\mu_X(\lambda,\gamma)}$ with $X = \tau$ for $\gamma < \hat{\gamma}$ and $X = Y$ for $\gamma > \hat{\gamma}$. It is then straightforward to check that ρ is higher than 1 and increasing in γ for $\gamma \in (0, \hat{\gamma})$, has a downward discontinuity at $\gamma = \hat{\gamma}$, and is lower than 1 and increasing in γ for $\gamma \in (\hat{\gamma}, 1)$; and that the converse properties hold for π. The discontinuity at $\gamma = \hat{\gamma}$ reflects the switch from the one form of forward guidance to the other. When $\gamma < 1$, the policymaker engages in instrument communication, the friction causes attenuation, and the optimal policy moderates the distortion by letting τ move less than one to one with θ. When instead $\gamma > 1$, the policymaker engages in target communication, the friction causes amplification, and the optimal policy moderates the distortion by letting τ move less than one to one with θ.

5.6 Comparative statics

Because the model is highly tractable, we can characterize the dependence of the optimal communication strategy on all model parameters.[13]

The effect of χ is obvious: raising the policymaker's concern about the output gap expands the range of γ for which target communication is optimal.

Consider next α. As α approaches 1, τ has a vanishingly small effect on Y for given K. The policymaker may therefore need to make very large adjustments in τ to hit a stated target for Y. This explains why target communication becomes less desirable as α increases.

Finally, consider a change in λ. For any given (α, γ), raising the belief friction (lowering λ) intensifies the distortion under both modes of communication. As shown next, however, the additional friction "bites harder" with target communication than under instrument communication:

Proposition 7. *For fixed (α, χ), the threshold $\hat{\gamma}$ is a decreasing function of λ. That is, the range of γ for which target communication is optimal increases as the friction gets smaller.*

Proof. See Appendix A.6. □

Furthermore, as the friction vanishes, the threshold $\hat{\gamma}$ has a well-defined limit given by

$$\lim_{\lambda \uparrow 1} \hat{\gamma} = \frac{1}{2-\alpha} \in \left(\frac{1}{2}, 1\right)$$

Whereas *exact* rational expectations (nested as $\lambda = 1$) leaves optimal communication indeterminate, *near* rational expectations (i.e., λ arbitrarily close to, but strictly lower than, 1) gives a non-trivial result. Put differently, a policymaker with small uncertainty—in either a Bayesian or a Knightian sense—about the parameter λ around the benchmark $\lambda = 1$ may reach qualitatively similar conclusions to the policymaker who is confident that the friction is large.

6 Level-k Thinking

The key mechanism in the previous section is agents' under-forecasting of others' responses to an announcement: as seen from conditions (15) and (16), $\bar{\mathbb{E}}[K]$ moves less than K in response to variation in $\hat{X}$. One could recast this as the consequence of agents' bounded ability to calculate others' responses or to comprehend the GE effects of the policy.

A simple formalization of such cognitive or computational bounds is Level-k Thinking. This concept represents a relaxation of the part of Assumption 2 that imposes common knowledge of rationality: agents are allowed to question the rationality of others. In particular, this concept is defined recursively by letting the level-0 agent make an exogenously specified choice (this is the completely irrational agent), the level-1 agent play optimally given the belief that others are level-0 (this agent is

[13] All the properties stated here are verified formally in Appendix (A.6).

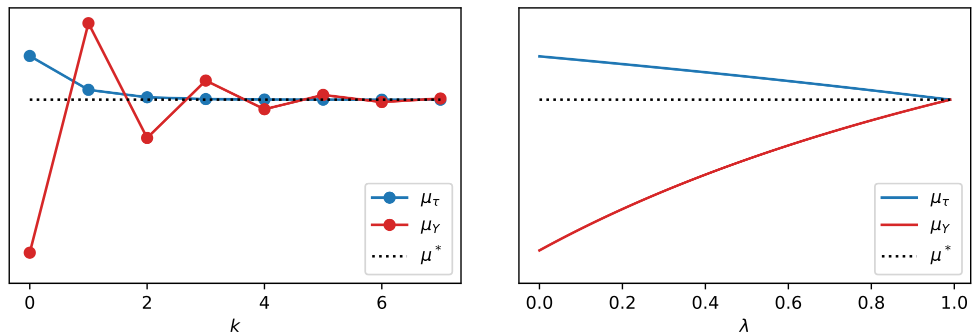

Figure 1: The implementability coefficients (μ_τ, μ_Y) under Level-k Thinking (left) and anchored beliefs (right).

rational but believes that others are irrational), the level-2 agent play optimally given the belief that others are level-1, and so on, up to some finite order k.

To see the implications of this concept in our context, assume all agents think to the same order $k \geq 1$ and let the "base case" (level-0) correspond to $K = 0$. Because every agent believes that all other agents are of cognitive order $k-1$, the expectation of K is now given by[14]

$$\bar{\mathbb{E}}[K] = \left((1-\delta_X) \sum_{h=0}^{k-1} \delta_X^h \right) \hat{X} \tag{17}$$

Comparing this expression to (13), which gave expected investment as a function of higher-order beliefs about X, reveals that Level-k Thinking is isomorphic to the following belief hierarchy about the policy announcement:

$$\bar{\mathbb{E}}^h[X] = \hat{X}, \ \ \forall h < k \qquad \text{and} \qquad \bar{\mathbb{E}}^j[X] = 0, \ \ \forall h \geq k$$

That is, it is *as if* agents know that others know that... others have heard the announcement only up to order k;beyond that order, beliefs are pegged at zero.

This is similar to the structure of higher-order beliefs considered in Section 5. Both approaches allow the announcement's effect on the h-th order belief to decay with h. Before, the decay was exponential in h; now it is a step function jumping from 1 to 0 at the specific order $h = k$.

This similarity suggests that the lessons derived earlier extend to Level-k Thinking. Indeed, $k = 1$ corresponds exactly to $\lambda = 0$ in our earlier analysis. Furthermore, for any odd number $k \geq 3$, one can find a $\lambda \in (0, 1)$ such that the implementability sets under Level-k Thinking coincide with those in our earlier analysis (see Appendix C for the exact construction).

The equivalence, however, breaks down for even k because of the concept's "oscillatory" behavior in games of strategic substitutability. Consider the game following target communication. For any given announcement, an agent wants to invest more when he expects others to investment less. Because the level-0 agent is assumed to be completely unresponsive, a level-1 agent expects K to move

[14] The formula applies for $k \geq 2$; for $k = 1$, $\bar{\mathbb{E}}[K] = 0$.

less than in the frictionless benchmark and thus moves *more*. A level-2 agent expects K to move *more* than in the frictionless benchmark and therefore chooses to move *less* himself. Whereas $k = 1$ amplifies the actual response of investment, $k = 2$ attenuates. The left panel of Figure 1 shows that this oscillatory pattern continues for higher k, and that this oscillation with target communication is the only qualitative difference between the present specification and that studied in Section 5.

We find this oscillatory, non-monotone pattern to be conceptually unappealing and suspect that it is an unintended consequence of a particular formalization that was developed in the experimental literature for games of complements, but may not be ideal for games of substitutes. Seen from this perspective, our heterogeneous-prior formalization captures the essence of Level-k Thinking while bypassing this "pathological" feature. In Appendix C, we show that the same goal can be achieved with a "smooth" version of Level-k Thinking along the lines of Garcıa-Schmidt and Woodford (2018). With these qualifications in mind, these alternative approaches to bounded rationality are essentially interchangeable in our model.

7 Forward Guidance for Monetary Policy

Consider the question of how aggregate demand responds to forward guidance when the latter takes the form of a unconditional commitment for keeping interest rates low after the economy has exited a liquidity trap. This question has already been addressed by Angeletos and Lian (2018), Garcıa-Schmidt and Woodford (2018), and Farhi and Werning (2016).[15] Our paper inverts the question: should the policymaker engage in this type of forward guidance, or should she instead commit to a credible target for GDP and unemployment?

Even though our framework is too stylized to nest the New Keynesian model, our results suggest that the answer to the above question depends critically on the strength of the underlying GE feedback mechanisms. But what are these mechanisms and what determines their strength?

Three such mechanisms are at work in the context of the baseline New Keynesian model: the positive feedback between aggregate income and aggregate spending, or the Keynesian cross, which underlies the Dynamic IS curve; the dynamic strategic complementarity in the firms' price-setting decisions, which underlies the New Keynesian Philips cure; and the inflation-spending feedback that is captured by the interaction of the Dynamic IS curve and the New Keynesian Philips cure.[16]

All these effects are purely forward-looking: the feedback goes in one direction from expectations of future outcomes to current behavior. Adding capital, habit persistence or commitments in con-

[15] The first paper models the friction as higher-order uncertainty in a rational-expectations, common-prior setting, the other two model it as Level-k Thinking. In line with our results regarding instrument communication, these papers find that the belief distortion reduces the power of the aforementioned type of forward guidance.

[16] Del Negro, Giannoni and Patterson (2015) and McKay, Nakamura and Steinsson (2016) contain detailed explanations of these mechanisms, as well as resolutions of the forward guidance puzzle outside the realm of beliefs. Angeletos and Lian (2018), on the other hand, develop a game-theoretic representation of the micro-founded New Keynesian model that roughly maps to the more abstract framework used here.

sumption, or wealth effects can introduce an opposite-direction feedback from current behavior to future outcomes. The combination delivers a two-way interaction similar to that stylized by our abstract framework, with K corresponding to current aggregate spending and Y to economic conditions in the not-so-far future.

Finally, building on the results of Angeletos and Lian (2018), one may expect the combined effect of all these mechanisms, proxied by γ in our framework, to increase with the horizon of forward guidance. The reason is that longer horizons entail longer chains of dynamic feedback effects. The results of Farhi and Werning (2016), on the other hand, suggest that the effective γ may increase with liquidity constraints, insofar as such constraints map to a large income-spending multiplier.

Combining these insights with the results of our paper suggests the following policy implication. Consider a situation in which the liquidity trap is expected to be sufficiently long and/or the Keynesian cross is sufficiently steep. It is precisely in this situation that traditional forward guidance is severely constrained, as argued by the aforementioned papers. But it is also then that the policymaker may bypass the friction by communicating, and committing to, a path for future employment and GDP rather than a path for the policy rate.[17]

We corroborate these intuitions in Appendix B.2, within the context of a stylized New Keynesian economy. We take a few shortcuts in order to keep the analysis tractable and to nest the economy to our abstract framework. These shortcuts do not necessarily drive our results, but preclude a quantitative evaluation or a richer understanding of the determinants of the optimal forward guidance. The micro-foundation of the welfare objective is another task left for future investigation.[18]

8 Erratic Beliefs

The analysis so far has abstracted from the possibility that bounded rationality is the source of random shifts in "market psychology" or of random mistakes in the agents' reasoning about the GE consequences of the policy. We now capture this possibility by considering the following, different relaxation of Assumption 2.

Assumption 4 (Erratic beliefs)**.** *Every agent believes that the other agents are rational but worries that a fraction $1-\sigma$ of them receives a randomly distorted message and is unaware of the distortion. In particular, every i believes that, for every $j \neq i$, $\mathbb{E}_j[X] = \hat{X}$ with probability σ and $\mathbb{E}_j[X] = \hat{X} + \varepsilon$ with probability $1-\sigma$, where $\sigma \in (0,1)$ is fixed scalar and ε is a random shock, drawn from a Normal distribution with mean zero and variance one, orthogonal to θ, and unobserved by the policymaker at the moment of his announcement. These facts and the value of σ are common knowledge.*

[17] We let the reader decide whether this is good news for policymakers, in the sense that the appropriate form of communication can indeed contain the friction under consideration, or bad news for the aforementioned papers, in the sense that the forward guidance puzzle has been translated to a different dimension rather than been truly resolved.

[18] Our analysis also abstracts from shocks, or constraints, that may interfere with the policymaker's control of the targeted outcome. As explained in Appendix F, such considerations tilt the balance in favor of instrument communication for any given γ, but do not affect our result that a higher γ (stronger GE effect) increases the relative value of target communication.

To understand what this assumption does, recall our earlier characterization of the expectations of K in terms of the higher-order beliefs of X:

$$\mathbb{E}_i[K] = \mathbb{E}_i\left[(1-\delta_X)\sum_{k=1}^{\infty}(\delta_X)^{k-1}\bar{\mathbb{E}}^k[X]\right] \tag{18}$$

In the scenario studied in Section 5, Assumption 3 allows the higher-order beliefs to move less than one-to-one with first-order beliefs ($\bar{\mathbb{E}}^k[X] = \lambda^{k-1}\hat{X}$), but rules out any orthogonal variation in the gap between first- and higher-order beliefs, thus giving $\mathbb{E}_i[K] = b\hat{X}$ for some $b < 1$. This captures "anchored beliefs" but rules out "erratic beliefs." The scenario studied now does the opposite: it lets ε drive random variation in higher-order beliefs and thereby in the response of $\mathbb{E}_i[K]$ to $\hat{X}$.

Indeed, because the typical agent believes that only a fraction σ of the population heard the actual message $\hat{X}$, whereas the remaining fraction heard the distorted message $\hat{X}+\varepsilon$, her second-order belief is given by $\mathbb{E}_i[\bar{\mathbb{E}}[X]] = \hat{X} + (1-\sigma)\varepsilon$. By induction, the h-th order average belief is given by

$$\bar{\mathbb{E}}^h[X] = \hat{X} + a_h\varepsilon \tag{19}$$

with $a_1 = 0$ and $a_h = \sigma a_{h-1} + (1-\sigma)$ for $h \geq 2$. Finally, combining the above with (18) yields

$$\mathbb{E}_i[K] = \hat{X} + \frac{1-\sigma}{1-\sigma\delta_X}\varepsilon, \tag{20}$$

which verifies that ε introduces waves of optimism and pessimism about the response of others to the policy announcement.

Similar formulations have been used before by Angeletos, Collard and Dellas (2018) and Huo and Takayama (2015) to study, and quantify, the role of "sentiments" in business cycles. One can think of the present setting as a stylized version of the richer models used in those papers. Similar belief fluctuations can also be produced in an extension of the Level-k setting of Section 6 that features a random default (level-0) point or a random cognitive order k. In this sense, the present setting help capture random mistakes in how agents reason about the behavior of others, or the policymaker's uncertainty about the agents' sophistication and depth of reasoning. We welcome all these interpretations.

How does the form of forward guidance interact with such erratic beliefs? Using condition (20) in the best-response condition (10), we get

$$K = \hat{X} + \frac{\delta_X(1-\sigma)}{1-\sigma\delta_X}\varepsilon, \tag{21}$$

from which it is evident that the form of forward guidance regulates not only the magnitude but also the sign of the effect of ε on K.[19] Proceeding in a similar manner as in Section 5, we then reach the following characterization of the implementability constraints faced by the policymaker.

[19]With instrument communication, the agents play a game of strategic complementarity and optimism about the choices of others feeds to more investment. With target communication, the agents play a game of strategic substitutability and the same optimism feeds to less investment. This echoes the differential effect of anchored beliefs documented earlier on.

Proposition 8. *A pair (τ, Y) is implementable if and only if*

$$\tau = Y + \psi_X \varepsilon$$

where $X \in \{\tau, Y\}$ indexes the mode of communication and where

$$\psi_\tau \equiv -\frac{\alpha^2\gamma(1-\sigma)}{1-\sigma\alpha\gamma} \leq 0 \qquad \text{and} \qquad \psi_Y \equiv \frac{\alpha(1-\alpha)(1-\gamma)(1-\sigma)}{1-\alpha(1-\sigma(1-\gamma)))} \geq 0.$$

In the scenario with anchored beliefs, the mode of communication regulated the slope of the implementability restriction between τ and Y. In the present scenario, this slope is pegged to 1, as in the REE benchmark, but the implementability constraint is perturbed away from that benchmark by the shock ε. The mode of communication now regulates the impact of this shock.

The next result sheds further light on the implementability constraints faced by the policymaker by studying the comparative statics of ψ_τ and ψ_Y with respect to γ.

Proposition 9. *(i) ψ_τ is non-positive and strictly decreasing in γ, and equals zero at $\gamma = 0$.*
(ii) ψ_Y is non-negative and strictly decreasing in γ, and equals zero at $\gamma = 1$.

These properties have a similar flavor as those documented in the case with anchored beliefs: a stronger GE feedback reduces the impact of erratic beliefs under target communication, and increases it under instrument communication. These properties indeed guarantee that our take-home lesson extends from the case with anchored beliefs to the case studied here.

Theorem 2. *For any $\sigma > 0$, there exists a threshold $\tilde{\gamma} \in (0,1)$ such that the following is true: for $\gamma \in [0, \tilde{\gamma})$, instrument communication is strictly optimal for all realizations of θ; and for $\gamma \in (\tilde{\gamma}, 1]$, target communication is strictly optimal for all realizations of θ.*

The common thread behind this result and the corresponding result for the case with anchored beliefs (Theorem 1) is how the GE feedback and the mode of communication interact in shaping the nature and the strength of the strategic interaction among the agents. The sharpest possible version of this point is made by considering, once again, the extremes in which $\gamma = 0$ and $\gamma = 1$. When $\gamma = 0$, instrument communication eliminates the impact of *either* kind of belief distortion simply by guaranteeing that the behavior of each agent is independent of her beliefs of the behavior of other agents. When instead $\gamma = 1$, the exact same thing is achieved by target communication. Finally, in between these two extremes, the impact of either distortion is non-zero under both modes of communication, but the basic logic survives in the sense that a higher γ tilts the balance in favor of target communication.

9 Other Policy Strategies

So far, we have focused on two policy options: "simple" forward guidance about either the policy instrument or the policy target. We now broaden the scope to a larger toolkit for managing expectations.

First, we show that attempting forward guidance about either the aggregate action K or the policy parameter θ is ineffective—or, at least, ill-posed—in our model. Next, we show that the consider forms of forward guidance may be dominated by a more sophisticated one that has the policymaker commit to a linear policy rule that links the instrument with the outcome. Such a strategy allows the policymaker substantially more flexibility in determining the relationship *between* different expectations, which in turn helps further contain any belief distortion. Our discussion thus provides a new perspective on the function that policy rules, including Taylor rules for monetary policy, may serve in the presence of bounded rationality or higher-order uncertainty.

9.1 Communicating θ or K

Our initial focus on communicating τ or Y seemed natural for applications. But, for completeness, we should also check whether it would be wiser either to communicate directly the realized value of θ, or to commit to a target for the aggregate action K.

Consider the first scenario. In this scenario, the policymaker is picking, and committing on, a mapping from θ to τ or Y, but does not tell this mapping to the agents. Instead, she only tells them what θ is. In other words, the policymaker tells the agents what he would like to achieve, but not the way she is going after it.

As already noted, such communication implements the first best under rational expectations. Because REE imposes a unique mapping from θ to both τ and Y, and the agents know that mapping, there is no need for the policymaker to communicate it. Away from that benchmark, however, many such mappings can be part of an equilibrium and, as a result, communicating merely θ does not necessarily pin down the agents' beliefs about either the policy or the outcome. In particular, there exists an equilibrium that replicates instrument communication, as well as an equilibrium that replicates target communication.

Consider next the scenario in which the policymaker communicates a target for K. This option may be impractical if K stands for a complex set of decisions that is hard to measure. But even abstracting from such measurement issues, this option may not be viable—or at least it is not well-posed in our model. Consider in particular the specification studied in Section 5 and let the policymaker announce and commit to a value $\hat{K}$ for aggregate investment. In Appendix E, we show that there exists a system of beliefs about τ and Y that is consistent with the belief that K will equal $\hat{K}$ if and only if $\lambda = 1$ (i.e., rational expectations). When instead $\lambda < 1$, there does not exist an equilibrium in which the policymaker fixes a target $\hat{K}$ for aggregate investment. The reason is that, unlike in the case of a Y target, the policymaker has not have the power to persuade the agents that she can attain a K target "no matter what."

We alluded to this kind of problem when we noted the necessity of letting τ have a direct, mechanical effect on Y (or the fragility of target communication as $\alpha \to 1$). The same basic logic applies. To make sense of commitments on K, we would have to add a new policy instrument that can *directly* control the investment decisions of the firms. That is, we would have to modify (2) to

$$k_i = (1-\alpha)\mathbb{E}_i[\tau] + \alpha\mathbb{E}_i[Y] + z,$$

where z is the new policy instrument. But this could bypass the issue of interest: instead of trying to influence K by manipulating the expectations of τ and Y, the policymaker could just use z to directly control K regardless of these expectations.

Put differently, it is precisely the absence of such an instrument that justifies the focus on "managing expectations." In the context of forward guidance studied in Section 7, this simply means the following: if the central bank could use current interest rates (the analogue of z) to control aggregate demand, there would naturally be less need for engaging in forward guidance of any type.

9.2 Policy rules

The choice between instrument and target communication remains a choice of "extremes." One could imagine a more sophisticated strategy in which the policy maker announces and commits to a policy rule of the following type:

$$\tau = A - BY \tag{22}$$

where (A, B) are free parameters. In the context of monetary policy, of course, this expression is a familiar Taylor rule.

Instrument communication can then be nested with $B = 0$ and $A = \hat{\tau}$, for arbitrary $\hat{\tau}$; and target communication can be though as the limit in which $B \to \infty$ and $A/B \to \hat{Y}$, for arbitrary $\hat{Y}$. Away from these two extremes, the policymaker's strategy is indexed by the pair (A, B) and policy communication amounts to the announcement of this pair, as opposed to a fixed value for either τ or Y.

For reasons outside our model, such feedback rules may be hard for the agents to comprehend and may therefore be less effective than the two extremes considered so far. We suspect that, in many real-world situations, there is a gain in conveying a sharp policy message of the form "we will keep interest rates at zero for 8 quarters" or "we will do whatever it takes to bring unemployment down to 4%," as opposed to communicating a complicated feedback rule. This explains why we a priori found it more interesting to focus on the two extremes.

Having said that, it is useful to explore how such policy rules work within our model. The key insights survive and, in fact, their scope expands: once one deviates from rational expectations, such policy rules play a function not previously identified in the literature and akin to that identified in the preceding analysis.

Consider first the rational expectations benchmark (as in Section 3). In this benchmark, the additional flexibility afforded by this class of policy rules is entirely useless, because the first best was already attained by the two extremes. Furthermore, our earlier irrelevance result directly extends: not only for the first best, but also for any other point in $\mathcal{A}^*$, there exist a continuum of values for (A, B) that implement it as part of an REE. The only subtlety worth mentioning is that such an REE may fail to be the unique equilibrium if $B < -1$. The logic is similar to the one underlying the Taylor principle.

To understand these properties, solve (22) and (1) jointly for τ and Y and substitute the solution into (2) to obtain the following game representation:

$$k_i = \zeta(A, B; \alpha, \gamma) + \delta(B; \alpha, \gamma)\mathbb{E}_i[K] \tag{23}$$

where

$$\zeta(A, B; \alpha, \gamma) \equiv \frac{(1 - \alpha\gamma)A}{1 + (1 - \alpha)B} \qquad \text{and} \qquad \delta(B; \alpha, \gamma) \equiv \frac{\alpha(\gamma - B(1 - \gamma))}{1 + (1 - \alpha)B}.$$

It is then evident that B controls the slope of the best responses and A their intercept. When $B < -1$, the policy induces a game of strategic complementarity in which the slope exceeds 1, opening the door to multiple equilibria. When instead $B \in (-1, \frac{\gamma}{1-\gamma})$, the slope is positive but less than one. And when $B > \frac{\gamma}{1-\gamma}$, the slope becomes negative, which means that the policy rule induces a game of strategic complementarity. Finally, it is clear that, for any value of K, there exist a continuum of (A, B) that induces this K as the fixed point of (23).

Consider now the case with anchored beliefs (as in Section 5). The extra flexibility afforded by the policy rules now becomes relevant: by varying A and B, the planner can induce a wide range of outcomes beyond those contained in $\mathcal{A}_\tau$ and $\mathcal{A}_Y$. What is more, there actually exist a subclass of policy rule that replicates $\mathcal{A}^*$, namely the set of outcomes that are attained under rational expectations. This subclass is given by setting B such that $\delta(B; \alpha, \gamma) = 0$, or equivalently $B = \frac{\gamma}{1-\gamma}$, and letting A vary in $\mathbb{R}$. Intuitively, setting B so that $\delta(B; \alpha, \gamma) = 0$ completely eliminates the need for the agents to forecast, or calculate, the behavior of others, which in turn guarantees that the distortion on the set of implementable vanishes regardless of λ. By varying A, the policymaker can then span the set $\mathcal{A}^*$. And by picking A so that $\zeta(A, B; \alpha, \gamma) = \theta$, she can implement the first best.[20]

We summarize these lessons in the following result.

Proposition 10. *Suppose that the policymaker can announce and commit on a policy rule as in (22) and let Assumptions 1 and 3 hold with $X = (A, B)$.*

When $\lambda = 1$ (rational expectations), the first best is implemented with any (A, B) such that $B > -1$ and $A = (1 + B)\theta$.

When instead $\lambda < 1$ (anchored beliefs), the first best is implemented if and only

$$B = \frac{\gamma}{1 - \gamma} \qquad \textit{and} \qquad A = \frac{\theta}{1 - \gamma}.$$

At first glance, this result may appear to dilute our take-home message: a more sophisticated strategy than the ones studied in the main body of our paper completely eliminates the problem. However, this property is fragile in the following sense. When the policymaker is uncertain about the structure of the economy, in particular about the values of γ, the values of B and A obtained above are also uncertain. The first best is therefore unattainable when $\lambda < 1$, even though it remains attainable under rational expectations.

[20]Clearly, this logic extends to the variants with Level-k Thinking and erratic beliefs.

Most importantly, our take-home message survives in the following two keys senses. First, the optimal strategy is indeterminate under rational expectations ($\lambda = 1$), whereas it is determinate with anchored beliefs ($\lambda < 1$). And second, for any $\lambda < 1$, a stronger GE effects calls for a policy rule that has a steeper slope with respect to Y and, in this sense, looks closer to target communication. In fact, in the limit as $\gamma \to 1$, the optimal policy rule has $B \to -\infty$ and $B/A \to \theta$, which is the same as the target communication with $\hat{Y} = \theta$.

We thus interpret Proposition 10 as a complement to our main analysis, not a sign that the choice between instrument and target communication was too narrowly framed. Proposition 10 also offers a new perspective on Taylor rules. The pertinent literature has focused on two functions: how the slope of the Taylor rule can induce a unique equilibrium; and how it must be designed if the policymaker cannot directly condition the intercept of the Taylor rule on the underlying fundamentals. The first issue maps to our discussion above about setting $B > -1$ as is know as the Taylor principle. The second issue is a modern variant of Poole (1970). Our own result brings up a completely different function: the role of such rules in regulating the distortionary effects of bounded rationality.

This function extends to common-prior settings that maintain rational expectations but allow for higher-order uncertainty. This is because policy rules that regulate the agents' strategic interaction also regulate the impact that any "belief wedge" (any gap between first- and higher-order beliefs) has on actual outcomes regardless of whether this wedge represents a departure from rational expectations or a rich enough informational friction. We view this point as another facet of the insights developed in the earlier sections of our paper.

10 Conclusion

What is the best way to manage expectations? Should a policymaker announce and commit to the intended value of the available policy instrument, such as the Federal Funds rate, or the target for the relevant economic outcome, such as employment?

We pose this question in a stylized model in which agents form mis-specified beliefs, either anchored to a reference point or subject to erratic impulses. Our main result is a sharp dependence of the optimal communication strategy on the GE feedback between aggregate outcomes and individual actions. Fixing outcomes instead of instruments is optimal if and only if this feedback is sufficiently high, as in a model of high aggregate demand externalities or a steep Keynesian cross.

Why? Instrument communication pins down expectations of the policy instrument itself, but leaves agents to predict, or reason about, the determination of aggregate outcomes. Target communication does the opposite, leaving agents to predict what policy will support the announced outcome. Which strategy is preferred depends on the relative cost of mistakes for each type of reasoning. High GE feedbacks, which make outcome expectations more essential for decisions (and associated mistakes more costly), tilt the balance toward directly communicating those outcomes.

Put more succinctly, the optimal form of forward guidance minimizes agents' need to "reason about the economy" precisely because this reasoning produces distortions.

Along the way, we uncovered additional insights, such as how Taylor rules can play a new role in regulating the bite of bounded rationality, or how the latter may itself be the source of a commitment problem. In all these cases, our analysis suggested interesting trade-offs but remained too stylized to give fully satisfying answers. We also took for granted the desirability of minimizing the distance of the equilibrium outcomes from their rational-expectations counterparts. But one could imagine situations with one distortion offsetting another— for instance, anchored beliefs offsetting financial amplification. Each of these issues merits a more complete investigation.

References

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

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

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 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, Fabrice Collard, and Harris Dellas. 2018. "Quantifying Confidence." *Econometrica*, 86(5): 1689–1726.

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

Ben Bernanke. 2015. "Inaugurating a New Blog." https://www.brookings.edu/blog/ben-bernanke/2015/03/30/inaugurating-a-new-blog/, Accessed: 2018-12-10.

Benhabib, Jess, Pengfei Wang, and Yi Wen. 2015. "Sentiments and Aggregate Demand Fluctuations." *Econometrica*, 83(2): 549–585.

Bergemann, Dirk, and Stephen Morris. 2013. "Robust Predictions in Games with Incomplete Information." *Econometrica*, 81(4): 1251–1308.

Bergemann, Dirk, and Stephen Morris. 2018. "Information Design: A Unified Perspective." *Cowles Foundation Discussion Paper No. 2075R3*.

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.

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, Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert. 2018. "Do You Know That I Know That You Know...? Higher-Order Beliefs in Survey Data." *NBER Working Paper No. 24987*.

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 mimeo*.

Evans, Georges, and Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics.* Princeton, NJ:Princeton University Press.

Farhi, Emmanuel, and Ivan Werning. 2016. "Fiscal Multipliers: Liquidity Traps and Currency Unions." *Handbook of Macroeconomics*, 2: 2417–2492.

Farhi, Emmanuel, and Ivan Werning. 2017. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *NBER Working Paper No. 23281*.

Gabaix, Xavier. 2017. "A Behavioral New Keynesian Model." *Harvard mimeo*.

Garcıa-Schmidt, Mariana, and Michael Woodford. 2018. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review*, forthcoming.

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

Huo, Zen, and Naoki Takayama. 2015. "Higher Order Beliefs, Confidence, and Business Cycles." Yale University miméo.

Inostroza, Nicolas, and Alessandro Pavan. 2018. "Persuasion in Global Games with Application to Stress Testing." *Northwestern University mimeo*.

Iovino, Luigi, and Dmitriy Sergeyev. 2017. "Quantitative Easing without Rational Expectations." *Work in progress*.

James, Jonathan G., and Phillip Lawler. 2011. "Optimal Policy Intervention and the Social Value of Public Information." *American Economic Review*, 101(4): 1561–74.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Marcet, Albert, and Juan P. Nicolini. 2003. "Recurrent Hyperinflations and Learning." *American Economic Review*, 93(5): 1476–1498.

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.

Morris, Stephen, and Hyun Song Shin. 2006. "Inertia of Forward-looking Expectations." *The American Economic Review*, 152–157.

Morris, Stephen, and Hyun Song Shin. 2007. "Optimal Communication." *Journal of the European Economic Association*, 5(2-3): 594–602.

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.

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

Sargent, Thomas J. 1993. *Bounded rationality in macroeconomics.* Oxford University Press.

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

Svensson, Lars E. O. 2006. "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con." *American Economic Review*, 96(1): 448–452.

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 mimeo*.

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.

A Proofs

A.1 Slopes of Budget Lines

A.1.1 Proof of Proposition 5

The relationship between action K and announcement $\hat{X}$, as derived in the main text, is the following:

$$K = \frac{1-\delta_X}{1-\lambda\delta_X}\hat{X}$$

Instrument communication. As shown in Proposition 5,

$$\mu_\tau = \left((1-\alpha) + \alpha\frac{1-\delta_\tau}{1-\lambda\delta_\tau}\right)^{-1} \tag{24}$$

Clearly, for $\delta_\tau \equiv \alpha\gamma \in (0,1)$, as implied by $\gamma \in [0,1]$ and $\alpha \in (0,1)$, $(1-\delta_\tau)/(1-\lambda\delta_\tau) \in [0,1]$ and $\mu_\tau^{-1} \in [0,1]$ and $\mu_\tau \geq 1$.

Further, $\partial\mu_\tau^{-1}/\partial\lambda > 0$ given $\delta_\tau \in (0,1)$ and $\partial\mu_\tau/\partial\lambda = -(\mu_\tau)^{-2}\partial\mu_\tau^{-1}/\partial\lambda < 0$.

When $\delta_\tau < 0$, we can have $\mu_\tau < 1$. A sufficient condition for this is $\gamma < 0$, or negative GE feedback.

Target communication. Let b denote the responsiveness of the action to the announcement, $\partial K/\partial\hat{Y}$. In general, the slope of the implementability constraint is

$$\mu_Y = \frac{1-\alpha b}{1-\alpha} = \frac{1-\lambda\delta_y - \alpha(1-\delta_y)}{(1-\alpha)(1-\lambda\delta_y)} \tag{25}$$

Given that $\delta_y \leq 0$, we know that $b \geq 1$ and hence $\mu_Y \leq 1$.

To check the derivative with respect to λ, note that

$$\frac{\partial b}{\partial\delta_y} = -\frac{\delta_y(\delta_y - 1)}{(1-\lambda\delta_y)^2} > 0$$

and $\partial\delta_y/\partial\gamma = \alpha/(1-\alpha) > 0$ and $\partial\mu_Y/\partial b = -\alpha/(1-\alpha) < 0$. Thus, by the chain rule, $\partial\mu_Y/\partial\gamma < 0$.

A.1.2 Further results

Lemma 4 (Sign of μ_Y). *$\mu_Y > 0$ if and only if $\lambda \geq \alpha$ or $\gamma > \frac{1+\alpha(\lambda-2)}{\alpha(\lambda-\alpha)}$.*

Proof. Note that $\mu_Y \in [0,1]$ when $b \in [1, 1/\alpha]$ and $\mu_Y < 0$ when $b > 1/\alpha$. This reduces to to

$$\gamma\alpha(\lambda-\alpha) < 1 - \alpha(2-\lambda)$$

Let's consider three cases of this. First, assume that $\lambda > \alpha$. Some algebraic manipulation yields the condition

$$\gamma < 1 + \frac{(1-\alpha)^2}{\alpha(\lambda-\alpha)}$$

which is obviously true for any $\gamma < 1$. Thus no more restrictions are required.

Next, consider $\lambda = \alpha$. The condition becomes

$$\alpha(2-\alpha) < 1$$

which is always true for $\alpha = \lambda \in (0,1)$.

Finally, consider $\lambda < \alpha$. In this case, the condition is

$$\gamma > \frac{1+\alpha(\lambda-2)}{\alpha(\lambda-\alpha)}$$

Note that the right-hand-side is less than 0 if $\lambda > 2 - \frac{1}{\alpha}$. Hence we used this as a sufficient condition for $\mu_Y > 0$ for all $\gamma \geq 0$. □

Lemma 5. *Assume that $\mu_Y > 0$ and $\alpha\gamma < 1$. Then $\mu_\tau > \mu_Y$.*

Proof. As long as $\mu_Y > 0$, we can show that $\mu_\tau > \mu_Y$. Written out in terms of parameters, this condition is:

$$\frac{1-\lambda\alpha\gamma}{(1-\alpha)(1-\lambda\alpha\gamma)+\alpha(1-\alpha\gamma)} \geq \frac{1+\frac{\lambda\alpha(1-\gamma\}}{1-\alpha}-\alpha\frac{1-\alpha\gamma}{1-\alpha}}{1-\alpha+\lambda\alpha(1-\gamma)}$$

Given that $\mu_Y > 0$, the left denominator is positive. The other three terms are necessarily positive. Thus an equivalent statement, after cross-multiplying, is the following:

$$(1-\lambda\alpha\gamma)(1-\alpha+\lambda\alpha(1-\gamma)) \geq \left((1-\lambda\alpha\gamma)+\frac{\alpha(1-\alpha\gamma)}{1-\alpha}\right)(1-\alpha+\lambda\alpha(1-\gamma)-\alpha(1-\alpha\gamma)$$

Subtracting like terms from each side, and dividing by $\alpha > 0$, yields the following condition:

$$(1-\lambda)(1-\alpha\gamma) \geq 0$$

Hence $\lambda < 1$ and $\alpha\gamma < 1$ are a sufficient condition for $\mu_\tau > \mu_Y$, and either $\lambda = 1$ or $\alpha\gamma = 1$ are a sufficient condition for $\mu_\tau = \mu_Y$. □

A.2 Proof of Proposition 6

Limit cases. At $\gamma = 1$, the slope given instrument communication is

$$\mu_{\tau,1} = \left((1-\alpha)+\alpha\frac{1-0}{1-\lambda\cdot 0}\right)^{-1}$$
$$= \frac{1}{1-\alpha} > 1$$

Meanwhile, the slope with target communication is

$$\mu_{y,1} = 1$$

At the other extreme $\gamma = 0$, the slope given target communication is

$$\mu_{y,0} = \frac{1-\alpha\frac{1-\lambda}{1-\alpha}}{1-\alpha(1-\alpha)}$$

This is less than one if and only if $1-\alpha < (1-\lambda)/(1-\alpha) < \alpha^{-1}$ or $(1-\alpha)^2 < 1-\lambda < (1-\alpha)\alpha$. This is implied by the arguments of Proposition 5.

With instrument communication at $\gamma = 0$, the slope is $\mu_{\tau,1} = ((1-\alpha)+\alpha\cdot 1)^{-1} = 1$.

Derivative of μ_τ with respect to γ. For fixed λ, we can calculate first a derivative of the inverse slope with respect to the interaction parameter

$$\frac{\partial\mu_\tau^{-1}(\lambda)}{\partial\delta_\tau} = -\frac{\alpha(1-\lambda)}{(1-\lambda\gamma)^2}$$

which is unambiguously negative for $\lambda < 1$. The interaction parameter $\delta_\tau := \alpha\gamma$ increases with γ. Thus, by the chain rule, $\partial\mu_\tau/\partial\delta_\tau = -(\mu_\tau)^{-2}(\partial\mu_\tau^{-1}/\partial\delta_\tau)(\partial\delta_\tau/\partial\gamma) > 0$.

Derivative of μ_Y with respect to γ. For fixed λ, the partial derivative with respect to interaction δ_y is

$$\frac{\partial\mu_Y(\lambda)}{\partial\delta_y} = \frac{\alpha(1-\lambda)}{(1-\alpha)(1-\lambda\delta_y)^2} > 0$$

The interaction parameter $\delta_y := (\gamma-1)\alpha/(1-\alpha)$ increases with γ. Hence $\partial\mu_Y(\lambda)/\partial\gamma > 0$. Note that this argument made no reference to the fact that $\mu_Y \geq 0$.

A.3 Proof of Theorem 1

Let $r \equiv Y/\theta$. The problem is, up to scale,

$$\min_{\mu\in\{\mu_\tau(\lambda),\mu_Y(\lambda)\},r\in\mathbb{R}} (1-\chi)(r-1)^2 + \chi(r/\mu-1)^2$$

We can concentrate out the parameter r with the following first-order condition

$$r^*(\mu) := \frac{\mu^2(1-\chi)+\mu\chi}{\mu^2(1-\chi)+\chi} \tag{26}$$

In this quadratic problem, the first-order condition is sufficient. 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 $\mu > 0$.

Let $\mathcal{L}(\mu)$ denote the loss function evaluated at this optimal r^*. Note that, from the envelope theorem, $\partial\mathcal{L}/\mu = -2\cdot\chi\cdot r^*\cdot(r^*/\mu-1)/\mu^2$. Combined with the previous expression for r^*, this suggests that $\partial\mathcal{L}/\mu = 0$ when $\mu = 1$, $\partial\mathcal{L}/\partial\mu > 0$ when $\mu > 1$, and $\partial\mathcal{L}/\partial\mu < 0$ when $\mu \in [0,1]$.

Finally, let $\mathcal{L}_\tau$ and $\mathcal{L}_Y$ denote the value of the loss function evaluated at $r^*(\mu)$ and, respectively, μ_τ and μ_Y. For fixed λ and α, we let $\mathcal{L}_\tau(\gamma)$ and $\mathcal{L}_Y(\gamma)$ denote these losses as function of γ. Note that, by the chain rule, $\partial \mathcal{L}_\tau / \partial \gamma = \partial \mathcal{L}/\partial \mu \cdot \partial \mu_\tau / \partial \gamma$ and $\partial \mathcal{L}_Y / \partial \gamma = \partial \mathcal{L}/\partial \mu \cdot \partial \mu_Y / \partial \gamma$. We will argue that these functions cross exactly once at some $\hat{\gamma}$, the critical threshold of GE feedback.

From here, we branch off the analysis for different domains of the parameters.

Simplest case. Consider the first parameter case covered in Lemma 4.

Note that $\mathcal{L}_\tau(0) = \mathcal{L}_Y(1) = 0$ and both functions are strictly positive elsewhere, by normalization. Since these functions are continuous, there exists (at least one) crossing point $\hat{\gamma} \in [0, 1]$ such that $\mathcal{L}_\tau(\hat{\gamma}) = \mathcal{L}_Y(\hat{\gamma})$.

In particular, $\mathcal{L}_\tau(\gamma)$ is strictly increasing and $\mathcal{L}_Y(\gamma)$ is strictly decreasing on the domain $\gamma \in (0, 1)$. By the previous argument, to show $\partial \mathcal{L}_\tau / \partial \gamma > 0$ and $\partial \mathcal{L}_Y / \partial \gamma < 0$, it suffices to show that $\partial \mu_\tau / \partial \gamma > 0$, $\partial \mu_Y / \partial \gamma > 0$, and $\mu_\tau > 1 > \mu_Y$. All three are established in Proposition 5.

Possibility of $\mu_Y < 0$. Now let us assume $\lambda < 2 - 1/\alpha$. There now exists a threshold

$$\underaccent{\smile}{\gamma} \equiv \frac{1 + \alpha(\lambda - 2)}{\alpha(\lambda - \alpha)} \in [0, 1)$$

such that, for $\gamma < \underaccent{\smile}{\gamma}$, $\mu_Y < 0$. For $\gamma \in [\underaccent{\smile}{\gamma}, 1]$, we can apply the same logic as previously. It remains to show that instrument communication is optimal for $\gamma \in [0, \underaccent{\smile}{\gamma})$.

First, note that $\partial \mathcal{L}_Y / \partial \gamma \leq 0$ as long as $r^*(\mu_Y) \geq 0$. The latter is true as long as $\mu_Y \geq -\chi/(1-\chi)$, which also implicitly defines a threshold $\check{\gamma}$ since μ_Y increases strictly in γ. Clearly the previous argument works for $\gamma \in [\check{\gamma}, 1]$, and it remains only to check $\gamma \in [0, \check{\gamma})$.

On this domain, $\partial \mathcal{L}/\partial \mu > 0$ since $r^*(\mu_Y) < 0$. But we also know that $\lim_{\mu \to -\infty} \mathcal{L}(\mu) = \chi$. This can be verified by direct calculation, or intuited by noticing that $\lim_{\mu \to -\infty} r^*(\mu) = 1$. Since μ_Y strictly increases in γ, it follows that $\mathcal{L}_y(\gamma) > \chi$ for $\gamma \in (-\infty, \check{\gamma}]$. Meanwhile, a similar argument for $\mu > 1$ (with $\lim_{\mu \to \infty} \mathcal{L}(\mu) = \chi$ and $\partial \mathcal{L}/\partial \mu > 0$) suggests that $\mathcal{L}_\tau(\gamma) < \chi$ for $\gamma \geq 0$. This shows that $\mathcal{L}_y(\gamma) > \chi > \mathcal{L}_\tau(\gamma)$ on this domain and thus instrument communication is strictly preferred.

It is worth pointing out that the limiting arguments for μ are "loose," since both μ_τ and μ_Y have finite limits:

$$\lim_{\gamma \to -\infty} \mu_\tau = \mu_{\tau, -\infty} \equiv \frac{\lambda}{\lambda + (1-\lambda)\alpha} \in (0, 1) \tag{27}$$

and

$$\lim_{\gamma \to -\infty} \mu_Y = \mu_{Y, -\infty} \equiv \frac{\lambda(1 - \alpha/\lambda)}{\lambda(1-\alpha)} \tag{28}$$

A.3.1 Other results

Theorem 3. *For any $\lambda < 1$, there exists some threshold $\dot{\gamma} < 0$ such that instrument communication is strictly preferred for $\gamma \in [\dot{\gamma}, 0]$. Further, if $\mu_Y > 0$ (as per the conditions of Lemma 4) or $\chi < 1/2$, $\dot{\gamma} = -\infty$.*

Proof. First, maintain Lemma 4 and its assumptions. Note that the second case ("more general") of the proof of the previous section does not use $\gamma > 0$. Hence the result is proved for $\dot{\gamma} = -\infty$ in this case.

Now relax those assumptions. Our best bound on the loss with target communication, for $\mu_Y < 0$, is $\min\{\mathcal{L}(\mu_{y,-\infty}), 1-\chi\}$, or the minimum loss between the $\gamma \to -\infty$ limit and the $\mu = 0$ extreme. $\mathcal{L}_\tau(\gamma)$ decreases smoothly on $\gamma \in (-\infty, 0]$ and is bounded above by $\mathcal{L}(0) = 1-\chi$. If $\mathcal{L}(\mu_{y,-\infty}) > 1-\chi$, it follows that $\underset{\cdot}{\gamma} = -\infty$ again. Since $\mathcal{L}(\mu_{y,-\infty}) > \lim_{\mu \to -\infty} \mathcal{L}(\mu) = \chi$, it follows that sufficient condition is $\chi > 1-\chi$ or $\chi > 1/2$.

Otherwise there must exist some $\dot{\gamma} < 0$ above which $\mathcal{L}_\tau(\gamma) < \chi$ and below which $\mathcal{L}_y(\gamma) > \chi$. We know for sure that instrument communication is optimal for $\gamma > \dot{\gamma}$ and target communication is optimal for $\gamma \in (-\infty, \dot{\gamma})$. □

A.4 Proof of Proposition 9

Instrument communication. Recall that

$$\psi_\tau = -\frac{\alpha^2\gamma(1-\sigma)}{1-\sigma\alpha\gamma}$$

For $\gamma \in [0,1]$ and $\sigma \in [0,1]$, the numerator is non-negative. Additionally, given $\alpha \in (0,1)$, the denominator is strictly positive. Hence $\psi_\tau \leq 0$ on this domain.

The partial derivative with respect to γ is the following:

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

so this function is decreasing for all values of γ. More transparently, the numerator of $|\psi_\tau|$ always increases and the denominator always decreases as γ increases.

Target communication. Recall that

$$\psi_Y = \frac{\alpha(1-\alpha)(1-\gamma)(1-\sigma)}{1-\alpha(1-\sigma(1-\gamma))}$$

and the derivative is

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

A.4.1 Other results

Proposition 11. *For any values of $\alpha \in (0,1)$, $\sigma \in [0,1)$, and $\gamma \leq 0$, $\psi_Y > \psi_\tau > 0$.*

Proof. It is obvious from the expressions why the values are positive. To see their relative size, note that $\psi_\tau = -\alpha g(\delta_\tau)$ and $\psi_Y = -\alpha g(\delta_Y)/(1-\alpha)$ for $g(\delta) \equiv \delta(1-\sigma)/(1-\sigma\delta)$. Note that $g(\delta)$ is non-positive and increasing for $\delta < 0$, and $\delta_Y < \delta_\tau \leq 0$ for $\gamma \leq 0$. Thus $\delta_\tau = -\alpha g(\delta_\tau) < -\alpha g(\delta_Y) < -\alpha g(\delta_y)/(1-\alpha) = \psi_Y$. □

A.5 Proof of Theorem 2

Let $\mathcal{L}_\tau(\gamma)$ and $\mathcal{L}_Y(\gamma)$ denote the policymaker's loss evaluated at the optimal message as a function of GE feedback parameter γ. As mentioned in the main text, $\mathcal{L}_\tau(\gamma) = \chi\psi_\tau^2(\gamma)$ and $\mathcal{L}_Y(\gamma) = (1-\chi)\psi_Y^2(\gamma)$. It is straightforward to deduce from the expressions for (ψ_τ, ψ_Y) and from Proposition 9 that the following are true:

1. $\partial\mathcal{L}_\tau/\partial\gamma = 2\chi\psi_\tau(\partial\psi_\tau/\partial\gamma) \geq 0$, $\mathcal{L}_\tau(0) = 0$, and $\mathcal{L}_\tau(1) > 0$.
2. $\partial\mathcal{L}_Y/\partial\gamma = 2(1-\chi)\psi_Y(\partial\psi_Y/\partial\gamma) \leq 0$, $\mathcal{L}_Y(0) > 0$, and $\mathcal{L}_Y(1) = 0$.

It follows that there exists as single crossing point $\hat{\gamma} \in (0,1)$ such that instrument communication is preferred for lower γ and target communication is preferred for higher γ.

A.6 Proof of Comparative Statics

The critical GE feedback threshold satisfies $\mathcal{L}_\tau(\hat{\gamma}) = \mathcal{L}_Y(\hat{\gamma})$. Plugging directly into the loss function produces a quadratic equation for the threshold. Of the two roots, the following one is in the correct domain $\gamma \in [0,1]$:

$$\hat{\gamma} = \Big(1 - \alpha(1-\chi\alpha)(1-\lambda) + \big(\alpha(\alpha - 2\lambda - 2\alpha(1-\lambda)\chi + (1-\alpha(1-\lambda)(1-\alpha\chi))^2\big)^{\frac{1}{2}}\Big)^{-1}$$

With this expression, we can do analytical comparative statics.

A.6.1 Policy parameter α

Proposition 12. *For fixed (λ, χ), the threshold $\hat{\gamma}$ is an increasing function of α. That is, the range of γ for which target communication is optimal decreases as $\partial Y/\partial K$ gets larger.*

Proof. The partial derivative $\partial\hat{\gamma}/\partial\alpha$, up to a strictly positive constant C, is

$$\frac{\partial\hat{\gamma}}{\partial\alpha} \cdot C = (1-2\alpha\chi)\left(1 - \frac{\alpha(1-\lambda)(1-\alpha\chi)}{\sqrt{(\alpha(1-\lambda)(1-\alpha\chi))^2 + (1-\alpha)^2}}\right) + \frac{1-\alpha}{\sqrt{(\alpha(1-\lambda)(1-\alpha\chi))^2 + (1-\alpha)^2}}$$

First, consider the case of $2\alpha\chi < 1$. It remains to show that the term in parenthesis is positive. A sufficient condition for this is

$$1 - 2\alpha(1-\lambda)(1-\alpha\chi) - \alpha(2\alpha(1-\lambda)\chi + 2\lambda - \alpha) > 0$$

Canceling out like terms,

$$(1-\alpha)^2 > 0$$

which is trivial for $\alpha \in (0,1)$. Thus $\hat{\gamma}$ decreases with λ.

Next, consider the case $2\alpha\chi > 1$. We can re-write the expression as

$$\frac{\partial\hat{\gamma}}{\partial\alpha} \cdot C = (1-\alpha\chi)^2 \left(1 - \frac{\alpha(1-\lambda)(1-\alpha\chi)}{\sqrt{(\alpha(1-\lambda)(1-\alpha\chi))^2 + (1-\alpha)^2}}\right) + \frac{1-\alpha+(\alpha\chi)^2}{\sqrt{(\alpha(1-\lambda)(1-\alpha\chi))^2 + (1-\alpha)^2}} - (\alpha\chi)^2$$

Note that the large denominator is bounded by $\sqrt{\alpha^2 + (1-\alpha)^2}$ and also bounded by one. Thus we can show that all terms are positive, and $\partial\hat{\gamma}/\partial\alpha > 0$. □

A.6.2 Attentive fraction λ (Proposition 7)

Up to a (different) positive constant, the relevant partial derivative is

$$\frac{\partial\hat{\gamma}}{\partial\lambda} \cdot C = \frac{\alpha(1-\lambda)(1-\alpha\chi)}{\sqrt{(\alpha(1-\lambda)(1-\alpha\chi))^2 + (1-\alpha)^2}} - 1$$

By the intermediate step of the previous argument, this is negative and thus γ decreases with λ.

A.6.3 Output gap parameter χ

Proposition 13. *For fixed (α, λ), the threshold $\hat{\gamma}$ is an decreasing function of χ. That is, the range of γ for which target communication is optimal increases as the policymaker pays more attention to the output gap.*

Proof. The relevant partial derivative (up to a constant) is equal to the previous one:

$$\frac{\partial\hat{\gamma}}{\partial\chi} \cdot C = \frac{\partial\hat{\gamma}}{\partial\lambda}$$

Hence we know it is negative, and $\hat{\gamma}$ decreases with χ. □

B Micro-Foundations

In this appendix we spell out the details of two micro-foundations that can be nested in our framework. The first is the neoclassical economy introduced in the setup of our framework (Section 2). The second is the stylized New Keynesian economy mentioned in our discussion of forward guidance (Section 7).

B.1 A neoclassical economy with aggregate demand externalities

In this section we fill in the details of the micro-founded example discussed in the main text. The set up was described in Section 2, p.2. Here, we solve the model and explain how it is nested in our abstract framework.

B.1.1 Solution

It is easiest to solve this model backward in time.

Period 2. The final goods producer's demand for intermediates is the following:

$$p_i = \eta(1-r)YX^{\rho-1}x_i^{-\rho}$$

where X is the CES aggregator of the individual x_i. This implies that the revenue for the entrepreneur has the following form:

$$p_i \cdot x_i = \alpha(1-r)Y\left(\frac{x_i}{X}\right)^{1-\rho} = \alpha(1-r)X^{\eta+\rho-1}N^{1-\eta}x_i^{1-\rho}$$

Profits scale more with aggregate investment X when ρ is high (high complementarity and high demand externality).

Labor supply has the following form:

$$w = (1+\phi)N^{\phi}$$

Labor demand is set by the final-goods firm:

$$w = (1-\eta)(1-r)\frac{Y}{N}$$

which decreases in the tax rate (or increases in the subsidy).

Period 1. The entrepreneur invests until the marginal return on capital is one:

$$1 = \mathbb{E}_i\left[\frac{\partial(x_i \cdot p_i)}{x_i}\right]$$

The first-order condition re-arranges to

$$x_i^{\rho} = \eta(1-\rho)\mathbb{E}_i\left[(1-r)X^{\eta+\rho-1}N^{1-\eta}\right] \tag{29}$$

Investment solves this fixed-point equation.

REE benchmark. Assume rational expectations with no uncertainty. In equilibrium, the agent will conjecture that $x_{-i} = x_i \equiv X$. Since everything is now known, we can pull X out of the expectation and solve to get

$$X_i = X = (\eta(1-\rho))^{\frac{1}{1-\eta}}(1-r)^{\frac{1}{1-\eta}}N$$

It is immediate that output is linear in labor:

$$Y = X^{\eta}N^{1-\eta} = (\eta(1-\rho))^{\frac{\eta}{1-\eta}}(1-r)^{\frac{\eta}{1-\eta}}N$$

Setting labor supply to labor demand gives

$$N = \left(\frac{1-\eta}{1-\phi}\right)^{\frac{1}{1+\phi}} (1-r)^{\frac{1}{1+\phi}} Y^{\frac{1}{1+\phi}}$$

and plugging that back into the equation for output gives

$$Y = \left(\frac{1-\eta}{1-\phi}\right)^{\frac{1}{\phi}} (\eta(1-\rho))^{\frac{\eta}{1-\eta}\frac{1-\phi}{\phi}} (1-r)^{\frac{\eta}{1-\eta}\frac{1-\phi}{\phi}+\frac{1}{\phi}}$$

From this point, we can also solve for output as a function of investment X. Crucially, none of the exponents (i.e., elasticities) depend on the value of ρ: only the constants (levels) do.

B.1.2 Log-linear approximation

Now consider a more general model in which agents do not form rational expectations, because of either limited information or various behavioral biases. The fixed-point equation 29 can no longer be solved without expectations. To make progress, we will take log-linear approximations. Let lowercase letters denote the log deviation quantities. Further, let $k_i = \frac{1+\eta\phi}{1+\phi} \log x_i$ and $\tau = \frac{1+\eta\phi}{\phi(1-\eta)} \log(1-r)$ be convenient monotonic transformations of investment and the subsidy, respectively.

Aggregate production is exactly log-linear:

$$y = \eta x + (1-\eta)n = \frac{\eta(1+\phi)}{1+\eta\phi} K + (1-\eta)n$$

Equilibrium labor is

$$n = \frac{y}{1+\phi} + \frac{\log(1-r)}{1+\phi} = \frac{y}{1+\phi} + \frac{\phi(1-\eta)}{(1+\eta\phi)(1+\phi)}\tau$$

Combining these two expressions yields the following expression for output as a function of investment and policy:

$$y = (1-\alpha)\tau + \alpha K \tag{30}$$

with

$$\alpha \equiv \frac{\eta(1+\phi)^2}{(\eta+\phi)(1+\eta\phi)} \tag{31}$$

The direct effect of policy, with weight $1-\alpha$, comes entirely through the expansion of labor demand. Unsurprisingly, this effect is strongest when the capital share of output η is relatively small

Let us now turn to the investment decision (29). To a log-linear approximation, it is

$$x_i = \left(1 - \frac{1-\eta}{\rho}\right) \mathbb{E}_i[x] + \frac{1-\eta}{\rho} \mathbb{E}_i[n] + \frac{1}{\rho} \mathbb{E}_i[\log(1-r)]$$

After substituting in equilibrium labor and rescaling investment and taxes, we get

$$k_i = (1-\gamma)\, \mathbb{E}_i[\tau] + \gamma \mathbb{E}[Y] \tag{32}$$

for feedback parameter

$$\gamma \equiv \frac{(1+\eta\phi)(\rho(\eta+\phi)-\phi(1-\eta))}{\eta\rho(1+\phi)^2} \tag{33}$$

For all $\phi > 0$, $\rho \in (0,1)$, and $\eta \in (0,1)$, this parameter is in the relevant domain $(-\infty, 1]$. A higher aggregate demand externality always corresponds to a larger feedback:

$$\frac{\partial\gamma}{\partial\rho} = \frac{(1-\eta)(1+\eta\phi)\phi}{\eta\rho^2(1+\phi)^2} > 0$$

For fixed (ϕ, η), γ reaches its maximum value $\bar{\gamma} = (1+\eta\phi)/(1+\phi)$ when $\rho = 1$. This exactly corresponds with unscaled investment equalling expected output: $x_i = \mathbb{E}_i[Y]$.

The feedback pararmeter is positive if and only if

$$\rho > \frac{\phi(1-\eta)}{\phi+\eta}$$

This is more likely (true for a larger sub-domain of $\rho \in [0,1]$) when the capital share is relatively high or the disutility of labor is relatively low. In both cases, the competition between firms over scarce labor resources is less severe.

B.2 A New Keynesian economy

In this section we present an alternative example of a stylized New Keynesian economy.

Set-up. Let all variables be in log deviation from a steady state. Consider a three-period economy ($t \in \{0,1,2\}$) with two types of consumers. A "type d (discretionary)" agent consumes in periods 1 and 2. She earns income y only in period 1 and has an opportunity to save at interest rate r for period 2. The consumption $c_{i,d}$ for one such individual obeys the following Euler equation:

$$c_{i,d} = (1-\beta)y - \beta r$$

where β is a discount factor (and thus $1-\beta$ is a permanent-income marginal propensity to consume). There are mass χ of such agents.

Our "reduced-form" period 2 can be thought of as a substitute for the infinite future in a fully-specified model (see, for instance, Angeletos and Lian, 2018). The simplification made in our Euler equation is that, at least in expectation, all future income and interest rates are at the steady state level. Thus, when we substitute the lifetime budget constraint into the standard Euler equation, we get the previous simple expression.

A second type of agent, of mass $1-\chi$, also consumes only in periods $t \in \{1,2\}$, but precommits at $t=0$ to her consumption at $t=1$ ("type p"). She also earns income y in period 1 and faces interest rate (from period 1 to period 2) r. She tries in expectation to follow the same reduced-form Euler equation.

$$c_{i,p} = (1-\beta)\mathbb{E}_i[y] - \beta\mathbb{E}[r]$$

These agents understand the structure of the economy but may have mis-specified beliefs about each others' actions.

Output, and hence income, at $t = 1$ is determined by total demand: $y = \int_i c_i \,\mathrm{d}i = (1-\chi)c_p + \chi c_d$, where (c_p, c_d) denote the cross-sectional average consumption of each agent. This scenario might represent a liquidity trap, which the economy will escape in period 2.

Finally, at $t = 0$, the policymaker can make an announcement about the intended value of r or target for y.

Solution. Let us re-arrange the equations slightly to clarify the feedback between each group's consumption. Substituting the Euler equation into the market clearing expression yields the following expression for output as a function of pre-committed consumption and interest rates:

$$y = \frac{\chi\beta}{1-\chi(1-\beta)}(-r) + \frac{1-\chi}{1-\chi(1-\beta)}c_p \tag{34}$$

This fits our abstract framework with $\tau = -r$ (a renormalization) and

$$\alpha = \frac{1-\chi}{1-\chi(1-\beta)} \tag{35}$$

The "direct effect" of decreasing the interest rate, via the Euler equation of the discretionary fraction χ, is $\chi\beta$. In equilibrium, decreasing the interest rate has an "extra kick" over this direct effect ($1-\alpha > \chi\beta$) because of a familiar Keynesian cross: more consumption also produces more income, and then even more consumption. Increasing pre-committed consumption does something similar: it also increases income, and hence spending, for the discretionary consumers. In that way, any "sentiment shock" that causes more pre-committed consumption causes an economic boom.

Recall that pre-committed consumption already matches our abstract model's form:

$$c_p = \beta\bar{\mathbb{E}}[-r] + (1-\beta)\bar{\mathbb{E}}[y] \tag{36}$$

for $\gamma := (1-\beta)$. Now, the pre-committed consumption increases when agents are optimistic about policy and income at $t = 1$. To sum up, there are two-way feedbacks between pre-committed and discretionary consumption, going through the fact that output is demand determined.

We can think of comparative statics in β controlling feedback in this economy. A larger β may correspond to a longer horizon or larger "effective" discounting induced by, for instance, liquidity constraints. Both of these amount to having a larger "pre-committed" Keynesian cross.

In our baseline formulation, it is impossible to affect this GE feedback effect without affecting the second-period Keynesian cross as well. In particular, a higher β will correspond to a larger direct effect of interest rates on the economy. These forces could be "mechanically" decoupled by assuming that pre-commitment and discretionary consumers have different discount rates (and/or MPCs).

C Equivalence with "Smooth" Level-k Thinking

Let $K = b_{X,k}\hat{X}$ in the level-k economy in which fundamental $X \in \{\tau, Y\}$ is announced and all agents compute to order k. The coefficients $(b_{X,k})_{k=0}^{\infty}$ follow the following recursion:

$$b_{X,k} = 1 + \delta_X \left(b_{X,k-1} - 1\right) \tag{37}$$

and we recover that $\lambda_{X,k} := b_{X,k}/b_{X,k-1}$ as the ratio between realized K and expected K. It follows that we could plug $\lambda_{X,k}$ into the inertial beliefs model and recover the same relationship between K and $\hat{X}$. For instrument communication, the sequence $\lambda_{\tau,k}$ is bounded above by 1 and decreases in k. For target communication, the sequence $\lambda_{Y,k}$ switches sides of one for even ($\lambda_{Y,k} < 1$) and odd ($\lambda_{Y,k} > 1$) periods. This oscillatory behavior is a familiar behavior of level-k models with strategic substitutability. Our model agrees with the level-k model for only the even k.

This last property can be "smoothed out" in a more refined solution concept. Consider the following model based on the "reflective equilibrium" in Garcıa-Schmidt and Woodford (2018). Let $\mathcal{B}(b_X; X)$ map the conjecture $\mathbb{E}_i[K] = b_X X$ to the equilibrium relationship $K = b_X^* X$ for a given communicated fundamental. Assume that the agents' conjecture is a smooth function of "cognitive depth" t solving the following differential equation:

$$\begin{aligned} \frac{\mathrm{d}b_X(t)}{\mathrm{d}t} &= \mathcal{B}(b_X(t); X) - b_X(t) \\ b_X(0) &= b_{X,k=0} \end{aligned} \tag{38}$$

It is straightforward to verify that, given an initially conservative guess ($b_X(0) < \mathcal{B}(b_X(0); X)$), the solution is also conservative ($b_X(t) < \mathcal{B}(b_X(t); X)$ for all $t > 0$). The dampening parameter $\lambda_X(T) := b_X(T)/\mathcal{B}(b_X(t); X)$, which maps to λ in our model, remains less than one for both communication modes.

Theorem 1 remains true for fixed T. There is no disruption to the logic from having different dampening λ_X conditional on communication choice.

D Positive vs. Negative GE Feedback

The entire analysis has presumed a positive GE feedback ($\gamma > 0$). We now briefly discuss the case with a negative GE feedback ($\gamma < 0$). In this case, K depends negative on expectations of Y. This may capture situations in which agents compete for finite resources, with higher output corresponding to higher prices and hence lower consumption or investment (see the micro-foundation of Section 2 for an example).

Both modes of communication now induce a game of strategic substitutes. In particular, the game of substitutes is more "severe" under target communication, or $\delta_Y < \delta_\tau < 0$. This directly translates into a "harder" implementability constraint under target communication, whether bounded rationality takes the form of anchored beliefs (in which case we have $\mu_Y < \mu_\tau < 1$) or erratic beliefs (in which case we have $|\psi_Y| > |\psi_\tau|$).

If we make parameter assumptions to rule out the case $\mu_Y < 0$, which involves policy moving in the opposite direction of output, it is easy to show in the anchored beliefs model that instrument communication is strictly preferred to target communication for any $\gamma < 0$. To achieve the same result more generally, we need further assumptions on the loss function. Theorem 3 in the Appendix elaborates on the details.

In the model with erratic beliefs, the shape of the loss function always matters. The policymaker cares asymmetrically about variation in the instrument gap (as induced by target communication) versus variation in the policy gap (as induced by instrument communication), and lacks any tools (like the choice of r in the first model) to shift the distortions between the gaps. Thus, even though $\psi_Y^2 > \psi_\tau^2$ unambiguously for all $\gamma < 0$, there exists a large enough weight on the output gap (χ) such that target communication is still preferred. Of course if the weights are equal or lower on the output gap ($\chi \leq 1/2$), instrument communication will be strictly preferred.

In all cases, for sufficiently low γ, we rely on possibly undesirable unstable equilibria for which $\delta_X < -1$. We might strengthen the equilibrium concept for implementability to eliminate these cases. Then all the previous statements could only hold for $\gamma \geq -(\alpha\lambda)^{-1}$ or $\gamma \geq -(\alpha\sigma)^{-1}$ in the two models, respectively. Below this point, the optimal communication choice is ill-posed, since neither announcement induces an interpretable equilibrium to determine K.

E Communicating K

Imagine that a policymaker announces $\hat{K}$ and agents have the "inertial beliefs" friction of Assumption 3. Assume that first-order beliefs about investment are correct ($\bar{\mathbb{E}}[K] = \hat{K}$) and higher-order beliefs are anchored toward zero ($\bar{\mathbb{E}}^h[K] = \lambda^{h-1}\hat{K}$). For the announcement to be fulfilled in equilibrium, it must be the case that

$$\hat{K} = (1-\delta_X)\bar{\mathbb{E}}[X] + \delta_X\bar{\mathbb{E}}[K] = (1-\delta_X)\bar{\mathbb{E}}[X] + \delta_X\hat{K}$$

for either fundamental $X \in \{\tau, Y\}$. The only first-order beliefs compatible with this announcement, then, are $\bar{\mathbb{E}}[\tau] = \bar{\mathbb{E}}[Y] = \bar{\mathbb{E}}[K] = \hat{K}$: on average (and, in fact, uniformly), agents believe that equilibrium will be $\tau = Y = K$. This is an ideal scenario for the policymaker.

It turns out, however, that a rational agent who doubts the attentiveness of others will doubt that other agents play the announcement, or that $K = \hat{K}$. If a given agent i thinks that agent j plays $k_j = \hat{K}$, she is implicitly taking a stand on agent j's beliefs about τ and Y. Specifically, agent i believes that agent j is following her best response (here, written with $X = \tau$), namely

$$\mathbb{E}_i[k_j] = (1-\delta_\tau)\mathbb{E}_i\mathbb{E}_j[\tau] + \delta_\tau\mathbb{E}_i\mathbb{E}_j[K]$$

We have assumed that $\mathbb{E}_i[k_j] = \hat{K}$ and $\mathbb{E}_i\mathbb{E}_j[K] = \lambda\hat{K}$. This produces the following restriction on second-order beliefs about τ:

$$\mathbb{E}_i\mathbb{E}_j[\tau] = \frac{1-\lambda\delta_\tau}{1-\delta_\tau}\hat{K}.$$

This has a simple interpretation: to rationalize aggregate investment being $\hat{K}$ despite the fact that fraction $(1-\lambda)$ of agents were inattentive to the announcement, agent i thinks that a typical other agent has *over*-forecasted the policy instrument τ.

At the same time, agent i knows that, like himself, all attentive agents expect τ to coincide with $\hat{K}$. And since agent i believes that the fraction of attentive agents is λ, the following restriction of second-order beliefs also has to hold:

$$\mathbb{E}_i\mathbb{E}_j[\tau] = \lambda\hat{K}.$$

When $\lambda = 1$ (rational expectations), the above two restrictions are jointly satisfied for any $\hat{K}$. When instead $\lambda < 1$, this is true only for $\hat{K} = 0$. This proves the claim made in the text that, as long as $\lambda < 1$, there is no equilibrium in which is infeasible to announce and commit to any $\hat{K}$ other than 0 (the default point).

In a nutshell, as noted in the main text, the problem with communicating K is that the policymaker has no direct control over it. From this perspective, output communication worked precisely because the policymaker had some plausible commitment. Agents could rationalize $Y = \hat{Y}$ regardless of their beliefs about K because there always existed some level of τ that implemented $\hat{Y}$. We alluded to the failure of this mechanism as $\alpha \to 1$, and the direct effect of policy vanished, in our baseline model (Section 5.6).

Throughout this paper, we have not directly addressed the issue of credible commitment. The previous discussion highlights that our analysis may have subtle interactions with commitment problems. Indeed, agents' (higher-order) beliefs about commitment problems may be crucial. We leave the formal investigation of this topic to future work.

F Adding More Shocks

Our baseline model included exogenous shocks to the preferences of the policymaker but excluded such shocks from conditions (1) and (2). This is without loss of generality if the other shocks are common knowledge and observed by the policymaker. These assumptions are extreme, but common in the Ramsey policy paradigm. In our context, they guarantee that implementability results remain true provided that that the quantities (τ, Y) are re-defined to be "partialed out" from the extra shocks.

A more plausible scenario, perhaps, is that other shocks are unobserved and the policymaker cannot condition on them. This introduces into our analysis similar considerations as those in Poole (1970). The latter focused on how two different policies—fixing the interest rate or fixing the money supply—differed in their robustness to external shocks. Primitive shocks (to supply and demand) had different effects on the policy objective (output gap) depending on the slope of the model equations and the policy choice. Poole could do comparative statics of optimal policy in these slopes as well as the relative variance of the shocks.

Such “Poole considerations” can be inserted into our framework and will naturally affect the choice between fixing τ and fixing Y. However, such consideration matter even in the REE benchmark and, roughly speaking, are “separable” from the mechanism we have identified in our paper. We make this point clearer with a few examples in the sequel.

F.1 Shocks to output

Consider now a model in which output contains a random component:

$$Y = (1 - \alpha)\tau + \alpha K + u,$$

where u is drawn from a Normal distribution with mean 0 and variance σ_u^2, is orthogonal to θ, and is unobserved by both the policymaker and the private agents. In this case, announcing and committing to a value for Y stabilizes output at the expense of letting the tax distortion fluctuate with u. Conversely, announcing and committing to a value for τ stabilizes the tax distortion at the expense of letting output fluctuate with u. It follows that, even in the frictionless benchmark ($\lambda = 1$), the policymaker is no more indifferent between the two. In particular, target communication is preferable if and only if the welfare cost of the fluctuations in Y exceeds that of the fluctuations in τ, which is in turn is the case whenever χ is high enough.

The above scenario has maintained the assumption that the ideal level of output is $Y^{\text{fb}} = \theta$. What if instead we let $Y^{\text{fb}} = \theta + u$? This could correspond to a micro-founded business-cycle model in which technology shocks that have symmetric effects on equilibrium and first-best allocations. Under this scenario, it becomes desirable to let output fluctuate with u, which in turn implies that, in the frictionless benchmark, instrument communication always dominates target communication. A non-trivial trade off between the two could then be recovered by adding unobserved shocks to the tax distortion. The optimal strategy is then determined by the relative variance of the two unobserved shocks and the relative importance of the resulting fluctuations, along the lines of Poole (1970).

While these possibilities are interesting on their own right, they are orthogonal to the message of our paper. Indeed, the shock considered above does not affect the strategic interaction of the private agents under either mode of communication: Lemmas 1 and 2 remain intact. By the same token, when $\lambda = 1$, the sets of the implementable (τ, Y) pairs remain invariant to γ, even though they now depend on the realization of u. It then also follows that, as long as $\lambda = 1$, the optimal mode of communication does not depend on γ. But as soon as $\lambda < 1$, the implementability sets and the optimal mode of communication start depending on γ, for exactly the same reasons as those explained before: a higher γ increases the bite of strategic uncertainty under instrument communication and decreases it under target communication, thus also tilting the balance in favor of the latter as soon as one departs from the frictionless benchmark.

F.2 Measurement errors

The same logic as above applies if we introduce measurement errors in the policymaker's estimates of τ and Y. To see this, consider a variant of our framework that lets the policymaker observe only $(\tilde{\tau}, \tilde{Y})$, where

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

and the u's are independent Gaussian shocks, orthogonal to θ, and unpredictable by both the policymaker and the private agents. Accordingly, instrument communication now amounts to announcing and committing to a value for $\tilde{\tau}$, whereas target communication amounts to announcing and committing to a value for $\tilde{Y}$.

By combining the above with condition (1), we infer that, under both communication modes, the following restriction has to hold:

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

where

$$\tilde{u} \equiv -(1 - \alpha)u_\tau + u_Y.$$

At the same time, because the u's are unpredictable, the best response of the agents can be restated as

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

This maps directly to the version with unobserved shocks just discussed above if we simply reinterpret $\tilde{\tau}$, $\tilde{Y}$, and $\tilde{u}$ as, respectively, the actual tax rate, the actual level of output, and the unobserved output shock.

To sum up, the presence of unobserved shocks and measurement error can tilt the optimal strategy of the policymaker one way or another in manners already studied in the literature that has followed the lead of Poole (1970). This, however, does not interfere with the essence of our paper's main message regarding the choice of a communication strategy as a means for regulating the impact of strategic uncertainty and the bite of the considered forms of bounded rationality.