NBER WORKING PAPER SERIES

FORWARD GUIDANCE WITHOUT COMMON KNOWLEDGE

George-Marios Angeletos
Chen Lian

Working Paper 22785
http://www.nber.org/papers/w22785

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2016

We have no financial interests or research support to disclose. For helpful comments, we thank Adrien Auclert, Xavier Gabaix, Alessandro Pavan, and seminar participants at the 2016 NBER Summer Institute and the 2016 ESEM in Edinburgh. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Forward Guidance without Common Knowledge
George-Marios Angeletos and Chen Lian
NBER Working Paper No. 22785
October 2016, Revised February 2017
JEL No. C72,D82,E03,E32,E43,E52,E58

ABSTRACT

Forward guidance—and macroeconomic policy more generally—relies on shifting expectations, not only of future policy, but also of future economic outcomes such as income and inflation. These expectations matter through general-equilibrium mechanisms. Recasting these expectations and these mechanisms in terms of higher-order beliefs reveals how standard policy predictions hinge on the assumption of common knowledge. Relaxing this assumption anchors expectations and attenuates the associated general-equilibrium effects. In the context of interest, this helps lessen the forward-guidance puzzle, as well as the paradox of flexibility. More broadly, it helps operationalize the idea that policy makers may find it hard to shift expectations of economic outcomes even if they can easily shift expectations of policy.

George-Marios Angeletos
Department of Economics, E52-530
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
angelet@mit.edu

Chen Lian
Department of Economics
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
lianchen@mit.edu

1 Introduction

The effectiveness of macroeconomic stabilization depends on the ability of policy makers to shift market expectations of the relevant economic outcomes. For example, the Fed continuously strives to manage expectations of future interest rates, with the intent of influencing macroeconomic activity. The actual effect of the Fed's acts and communications depends on the joint response of a multitude of economic agents (firms, consumers, banks, etc). The response of each one of them, in turn, hinges on how that agent expects *other* agents to respond—or, equivalently, on the expected effect on aggregate spending, inflation, and labor- or credit-market conditions. As a result, the Fed's ability to steer the economy hinges on its ability to manage this kind of expectations, not just the expectations of future interest rates.

In this paper, we revisit the predictions that standard macroeconomic models make about the ability of policy makers to shift market expectations of the relevant economic outcomes. In particular, we argue that standard models "maximize" this kind of ability by imposing common knowledge of the policy itself and of the responses of economic agents to it; they therefore risk overstating the effectiveness of certain policies, such as that of forward guidance in the context of a liquidity trap. Conversely, by relaxing common knowledge, we operationalize the idea that policy makers may have difficulty in managing expectations of economic outcomes, even if they can perfectly manage expectations of policy.

Context. We demonstrate the above ideas within the following context: the power of forward guidance during a liquidity trap, that is, the power of the monetary authority to stimulate macroeconomic activity *today* by promising to follow a certain policy in the *future,* after the economy has exited the trap.[1]

More specifically, suppose that the economy is in a slump and that the zero-lower-bound (ZLB) binds up to a future date $t = T - 1$, for some $T \geq 2$. Because of this constraint, the monetary authority is unable to stimulate aggregate demand by reducing the current Federal funds rate. According to the New-Keynesian (NK) model, the same goal can be achieved by a credible promise to keep the interest rate below its "natural level" after the ZLB has ceased to bind, that is, at $t \geq T$. What is more, the larger T is and the further into the future forward guidance operates, the stronger is its effect on real economic activity and inflation.

Policy makers and macroeconomists alike find these predictions to be counterintuitive, which is why the issue is known as the "forward guidance puzzle".[2] Existing attempts to resolve the puzzle, such as McKay, Nakamura and Steinsson (2016b), hinge on modifying the micro-foundations of the NK model. The alternative we propose here shifts the focus to the formation of expectations.[3]

Preview. We take as given the monetary authority's ability to shift the expectations of the interest-rate policy that will be in place after the economy exits the liquidity trap.[4] Without serious loss of generality,

---

[1]Our insights apply more generally. The specific context, however, helps for concreteness, plus it is topical.

[2]The term was coined by Del Negro, Giannoni, and Patterson (2012); see also Eggertsson and Woodford (2003), Carlstrom, Fuerst, and Paustian (2012), and McKay, Nakamura, and Steinsson (2016b).

[3]Throughout, by "micro-foundations" we refer to the model ingredients that determine the best responses of the relevant "players" in the economy. This includes the specification of preferences, technologies, idiosyncratic risks, and liquidity constraints—which is McKay, Nakamura and Steinsson (2016b) come in—but leaves out the formation of the beliefs these players form about the actions of others—which is where our contribution comes in.

[4]Throughout, we abstract from the commitment problems discussed in, inter alia, Bassetto (2015) and Woodford (2012). Nev-

we suppose that policy maker uses an announcement or other means to vary only expectations of R_T, the interest rate *right after* the ZLB has ceased to bind, and thereafter replicates flexible-price outcomes.[5] We then seek to investigate how the economy responds to such a shift.

Before tackling this issue, it is useful to make a digression from macroeconomics to microeconomics. Consider the following question: how does an individual consumer responds to news of a future interest-rate change that applies *only to herself*, as opposed to the entire economy? Answering this question helps isolate the partial-equilibrium (PE) effect of forward guidance from its general-equilibrium (GE) effects.

Because agents discount the future, the PE effect diminishes with the horizon, T, and becomes vanishingly small as $T \to \infty$. This elementary observation, which is essentially a corollary of the Permanent Income Hypothesis, underscores that the forward-guidance puzzle is *exclusively* about GE mechanisms.

What are these GE mechanisms? The most crucial one is the feedback loop between aggregate spending and inflation. Reducing the nominal interest rate at $t = T$ causes inflation at $t = T$. Because the nominal interest rate is pegged prior to T, this translates to a low *real* interest rate between $T-1$ and T. This stimulates aggregate spending at $T-1$, contributing to even higher inflation at $T-1$, which in turn feeds to even higher spending at $T-2$, and so on. Clearly, the cumulative effect at $t = 0$ increases with T, which explains why the power of forward guidance also increases with T—indeed without bound—according to the NK model.

The aforementioned feedback loop is the flip-side of the "deflationary spiral" that helps the NK model generate a sizable recession during a liquidity trap. The initial trigger is different, and the sign is reversed, but the GE mechanism is the same. In the standard model, this mechanism is captured by the *interaction* of the representative household's Euler condition with the New-Keynesian Philips Curve (NKPC).[6]

But there are two additional GE mechanisms, buried underneath these equations. The one has to do with the feedback from future inflation to current inflation: for given *real* marginal costs, the individual firm is more willing to raise its *nominal* price today if she expects other firms to do the same in the future. The other is a modern, and dynamic, variant of what was known as the "income multiplier" in the IS-LM model: when the individual consumer expects other consumers to spend more in the *future*, she is encouraged to spend more herself *today*, because her own income increases with aggregate consumption.

Our main result is that relaxing common knowledge attenuates *all* the aforementioned GE effects by anchoring the expectations of future income and inflation. An integral step to this result is the development of a game-theoretic representation of the NK model, which helps unearth the role of higher-order beliefs.

Once we depart from the common-knowledge benchmark, the demand block of the NK model (the Euler condition) becomes a dynamic beauty contest among the consumers, whereas its supply block (the NKPC) becomes a dynamic beauty contest among the firms. Each beauty contest *alone* captures the GE interaction—equivalently, the strategic complementarity—that operates *within* the corresponding block of the model. That is, the one beauty contest captures the strategic complementarity in the spending decisions

ertheless, the friction we accommodate could also reflect lack of common knowledge about policy maker's commitment.

[5]Our results can readily be extended to the case in which the nominal rate is lowered below the natural rate for Δ periods after exiting the trap, for any finite $\Delta \geq 1$. We let $\Delta = 1$ only to simplify the exposition.

[6]For the NK version of the liquidity trap and the operation of the aforementioned mechanism under common knowledge, see, inter alia, Krugman (1998), Eggertsson and Woodford (2003), and Werning (2012).

of the consumers, while the other captures the strategic complementarity in the pricing decisions of the firms. The feedback loop between aggregate spending and inflation can then be recast as a higher-layer beauty contest that is played *between* the two groups of agents.[7]

Under these lenses, it becomes abundantly clear how the predictions of the theory depend, not only on the assumed micro-foundations, but also—and indeed quite crucially—on an additional set of assumptions that help pin down *higher-order* beliefs (i.e., beliefs about the beliefs of others).

Some of these assumptions are embedded in the solution concept: similarly to Perfect Bayesian Equilibrium (PBE), Rational-Expectations Equilibrium (REE) imposes common knowledge of the rationality and of the strategies of all agents. The rest come in the form of a strong informational assumption: all agents share the same information at all times, and this property is itself common knowledge. In general, the combination of these assumptions imposes that the agents are able to perfectly coordinate the adjustment of their beliefs and of their actions to any exogenous change in the environment. In our context, it boils down to *maximizing the capacity of the policy maker to control the agents' expectations of inflation and real economic activity,* in a sense that we make precise in due course.

In our view, the puzzle is squarely about this issue. We therefore seek to loosen the policy maker's grip on the market's expectations of future inflation and future economic activity.

We achieve this by letting agents have incomplete information and therefore lack common knowledge of the policy and of one another's response to it. We thus build on a growing literature that uses higher-order uncertainty as a modeling device for accommodating a realistic friction in the coordination and the adjustment of equilibrium beliefs, while maintaining the standard, rational-expectations, solution concept. See Morris and Shin (1998, 2002) and Woodford (2003a) for key early contributions; Angeletos and Lian (2016b) for a survey; and Coibion and Gorodnichenko (2012, 2015) for corroborating evidence.

In the presence of the aforementioned friction, higher-order beliefs tend to move less so than lower-order beliefs in response to news about future policy. By the same token, expectations of future inflation and income can be anchored *even if* expectations of future nominal interest rates are not. It follows that all the aforementioned GE effects are attenuated—and so does the power of forward guidance. What is more, we show that longer horizons raise the relative importance of higher-order beliefs. It follows that the power of forward guidance is diminished more when the policy operates further into the future. This provides more broadly a rationale for "acting now rather than later", or for front-loading policy interventions.

We develop the key ideas under fairly general specifications of the information structure, allowing, inter alia, for endogenous learning through market signals.[8] This helps clarify the robustness of our insights, but risks obscuring the take-home message. We thus also illustrate the crux of our ideas with a stark information structure that kills learning, parameterizes the degree of common knowledge by a scalar $\lambda \in (0, 1]$, and facilitates sharp comparative statics. When $\lambda = 1$, the standard NK model is nested. When $\lambda < 1$, common

[7] Like Morris and Shin (2002), Angeletos and Pavan (2007), Bergemann and Morris (2013), and others, by "beauty contests" we refer to a class of linear-quadratic games that feature strategic complementarity and incomplete information.

[8] In general, such signals can shape the level, persistence, and propagation of higher-order uncertainty. See, e.g., Amador and Weill (2010), Angeletos and La'O (2013), Benhabib, Wang, and Wen (2015), Charhour and Gabbalo (2016), Hellwig and Venkateswaran (2009), Huo and Takayama (2015), and Veldkamp (2011). They are, however, of lesser importance for our purposes.

knowledge is relaxed. The lower λ is, the larger the departure from common knowledge and, as we show, the weaker the policy maker's control over expectations of income and inflation.

Under this specification, we can measure the power of forward guidance by the elasticity of aggregate output at $t = 0$ ("now") with respect to the period-0 average expectation of R_T ("the policy after the liquidity trap has ended"). Let $\phi = \phi(\lambda, T)$ denote the absolute value of this elasticity as a function of the degree of common knowledge, λ, and of the horizon, T. The standard NK model is nested with $\lambda = 1$; the power of forward guidance is then given by by $\phi^* = \phi(1, T)$; and the puzzle relates to the property that ϕ^* increases with T and explodes as $T \to \infty$. Relative to this benchmark, we show that, for any T, ϕ is strictly decreasing in λ; that is, the attenuation of the power of forward guidance is stronger when the degree of common knowledge is smaller. Furthermore, the attenuation increases without bound as forward guidance operates further into the future: for any $\lambda < 1$, the ratio ϕ/ϕ^* decreases with T and vanishes as $T \to \infty$. As noted earlier on, this is because longer horizons raise the relative importance of higher-order beliefs, which translates to stronger anchoring of expectations of inflation and real economic activity. Finally, the documented attenuation effect can be quantitatively significant: for a plausible parameterization and a relatively modest level of higher-order uncertainty, ϕ is about one quarter of ϕ^* at horizon of 5 years.

We complement these results with one that relates to the "paradox of flexibility." The latter refers to the prediction that greater price flexibility amplifies the recession during a liquidity trap—or, more generally, the volatility in the output gap when the current interest rate is pegged.[9] The flip-side is that monetary policy, in the form of forward guidance, becomes *more* potent when there is *less* nominal rigidity. Our own result is that the attenuation is itself stronger when prices are more flexible. This is because the same GE mechanisms that underlie the forward-guidance puzzle also underlie the paradox of flexibility. By anchoring expectations of inflation and income, lack of common knowledge helps contain both "anomalies".[10]

We conclude the paper with an "as if" result that recasts the lack of common knowledge as a form of discounting of the forward-looking terms of the Euler condition and the NKPC of the otherwise-standard, representative-agent, NK model. This illustrates that the assumed friction can be re-interpreted as a relaxation of solution concept, echoing the observations made in Angeletos and Lian (2016a,b) and complementing the results of Garcia-Schmidt and Woodford (2015) and Farhi and Werning (2016). It also builds an intriguing connection to recent work on "discounted Euler conditions", which we discuss in the sequel.

Layout. The remainder of the paper is organized as follows. Section 2 expands on the relation of our paper to the literature. Section 3 introduces our framework. Section 4 reviews the standard, complete-information, version of puzzle. Section 5 removes common knowledge, develops our beauty-contest representation, and reveals the role of higher-order beliefs. Section 6 illustrates the key results with a stark information structure. Sections 7 and 8 elaborate on the robustness of our insights and on the role of longer horizons. Section 9 touches on the paradox of flexibility and provides the aforementioned "as if" result. Section 10 contains the conclusions. The Appendix contains the proofs.

[9] See, inter alia, Eggertsson and Krugman (2012), Werning (2012), and Kocherlakota (2016).

[10] The same logic is likely to apply to other anomalies, such as the "paradox of toil" (Eggertsson, 2010, Wieland, 2014).

2 Related Literature

Our paper builds heavily on the macroeconomic literature on incomplete information and beauty contests. Important precedents include Morris and Shin (2002, 2006), Woodford (2003a), Angeletos and Pavan (2007), Nimark (2008, 2011), Angeletos and La'O (2010, 2013), and Bergemann and Morris (2013); see Angeletos and Lian (2016b) for a survey. Our marginal contribution rests on the representation of the two blocks of the NK model as beauty contests; on the specific dynamic structure of these games, which requires the development of new formal arguments; and on the lessons delivered for the application of interest.

The "as if" result that maps our economy to a representative-agent model with a discounted Euler condition and a discounted NKPC resembles results from McKay, Nakamura and Steinsson (2016a), Werning (2015), and Gabaix (2016). In contrast to these papers, however, the distortion of the equilibrium conditions at the aggregate level does *not* originate from a distortion of the optimality conditions at the individual level (or, equivalently, of the structure of best responses).[11] This is because our paper holds constant the micro-foundations and, instead, uses lack of common knowledge to anchor equilibrium expectations and to attenuate the given GE interactions. The assumed friction is therefore distinct, not only conceptually, but also empirically. Furthermore, it appears to be consistent with the evidence on inflation expectations documented in Coibion and Gorodnichenko (2012, 2015).[12]

The idea that lack of common knowledge attenuates GE effects is not limited to the context of this paper. In Angeletos and Lian (2016a), we have formalized this idea within a more abstract framework by establishing two closely related results: an observational equivalence between the introduction of incomplete information and certain relaxations of the solution concept, including Tatonnement and the reflective-equilibrium concept of Garcia-Schmidt and Woodford (2015); and a result that bridges the gap between micro and macro elasticities. The present paper can be seen as an application of this broader idea.[13]

The two works that are closest to the present paper are Farhi and Werning (2016) and Wiederholt (2016). Farhi and Werning (2016) share our objective of attenuating the feedback mechanism between aggregate income and individual spending, but attain this objective by replacing the standard equilibrium concept with level-k reasoning as opposed to introducing incomplete information. They also abstract from the feedback loop between aggregate spending and inflation, which, as we explain, turns out to be of central importance from quantitative perspective. Instead, they focus on the interaction between level-k reasoning and liquidity constraints, a friction from which we abstract entirely. Wiederholt (2016), on the other hand, shares our insight that incomplete information helps attenuate the feedback loop between aggregate spending and

[11] In McKay, Nakamura and Steinsson (2016a) and Werning (2015), the distortion originates in liquidity constraints; in Gabaix (2016), it obtains because agents are not as proficient in maximization and forecasting as mainstream economics impose.

[12] By recasting the NK model as a beauty contest and introducing a friction in the formation of the beliefs of the actions of others, we also connect to the experimental literature on beauty contests (e.g., Nagel, 1995, Duffy and Nagel, 1997), which has also emphasized the relevance of departing from the common-knowledge, fully-rational, benchmark.

[13] In addition to the specific lessons we deliver for the context of interest, a novelty in the present paper is the emphasis given to forward-looking expectations. Related in this respect are also Allen, Morris and Shin (2006), Bachetta and Wincoop (2006), and Morris and Shin (2006), who emphasize how incomplete information induces inertia in the response of forward-looking expectations to innovations in fundamentals in asset-pricing models.

inflation, but abstracts from the one between aggregate income and individual spending. Finally, neither of these works contains our results about either the paradox of flexibility or how the horizon increases the relative importance of higher-order beliefs.[14]

The feedback mechanism between aggregate and individual spending plays a prominent role in a number of other papers, although in manners that are orthogonal to the contribution of the present paper. For instance, Gali et al. (2007) and Farhi and Werning (2012) use this mechanism to increase the GE effects of fiscal stimuli in the presence of liquidity-constrained, or "hand-to-mouth", consumers, whereas Werning (2015) uses it to explain why liquidity constraints may *increase* the GE response of aggregate consumption to interest rates, in contrast to the case made by McKay, Nakamura, and Steinsson (2016a,b). These papers work with the NK framework and rest on the interaction of the aforementioned mechanism with nominal rigidity. Angeletos and Lian (2016c), on the other hand, combine this mechanism with rational confusion to augment an otherwise standard RBC model with effects that resemble Keynesian multipliers and that help generate realistic, demand-driven, business cycles in the absence of nominal rigidity.

3 Framework

In this section, we set up our framework. This is the same as the textbook NK model (Woodford, 2003b; Gali, 2008), except that we remain more flexible about the formation of expectations.

Consumers. There is a measure-one continuum of ex-ante identical, infinitely-lived, households, or consumers, in the economy, indexed by $i \in [0, 1]$. The preferences of consumer i are given by

$$\mathcal{U}_0 = \sum_{t=0}^{+\infty} \beta^t \left(\log c_{i,t} - \tfrac{1}{1+\epsilon} n_{i,t}^{1+\epsilon} \right), \tag{1}$$

where $c_{i,t}$ and $n_{i,t}$ denotes her consumption and labor supply at period t, $\beta \equiv e^{-\rho} \in (0, 1)$ is the discount factor, $\rho > 0$ is the discount rate, and $\epsilon > 0$ is the inverse of the Frisch elasticity.[15] The budget constraint in period t is given, in real terms, by the following:

$$c_{i,t} + s_{i,t} = a_{i,t} + w_{i,t} n_{i,t} + e_{i,t}, \tag{2}$$

where $s_{i,t}$ is the consumer's saving in period t, $a_{i,t} = R_{t-1} s_{i,t-1}/\pi_t$ is her initial asset position, R_{t-1} is the gross nominal interest rate between $t-1$ and t, $\pi_t \equiv p_t/p_{t-1}$ is the gross inflation rate, p_t is the aggregate price level at t, and $w_{i,t}$ and $e_{i,t}$ are the real wage and the real dividends received by the consumer.

Firms. The final good is produced by a competitive sector, using a continuum of intermediate-good

[14] Farhi and Werning's work was developed concurrently with ours; Wiederholt's appeared earlier. Somewhat related are also the following papers: Andrade et al. (2015), which studies how agents with different prior information can give different interpretations to the same forward-guidance announcement; Gaballo (2016), which studies the social value of news about future policy; and Kiley (2016), which argues that some of the paradoxical predictions of the NK model are not shared by variant models that attribute the nominal rigidity to informational frictions.

[15] To economize on notation, we have set the elasticity of intertemporal substitution to one.

varieties, indexed by $j \in [0,1]$, and a CES technology with elasticity ς. Aggregate output is thus given by

$$y_t = \left(\int_0^1 \left(y_t^j \right)^{\frac{\varsigma-1}{\varsigma}} dj \right)^{\frac{\varsigma}{\varsigma-1}}, \tag{3}$$

where y_t^j is the intermediate-good input from variety j. Each variety j is in turn produced by a monopolistic firm (also indexed by j), using labor under a linear technology:

$$y_t^j = l_t^j, \tag{4}$$

where l_t^j is the labor input of monopolist j.

We introduce nominal rigidity in the usual, Calvo-like, fashion: in each period, a randomly selected fraction $\theta \in (0,1]$ of the firms must keep their prices unchanged, while the rest can reset them. In addition, we assume that each resetting firm faces a firm-specific markup shock, which we denote by μ_t^j. The latter is i.i.d. over time but correlated across j. The aggregate markup shock is denoted by μ_t. Its modeling role is to limit the information that agents can extract from observed inflation.[16] Finally, we let each firm pay out all the profits it earns in each period, e_t^j, as dividends to the consumers.

Idiosyncratic Shocks. To limit the aggregation of information through market signals, we introduce idiosyncratic shocks to wages and dividends. The real wage paid by firm j at t is $w_t^j = w_t u_t^j$, where w_t is the average wage in the economy and u_t^j is i.i.d. across j and t. The real wage received by consumer i at t is $w_{i,t} = w_t \xi_{i,t}$, where $\xi_{i,t}$ is i.i.d across i and t. Finally, the dividend received by consumer i is $e_{i,t} = e_t \zeta_{i,t}$, where e_t denotes aggregate profits and $\zeta_{i,t}$ is i.i.d across i and t.[17]

Log-linearization. To keep the analysis tractable, we work with the log-linearized conditions of the model. We let a tilde over a variable denote the log-deviation of this variable from its unconditional mean. We also let the means of the (logs of the) idiosyncratic shocks be zero.

Equilibrium, Information, and Common Knowledge. With the exception of Proposition 2, which develops the more general beauty-contest representation of the NK model, we employ the standard solution concept, namely rational-expectations equilibrium (REE). We nevertheless depart from standard practice by relaxing the information structure. Let $\mathcal{I}_{i,t}$ and $\mathcal{I}_{j,t}^f$ denote the information set of, respectively, consumer i and firm j in period t. Next, let $\mathcal{I}_t \equiv (\cup_i \mathcal{I}_{i,t}) \cup (\cup_j \mathcal{I}_{j,t}^f)$ denote the union of the information sets of all agents. The standard model assumes complete information, or common knowledge, in the sense that $\mathcal{I}_{i,t} = \mathcal{I}_{j,t}^f = \mathcal{I}_t$ for all i, j, t and all states of Nature. We will instead allow for incomplete information, or lack of common

[16] It is possible to extend the model so as to micro-found the markup shocks in terms of preference shocks that shift the elasticity of the demand faced by each monopolist; see Angeletos, La'O and Iovino (2016) for an example.

[17] The aforementioned shocks can be justified in a similar manner as in Lucas (1972) and Lorenzoni (2009). Suppose that each household contains one worker and let $\eta_{i,t}$ be a shock to her labor disutility. Suppose further that, at each t, the economy is split in a large set of decentralized labor markets ("islands"). Workers and firms are randomly allocated in these markets, in such a manner that the these markets differ from one another only in terms of the average preference shock: the "average" worker in market m has taste shock equal to $\bar{\eta}_t^m$, where $\bar{\eta}_t^m$ washes out once we aggregate across all markets but is non-zero within the typical market. Then, it is possible to show the equilibrium wage in market m is given by $w_t \bar{\eta}_t^m$, which in turn pins down the wage shock of any given firm or any given consumer by the shock $\bar{\eta}_t^m$ of the market to which this agent have been allocated. One way or another, the sole modeling role of these shocks is to limit the extraction of information through market signals.

knowledge, in the sense that the aforementioned equalities cease to hold (in a non-trivial manner).

Auxiliary assumptions. To sharpen the exposition, we assume the following: at every t, the current values of R_t and of $(p_t^j)_{j\in[0,1]}$ are commonly known (and therefore so is the current price level); each consumer i has knowledge of the current values of $(w_{i,t}, e_{i,t})$; and each monopolist j has knowledge of the current values of (w_t^j, μ_t^j). These assumptions could be relaxed without affecting the essence of our results. They are nevertheless useful, not only because they simplify the exposition, but also because they help clarify that that our results do not require that the agents be inattentive to, or unaware of, their *current* economic condition. Instead, it suffices that they are uncertain about the future. We finally assume that the shocks to markups, wages, and endowments are unpredictable on the basis of past information.

Monetary Policy. With the exemption, once again, of Proposition 2, we assume the following.

Assumption 1 (Monetary Policy) *There exists a known $T \geq 2$ such that: (i) At any $t \leq T-1$, the interest rate is pegged at $R_t = \bar{R}$, where $\bar{R} \geq 1$ is known. (ii) At any $t \geq T+1$, monetary policy replicates flexible-price outcomes. (iii) The common prior about R_T is a lognormal distribution with mean ρ and variance $\sigma_R^2 > 0$.*

This assumption restricts monetary policy at all dates other than T, thus letting us identify forward guidance with expectations of R_T. Part (i) is automatically satisfied in a liquidity-trap context: a binding ZLB constraint is nested by letting $\bar{R} = 1$. Because we focus on the log-linearized version of the model, however, it makes no difference whether $\bar{R}$ is 1 or an arbitrary peg. To simplify the exposition, we set $\bar{R} = \frac{1}{\beta}$, which means, in terms of log-deviations, that $\tilde{R}_t = 0$ for all $t \leq T-1$.[18] Part (ii) permits us to identify forward guidance with expectations of the interest rate R_T that will obtain right after the ZLB has ceased to bind.[19] Finally, part (iii) allows R_T to be a random variable, so that we can accommodate variation in expectations of it, as well as higher-order uncertainty about it.

How can the policy maker influence expectations of R_T? We bypass this question and, instead, focus on understanding the extent to which the policy maker can influence expectations of future income and future inflation, taking as given the shift in expectations of future policy. To be concrete, however, it is useful to think of the policy maker as making an announcement at $t = 0$ that is informative about the likely value of R_T. We can then obtain the crux of our results by varying the degree of common knowledge about this announcement. This anticipates the exercise we conduct in Section 6. Before doing this, however, we review the forward-guidance puzzle in the context of the standard NK model.

4 The Standard NK Model and the Puzzle

The standard version of the NK model is nested in our framework by imposing the following assumption.

[18] Note that $\tilde{R}_t \equiv \log R_t - \log R^*$, where $R^* = \frac{1}{\beta}$ is the steady-state interest rate. Also note that we can readily introduce a discount-rate shock that forces the ZLB to bind for all $t < T$ and accordingly let $\bar{R} = 1$. All the results then continue to hold, subject to re-interpreting $\tilde{y}_t$ and $\tilde{\pi}_t$, for all $t < T$, as the gaps from the liquidity-trap path rather than from the steady state.

[19] The results extend if we assume that flexible-price allocations obtain after a commonly known date T', for any $T' \geq T+1$. Letting $T' = T+1$ is only for simplicity. Finally, the assumption that such a T' exists and is commonly known is consistent with standard treatments in the literature (e.g., Eggertsson and Woodford, 2003, Werning, 2012), but is not innocuous: it sidesteps the equilibrium selection issues emphasized in Cochrane (2016b) by pegging beliefs of outcomes at $t \geq T'$.

Assumption 2 (Complete Information) *All consumers and all firms share the same information at all dates and all states of Nature: $\mathcal{I}_{i,t} = \mathcal{I}^f_{j,t} = \mathcal{I}_t$ for all i, j, t and all states of Nature.*

This assumption allows the agents to be uncertain about future monetary policy and future economic outcomes. In combination with the REE concept, however, it rules out all kinds of *higher-order* uncertainty: all agents always share the same expectations, not only of future interest rates, but also of future income and future inflation, plus this fact is itself common knowledge. It is this kind of "perfection" in the coordination of beliefs that we find unrealistic and that we seek to relax in the sequel. For the time being, however, we can use this property to reach the following.[20]

Lemma 1 *Under Assumption 2, equilibrium output and inflation solve the following system:*

$$\tilde{y}_t = E_t\left[\tilde{y}_{t+1}\right] - \tilde{R}_t + E_t\left[\tilde{\pi}_{t+1}\right], \tag{5}$$

$$\tilde{\pi}_t = \beta E_t\left[\tilde{\pi}_{t+1}\right] + \kappa\tilde{y}_t + \kappa\tilde{\mu}_t, \tag{6}$$

where $\kappa \equiv \frac{(1-\theta)(1-\beta\theta)(\epsilon+1)}{\theta} \geq 0$ and where E_t denotes the rational expectation conditional on $\mathcal{I}_t$.

Condition (5) is the standard Euler condition. Condition (6) is the standard New-Keynesian Philips curve (NKPC). It is well known how to derive these conditions in the representative-agent version of the NK model. Lemma 1 establishes that, as long as information is complete, the same conditions characterize the log-linearized equilibrium of our model despite the presence of idiosyncratic risk.[21]

To study the predictions of the model for forward guidance, we bring in Assumption 1. From part (ii) of this assumption, $E_T[\tilde{y}_{T+1}] = E_T[\tilde{\pi}_{T+1}] = 0$. From part (i), $\tilde{R}_t = 0$ for all $t < T$. Using these facts, iterating (5) and (6) backwards from $t = T$ to $t = 0$, and using the Law of Iterated Expectations,[22] we can obtain the following characterization of equilibrium spending at $t = 0$.

Lemma 2 *There exists a function $\phi^* : \mathbb{N} \to \mathbb{R}_+$ such that, under Assumptions 1 and 2, equilibrium spending at $t = 0$ is given by*

$$\tilde{y}_0 = -\phi^*(T) \cdot E_0[\tilde{R}_T].$$

This result permits us to identify the "power of forward guidance" by the scalar $\phi^*(T)$: this scalar measures how much the policy maker can stimulate economic activity at $t = 0$ by shifting expectations of the interest rate at $t = T$. As noted earlier, we can think of the policy maker influencing $E_0[\tilde{R}_T]$ through a policy announcement. By defining the power of forward guidance in terms of the elasticity of $\tilde{y}_0$ with respect to $E_0[\tilde{R}_T]$, as opposed to its elasticity with respect to the announcement itself, we sidestep the question of how credible or noisy that announcement might be. While this is an important question on its own right, it is not

[20] To simplify the notation, we henceforth rescale the markup shock by $(1 + \epsilon)$.

[21] This clarifies, first, that the sole modeling role of this risk is to limit the revelation of information and, second, that the benchmark of comparison is indeed the textbook, representative-agent, model, with its familiar equations.

[22] At the individual level, this law applies *always*. At the aggregate level, it applies if information is complete, but not otherwise.

the question of interest for us. Instead, the question of interest is how much influence the policy maker has on economic activity *conditional* on her influence on expectations of interest rates. In the standard model, the answer to this question is given by ϕ^*.[23]

With this in mind, we can now formulate our version of the "forward-guidance puzzle" as follows.

Proposition 1 (Benchmark) *The scalar $\phi^*(T)$, which measures the power of forward guidance under common knowledge, satisfies the following:*

(i) If $\kappa = 0$ (equivalently, $\theta = 1$), then $\phi^*(T) = 1$ for all T.

(ii) If $\kappa > 0$ (equivalently, $\theta < 1$), then $\phi^*(T)$ is strictly higher than 1 for all T, is strictly increasing in T, and explodes to infinity as $T \to \infty$.

Part (i) considers the pedagogically useful case in which prices are *infinitely* sticky (i.e., $\theta = 1$), so that inflation is completely unresponsive to aggregate demand (i.e., $\kappa = 0$). In this case, the power of forward guidance is the same regardless of the horizon at which it operates. Part (ii) considers the more relevant case in which inflation is responsive (i.e., $\kappa > 0$). In this case, the power of forward guidance increases without bound as the time of action is pushed further and further into the future.

To understand part (i), it suffices to inspect the Euler condition *alone*. When prices are totally sticky ($\theta = 1$), inflation is fixed at zero. Iterating the Euler condition between $t = 0$ and $t = T$ gives

$$\tilde{y}_0 = -\tilde{R}_0 - \sum_{t=1}^{T} E_0[\tilde{R}_t] + E_0[\tilde{y}_{T+1}].$$

This reveals an important property: if we fix expectations of inflation, the effect of future interest rates on current aggregate spending is the same as that of current interest rates, regardless of how far in the future we look at. Note that this marginal effect is 1 here because we have fixed the elasticity of intertemporal substitution to 1; allowing the latter to be another number is completely inconsequential for our purposes: it merely scales $\phi^*(T)$ by that number for all T. Part (i) is a direct implication of this property along with the fact that Assumption 1 pegs the interest rate before T and kills the output gap at $T + 1$.

Part (ii) rests on the *interaction* of the Euler equation and the NKPC, that is, on the feedback loop between aggregate spending and inflation. Reducing the interest rate at $t = T$ causes inflation at $t = T$. Because the nominal interest rate is pegged prior to T, this translates to a low *real* interest rate between $T - 1$ and T. This stimulates demand at $T - 1$, contributing to even higher inflation at $T - 1$, which in turn feeds to even

[23] To shed further light on the meaning of ϕ^*, drop Assumption 1 and consider, instead, the following exercise. Starting from an arbitrary equilibrium path, consider a variation in monetary policy that changes expectations of interest rates at $t \leq T$, but not after T, for an arbitrary $T \geq 1$. Then, for all $t \leq T$,

$$\Delta\tilde{c}_t = -\sum_{\tau=0}^{T-t} \phi^*(\tau) \cdot \Delta E_t[\tilde{R}_{t+\tau}],$$

where $\Delta E_t[\tilde{R}_{t+\tau}]$ is the aforementioned change in expectations of interest rates, $\Delta\tilde{c}_t$ is the resulting change in equilibrium spending, and ϕ^* is the same function as the one in Lemma 1. Therefore, even if we move outside the liquidity-trap context, we can still interpret $\phi^*(\tau)$ as the partial elasticity of equilibrium spending with respect to the expected interest rate τ periods ahead, where "partial" means holding expectations of other interest rates constant.

higher demand at $T-2$, and so on. Clearly, the cumulative effect at $t=0$ increases with T, which explains why the power of forward guidance also increases with T.

We henceforth refer to this feedback loop as the "inflationary spiral" caused by the promise of lax monetary policy in the future. As the term suggests, this is the flip-side of the "deflationary spiral" that helps the NK model generate a large recession out of a small drop in aggregate demand (e.g., an exogenous discount-factor shock). The trigger is different, and the sign is reversed, but the mechanism is the same. By the same token, the insights we develop in the sequel regarding the attenuation of this mechanism apply, not only to forward guidance, but also to demand shocks whose magnitude or persistence is not common knowledge.

In the standard model, the aforementioned spiral causes the power of forward guidance to increase without bound as T increases—which is what part (ii) of Proposition 1 states. The "forward-guidance puzzle" pertains to this theoretical prediction and to its quantitative counterparts: plausible calibrations of the model deliver a quantitatively large $\phi^*(T)$ even if T is relatively small. For more details, see Carlstrom, Fuerst, and Paustian (2012), Del Negro, Giannoni, and Patterson (2012), and McKay, Nakamura, and Steinsson (2016b).

This completes our review of the standard NK model. In the sequel, we show how relaxing the common-knowledge requirements of the model reduces the ferocity of the inflationary spiral and, in so doing, helps diminish the puzzle. But this is only one part of the story. Even if we shut down this spiral by brute force (i.e., by setting $\kappa = 0$ and forcing inflation to be unresponsive), we still get an attenuation effect relative to the standard model: part (i) of Proposition 1 ceases to hold. As we explain, this is because there is another important GE mechanism at work, which is hidden underneath the Euler condition and which is also attenuated once information is incomplete; this mechanism is the modern reincarnation of what was known as the "income multiplier" in the IS-LM model.

5 Deconstructing the NK Model into Two Beauty Contests

In this section, we develop a game-theoretic representation of the NK model, which clarifies the relevant general-equilibrium interactions and unearths the role of higher-order beliefs. To this goal, we momentarily drop, not only the assumption of complete information, but also the REE solution concept. This permits us to characterize the optimal behavior of the consumers and the firms as functions of their *subjective* expectations of future income and inflation or, equivalently, of the future actions of other agents. Once this is accomplished, it becomes evident how the predictions of the standard NK model hinge on presuming a certain "perfection" in the adjustment of this kind of *endogenous* expectations to any *exogenous* change in the environment—and, conversely, how the friction we accommodate in this paper rationalizes a smaller adjustment in these expectations, thereby also attenuating the GE effects of forward guidance.

Expectations and Beauty Contests

We start by characterizing the optimal behavior of the consumers and of the firms, while allowing them to hold arbitrary subjective beliefs of future income and future inflation. This leads to our first key result: we represent the NK model as a pair of beauty contests, one for the demand block and another for the supply block of the model. This result is instrumental for the subsequent analysis, but it is also interesting on its own right: its applicability extends well beyond the context of forward guidance.

Consumers. Consider an arbitrary consumer i. As usual, the following intertemporal budget constraint holds in all periods and all states of Nature:[24]

$$\sum_{k=0}^{+\infty}\left\{\Pi_{j=1}^{k}\left(\frac{R_{t+j-1}}{\pi_{t+j}}\right)^{-1}c_{i,t+k}\right\}=a_{i,t}+\sum_{k=0}^{+\infty}\left\{\Pi_{j=1}^{k}\left(\frac{R_{t+j-1}}{\pi_{t+j}}\right)^{-1}\left(w_{i,t+k}n_{i,t+k}+e_{i,t+k}\right)\right\} \quad . \tag{7}$$

Taking the log-linear approximation of the above around the steady state,[25] we get the following:

$$\sum_{k=0}^{+\infty}\beta^k\tilde{c}_{i,t+k}=\tilde{a}_{i,t}+\sum_{k=0}^{+\infty}\beta^k\left\{\Omega\left(\tilde{w}_{i,t+k}+\tilde{n}_{i,t+1}\right)+\left(1-\Omega\right)\tilde{e}_{i,t+k}\right\}, \tag{8}$$

where Ω is the ratio of labor income to total income in steady state. The consumer's optimality conditions, on the other hand, can be expressed as follows:

$$\tilde{n}_{i,t}=\frac{1}{\epsilon}\left(\tilde{w}_{i,t}-\tilde{c}_{i,t}\right), \tag{9}$$

$$\tilde{c}_{i,t}=E_{i,t}\left[\tilde{c}_{i,t+1}-\tilde{R}_t+\tilde{\pi}_{t+1}\right], \tag{10}$$

where $E_{i,t}$ is the expectation of consumer i in period t. The first condition describes optimal labor supply; the second is the *individual-level* Euler condition, which describes optimal consumption and saving.

At this point, it is worth emphasizing that our analysis preserves the *standard* Euler condition at the individual level. This contrasts with McKay, Nakamura, and Steinsson (2016a,b) and Werning (2015), where liquidity constraints cause this condition to be violated for *some* agents, as well as with Gabaix (2016), where a cognitive friction causes this condition to be violated for *every* agent. We revisit this point in Section 9, when we show that our analysis rationalizes a *discounted* Euler condition at the aggregate level, *in spite of* the preservation of the standard condition at the individual level.

Combining conditions (8), (9) and (10), we obtain the optimal expenditure of consumer i in period t as a function of the current and the expected future values of wages, dividends, and real interest rates:

$$\begin{aligned}\tilde{c}_{i,t}=\tfrac{\epsilon(1-\beta)}{\epsilon+\Omega}\tilde{a}_{i,t}-\sum_{k=1}^{+\infty}\beta^k\left(E_{i,t}\left[\tilde{R}_{t+k-1}-\tilde{\pi}_{t+k}\right]\right)\\+\left(1-\beta\right)\left[\tfrac{(\epsilon+1)\Omega}{\epsilon+\Omega}\tilde{w}_{i,t}+\tfrac{\epsilon(1-\Omega)}{\epsilon+\Omega}\tilde{e}_{i,t}\right]+\left(1-\beta\right)\sum_{k=1}^{+\infty}\beta^kE_{i,t}\left[\tfrac{(\epsilon+1)\Omega}{\epsilon+\Omega}\tilde{w}_{i,t+k}+\tfrac{\epsilon(1-\Omega)}{\epsilon+\Omega}\tilde{e}_{i,t+k}\right],\end{aligned} \tag{11}$$

[24] For now, one should think of the state of Nature as a realization the exogenous payoff relevant shocks along with the entire cross-sectional distribution of subjective beliefs in the population. Once we impose the REE concept, the latter can be replaced with the cross-sectional distribution of the exogenous signals (information) received by the agents.

[25] In a symmetric steady state, $a_{it}=a^*=0$. For this reason, we let $\tilde{a}_{i,t}\equiv\frac{a_{i,t}}{c^*}$, where c^* is steady-state consumption.

This condition, which is a variant of the consumption function seen in textbook treatments of the Permanent Income Hypothesis,[26] contains two elementary insights. First, all future variables—wages, dividends, and real interest rates—are discounted. Second, the *current* spending of a consumer depends on the present value of her income, which in turn depends, in equilibrium, on the *future* spending of other consumers.

The first property guarantees that the decision-theoretic, or partial-equilibrium, effect of forward guidance diminishes with the horizon at which interest rates are changed; the second represents a dynamic strategic complementarity, which is the modern reincarnation what was known as the "income multiplier" in the IS-LM framework. We elaborate on these two points in the sequel. For the time being, we aggregate condition (11), and use the facts that assets average to zero and that future idiosyncratic shocks are unpredictable, to obtain the following condition for aggregate spending:

$$\begin{aligned}\tilde{c}_t = -\sum_{k=1}^{+\infty} \beta^k \bar{E}_t \left[\tilde{R}_{t+k-1} - \tilde{\pi}_{t+k}\right] + (1-\beta)\left[\tfrac{(\epsilon+1)\Omega}{\epsilon+\Omega}\tilde{w}_t + \tfrac{\epsilon(1-\Omega)}{\epsilon+\Omega}\tilde{e}_t\right] \\ + (1-\beta)\sum_{k=1}^{+\infty}\beta^k \bar{E}_t\left[\tfrac{(\epsilon+1)\Omega}{\epsilon+\Omega}\tilde{w}_{t+k} + \tfrac{\epsilon(1-\Omega)}{\epsilon+\Omega}\tilde{e}_{t+k}\right]\end{aligned} \tag{12}$$

where $\bar{E}_t[\cdot]$ henceforth denotes the *average* expectation of the consumers in period t.

Firms. Consider a firm j that gets the chance to reset its price during period t. The optimal reset price, denoted by p_t^{j*}, is given by the following:[27]

$$\tilde{p}_t^{j*} = (1-\beta\theta)\left\{(\tilde{w}_t^j + \tilde{p}_t) + \sum_{k=1}^{+\infty}(\beta\theta)^k E_{j,t}^f\left[\tilde{w}_{t+k}^j + \tilde{p}_{t+k}\right]\right\} + (1+\epsilon)(1-\beta\theta)\tilde{\mu}_t^j, \tag{13}$$

where $E_{j,t}^f$ denotes the firm's expectations in period t and $\tilde{\mu}_t^j$ is the corresponding markup shock. The interpretation of this condition is familiar: the optimal "reset" price is given by the expected nominal marginal cost over the expected lifespan of the new price, plus the markup. Aggregating the above condition, using the fact that the past price level is known and that inflation is given by $\tilde{\pi}_t = (1-\theta)(\tilde{p}_t^* - \tilde{p}_{t-1})$, where $\tilde{p}_t^* \equiv \int \tilde{p}_t^{j*}dj$, we obtain the following condition for the level of inflation in period t:

$$\tilde{\pi}_t = \tfrac{(1-\theta)(1-\beta\theta)}{\theta}\tilde{w}_t + \tfrac{(1-\theta)(1-\beta\theta)}{\theta}\sum_{k=1}^{+\infty}(\beta\theta)^k \bar{E}_t^f[\tilde{w}_{t+k}] + \tfrac{1-\theta}{\theta}\sum_{k=1}^{+\infty}(\beta\theta)^k \bar{E}_t^f[\tilde{\pi}_{t+k}] + \kappa\tilde{\mu}_t, \tag{14}$$

where $\bar{E}_t^f[\cdot]$ henceforth denotes the average expectation of the firms. The latter may or may not be the same as the average expectation of the consumers.

[26]To see this more clearly, suppose that initial assets are zero, that the real interest rate is expected to equal the subjective discount rate at all periods, and that labor supply is fixed ($\epsilon \to \infty$). Condition (11) then reduces to $\tilde{c}_{i,t} = (1-\beta)[\Omega\tilde{w}_{i,t} + (1-\Omega)\tilde{e}_{i,t}] + (1-\beta)\sum_{k=1}^{+\infty}\beta^k E_{i,t}[\Omega\tilde{w}_{i,t+k} + (1-\Omega)\tilde{e}_{i,t+k}]$, which means that optimal consumption equals "permanent income" (the annuity value of current and future income). Relative to this benchmark, condition (11) adjusts for three factors: for the endogeneity of labor supply, which explains the different weights on wages and dividends; for initial assets, which explains the first term in condition (11); and for the potential gap between the real interest rate and the subjective discount rate, which explains the second term.

[27]To economize notation, we rescale the markup shocks: we replace $\tilde{\mu}_t^j$ and $\tilde{\mu}_t$ with $(1+\epsilon)\tilde{\mu}_t^j$ and $(1+\epsilon)\tilde{\mu}_t$.

Market Clearing, Wages, and Profits. Because the final-good sector is competitive and the technology satisfies (3) and (4), we have that $\tilde{p}_t = \int \tilde{p}_t^j dj$ and $\tilde{y}_t = \int \tilde{y}_t^j dj = \int \tilde{l}_t^j dj$. The latter, together with market clearing in the labor market, gives $\tilde{y}_t = \tilde{n}_t \equiv \int \tilde{n}_{i,t} di$. Market clearing in the market for the final good, on the other hand, gives

$$\tilde{y}_t = \tilde{c}_t \equiv \int \tilde{c}_{i,t} di.$$

Finally, note that the real profit of monopolist j at period t is given by $e_t^j = \left(\frac{p_t^j}{p_t} - w_t^j\right) y_t^j$. Log-linearizing and aggregating the latter gives $\tilde{e}_t = -\frac{\Omega}{1-\Omega}\tilde{w}_t + \tilde{y}_t$. Combining all these facts with (9), the optimality condition for labor supply, we arrive at the following characterization of the aggregate wages and the profits:

$$\tilde{w}_t = (\epsilon + 1)\, \tilde{y}_t \qquad \text{and} \qquad \tilde{e}_t = \left[1 - \tfrac{\Omega(\epsilon+1)}{1-\Omega}\right] \tilde{y}_t. \tag{15}$$

Beauty Contests. Condition (12), which follows merely from consumer optimality, pins down aggregate spending as a function of the average subjective beliefs of wages, profits, interest rates, and inflation. If a consumer knows that markets clear and that other consumers are rational, she can infer that (15) holds, which means that her subjective beliefs of wages and profits are pinned down by her subjective beliefs of aggregate spending according to (15). Aggregate spending can then be expressed as a function of the consumers' average beliefs of interest rates, of inflation, and of aggregate spending itself. Similarly, combining (14) and (15), we can express aggregate inflation as a function of the firms' average beliefs of aggregate spending and of inflation itself. We therefore reach the following result.

Proposition 2 (Beauty Contests) *Suppose that agents have (first-order) knowledge of the structure of the economy and of the rationality of other agents. Then, aggregate spending satisfies*

$$\tilde{y}_t = -\sum_{k=1}^{+\infty} \beta^{k-1} \left\{ \bar{E}_t[\tilde{R}_{t+k-1}] - \bar{E}_t\left[\tilde{\pi}_{t+k}\right] \right\} + (1-\beta) \left\{ \sum_{k=1}^{+\infty} \beta^{k-1} \bar{E}_t\left[\tilde{y}_{t+k}\right] \right\}, \tag{16}$$

where $\bar{E}_t$ denotes the average expectation of the consumers. Inflation, on the other hand, satisfies

$$\tilde{\pi}_t = \kappa \tilde{y}_t + \kappa \sum_{k=1}^{+\infty} (\beta\theta)^k\, \bar{E}_t^f\left[\tilde{y}_{t+k}\right] + \tfrac{1-\theta}{\theta} \sum_{k=1}^{+\infty} (\beta\theta)^k\, \bar{E}_t^f\left[\tilde{\pi}_{t+k}\right] + \kappa \tilde{\mu}_t, \tag{17}$$

where $\bar{E}_t^f$ denotes the average expectation of the firms.

This result, which has been obtained *without* the use of either the equilibrium concept or any particular information structure, permits us to interpret the NK model as the interaction of two dynamic beauty contests, one for the "demand block" of the model and another for the "supply block".

Condition (16) defines a dynamic beauty contest among the consumers. In this game, expectations of inflation shape expectations of *real* interest rates. Conditional on the latter, the optimal spending of a consumer increases with current aggregate spending as well as with her expectations of future aggregate spending, because the spending of other consumers determines her own income. Condition (16) therefore isolates the strategic complementarity that operates within the demand block—or, equivalently, the GE mechanism that we refer to as the "income multiplier".

Condition (17) defines a dynamic beauty contest among the firms. In this game, expectations of aggregate spending shape expectations of *real* marginal costs. Conditional on these expectations, the optimal reset price of a firm depends, not only on past inflation, but also on her expectation of future inflation, because the latter shapes the firm's *nominal* marginal costs in the future. Condition (17) therefore isolates the strategic complementarity that operates within the supply block.

To the best of our knowledge, our representation of the demand block of the NK model as a "consumption beauty contest" is novel to the literature.[28] This representation is instrumental for the subsequent analysis. Its usefulness, however, extends beyond the scope of this paper: by recasting the income multiplier that is hidden behind the representative-agent Euler condition as a form of strategic complementarity, we indicate more generally how lack of common knowledge—whether of the state of Nature or of the rationality of agents—can attenuate the related GE effects of all kinds of policy or other shocks.

The "inflation beauty contest" in condition (17), on the other hand, builds a bridge between our work and that of Woodford (2003a). Similarly to that paper, the fact that higher-order beliefs are anchored helps rationalize inertia in the response of inflation to changes in aggregate demand. In contrast to that paper, however, the change in aggregate demand is itself the product of the endogenous behavior of the consumers, a property that adds an extra layer of higher-order beliefs. Furthermore, the firms in our framework face no uncertainty about *current* marginal costs. Instead, they only face uncertainty about *future* marginal costs. This uncertainty matters because the Calvo friction causes the firms' pricing-setting behavior to be forward-looking, a property that is absent in Woodford (2003a). By the same token, the complementarity captured in condition (17) is distinct from the one emphasized in Woodford (2003a): whereas the latter concerns the static relation between aggregate income and real marginal costs, ours concerns the degree of price stickiness and is the game-theoretic analogue of the forward-looking aspect of the NKPC.[29]

Each of the aforementioned beauty contests describes the behavior of one group of agents—the consumers or the firms—taking as given this group's expectations of the other group's behavior. The *interaction* of the two groups is itself a higher-layer beauty contest, one that captures the feedback loop between aggregate spending and inflation, that is, the GE mechanism that we refer to as the "inflationary spiral".

Remark. The beauty contests obtained above involve expectations of the actions of others in *future* periods. In this regard, they are reminiscent of the beauty contests found in the literature on incomplete-

[28] As already explained, this representation follows, essentially, from the Permanent Income Hypothesis together with the fact that aggregate output equals aggregate consumption; the novelty rests in the game-theoretic interpretation and the related implications, which will be come clear in the sequel. Farhi and Werning (2016) have concurrently developed a similar representation, which is used to study the interaction of level-k reasoning with incomplete markets. The same kind of beauty contest appears also in Angeletos and Lian (2016c), although in the context of a model that abstracts from nominal rigidity and aims at different goals. Finally, Wiederholt (2016), contains a *different* beauty-contest representation, one that concerns the interaction of the Euler condition and the NKPC as opposed to opening up the Euler condition.

[29] By letting firms face both incomplete information and a Calvo friction, Nimark (2008) also touches on the role of forward-looking higher-order beliefs in price-setting behavior, although for different purposes than our paper.

information asset-pricing models, such as Singleton (1987), Allen, Morris and Shin (2006), and Bacchetta and Van Wincoop (2006). There is, however, a certain difference. In the aforementioned papers, outcomes in period t depend merely on expectations of outcomes in period $t+1$. In our setting, by contrast, expectations of *all* periods after t matter. This explains the additional complexity in the structure of the hierarchy of beliefs that is relevant in our paper, as well as the specific horizon effect we document.

Higher-Order Beliefs and Equilibrium Expectations

The preceding analysis imposed individual rationality together with market clearing, but did not impose either an equilibrium concept or common knowledge of rationality. It thus obtained aggregate outcomes as functions of *subjective* expectations of future nominal interest rates, future inflation, and future income. Throughout our analysis, the expectations of nominal interest rates are treated as exogenous. The expectations of inflation and income, instead, are crucial endogenous objects. To pin down this kind of expectations, the macroeconomic theorist must specify the solution concept and the information structure.

This is a delicate choice, which the standard practice resolves in rather lighthearted manner. By imposing the REE concept along with complete information, standard macroeconomic models impose common knowledge of rationality, of strategies, and of the state of Nature. These are strong assumptions that together impose—indeed define—a certain kind of perfection in the response of this kind of subjective expectations to exogenous impulses. In this paper, by contrast, we accommodate a realistic imperfection in the aforementioned response by allowing information to be incomplete.

For the rest of the paper, we thus bring back the REE concept, as well as Assumption 1,[30] but refrain from imposing Assumption 2. We thus show how the standard models imposes a certain restriction on the response of expectations of income and inflation to forward guidance—and, conversely, how the friction we accommodate in this paper helps modify this response in an empirically plausible manner.

Whatever the information structure may happen to be, once we have imposed the REE concept, we can express the relevant expectations of future income and inflation as a function of the *higher-order beliefs* of $\tilde{R}_T$. To see this more clearly, suppose momentarily that prices are infinitely sticky ($\theta = 1$ and, equivalently, $\kappa = 0$), which kills inflation and let us concentrate on the demand side.

Since prices cannot change, inflation expectations satisfy $\bar{E}_t\left[\tilde{\pi}_\tau\right] = 0$ for all (t, τ). By Assumption 1, on the other hand, income expectations satisfy $\bar{E}_t\left[\tilde{y}_\tau\right] = 0$ for all (t, τ) such that $t \leq T < \tau$, whereas interest-rate expectations satisfy $\bar{E}_t[\tilde{R}_\tau] = 0$ for all (t, τ) such that $t \leq \tau < T$. It follows that, for any $t < T$, condition (16) reduces to the following:

$$\tilde{y}_t = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + (1-\beta)\left\{\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right]\right\}$$

This is a simplified version of our "consumption beauty contest", where inflation expectations have dropped

[30] For our purposes, it suffices to employ the weaker solution concept of Rationalizability: once Assumption 1 is imposed, our results follow directly from iterating conditions (16) and (17), that is, from assuming common knowledge of rationality.

out, leaving the "income multiplier" among the consumers as the only relevant GE mechanism.

Now, pick any (t,τ) such that $\tau < t < T$ and take the period-τ average expectation of both sides of the above condition to obtain the following representation of $\bar{E}_\tau[\tilde{y}_t]$:

$$\bar{E}_\tau[\tilde{y}_t] = -\beta^{T-t}\bar{E}_\tau\bar{E}_t[\tilde{R}_T] + (1-\beta)\left\{\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_\tau\bar{E}_t\left[\tilde{y}_{t+k}\right]\right\}.$$

This gives the consumers' *first-order* equilibrium beliefs (i.e., the equilibrium expectations) of future income as a function of their *second-order* beliefs of interest rates and income. Next, provided that $\tau > 1$, for any $s < \tau$ we can take the period-s average expectation of both sides of the above to obtain $\bar{E}_s\bar{E}_\tau[\tilde{y}_t]$, the second-order beliefs of income, as a function of *third-order* beliefs of interest rates and income. Using this argument iteratively, we can ultimately express the consumers expectations' of income at any given period as a function of their hierarchy of beliefs of future interest rates (i.e., of k-order beliefs for $k = 1, 2, 3$,...).

It is worth noting that this iterative procedure assumes higher and higher orders of knowledge of the rationality of others. Furthermore, *any* such order is implied by common knowledge of rationality, which is itself imposed by the REE concept. This explains why this procedure helps, not only characterize, but also dissect the expectations that obtain under the REE concept.

The same basic logic applies if we let $\kappa > 0$, except for two differences. First, higher-order beliefs pin down, not only expectations of income, but also expectations of inflation. Second, the relevant higher-order beliefs get richer, because they involve, not only what the consumers think about other consumers, but also what the firms think about other firms, as well as what the consumers think about the firms and vice versa. The consumers' beliefs about other consumers matter through the income multiplier; the firms' beliefs about other firms matter through the pricing complementarity; and the one group's beliefs about the other group matter through the inflationary spiral.

We summarize these observations in the following.

Corollary 1 *Regardless of the information structure, the* equilibrium *expectations of future income and future inflation are pinned down by the hierarchy of beliefs about $\tilde{R}_T$.*

For our purposes, higher-order beliefs of $\tilde{R}_T$ are therefore *synonymous* to expectations of future income and inflation and are directly tied to the GE effects of forward guidance. It is then important to recognize how different informational assumptions map to different restrictions on these objects.

By imposing complete information along with the REE concept, the standard version of the NK imposes that all kinds of higher-order beliefs collapse to first-order beliefs. Although this makes it impossible to detect the role of higher-order beliefs, this does not mean that higher-order beliefs don't matter. It only means that higher-order beliefs have been "swept under the carpet".

From this perspective, the added value of Proposition 2 is to "open up" the two fundamental equations of the NK model and to reveal the higher-order beliefs that operate both within and across the demand and the supply blocks of the model. The next step is to show how incomplete information helps anchors all these higher-order beliefs and, in so doing, also anchors the *equilibrium* expectations that consumers and firms form about future income and future inflation. We complete this step in the next sections.

PE vs GE, and Relaxing Rational Expectations

We conclude the present section with few additional points regarding the crucial distinction between GE and PE effects and the possible interpretation of the friction we introduce in this paper.

Inspecting condition (16) reveals the following basic point: if we were to hold constant expectations of inflation and aggregate spending, a one-unit reduction in $\bar{E}_t \tilde{R}_T$ would have raised $\tilde{y}_t$ by β^{T-t}. This effect can be interpreted as the PE effect of forward guidance.[31] Clearly, this effect is bounded, decreases with the distance between t and T, and vanishes as this distance goes to infinity. This clarifies that the puzzle is exclusively about the GE effects of forward guidance. But now note the following: any mechanism that anchors expectations of future inflation and income also attenuates the relevant GE effects. This facilitates the following complementary interpretation of what we do in the sequel: by accommodating incomplete information, we operationalize the notion that GE effects operate in "less-than maximum capacity".

Another complementary interpretation is the following. As noted before, conditions (16) and (17) are valid for arbitrary *subjective* expectations of future income and future income. Imposing the REE concept along with complete information—which is the standard practice—restricts these expectations in a very specific manner, ultimately leading to the predictions we reviewed in the previous section. Relative to this benchmark, relaxing *either* the REE concept *or* the assumption of complete information help "free up" expectations. This explains the sense in which incomplete information can be seen as a *substitute* for relaxing rational expectations. That said, it is not the case that "anything goes" once we allow for incomplete information. Rather, we obtain a structured, or "disciplined", departure from the standard model. The precise structure is made clear in the rest of the paper.[32]

6 Revisiting the Power of Forward Guidance

In the previous section, we represented the NK model as a pair of beauty contests and used this representation to recast the equilibrium expectations of income and inflation, and the relevant GE effects, in terms of higher-order beliefs. To complete our contribution, we need to show how lack of common knowledge anchors this kind of expectations and, in so doing, also attenuates the GE effects of forward guidance. In this section,

[31] Going back to condition (11), we see that the elasticity of individual consumption with respect to the individual's own expectation of $\tilde{R}_T$, holding constant her expectations of *future* income and inflation, as well as her *current* income, is β^{T-t+1}. The difference between this number and the one reported in the main text is due to the fact that (11) has already incorporated the GE effect on *contemporaneous* income. This difference, however, is an artifact of the discrete-time specification of the model: as the length of the time interval shrinks to zero, the effect of contemporaneous income on spending becomes vanishingly small. This justifies the interpretation given in the main text.

[32] This discussion echoes the related broader observations we have made in Angeletos and Lian (2016a,b). It also highlights the complementarity between the approach taken in this paper and the one taken in Farhi and Werning (2016), which capture bounded rationality with a weaker solution concept, namely, level-k reasoning.

we illustrate this point with the help of a stark specification of the information structure, namely under the following assumption.

Assumption 3 *(i) Monetary policy is specified as in Assumption 1 and the period-0 policy announcement is given by* $z \sim \mathcal{N}(\tilde{R}_T, \sigma_z^2)$

(ii) Each consumer and each firm receives a private signal of z at $t = 0$. This signal is given by $x_i = z + v_i$ for consumer i and by $x^j = z + v^j$ for firm j, where $v_i \sim N(0, \sigma^2)$ and $v^j \sim N(0, \sigma_f^2)$ are independent of each other, independent of $\tilde{R}_T$, and i.i.d. across, respectively, i and j.

(iii) No other information arrives exogenously at any $t \leq T$.

(iv) The volatilities of the aggregate markup shock and of all the idiosyncratic shocks are infinite.

Part (i) facilitates a concrete interpretation of how forward guidance is conducted: through the announcement z. Accordingly, σ_z^2 can be interpreted as (the inverse of) the precision, or the "trustworthiness", of the policy announcement.

Part (ii) removes common knowledge of the policy announcement: agents observe z with idiosyncratic noise. It is possible to micro-found the idiosyncratic noise as the product of rational inattention or costly information processing, as, e.g., in Sims (2003) or Myatt and Wallace (2012). We have also considered a variant (available upon request) that allows the agents to become gradually aware of the policy announcement, as in Mankiw and Reis (2002), and have found that this delivers similar results as the simple case considered here. With this in mind, one may interpret σ as a measure of how "inattentive" consumers are, and similarly σ^f as the corresponding measure for the firms. That said, as it will be clear in the sequel, the key is that letting $\sigma > 0$ and $\sigma^f > 0$ introduces *higher-order* uncertainty about the policy announcement (i.e., uncertainty about the beliefs of others, the beliefs of the beliefs of others, etc). Accordingly, it is best to interpret σ and σ^f as measures of the departure from common knowledge.

Finally, part (iii) shuts down any other *exogenous* source of information, while part (iv) shuts down *endogenous* learning from all kinds of market outcomes. These are stark assumptions, which we relax in the next section.[33] Here, we employ them in order to guarantee a tractable belief hierarchy and to deliver a sharp illustration of the more general insights of our paper.

Remark. The assumption that the agents lack common knowledge of the policy announcement should not be taken too literally. As already noted, the idiosyncratic noise in the observation of the policy announcement can be the product of rational inattention. Furthermore, even if the announcement itself is perfectly and publicly observed, different agents may give different interpretations to it because they have different private information about the other elements of the state of Nature, thus giving rise once again to higher-order uncertainty about $\tilde{R}_T$. Last but not least, what is ultimately of essence is not the uncertainty about the policy itself, but rather the uncertainty about the responses of other agents.

[33] That said, part (iv) is not necessarily as unrealistic as it may appear at first glance: first, idiosyncratic shocks are at least an order of magnitude bigger than aggregate shocks; and second, estimated DSGE models and VAR-based empirical work alike attribute most of the variation in inflation to markup shocks, or other residuals, which are orthogonal to identified monetary shocks and contribute very little to the observed variation in aggregate employment and output.

Higher Order Beliefs

Recall from Section 5 that aggregate spending at any date $t < T$ depends on expectations of future income and inflation, which in turn can be expressed as functions of higher-order beliefs. With this in mind, we now explain how Assumption 3 pins down the relevant higher-order beliefs.

To simplify the exposition, we momentarily shut down the friction among the firms ($\sigma^f = 0$) and focus on the friction among the consumers ($\sigma > 0$). First, note that the information set of a consumer remains the same for all $t \in \{0, ..., T-1\}$. This means, not only that a consumer's belief of $\tilde{R}_T$ stays constant over time, but also that her beliefs of the *future* beliefs of other consumers collapse to her beliefs of the *current* beliefs of other consumers.[34] Second, note that $E[\tilde{R}_T|z] = \delta z$, where $\delta \equiv \frac{\sigma_z^{-2}}{\sigma_R^{-2}+\sigma_z^{-2}} \in (0,1)$ and, since z contains more information about $\tilde{R}_T$ than the information of any individual consumer, $E_{i,t}[\tilde{R}_T] = E_{i,t}[E[\tilde{R}_T|z]] = \delta E_{i,t}[z]$ and therefore $\bar{E}_t[\tilde{R}_T] = \delta \bar{E}_t[z]$. Finally, note that $E_{i,t}[z] = \lambda x_i$, where x_i is i's signal about z and $\lambda \equiv \frac{\sigma^{-2}}{\sigma^{-2}+\left(\sigma_z^2+\sigma_R^2\right)^{-1}} \in (0,1)$, and therefore $\bar{E}_t[z] = \lambda z$. It follows that the average first-order belief of $\tilde{R}_T$ is given by

$$\bar{E}_t[\tilde{R}_T] = \bar{E}_0[\tilde{R}_T] = \lambda\delta z. \tag{18}$$

Iterating the above then gives the following expression for higher-order beliefs:

$$\bar{E}_0\left[\bar{E}_{\tau_1}\left[...\bar{E}_{\tau_k}\left[\tilde{R}_T\right]...\right]\right] = \lambda^{k+1}\delta z = \lambda^k \bar{E}_0[\tilde{R}_T] \tag{19}$$

for all $k \geq 1$ and all $\{\tau_1, ...\tau_k\}$ such that $0 \leq \tau_1 \leq ... \leq \tau_k \leq T-1$.

For any realization of z, the scalar δ, which relates to the precision, or the "trustworthiness", of the policy announcement, has the same proportional effect on beliefs of all orders. By contrast, the scalar λ, which parameterizes the lack of common knowledge among the consumers,[35] has a *differential* effect across the belief hierarchy. In particular, for any $\lambda < 1$, higher-order beliefs vary less than lower-order beliefs with any given variation in z. Furthermore, the larger the departure from common knowledge is (i.e., the smaller λ is), the faster higher-order beliefs converge to the common prior, whose mean is zero, as $k \to \infty$. Finally, even when the departure from common knowledge is arbitrarily small, in the sense that λ is arbitrarily close to 1, beliefs of sufficiently high order remain arbitrarily anchored to the common prior: for any $\lambda < 1$, the average k-th order belief of $\tilde{R}_T$ converges to 0 as $k \to \infty$, regardless of how close λ is to 1.

In a moment we will see how these properties rationalize why the consumers' expectations of income and inflation can be anchored and, in so doing, also attenuate the power of forward guidance. For now, we would like to emphasize that these properties are not unduly sensitive to the stark informational assumptions we have made. The property that higher-order beliefs are anchored to the common prior is true for *generic* information structures and is indeed the hallmark of the friction we consider in this paper, that is, of the lack of common knowledge.[36]

[34] It is then *as if* the agents play a static beauty contest in each period, with an important twist: the degree of complementarity in this "as if" static game is increasing in T.

[35] Indeed, λ is directly related to the level of "common-p belief" (Kajii and Morris, 1997).

[36] See Samet (1998) for a formalization of this point.

We finally note that, once we allow for a similar friction on the firm side ($\sigma^f > 0$), we obtain similar properties not only about the higher-order beliefs that operate within the group of firms, but also about those that transcend the two groups. By the former we refer to beliefs of the form $\bar{E}_t^f \bar{E}_\tau^f$: this kind of belief satisfy (19) with $\bar{E}_t^f$ in place of $\bar{E}_t$ and $\lambda^f \equiv \frac{\sigma_f^{-2}}{\sigma_f^{-2}+\left(\sigma_z^2+\sigma_R^2\right)^{-1}} \in (0,1)$ in place of λ. By the latter we refer to beliefs of the form $\bar{E}_t \bar{E}_\tau^f$ or $\bar{E}_\tau^f \bar{E}_t$: this kind of belief satisfy a similar condition like (19), except that now λ and λ^f matter jointly.[37] Notwithstanding this additional complexity, the essence remains the same: whether we look at the consumers or the firms, lack of common knowledge anchors higher-order beliefs and therefore also anchors expectations of income and inflation.

The Power of Forward Guidance

The preceding analysis explained why λ and λ^f measure the departure from common knowledge or, equivalently, the degree to which higher-oder beliefs are anchored.[38] The standard model, which we reviewed in Section 4, is nested by letting $\lambda = \lambda^f = 1$. We now extend the results of that section to the case in which λ and λ^f are less than one, that is, to the case in which the consumers and/or the firms lack common knowledge.

Lemma 3 *Under Assumption 3, there exists a function $\phi : (0,1]^2 \times \mathbb{N} \to \mathbb{R}_+$ such that*

$$\tilde{y}_0 = -\phi\left(\lambda, \lambda^f, T\right) \cdot \bar{E}_0[\tilde{R}_T],$$

where λ and λ^f parameterize the informational friction of, respectively, the consumers and the firms.

This lemma generalizes Lemma 2 from the standard model. To see why this lemma is true, recall first that, regardless of the information structure, Proposition 2 permits us to express that aggregate income—and inflation as well—as a linear function of higher-order beliefs of $\tilde{R}_T$. Next, note that, under the assumed information structure, higher-order beliefs are co-linear to first-order beliefs. It follows that we can express aggregate income as a linear function of $\bar{E}_0[\tilde{R}_T]$, with a slope coefficient ϕ that itself depends on the pair (λ, λ^f), on the horizon T, and, of course, on the underlying preference and technologies parameters.[39]

Just as Lemma 2 permitted us to study the power of forward guidance in the standard, common-knowledge benchmark by studying the function ϕ^*, the new lemma permits us to study the power of forward guidance *away* from that benchmark by studying the function ϕ. This is done in the next proposition, which extends Proposition 1 from that benchmark.

Proposition 3 (Attenuation) *The scalar $\phi = \phi(\lambda, \lambda^f, T)$, which measures that power of forward guidance under Assumption 3, satisfies the following properties.*

[37] For example, $\bar{E}_t[\bar{E}_\tau^f[\bar{E}_{\tau'}[\tilde{R}_T]]] = \lambda\lambda^f \bar{E}_t[\tilde{R}_T]$.

[38] By varying σ_z, σ, and σ^f at the same time, we can vary λ and λ^f while keeping $Var(\bar{E}_0[\tilde{R}_T])$ constant. It follows that we can indeed interpret λ and λ^f as scalars that parameterize how anchored higher-order beliefs are, *while holding constant the variation in first-order beliefs*.

[39] To simplify the notation, we suppress the dependence of ϕ on the aforementioned parameters.

(i) When $\kappa = 0$, ϕ is strictly increasing in λ but invariant in λ^f. Furthermore, whenever $\lambda < 1$, ϕ is strictly less than 1, *is strictly decreasing in T, and converges to* 0 *as $T \to \infty$.*

(ii) When instead $\kappa > 0$, ϕ is strictly increasing in both λ and λ^f. Furthermore, whenever $\lambda < 1$ and/or $\lambda^f < 1$, the ratio ϕ/ϕ^* is strictly less than 1, *is strictly decreasing in T, and converges to* 0 *as $T \to \infty$. Finally, when λ is sufficiently low, ϕ also converges to* 0 *as $T \to \infty$.*

To interpret this result, keep in mind that $\phi(1, 1, T) = \phi^*(T)$, where ϕ^* is the function obtained in Proposition 1. This simply means that the common-knowledge benchmark is nested with $\lambda = \lambda^f = 1$.

Part (i) sets $\kappa = 0$ and shuts down the inflation response, thus concentrating on the demand block of the model. Recall from Proposition 1 that, in this case, the standard model predicts that $\phi^* = 1$, regardless of how far in the future forward guidance operates. Here, instead, as long as the consumers lack common knowledge ($\lambda < 1$), we have that $\phi < 1$, it decreases with the horizon T, and it vanishes as T gets larger and larger. In short, the qualitative effect resembles the PE effect of forward guidance. As we explain further in the next section, this is because lack of common knowledge attenuates the income multiplier (which is the only relevant GE effect when $\kappa = 0$).

Part (ii) considers the more general case in which $\kappa > 0$ and inflation is responsive, so that the demand and supply blocks interact with each other. In this case, the standard model predicts that ϕ^* increases monotonically with T and explodes to infinity as $T \to \infty$, due to the feedback loop between aggregate spending and inflation. Relative to this benchmark, lack of common knowledge reduces, not only the level of ϕ, but also its slope with respect to T. That is, not only is the power of forward guidance attenuated, but also the attenuation is stronger the longer the horizon. What is more, the attenuation effect itself explodes with T, in the sense that ϕ becomes vanishingly small relative its common-benchmark counterpart. Finally, if λ is small enough, ϕ becomes vanishingly small, not only in relative to ϕ^*, but also in absolute magnitude.

In the next two sections, we elaborate on the robustness of these findings, on the channels through which the lack of common knowledge operates, and on the precise reason for which the attenuation effect increases with T. For now, we point out that the result stated in the Introduction is a corollary of part (ii) of Proposition 3, obtained under the restriction $\lambda^f = \lambda$. By allowing $\lambda^f \neq \lambda$, we have clarified that it is not necessary to anchor the expectations of *both* the consumers and the firms: anchoring the consumer side is always sufficient by itself, and anchoring the firm side is also sufficient by itself provided that $\kappa > 0$. That said, it is also worth noting that, as long as $\kappa > 0$, the cross-partial derivative of ϕ with respect to λ and λ^f is strictly positive: the friction on the demand side and the friction on the supply side reinforce the marginal effects of each other. Finally, although we have documented the attenuation effect only it terms of aggregate spending, a similar result applies to inflation as well.

A Numerical Illustration

The quantitative evaluation of the power of forward guidance is beyond the scope of this paper. We nevertheless find it useful to illustrate the attenuation effect under a plausible parameterization.

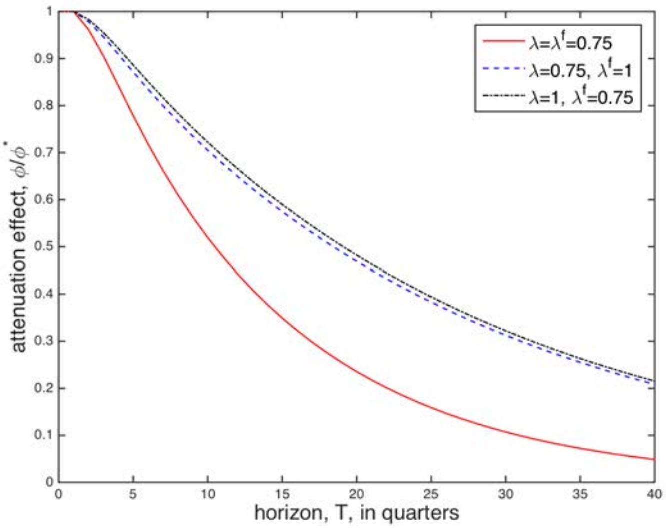

Figure 1: Attenuation effect, relative to common-knowledge benchmark, at different horizons.

For the parameters that are common to the standard NK model, we adopt the same parameterization as the one in McKay, Nakamura, and Steinsson (2016b). We thus interpret the period length as a quarter and set the discount factor, β, to 0.986, targeting an 2% annual interest rate in steady state. We next set the Frisch elasticity to $1/2$, meaning $\epsilon = 2$. Following Christiano, Eichenbaum, and Rebelo (2011), we finally let the price revision rate, $1-\theta$, be 0.15. These values imply $\kappa \approx 0.09$, which corresponds to a relatively flat NKPC and helps the NK model generate realistic volatility in inflation.

What remains is the parameterization of λ and λ^f, which measure the departure from common knowledge. Unfortunately, the existing literature offers little guidance on how to what values may be empirically relevant.[40] In want of a better alternative, we make a plausible yet somewhat arbitrary guess for the degree of departure from common knowledge: we set $\lambda = \lambda^f = 0.75$. One can think of this as a situation in which every agent who has heard the policy announcement believes that any other agent has failed to hear, or "trust", the announcement with a probability equal to 25%. This is arguably a modest "grain of doubt" in the minds of people about their ability to coordinate the adjustment in their beliefs and their behavior.

The solid red line in Figure 1 plots the resulting attenuation effect, as measured by the ratio ϕ/ϕ^*, against the horizon length, T. By setting $\lambda = \lambda^f$, this line assumes that the consumers and the firms are subject to the same level of informational friction. The remaining two lines in the figure shut down the friction in the one side of the economy (that is, we set either $\lambda_f = 1$ or $\lambda = 1$), isolating the role of the friction in the other side: the dashed blue line isolates the friction in the consumer side, the dotted black line isolated the friction in the firm side.

As expected, the attenuation effect is largest when the friction is present in both sides, which each side contributing to about one half of the overall attenuation effect. In addition, the effect appears to quantitatively

[40] For instance, although Coibion and Gorodnichenko (2012, 2015) provide evidence in support of the type of informational friction we have sought to accommodate in this paper, this evidence regards very different types of shocks and circumstances than those considered in our paper, making it hard to extrapolate from their empirical findings to our setting.

significant. For example, at a horizon of 5 years (i.e, $T = 20$), the power of forward guidance is about one quarter of its common-knowledge counterpart. Importantly, the aforementioned reduction in the power of forward guidance is on top of any mechanical effect that the informational friction might have on the extent by which forward guidance shifts expectations of future interest rates: by construction, the documented attenuation effect is normalized by the variation in first-order beliefs of $\tilde{R}_T$.

Expectations of Policy vs Expectations of Income and Inflation

The preceding analysis has treated the ability of the policy maker to shift the expectations of $\tilde{R}_T$ as given. As noted before, this is because we wish to disentangle expectations of income and inflation from expectations of interest rates—a disentangling that becomes meaningful once we have departed from common knowledge. In line with this, we now state the following result, which helps recast the scalars λ and λ^f as measures of the policy maker's capacity to shift expectations of income and inflation.[41]

Proposition 4 (Income and Inflation Expectations) *Under Assumption 3, there exist vectors $(\delta_t^y)_{t=1}^T$ and $(\delta_t^\pi)_{t=1}^T$, which are functions of (λ, λ^f), such that, for all $t \in \{1, ..., T\}$,*

$$\bar{E}_0[\tilde{y}_t] = -\delta_t^y \cdot \bar{E}_0[\tilde{R}_T] \qquad \text{and} \qquad \bar{E}_0[\tilde{\pi}_t] = -\delta_t^\pi \cdot \bar{E}_0[\tilde{R}_T].$$

Furthermore, for any $t < T$, δ_t^y and δ_t^π are strictly increasing in λ and, as long as $\kappa > 0$, they are also strictly increasing in λ^f.

To interpret this result, note that the vectors $(\delta_t^y)_{t=1}^T$ and $(\delta_t^\pi)_{t=1}^T$ measure the shift in the "term structure" of expectations of, respectively, future income and future inflation caused by a one-unit shift in expectations of $\tilde{R}_T$. In simple words, a larger departure from common knowledge decreases the shift in this kind of expectations for any given shift in expectations of future monetary policy.

This also formalizes the following notion, which we put forward in the Introduction.

Corollary 2 *The standard model "maximizes" the ability of the policy maker to control expectations of future income and future inflation in the following sense: the standard model is nested by $\lambda = \lambda^f = 1$, which maximizes the values of δ_t^y and δ_t^π for every $t < T$.*

By the same token, one can interpret the exercise conducted in this paper as a method for operationalizing the notion that the policy maker may face significant difficulty in managing expectations of income and inflation, even if she can easily move expectations of interest rates.

This pegs the following question: can the policy maker influence the values of λ and λ^f by the manner in which she talks and communicates her intensions about future monetary policy? While this question is beyond the scope of this paper, we wish to highlight the following possibility: it may be that the policy maker

[41] Proposition 4 follows from the same arguments as Proposition 3, yet it offers a new perspective.

faces a trade off between the precision and the transparency of her communications.[42] By "precision" we have in mind how informative the communicated signals are about future policy; and by "transparency" we have in mind the degree to which the agents in the economy can reach a common understanding and a common interpretation of the communicated signals. In the present context, the former maps to the reciprocal of σ_z^2, the latter maps to λ and λ^f. A corollary of our results is then that longer horizons raise the value of transparency relative to that of precision.

7 GE Attenuation and Expectation Anchoring

In the previous section we presented the crux of our ideas under a stark specification of the information structure. We now examine the robustness of our insights to friction to decrease with time, as the product of learning via either the endogenous market signals or the arrival of new exogenous information. Most importantly, we elaborate on the intuition behind our results, including how lack of common knowledge attenuates the GE mechanisms discussed earlier on. We finally offer a first pass in evaluating the quantitative importance of anchoring expectations of inflation relative to that of anchoring expectations of output.

Information and Higher-Order Beliefs

For the purposes of this section, we replace Assumption 3 with the following.

Assumption 4 *(i) Monetary policy is specified as in Assumption 1 and the announcement is* $z \sim \mathcal{N}(\tilde{R}_T, \sigma_z^2)$.

(ii) In each period $t \leq T-1$, *each consumer receives a new private signal of the policy announcement* z, *given by* $x_{it} = z + v_{it}$, *where* $v_{it} \sim N(0, \sigma_t^2)$ *is i.i.d. across* i *and* t *and* $\sigma_t > 0$. *In addition, consumers learn from market signals available to them, that is, from prices,* $\tilde{p}_t$ *wages,* $\tilde{w}_{it}$ *and dividends* $\tilde{e}_{it}$.

(iii) The firms have complete information: $\mathcal{I}_{j,t}^f = \mathcal{I}_t$ *for all* j, t *and all states of Nature.*

Part (i) maintains the dual interpretation of z as the policy instrument through which the policy maker attempts to shift expectations of future monetary policy, and as the random variable about which agents lack common knowledge.[43] Part (ii) allows the lack of common knowledge to ease over time, as the consumers can now accumulate more information about z over time both exogenously and through the observation of (non-trivial) market signals. Part (iii) shuts down the informational friction on the firm side. Although this shuts down part of the attenuation effects, it sharpens the exposition by letting us reduce the equilibrium of the economy to a single beauty contest among the consumers *even when* we incorporate the response of inflation (see Lemma 6 in the sequel).

[42] This possibility has been emphasized before by Angeletos and Pavan (2007), Morris and Shin (2007), and Charhour (2014), although not in the context we study here.

[43] Note, in particular, that $E[\tilde{R}_T|\mathcal{I}_t] = E[\tilde{R}_T|z]$ for all $t \in \{1, ..., T-1\}$: if agents had been able to share information, they would, not only reach common knowledge of z, but also use z as the only signal for predicting $\tilde{R}_T$. This property limits the complexity of the hierarchy of beliefs, but is not strictly needed.

Because of the endogeneity of the information contained in the market signals and also because of the continuous arrival of exogenous information, the characterization of the hierarchy of beliefs is significantly more complicated than in the previous section. We can nevertheless make progress by noting the following. Let $h_t \equiv \{\tilde{p}_0, ...\tilde{p}_t\}$ and note that this is a sufficient statistic of the public information in period t. For convenience, let also $h_{-1} = \varnothing$. The following is true.[44]

Lemma 4 *Under Assumption 4, for all $t \leq T-1$, $\tilde{y}_t$ is a linear function of (z, h_t) and $\tilde{p}_t$ is a linear function of $(z, h_{t-1}, \tilde{\mu}_t)$.*

Since h_t is publicly known, this lemma implies that, conditional on h_t, the observation of $\tilde{w}_{it}$ and $\tilde{e}_{it}$ is equivalent to the observation of two private signals about z; and conditional on h_{t-1}, the observation of p_t is equivalent to the observation of a public signal about z. We conclude that it is *as if* the consumers play a dynamic beauty-contest game in which they receive an exogenous sequence of private and public signals about z. The fact that the precisions of these signals are actually endogenous to the equilibrium does not matter for the properties we document below.[45]

This equivalence helps characterize the belief hierarchy. For any $t \leq T-1$ and any $k \in \{1, ..., T-t\}$, we henceforth let B_t^k collect all the relevant k-order beliefs, as of period t :

$$B_t^k \equiv \left\{ x \ : \ \exists (t_1, t_2, ..., t_k), \text{ with } t = t_1 < t_2 < ... < t_k \leq T-1, \text{ such that } x = \bar{E}_{t_1}[\bar{E}_{t_2}[\cdots \bar{E}_{t_k}[\tilde{R}_T] \cdots]] \right\}$$

For convenience, we also let $B_t^0 \equiv \{\tilde{R}_T\}$. We can then show the following.

Lemma 5 *Under Assumption 4, $Var(x) < Var(\bar{E}_t[\tilde{R}_T])$ for all $x \in B_t^k$, $k \in \{2, ..., T-t\}$, $t \leq T-2$.*

Corollary 3 *Pick any $t \leq T-2$, any $k \in \{2, ..., T-t\}$, and any $x \in B_t^k$, and consider the coefficient of the projection of x on $\bar{E}_t[\tilde{R}_T]$. This coefficient is strictly less than one under Assumption 4, whereas it is equal to one under complete information.*

These results verify that the kind of forward-looking higher-order beliefs that matter in our setting are less volatile than the corresponding first-order beliefs in two complementary senses: their total variance is lower; and their "slope" with respect to first-order beliefs, as measured by the aforementioned projection coefficient, is less than one. This is similar to what we had in the previous section, except that now higher-order belief need not be perfectly collinear with first-order beliefs, due to correlated noise introduced by the markup shock.

[44] The intuition for Lemma 4 is as follows. At any t, the information set of consumer i is given by

$$\mathcal{I}_{it} = \mathcal{I}_{it-1} \cup \{x_{it}, \tilde{w}_{it}, \tilde{e}_{it}, \tilde{p}_t\} = \{x_{i0}..., x_{it}, \tilde{w}_{i0}, ...\tilde{w}_{it}, \tilde{e}_{i0}, ...\tilde{e}_{it}\} \cup h_t.$$

Since an individual's consumption is measurable in h_t and her private signals, aggregate consumption—and hence $\tilde{y}_t$ as well—is measurable in h_t and in the cross-sectional averages of the private signals. The cross-sectional average of x_{it}, the *exogenous* private signals, is given by z, whereas the cross-sectional averages of $\tilde{w}_{it}$ and $\tilde{e}_{it}$, the *endogenous* private signals, are pinned down by $\tilde{y}_t$. It follows that the variation in aggregate income is necessarily spanned by (z, h_t). By the NKPC, we can then show that the variation in π_t, and hence also the one in $\tilde{p}_t$, is spanned by $(z, h_{t-1}, \tilde{\mu}_t)$.

[45] By the same token, even if there happen to exist multiple equilibria due to the endogeneity of the information, the results that follow are valid for *any* equilibrium, provided at least that we rule out sunspots.

The results in the rest of this section use *only* the property that higher-order beliefs move less than one-to-one with first-order beliefs. It is well known in the literature that this property holds for *arbitrary* information structures as long as one restrict attention to *static* beauty contests or, equivalently, to *within-period* higher-order beliefs of the form $\bar{E}_t\left[\bar{E}_t\left[\cdots\bar{E}_t\left[\cdot\right]\cdots\right]\right]$.[46] Unfortunately, a similar result is not readily available for *dynamic* beauty contests and for the type of forward-looking beliefs we are interested in. In fact, we can engineer "pathological" examples that violate this property.[47] This explains, not only why the result stated above does not follow from existing results in the literature, but also why we could not have accommodated an entirely arbitrary specification of the information structure. That said, Assumption 4 is far from necessary for the desired property to hold.

Special case with $\kappa = 0$: Attenuating the Income Multiplier

We are now ready to elaborate on how the property that higher-order beliefs are anchored, in the sense of Corollary 3, translates to attenuation of the GE effects of forward guidance. To develop insight, we start with the special case in which prices are infinitely sticky; that is, we set $\theta = 1$ and, equivalently, $\kappa = 0$. By shutting down the response of inflation, this case isolates the GE mechanism (or, equivalently, the strategic complementarity) that operates within the demand block of the economy.

In this case, the "consumption beauty contest" obtained in condition (16) reduces to the following:

$$\tilde{y}_t = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t[\tilde{y}_{t+k}]. \tag{20}$$

Expectations of inflation have disappeared simply because $\kappa = 0$ implies $\pi_t = 0$ for all t. By the same token, two of the relevant GE effects, namely the inflationary spiral and the strategic complementarity in the firms' price-setting decisions, have been shut down. Expectations of $\tilde{R}_T$ and $\tilde{y}_{t+k}$, on the other hand, continue to appear. The first capture the PE effects of forward guidance; and latter capture the remaining GE effect, namely the strategic complementarity in the spending decisions of the consumers.

As noted before, we treat $\bar{E}_t[\tilde{R}_T]$, the expectations of future monetary policy, as exogenous. By contrast, $\bar{E}_t\left[\tilde{y}_{t+k}\right]$, the expectations of future income, are endogenously determined as an integral part of the equilibrium—and this is where higher-order beliefs come into the picture.

When information is complete, higher-order beliefs of $\tilde{R}_T$ collapse to first-order beliefs, implying that expectations of income move one to one with expectations of monetary policy. This is the key to understanding part (i) of Proposition 2. As evident in (20), an increase in the horizon shifts weight from expectations of $\tilde{R}_T$ to expectations of future income. Under complete information, expectations of future income move one

[46] See, e.g., Morris and Shin (2002) and Bergemann and Morris (2013).

[47] Suppose that there are two periods, $t = 0$ and $t = 1$, and two types of consumers, type A and type B. Suppose type A gets a signal $x_A = z + u_A$ at $t = 1$ and no other information either at $t = 0$ or at $t = 1$. Suppose further that type B gets a signal $x_B = u_A + u_B$ at $t = 0$ and no other information either at $t = 0$ or at $t = 1$. Then, clearly $Var(\bar{E}_0[\tilde{R}_T]) = 0$, because no agent has any information about $\tilde{R}_T$ at $t = 0$, and yet $Var(\bar{E}_0[\bar{E}_1[\tilde{R}_T]]) > 0$, because type-A consumers have information at $t = 0$ of the error that type-B consumer will make at $t = 1$.

to one with expectations of $\tilde{R}_T$. It follows that shifting weight from the one to the other has no consequence for the overall effect on aggregate spending.

By contrast, once information is incomplete, the aforementioned tight relation between expectations of income and expectations of monetary policy is broken: Corollary 3 means, in effect, that incomplete information reduces the "slope" of the expectations of future income with respect the expectations of monetary policy. By the same token, the following is true.

Proposition 5 (Attenuating the Income Multiplier) *Let ϕ_t denote the absolute value of the coefficient of the projection of $\tilde{y}_t$ on $\bar{E}_t[\tilde{R}_T]$, for any $t \leq T-1$, and suppose that that Assumption 4 holds and that $\kappa = 0$. Then, $\phi_t < 1$ for all $t < T-1$.*

This result generalizes part (i) of Proposition 3. There, $\tilde{y}_t$ was proportional to $\bar{E}_0[\tilde{R}_T]$, because first- and higher-order beliefs were *perfectly* collinear. Now, we have $\tilde{y}_t = -\phi_t \bar{E}_0[\tilde{R}_T] + residual_t$, where $residual_t$ is a random variable that captures the combined variation in higher-order beliefs that is orthogonal to the variation in first-order beliefs.[48] Notwithstanding this change, the following key lesson remains: whereas complete information imposes that $\phi_t = 1$ when $\kappa = 0$, we have $\phi_t < 1$ under incomplete information. As already explained, this is because lack of common knowledge among the consumers attenuates the income multiplier (which is the only GE mechanism at work when $\kappa = 0$).

General case with $\kappa > 0$: Attenuating also the Inflation Spiral

We now add back the feedback mechanism between aggregate spending and inflation. Quantitatively, this mechanism is the most important component of the forward-guidance puzzle: the standard prediction that ϕ^* increases with T and explodes to infinity as $T \to \infty$ is driven precisely by this mechanism. Attenuating this mechanism is therefore even more essential than attenuating the "income multiplier". We now explain why this can be achieved by removing common knowledge among the consumers, even if we maintain common knowledge among the firms.

Whenever $\kappa > 0$, aggregate demand and inflation are codetermined by the interaction of two beauty contests in Proposition 2. However, insofar as firms have complete information, we can reduce this interaction to a single "meta-game" that is played only among the consumers and that nevertheless encompasses both the income multiplier and the inflationary spiral.

To see why, recall that the standard NKPC applies once the firms have complete information. For any $t \leq T$, the period-t inflation can thus be expressed as follows:

$$\tilde{\pi}_t = \kappa E_t \left[\sum_{k=0}^{T-t} \beta^k \tilde{y}_{t+k} \right] + \kappa \tilde{\mu}_t, \tag{21}$$

[48]Under Assumption 4, this residual can be expressed as a linear combination of the current and the past values of the markup shock. More generally, this residual may capture any kind of noise in information that happens to be correlated across agents, such as the one resulting from communication or "sentiment shocks" (Angeletos and La'O, 2013).

where E_t is the expectation condition on $\mathcal{I}_t$, the entire information in the economy. Furthermore, because $\mathcal{I}_{it}$ is a subset of $\mathcal{I}_t$ for all consumers and all t, we have that $\bar{E}_t E_t = \bar{E}_t$ for all t. Using these facts with (16), we reach the following result, which, as anticipated, permits us to represent the entire equilibrium as the solution to a single beauty contest among the consumers.

Lemma 6 *When firms have complete information and monetary policy satisfies Assumption 1, equilibrium income is given by the solution to the following beauty contest:*

$$\tilde{y}_t = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \sum_{k=1}^{T-t} (1-\beta+k\kappa)\,\beta^{k-1}\bar{E}_t[\tilde{y}_{t+k}]. \tag{22}$$

By comparing condition (22) to (20), we see that the feedback between inflation and aggregate demand increases the effective degree of strategic complementarity between the consumers: the sensitivity of current spending to expectations of spending k periods latter has increased from $(1-\beta)\beta^{k-1}$ to $(1-\beta+k\kappa)\beta^{k-1}$; the term $k\kappa$ captures the cumulative effect that expectations of future spending have on current spending through inflation and real interest rates. Notwithstanding this point, the nature of the beauty contest remains the same. It is thus evident that removing common knowledge from the demand block of the economy *alone* helps attenuate, not only the income multiplier, but also the inflationary spiral. Indeed, from the perspective of condition (22), the two channels are indistinguishable: what mattes is only the overall strategic complementarity, as measured by the coefficients $(1-\beta+k\kappa)\beta^{k-1}$, not its decomposition between the two GE mechanisms.

We can therefore reach the following result, which generalizes part (ii) of Proposition 3.

Proposition 6 (Attenuating also the Inflationary Spiral) *Let ϕ_t denote the absolute value of the coefficient of the projection of $\tilde{y}_t$ on $\bar{E}_t[\tilde{R}_T]$, for any $t \leq T-1$. Under Assumption 4, we have $\phi_t < \phi_t^*$ for all $t < T-1$, where $\phi_t^* \equiv \phi^*(T-t)$ is the complete-information counterpart.*

The logic can be summarized as follows. Regardless of the information structure, the inflationary pressures triggered by the anticipation of lax future monetary policy raise the value of ϕ_t that obtains when $\kappa > 0$ relative to the one that obtains when $\kappa = 0$. These pressures, however, are attenuated when consumers lack common knowledge. Importantly, the attenuation is self-fulfilling in the following sense: when consumers "fail" to adjust their inflation expectations as predicted by the standard model, they also "fail" to spend as much as predicted by the standard model, which in turn justifies a lower inflation response in the supply block of the economy.

A decomposition

Throughout this section, we allowed for higher-order uncertainty among the consumers but not among the firms. This highlights that removing common knowledge from the one block of the model attenuates, not only the strategic complementarity that operates *within* that block, but also the one that operates *across* the

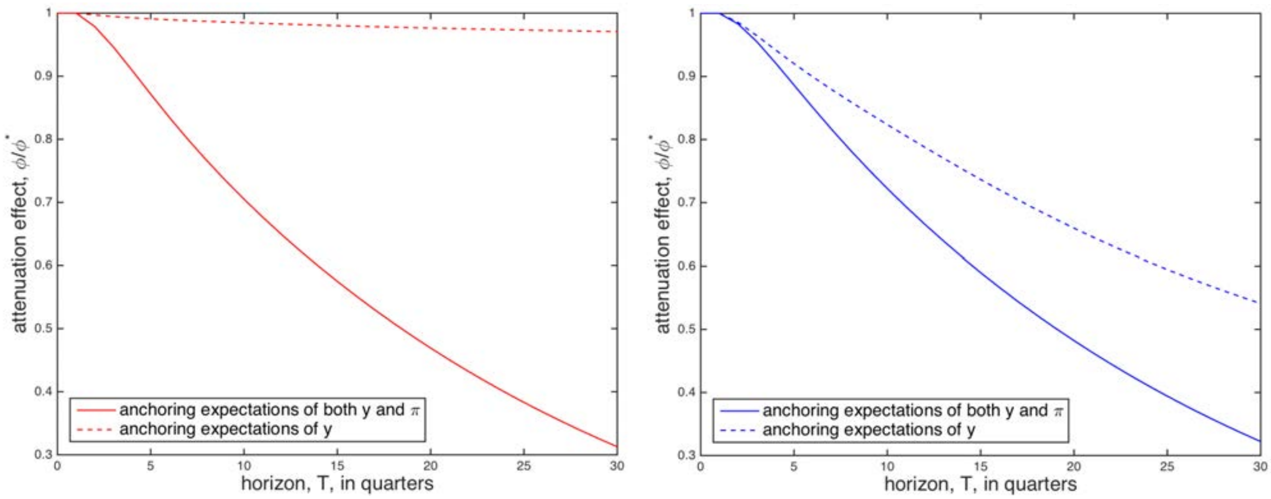

Figure 2: Decomposing the effect of different expectations.

two blocks. Not surprisingly, allowing the friction to be present in both blocks at once has a reinforcing effect, as illustrated in the previous section. Indeed, not only was ϕ increasing in *both* λ and λ^f under Assumption 3, but also the cross partial was positive: adding the friction in the one block of the model increases the marginal attenuation effect of the friction that is present in the other block.

Figure 1 also suggests that, for a plausible parameterization, the importance of the two blocks is comparable: letting the informational friction inflict only the consumer side of the economy appears to be of similar quantitative importance as letting it inflict only the firm side. Regardless of which side we focus on, however, attenuation obtains for two conceptually separate forces: the anchoring of inflation expectations; and the anchoring of output expectations. (Keep in mind that expectations of output translate to expectations of income in the eyes of the consumers, and to expectations of marginal costs in the eyes of the firms.) How much does each of these two forces contribute to the attenuation effects seen in Figure 1?

We address this question in Figure 2, under the same information structure and same parameterization as the one used in Figure 1. The left panel focuses on the consumer side; the right panel shifts attention to the firm side. In both cases, we decompose the difference between ϕ and ϕ^* into two parts: one accounted by the change in inflation expectations; and another accounted by the change in output expectations.

Consider first the left panel. The solid red line represents the value of ϕ, normalized by ϕ^*, that obtains *in equilibrium* when $\lambda = .75$ and $\lambda_f = 1$, that is, when only the consumers are subject to the friction. The dashed red line represents the value of ϕ, normalized once again by ϕ^*, that obtains in an ad-hoc, *off-equilibrium*, variant in which the following properties happen to be true: the consumers' subjective expectations of *income* remain the same as in the aforementioned *incomplete-information* equilibrium; their subjective expectations of *inflation* instead coincide with those in the *complete-information* equilibrium; and consumers choose their spending optimally given the aforementioned subjective expectations. By construction, this variant therefore isolates the part of the attenuation effect that reflects the anchoring of the consumers' expectations of income from the part that reflects the anchoring of their expectations of inflation. As evident in the figure, the anchoring of inflation expectations accounts for the lion's share of the attenuation effect in the consumer side.

Consider now the right panel, which regards the firm side. The solid blue line represents the value of ϕ, normalized by ϕ^*, that obtains *in equilibrium* when $\lambda = 1$ and $\lambda_f = .75$, that is, when only the firms are subject to the friction. The dashed blue line represents the value that obtains in the following ad-hoc variant: the firms' subjective expectations of *output*, or of real marginal costs, remain the same as those in the *incomplete-information* equilibrium; their subjective expectations of *inflation* coincide with those in the *complete-information* equilibrium; firms set prices optimally given the aforementioned subjective expectations; and finally the consumers form rational expectations and spend optimally, knowing that the firms behave in the aforementioned manner. By construction, this variant therefore isolates the part of the attenuation effect that is due to anchoring the firms' expectations of output from the part that is due to anchoring their expectations of inflation. As evident in the figure, the latter part is now a bit less significant than the former.

These findings are, of course, sensitive to the chosen parameterization, especially with regard to the parameters that determine the size of the inflation movements. For instance, by setting $\theta = 1$ (equivalently, $\kappa = 0$), we can force inflation to be irresponsive to forward guidance regardless of the information structure; this renders the firms' expectations irrelevant, pegs the consumers' expectations of inflation to zero, and mechanically guarantees that the entire attenuation effect is accounted by anchoring the consumers' expectations of income. That said, the aforementioned findings appear to be robust to a reasonable range of plausible parameterizations. Indeed, it is interesting to note that the consumers' expectations of inflation turn out to play a dominant role in our benchmark parameterization despite the fact this parameterization imposes a rather high degree of price stickiness (namely, $\theta = .85$ and $\kappa \approx 0.09$). We conclude that, when looking at the consumer side, most of the attenuation is accounted by anchoring of inflation expectations. When looking at the firm side, on the other hand, anchoring expectations of inflation and anchoring expectations of output and real marginal costs are of roughly equal quantitative importance.

8 Horizons and HOB

In the previous section, we elaborated on the robustness of the insight that lack of common knowledge attenuates the GE mechanisms of forward guidance, but did not investigate how this attenuation depends on the horizon at which forward guidance operates. We now show that longer horizons increase the relative importance of higher-order beliefs, contributing towards more attenuation. The crux of the argument is the following. As noted before, longer horizons increase the GE effects of forward guidance under common knowledge because they increase the number of loops from future spending and inflation to current spending and inflation. But when one increases the number of loops, one is effectively walking down the hierarchy of beliefs, which in turn explains why the attenuation implied by lack of common knowledge itself increases with the length of the horizon.

To illustrate the generality of this insight, we allow for an entirely arbitrary information structure.[49] We only bound the level of higher-order uncertainty away from zero, in a manner that we make precise momentarily. Although not strictly needed, we also assume that firms are completely informed; this simplifies the exposition by letting us focus on the meta-game among the consumers described in Lemma 6.[50]

Fix any $t \leq T-1$ and let $\tau = T-t$. Following our earlier results, the equilibrium value of output (and also of inflation) in period t can can be expressed as a linear function of beliefs of order up to τ. The weights on these beliefs depend on τ, the length of the time that remains till the "moment of action", but not on t per se. This is due to the recursive structure of the economy. For any $k \in \{1, ..., \tau\}$, let $\chi_{k,\tau} > 0$ be the absolute value of the total weight of $\tilde{y}_t$ on beliefs of order equal to k; and let $s_{k,\tau} \equiv \sum_{r=1}^{k} \chi_{r,\tau}$ be the total weight on beliefs of order up to, and including, k.

Proposition 7 *For any $k \geq 1$ and any $\tau \geq k$, let $s_{k,\tau}$ be defined as above. The ratio $s_{k,\tau}/s_{\tau,\tau}$, which measures the relative contribution of the first k orders of beliefs to aggregate spending, strictly decreases with the horizon τ and converges to 0 as $\tau \to \infty$.*

In simple words, longer horizons amplify the importance of higher-order uncertainty, because longer horizons map, in effect, to stronger complementarity. This is an elementary but important observation, whose implications may extend beyond the context of our paper. For instance, it could help reduce the indeterminacy problems of the NK model, for these problems are tied to expectations at very long horizons. In a nutshell, we suggest that one should trust less the predictions of the standard model—or, equivalently, worry less about its anomalies—when these regard the cumulation of GE effects over long horizons than when these regard mechanisms that operate in the short run.

Combining the above result with the fact that higher-order beliefs are more anchored than lower-beliefs suggests that the attenuation we have documented increases without bound as the horizon goes to ∞. We verify this intuition in Proposition 8 below, under the following assumption.

Assumption 5 (Non-Vanishing Higher-Order Uncertainty) *There exists an $\epsilon > 0$ such that:*

(i) For all T and all $t \in \{0, ..., T-2\}$, there exists at least a mass ϵ of consumers such that

$$Var\left(E_t[x]|\mathcal{I}_{it}\right) \geq \epsilon Var\left(E_t\left[x\right]\right)$$

for all $x \in B_{\tau}^{k}$, $\tau \in \{t+1, ..., T-1\}$, and $k \in \{0, 1, ..., T-\tau\}$.

(ii) $Var\left(\bar{E}_0[\tilde{R}_T]\right) \geq \epsilon$.

[49] Relative to Assumption 4, this means, not only that we allow an arbitrary correlation structure in the noise (both across agents and across time), but also that we no more need the exogenous private signals to be centered around the initial policy announcement. For instance, it could be that the policy maker makes new announcements over time and that the agents have arbitrary private information about each announcement. It could also be that the agents receive signals that are independent of the policy announcement but are otherwise informative about $\tilde{R}_T$; such signals could help the private views that different agents may have about the future policy regardless of the policy announcements.

[50] We have proved Propositions 7 and 8 also for the case where firms have incomplete information (the proof is available upon request). In this case, the relative importance of higher-order beliefs is actually increased: by letting firms have complete information, we have effectively collapsed certain beliefs of higher orders to beliefs of lower orders.

To interpret this assumption, note that complete information imposes that $E_t[x]$ is known to every agent, and therefore that $Var(E_t[x]]|\mathcal{I}_{it}) = 0$, regardless of how volatile $E_t[x]$ itself is. By contrast, letting $Var(E_t[x]|\mathcal{I}_{it}) > 0$ whenever $Var(E_t[x]) > 0$ is essentially *tautological* to assuming that agents have incomplete information or, equivalently, that they face higher-order uncertainty. Relative to this tautology, part (i) introduces an arbitrarily small bound on the level of higher-order uncertainty. This bound is needed in order to guarantee that the higher-order uncertainty does not vanish as we let T go to infinity. Part (ii), on the other hand, means simply that there is non-vanishing variation in first-order beliefs in the first place. The next result then follows from Proposition 7.

Proposition 8 (Horizon) *Allow $\kappa > 0$, let ϕ denote the absolute value of the coefficient of the projection of $\tilde{y}_0$ on $\bar{E}_0[\tilde{R}_T]$. Under Assumption 5, the ratio ϕ/ϕ^* converges to zero as $T \to \infty$.*

This result generalizes the lesson first derived in Proposition 3 that the attenuation is stronger when the horizon is longer.[51] As already explained, this is because longer horizons increase the number of loops from future outcomes to current outcomes, thus also increasing the effective degree of strategic complementarity.

To further appreciate what this means, note that, regardless of the horizon, aggregate spending in any given period depends primarily on expectations of income and inflation in the short- to medium run; expectations in the far future do not matter much because of discounting. But it is precisely the former kind of expectations that depend more heavily on higher-order beliefs. Indeed, for any triplet (t, τ_1, τ_2) such that $t < \tau_1 < \tau_2$, we have that $\bar{E}_t[\tilde{y}_{\tau_1}]$ involves beliefs of order up to $k_1 = T - \tau_1 + 1$, whereas $\bar{E}_t[\tilde{y}_{\tau_2}]$ involves beliefs of order only up to $k_2 = T - \tau_2 + 1 < k_1$. Furthermore, although beliefs of order $k \leq k_2$ show up in both $\bar{E}_t[\tilde{y}_{\tau_1}]$ and $\bar{E}_t[\tilde{y}_{\tau_2}]$, they enter with smaller relative weight in the former than in the latter (by implication of Proposition 7). To recap, raising T does not affect the relative importance of expectations of income and inflation at different horizons, and yet it raises the importance of higher-order beliefs in the determination of these expectations.

This insight seems relevant, not only in the context of the forward-guidance puzzle, but also for any other prediction that rests on general-equilibrium effects operating over long horizons. For instance, consider the neo-Fisherian prediction that pegging to a low interest rate for a long time can, perhaps paradoxically, contribute to disinflation (Cochrane, 2016a). This prediction appears to depend on iterating beliefs over long horizons and may thus be sensitive to relaxations of common knowledge. We leave this conjecture open for future research.

9 Additional Results

We conclude the paper with two additions results. The first one touches upon the "paradox of flexibility" and offers another example of the broader applicability of our insights. The second result recasts our

[51] One caveat is that, due to the level of generality we have sought to accommodate in this section, there is no guarantee that ϕ/ϕ^* decreases monotonically with T.

mechanism—the anchoring of equilibrium expectations—in terms of a discounted Euler condition and a discounted NKPC. This in turn highlight an intriguing connection to, but also a crucial difference from, the existing literature. Both of these results utilize the stark information structure introduced in Section 6, but the insights are more general.

On the Paradox of Flexibility

We now turn attention to another implication of our analysis. In the standard model, the effect of forward guidance increases with the degree of price flexibility: ϕ^* increases with κ.[52] As noted in the Introduction, this property is directly related to the "paradox of flexibility" (Eggerstsson and Krugman, 2013). The next result proves, in effect, that the mechanism identified in this paper helps diminish this paradox as well.

Proposition 9 (Price Flexibility) *Let ϕ be the scalar characterized in Proposition 3 and set $\lambda^f = 1$. We have $\frac{\partial \phi}{\partial \kappa} > 0$ and $\frac{\partial}{\partial \lambda}\left(\frac{\partial \phi}{\partial \kappa}\right) > 0$. That is, the power of forward guidance increases with the degree of price flexibility, but at a rate that is slower the greater the departure from common knowledge.*

Note that we have proved the above result only under the restriction $\lambda^f = 1$, which means that only the consumers lack common knowledge. Whenever $\lambda^f < 1$, there is a conflicting effect, which is that higher price flexibility reduces the strategic complementarity that operates within the supply block, thereby also reducing the role of λ^f itself. Numerical explorations, however, suggest that the overall effect of higher price flexibility is qualitatively the same whether $\lambda^f = 1$ or $\lambda^f = \lambda$.

We illustrate this in Figure 3. We let $\lambda^f = \lambda$, use the same parameter values as those used in Figure 1, and plot the relation between the ratio ϕ/ϕ^* and the horizon T under two values for θ. The solid red line is the same as in Figure 1. The dashed blue line corresponds to a lower value for θ, that is, to more price flexibility. As evident in the figure, more price flexibility maps, not only to a lower ratio ϕ/ϕ^* (i.e., stronger attenuation) for any given T, but also to a more rapid decay in that ratio as we raise T.

This finding is an example of how lack of common knowledge reduces the paradox of flexibility more generally. In the standard model, a higher degree of price flexibility raises the GE effects of all kinds of demand shocks—whether these come in the form of forward guidance, discount rates, or borrowing constraints—because it intensifies the feedback loop between aggregate spending and inflation. By intensifying this kind of macroeconomic complementarity, however, a higher degree of price flexibility also raises the relative importance of higher-order beliefs, which in turn contributes to stronger attenuation effects of the type we have documented in this paper. In a nutshell, the very same mechanism that creates the paradox of flexibility within the NK framework also helps contain the paradox once we relax the common-knowledge assumptions of that framework.

[52]Whenever we vary κ, we vary θ while keeping the Frisch elasticity constant, which means that variation in κ maps one-to-one to variation in the degree of price flexibility.

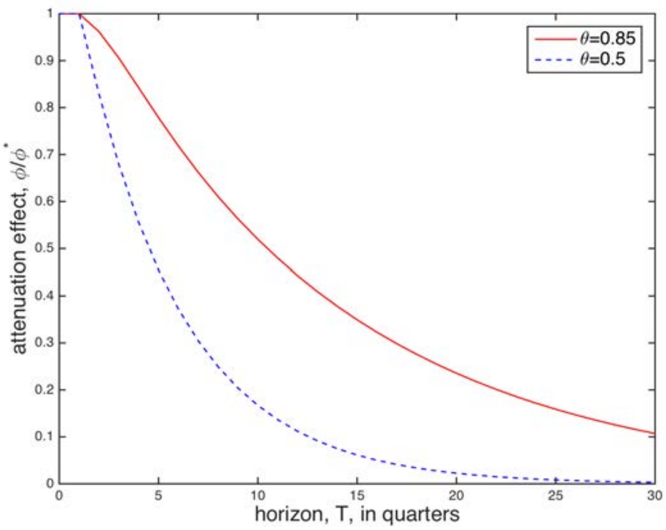

Figure 3: Varying the degree of price flexibility.

Discounted Euler Equation and Discounted NKPC

Let $E_t[x]$ denote the rational expectation of variable x conditional on $\mathcal{I}_t$, the union of the information sets in the economy. This is the same expectation operator as the one that shows up in Lemma 1, that is, in the Euler condition and the NKPC of the standard model. With this in mind, we now provide the following representation of the equilibrium that obtains under the information structure we introduced in Section 6.

Proposition 10 (Discounted Euler Equation and NKPC) *Under Assumption 3, the equilibrium dynamics for aggregate spending and inflation solve the following system, for all $t < T - 1$:*

$$\tilde{y}_t = \Lambda E_t[\tilde{y}_{t+1}] + \lambda E_t[\tilde{\pi}_{t+1}] \tag{23}$$

$$\tilde{\pi}_t = \beta M E_t[\tilde{\pi}_{t+1}] + \kappa m_t \tilde{y}_t + \kappa \tilde{\mu}_t \tag{24}$$

where $\Lambda \equiv \beta + (1-\beta)\lambda \in (\beta, 1)$, $M \equiv \theta + (1-\theta)\lambda^f \in (\theta, 1)$, and $m_t \equiv 1 - \beta\theta\left(1-\lambda^f\right)\frac{\phi(T-t-1)}{\phi(T-t)} \in (0,1)$.

This is an "as if" result that maps our heterogeneous-agent, incomplete-information model to a fictitious representative-agent, complete-information, model in which the equilibrium dynamics happen to solve conditions (23) and (24). Condition (23) reduces to the standard Euler equation when $\lambda = 1$, and similarly condition (24) reduces to the standard NKPC when $\lambda^f = 1$. Once we remove common knowledge among the consumers ($\lambda < 1$), it is *as if* the representative consumer discounts her expectations of next period's aggregate income and inflation by a factor equal to, respectively, Λ and λ. Accordingly, we can interpret (23) as a *discounted* Euler equation. Similarly, condition (24) represents a *discounted* NKPC: inflation responds less to both the rational expectation of inflation tomorrow and to the innovation in the current real marginal cost by a factor equal to, respectively, $M \in (\theta, 1)$ and $m_t \in (0, 1)$.[53]

[53] Note that Λ and M are bounded from below by, respectively, β and θ. These bounds have to do with the fact that lack of common knowledge attenuates only the GE effects.

These forms of discounting *must* reflect the property that equilibrium expectations (equivalently, higher-order beliefs) are anchored, for this is the *only* friction relative to the standard model. To shed further light on why the friction manifests as discounting in the fundamental equations of the NK model, let us first concentrate on condition (23).

Recall that, because we have abstracted from borrowing constraints and have rested on the familiar log-linearizations, the *standard* Euler equation holds at the *individual* level:

$$\tilde{c}_{i,t} = E_{i,t}\left[\tilde{c}_{i,t+1}\right] + \tilde{R}_t - E_{i,t}\left[\tilde{\pi}_{t+1}\right] \quad \forall i, t.$$

Note that $\tilde{R}_t$ shows up without an expectations operator in front of it, because $\tilde{R}_t$ is known in the beginning of t. By contrast, $\tilde{c}_{i,t+1}$ and $\tilde{\pi}_{t+1}$ are still uncertain. Integrating the above over i gives the following condition at the *aggregate* level:

$$\tilde{c}_t = \int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di + \tilde{R}_t - \bar{E}_t\left[\tilde{\pi}_{t+1}\right]. \tag{25}$$

Note here a delicate point: what shows up in the right-hand side is the average of the individuals' forecasts of their *own* consumption, not of *aggregate* consumption. Assumption 1 pegs $\tilde{R}_t = 0$ for all $t < T$. To reach condition (23), we therefore need to relate $\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right]$ with $E_t\left[\tilde{c}_{t+1}\right]$ and $\bar{E}_t[\tilde{\pi}_{t+1}]$ with $E_t[\tilde{\pi}_{t+1}]$.[54] It is at this point that lack of common knowledge manifests as a form of discounting.

Let us elaborate. Due to the uninsurable idiosyncratic risk, c_{it} need not equal c_t, regardless of the information structure. Nevertheless, as long as information is complete, $E_{it}\left[x\right] = \bar{E}_t\left[x\right] = E_t\left[x\right]$ for *any* random variable x. Furthermore, $\int_0^1 E_t\left[\tilde{c}_{i,t+1}\right] di = E_t\left[\tilde{c}_{t+1}\right]$, because $\tilde{c}_{i,t+1}$ is the sum of $\tilde{c}_{t+1}$ and of a purely idiosyncratic component. It follows that complete information guarantees both $\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di = E_t\left[\tilde{c}_{t+1}\right]$ and $\bar{E}_t\left[\tilde{\pi}_{t+1}\right] = E_t\left[\tilde{\pi}_{t+1}\right]$, which in turn imply that the *standard* Euler condition continues to hold, not only at the individual level, but also at the aggregate level.

Once information is incomplete, however, the aforementioned equalities break down. The property that higher-order beliefs vary less that lower-order beliefs implies both that $\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di$ varies less than $E_t\left[\tilde{c}_{t+1}\right]$ and that $\bar{E}_t\left[\tilde{\pi}_{t+1}\right]$ varies less than $E_t\left[\tilde{\pi}_{t+1}\right]$. This indicates that some kind of discounting is present at the aggregate level *regardless* of the specifics of the information structure. Under Assumption 3, this discounting take a particular stark form: we have

$$\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di = \Lambda E_t[\tilde{c}_{t+1}] \qquad \text{and} \qquad \bar{E}_t\left[\tilde{\pi}_{t+1}\right] = \lambda E_t[\tilde{\pi}_t],$$

which in turn gives condition (23). The same logic explains the discounted NKPC obtained in that proposition: by dampening the response of beliefs about future marginal costs and future inflation, higher-order uncertainty gives rise to a form of discounting in the NKPC.[55]

[54] Keep in mind that $\tilde{c}_t = \tilde{y}_t$ in equilibrium.

[55] We have considered an extension (available upon request) that allows for the information friction to diminish with time. This results to a generalized discounted Euler condition, in which the discounting factors reduce with time. This illustrates the robustness of the insight, but also clarifies that the stark form of discounting obtained in (23) rests on the stationarity of the friction.

Related forms of discounting can be found in McKay, Nakamura, and Steinsson (2016a,b) and Gabaix (2016). Yet, the underlying mechanism is different, and so are some of the testable implications. As already noted, these works obtain a discounted Euler condition at the aggregate level *only* by distorting to a commensurate degree the consumption-saving behavior at the individual level. McKay, Nakamura, and Steinsson (2016a,b) do so by letting agents be liquidity-constrained, Gabaix (2016) by letting agents apply a form of cognitive discount on feature state variables. In either case, the distortion that obtains in macro time series is tied to the one that can be detected in micro data. By contrast, our approach obtains the same "as if" discounting at the aggregate level, even in the absence of cognitive discounting or any other distortion at the individual level. Importantly, the mechanism that underlies this form of "as if" discounting seems to be consistent with the evidence on expectations provided in Coibion and Gorodnichenko (2012, 2015).

The same points distinguish our contribution from Werning (2015). That paper provides a thorough analysis of how liquidity constraints impact the aggregate-level Euler condition under common knowledge. It points out that, while the consumption of liquidity-constrained agents is less sensitive to interest rates, it is also more sensitive to income, which itself reacts in general equilibrium to interest rates. In an important benchmark, these two effects offset each other in such a manner that the aggregate-level Euler condition remains the same as the one in the representative-agent model. This underscores that the result of McKay, Nakamura, and Steinsson (2016a,b) is special. In fact, Werning argues that the empirically relevant case is the opposite than the one advocated by McKay et al. While this debate is interesting on its own right, it is orthogonal to our contribution: we shift the focus from liquidity constraints or any other aspects of the micro-foundations to the equilibrium adjustment in expectations and identify a mechanism that applies *regardless* of the underlying micro-foundations.

Notwithstanding these observations, the result obtained in Proposition 10 complements all the aforementioned works in that it provides an additional motivation for augmenting DSGE models with a discounted Euler equation and a discounted NKPC: the discounting can be a proxy for a variety of realistic frictions, whether these relate to liquidity constraints, cognitive limitations, or the type of anchored equilibrium expectations we have accommodated in this paper. Qualitatively, all these forms of frictions appear to contribute in the same direction; their combined quantitative importance remains an open question.

10 Conclusion

Modern macroeconomic models assign a crucial role to forward-looking expectations, such as consumer expectations of future income and future real interest rates or firm expectations of future inflation and future real marginal costs. This property seems both desirable and realistic. However, by assuming common knowledge of everything along with the rational-expectations solution concept, these models hardwire a certain kind of perfection in the ability of economic agents to coordinate their expectations and to synchronize their responses to any exogenous impulse. In so doing, they also maximize the capacity of policy makers to manage these expectations and to steer the economy.

In this paper, we argued that removing common knowledge of policy helps accommodate a realistic friction in the ability of policy makers to manage the aforementioned kind of expectations and to steer the economy. We illustrated these ideas in the context of the forward-guidance puzzle, working with an incomplete-information extension of the New-Keynesian model. More specifically, we assumed that the economy is in a liquidity trap, so that the zero lower bound (ZLB) on interest rates is currently binding, and that the policy maker tries to stimulate the economy by managing expectations of the monetary policy to be conducted after the economy has exited the trap. We then showed how lack of common knowledge anchors expectations of future income and inflation, attenuates the general-equilibrium effects of such policy commitments, and lessens the forward-guidance puzzle.

The insights and the techniques we have developed in this paper may find additional applications. We offered an example by touching on the paradox of flexibility. Another application could be the determinacy issues and the "neo-Fisherian" effects emphasized by Cochrane (2016a,b). Yet another application could be fiscal policy during a liquidity trap. The standard New-Keynesian model predicts that, in the presence of a binding ZLB constraint, fiscal stimuli are more effective when they last longer or when they are back-loaded.[56] The magnitude of the predicted effects hinge on the same general-equilibrium mechanisms and the same kind of expectations as those that determine the power of forward guidance; they therefore appear to be equally sensitive to relaxing the common-knowledge assumption.[57]

The broader policy lessons of our paper are subject to the following qualification. In the model under consideration, the relevant GE effects translate to strategic *complementarity*; equivalently, they *amplify* the PE effects. In other models, the relevant GE effects may translate to strategic *substitutability* and may therefore *mitigate* the PE effects. In this case, lack of common knowledge continues to anchor the expectations of macroeconomic outcomes and to attenuate the relevant GE effects. Yet, because these effects are working in the opposite direction to start with, the policy implications are also reversed: attenuating the GE effects now helps increase, not reduce, the effects of policy.[58] Notwithstanding this qualification, we believe that the first scenario is more relevant in the context of fiscal and monetary policy over the business cycle, not only because of the specific mechanisms we have studied in this paper, but also because of other forms of "macroeconomic complementarities", such as those associated with collateral constraints and the feedback loop between asset prices and economic activity. Applying our insights to such settings and working out the implications for, say, quantitative easing and macro-prudential policies is another intriguing direction for future research.

We conclude by iterating that the formalization we have adopted in this paper does not have to be taken too literally. We view incomplete information and higher-order uncertainty as modeling devices that permit the researcher to anchor the response of expectations to changes in policy, and thereby to attenuate

[56] See Christiano, Eichenbaum, and Rebelo (2011), Woodford (2011), and Werning (2012).

[57] This indicates more broadly the value of investigating how lack of common knowledge influences the macroeconomic effects of fiscal policy, not only in the New-Keynesian framework, but also in the RBC framework. We are investigating this question in ongoing work (Angeletos and Lian, 2016c).

[58] See Angeletos and Lian (2016) for a more abstract setting that allows the GE effects to be the source of either strategic complementarity or strategic substitutability and that shows that lack of common knowledge attenuates the GE effects in either case.

the relevant GE effects, while remaining inside the "comfort zone" of the rational-expectations hypothesis. But these modeling devices can also be seen a structured way for relaxing that hypothesis. Whether this represents a competitor or a complement to behavioral approaches is for the reader to decide. One way or another, the key feature of our approach is that it operationalizes the notion that policy makers may have less control on market expectations of economic outcomes than what is presumed in standard macroeconomic models, a property that seems both theoretically appealing and empirically relevant.[59]

[59]A related point is made in Angeletos and La'O (2013) and Angeletos, Collard and Dellas (2016): higher-order uncertainty is used as a modeling device to accommodate autonomous variation in expectations and forces akin to "animal spirits".

Appendix: Proofs

Proof of Lemma 1. It follows directly from Proposition 2, imposing complete information and equilibrium.

Proof of Lemma 2. We prove the following stronger result: there exists functions $\phi^*, \varpi^* : \mathbb{N} \to \mathbb{R}_+$ such that, whenever Assumptions 1 and 2 hold, equilibrium spending and inflation at any $t \leq T$ are given by

$$\tilde{y}_t = -\phi^* (T-t) \cdot E_t[\tilde{R}_T], \tag{26}$$

$$\tilde{\pi}_t = \kappa \tilde{y}_t + \kappa \tilde{\mu}_t - \varpi^* (T-t) \cdot E_t[\tilde{R}_T]. \tag{27}$$

We prove this result by induction, starting with $t = T$ and proceeding backwards.

When $t = T$, we have $\tilde{y}_T = -\tilde{R}_T$ and $\tilde{\pi}_T = \kappa \tilde{\mu}_T + \kappa \tilde{y}_T$ from Lemma 1. (Recall that $\tilde{y}_\tau = \tilde{\pi}_\tau = 0 \, \forall \tau > T$ by Assumption 1.) This verifies (26) and (27) for $t = T$, with $\phi^* (0) = 1$ and $\varpi^* (0) = 0$.

Now suppose that the result holds for arbitrary $t \in \{1, ..., T\}$ and let's prove that it also holds for $t-1$. By the assumption that (26) and (27) hold at t along with the Law of Iterated Expectations and the assumption that future markup shocks are unpredictable, we have

$$E_{t-1}[\tilde{y}_t] = -\phi^* (T-t) \cdot E_{t-1}[\tilde{R}_T],$$

$$E_{t-1}[\tilde{\pi}_t] = -\left(\kappa \phi^* (T-t) + \varpi^* (T-t)\right) \cdot E_{t-1}[\tilde{R}_T],$$

Using the above together with Lemma 1 verifies that (26) and (27) hold also for $t-1$, with

$$\phi^* (T-t+1) = (1+\kappa)\, \phi^* (T-t) + \varpi^* (T-t),$$

$$\varpi^* (T-t+1) = \beta \kappa \phi^* (T-t) + \beta \varpi^* (T-t).$$

This completes the proof and gives a recursive formula that can be used to compute $\phi^*(T)$.

Proof of Proposition 1. (i) If prices are infinitely sticky ($\theta = 1$ or, equivalently, $\kappa = 0$), from Lemma 1 we have $\pi_t = 0 \, \forall t$ and

$$\tilde{y}_0 = -\tilde{R}_0 - \sum_{t=1}^{T} E_0[\tilde{R}_T] + E_0\left[\tilde{c}_{T+1}\right].$$

By Assumption 1 then,

$$\tilde{y}_0 = -E_0[\tilde{R}_T],$$

which proves that $\phi^*(T) = 1$.

(ii) From the proof of Lemma 2, we have that

$$\phi^* (0) = 1 \;\text{ and }\; \varpi^* (0) = 0,$$

and, for every $\tau \geq 0$,

$$\phi^* (\tau + 1) = (1 + \kappa) \phi^* (\tau) + \varpi^* (\tau) \tag{28}$$

$$\varpi^* (\tau + 1) = \beta\kappa\phi^* (\tau) + \beta\varpi^* (\tau) \tag{29}$$

From condition (29), we know, when $\kappa > 0$, $\varpi^* (\tau) > 0 \; \forall \tau \geq 1$. Then, from equation (28), we have $\phi^*(\tau) > 1 \; \forall \tau \geq 1$, and $\phi^*(\tau)$ is strictly increasing in τ.

Now we prove that $\phi^* (\tau)$ explodes to infinity as $\tau \to \infty$. To this goal, we first prove the following equality:

$$\frac{\phi^* (\tau + 1)}{\phi^* (\tau)} + \beta \frac{\phi^* (\tau - 1)}{\phi^* (\tau)} = 1 + \beta + \kappa \quad \forall \tau \geq 1. \tag{30}$$

From condition (28), we have, for all $\tau \geq 1$,

$$\beta\phi^* (\tau) = \beta (1 + \kappa) \phi^* (\tau - 1) + \beta\varpi^* (\tau - 1).$$

Together with conditions (28) and (29), we arrive at condition (30).[60]

Second, we prove that, when $\kappa > 0$,

$$\frac{\phi^* (\tau)}{\phi^* (\tau - 1)} \text{ is strictly increasing in } \tau \geq 1. \tag{31}$$

From conditions (28) and (29), we have

$$\phi^* (1) = 1 + \kappa \;\text{ and }\; \phi^* (2) = (1 + \kappa)^2 + \kappa\beta.$$

As a result, when $\kappa > 0$,

$$\frac{\phi^* (2)}{\phi^* (1)} = 1 + \kappa + \frac{\kappa\beta}{1 + \kappa} > \frac{\phi^* (1)}{\phi^* (0)}.$$

Now we proceed by induction. Suppose that, for $\tau \geq 1$, we have

$$\frac{\phi^* (\tau + 1)}{\phi^* (\tau)} > \frac{\phi^* (\tau)}{\phi^* (\tau - 1)}. \tag{32}$$

From condition (30), we know

$$\frac{\phi^* (\tau + 2)}{\phi^* (\tau + 1)} + \beta \frac{\phi^* (\tau)}{\phi^* (\tau + 1)} = 1 + \beta + \kappa \quad \forall \tau \geq 1.$$

Together with condition (32) we have

$$\frac{\phi^* (\tau + 2)}{\phi^* (\tau + 1)} > \frac{\phi^* (\tau + 1)}{\phi^* (\tau)}.$$

[60] Note that equality (30) also holds for $\kappa = 0$, because conditions (28) and (29) also hold for $\kappa = 0$. This will be used later in the paper.

This proves (31) when $\kappa > 0$.

Finally, from condition (30), we know $\frac{\phi^*(\tau)}{\phi^*(\tau-1)}$ is bounded above. Together with (31), $\frac{\phi^*(\tau)}{\phi^*(\tau-1)}$ must converge to $\Gamma^* > 0$, as $\tau \to \infty$. From condition (30) again, we know Γ^* satisfy[61]

$$\Gamma^* + \beta \frac{1}{\Gamma^*} = 1 + \beta + \kappa. \tag{33}$$

As a result, $\Gamma^* > 1$ whenever $\kappa > 0$. Therefore, $\phi^*(\tau)$ explodes to infinity as $\tau \to \infty$ whenever $\kappa > 0$. The proof of Proposition 1 follows simply from letting $\tau = T$.

Proof of Proposition 2. Substituting condition (15) into condition (12) gives condition (16), the consumption beauty contest. Substituting condition (15) into condition (14) gives condition (17), the inflation beauty contest.

Proof of Lemma 3. The lemma follows directly from this claim: under Assumption 3, there exists functions $\phi, \varpi : (0,1] \times (0,1] \times \mathbb{N} \to \mathbb{R}_+$ such that, for any $t \leq T-1$,

$$\tilde{y}_t = -\phi\left(\lambda, \lambda^f, T-t\right) \cdot \bar{E}[\tilde{R}_T], \tag{34}$$

$$\tilde{\pi}_t = \kappa \tilde{y}_t + \kappa \tilde{\mu}_t - \varpi\left(\lambda, \lambda^f, T-t\right) \bar{E}_t^f[\tilde{R}_T] \tag{35}$$

We now prove this claim by induction.

First, consider $t = T-1$. From conditions (16) and (17) along with the fact that monetary policy replicates flexible-price allocations from $T+1$ and on, we have

$$\tilde{y}_T = -\tilde{R}_T \qquad \text{and} \qquad \tilde{\pi}_T = \kappa \tilde{y}_T + \kappa \tilde{\mu}_T,$$

and therefore

$$\tilde{y}_{T-1} = \bar{E}_{T-1}[\tilde{\pi}_T] - \beta \bar{E}_{T-1}\left[\tilde{R}_T\right] + (1-\beta)\bar{E}_{T-1}[\tilde{y}_T] = -(1+\kappa)\bar{E}_{T-1}[\tilde{R}_T],$$

$$\tilde{\pi}_{T-1} = \kappa \tilde{y}_{T-1} + \kappa \tilde{\mu}_{T-1} + \kappa\beta\theta \bar{E}_{T-1}^f[\tilde{y}_T] + (1-\theta)\beta \bar{E}_{T-1}^f[\tilde{\pi}_T] = \kappa \tilde{y}_{T-1} + \kappa \tilde{\mu}_{T-1} - \kappa\beta \bar{E}_{T-1}^f[\tilde{R}_T].$$

It follows that the claim holds for $t = T-1$ with

$$\phi(\lambda, \lambda^f, 1) = 1 + \kappa \quad \forall 0 < \lambda, \lambda^f \leq 1, \tag{36}$$

$$\varpi(\lambda, \lambda^f, 1) = \kappa\beta \quad \forall 0 < \lambda, \lambda^f \leq 1. \tag{37}$$

Now, pick an arbitrary $t \leq T-2$, assume that the claim holds for all $\tau \in \{t+1, ..., T-1\}$, and let us prove

[61] Note that equality (33) also holds for $\kappa = 0$, because, when $\kappa = 0$, $\frac{\phi^*(\tau)}{\phi^*(\tau-1)} = \tau \; \forall \; \tau \geq 1$. This will be used later in the paper.

that it also holds for t. Under Assumption 1, the consumption beauty contest in condition (16) becomes

$$\tilde{y}_t = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \sum_{k=1}^{T-t} \beta^{k-1}\bar{E}_t\left[\tilde{\pi}_{t+k}\right] + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right]. \tag{38}$$

Since the claim holds for $\tau \in \{t+1, ...T-1\}$, and since $\tilde{y}_T = -\tilde{R}_T$ and $\tilde{\pi}_T = -\kappa\tilde{R}_T + \kappa\tilde{\mu}_T$, we have

$$\begin{aligned}
\tilde{y}_t &= -\beta^{T-t-1}(1+\kappa)\,\bar{E}_t[\tilde{R}_T] - (1-\beta+\kappa)\sum_{k=1}^{T-t-1}\beta^{k-1}\phi\left(\lambda,\lambda^f,T-t-k\right)\bar{E}_t[\bar{E}_{t+k}[\tilde{R}_T]] \\
&\quad - \sum_{k=1}^{T-t-1}\beta^{k-1}\varpi\left(\lambda,\lambda^f,T-t-k\right)\bar{E}_t\left[\bar{E}^f_{t+k}[\tilde{R}_T]\right] \\
&= -\left\{\beta^{T-t-1}(1+\kappa) + \sum_{k=1}^{T-t-1}\beta^{k-1}\left[(1-\beta+\kappa)\,\lambda\phi\left(\lambda,\lambda^f,T-t-k\right) + \lambda^f\varpi\left(\lambda,\lambda^f,T-t-k\right)\right]\right\}\bar{E}_t[\tilde{R}_T].
\end{aligned}$$

where we have used the fact that, under Assumption 3, for $1 \le k \le T-t-1$,

$$\bar{E}_t[\bar{E}_{t+k}[\tilde{R}_T]] = \bar{E}_t[\lambda\delta z] = \lambda\delta\lambda z = \lambda\bar{E}_t[\tilde{R}_T] \quad \text{and} \quad \bar{E}_t[\bar{E}^f_{t+k}[\tilde{R}_T]] = \bar{E}_t[\lambda^f\delta z] = \lambda^f\delta\lambda z = \lambda^f\bar{E}_t[\tilde{R}_T].$$

This proves the part of the claim that regards output, condition (34), with

$$\phi\left(\lambda,\lambda^f,T-t\right) = \beta^{T-t-1}(1+\kappa) + \sum_{k=1}^{T-t-1}\beta^{k-1}\left[(1-\beta+\kappa)\,\lambda\phi\left(\lambda,\lambda^f,T-t-k\right) + \lambda^f\varpi\left(\lambda,\lambda^f,T-t-k\right)\right]. \tag{39}$$

Similarly, the inflation beauty contest in condition (17) gives

$$\begin{aligned}
\tilde{\pi}_t &= \kappa\tilde{y}_t + \kappa\tilde{\mu}_t - \left(\kappa + \kappa\tfrac{1-\theta}{\theta}\right)\sum_{k=1}^{T-t-1}(\beta\theta)^k\,\phi\left(\lambda,\lambda^f,T-t-k\right)\bar{E}^f_t[\bar{E}_{t+k}[\tilde{R}_T]] \\
&\quad - \tfrac{1-\theta}{\theta}\sum_{k=1}^{T-t-1}(\beta\theta)^k\,\varpi\left(\lambda,\lambda^f,T-t-k\right)\bar{E}^f_t[\bar{E}^f_{t+k}[\tilde{R}_T]] - \left(\kappa + \kappa\tfrac{1-\theta}{\theta}\right)(\beta\theta)^{T-t}\,\bar{E}^f_t[\tilde{R}_T] \\
&= \kappa\tilde{y}_t + \kappa\tilde{\mu}_t \\
&\quad - \left\{\tfrac{\kappa}{\theta}(\beta\theta)^{T-t} + \sum_{k=1}^{T-t-1}(\beta\theta)^k\left[\tfrac{\kappa\lambda}{\theta}\phi\left(\lambda,\lambda^f,T-t-k\right) + \tfrac{(1-\theta)\lambda^f}{\theta}\varpi\left(\lambda,\lambda^f,T-t-k\right)\right]\right\}\bar{E}^f_t[\tilde{R}_T].
\end{aligned}$$

where we have used the fact that, similarly to the consumers' case, for $1 \le k \le T-t-1$, $\bar{E}^f_t[\bar{E}_{t+k}[\tilde{R}_T]] = \lambda\bar{E}^f_t[\tilde{R}_T]$ and $\bar{E}^f_t[\bar{E}^f_{t+k}[\tilde{R}_T]] = \lambda^f\bar{E}_t[\tilde{R}_T]$. This proves the part of the claim that regards inflation, condition (35) with

$$\varpi\left(\lambda,\lambda^f,T-t\right) = \frac{\kappa}{\theta}(\beta\theta)^{T-t} + \sum_{k=1}^{T-t-1}(\beta\theta)^k\left(\tfrac{\kappa\lambda}{\theta}\phi\left(\lambda,\lambda^f,T-t-k\right) + \tfrac{(1-\theta)\lambda^f}{\theta}\varpi\left(\lambda,\lambda^f,T-t-k\right)\right). \tag{40}$$

We finally provide a recursive formula for computing $\phi(\lambda, \lambda^f, T-t)$ and $\varpi\left(\lambda, \lambda^f, T-t\right)$, which will be useful later. From condition (39), we have, for $t \leq T-2$,

$$\begin{aligned}\phi\left(\lambda, \lambda^f, T-t\right) &= \beta\phi\left(\lambda, \lambda^f, T-t-1\right) + (1-\beta+\kappa)\lambda\phi\left(\lambda, \lambda^f, T-t-1\right) + \lambda^f\varpi\left(\lambda, \lambda^f, T-t-1\right) \\ &= (\beta + (1-\beta+\kappa)\lambda)\phi\left(\lambda, \lambda^f, T-t-1\right) + \lambda^f\varpi\left(\lambda, \lambda^f, T-t-1\right). \qquad (41)\end{aligned}$$

Similarly, from condition (40), we have, for $t \leq T-2$,

$$\begin{aligned}\varpi\left(\lambda, \lambda^f, T-t\right) &= \beta\theta\varpi\left(\lambda, \lambda^f, T-t-1\right) + \beta\theta\left(\tfrac{\kappa\lambda}{\theta}\phi\left(\lambda, \lambda^f, T-t-1\right) + \tfrac{(1-\theta)\lambda^f}{\theta}\varpi\left(\lambda, \lambda^f, T-t-1\right)\right) \\ &= \kappa\beta\lambda\phi\left(\lambda, \lambda^f, T-t-1\right) + \beta\left[\theta + (1-\theta)\lambda^f\right]\varpi\left(\lambda, \lambda^f, T-t-1\right). \qquad (42)\end{aligned}$$

Proof of Proposition 3. To simplify notation, we use ϕ_τ and ϖ_τ as shortcuts for, respectively, $\phi\left(\lambda, \lambda^f, \tau\right)$ and $\varpi\left(\lambda, \lambda^f, \tau\right)$, the latter defined as in the proof of Lemma 3. Similarly, we use ϕ^*_τ as a shortcut for $\phi^*(\tau)$, the latter be defined as in Lemma 2.

(i) When $\kappa = 0$, from conditions (36), (37), (41) and (42), we have, for all $\tau \geq 1$,

$$\begin{aligned}&\varpi_\tau = 0, \\ &\phi_\tau = (\beta + (1-\beta)\lambda)^{\tau-1}.\end{aligned}$$

The proof of part (i) of Proposition 3 then follows simply from letting $\tau = T$.

(ii) When $\kappa > 0$, from conditions (36), (37), (41) and (42), we know that $\phi_\tau, \varpi_\tau > 0$ for all $\tau \geq 1$.

We will first prove, for $\tau \geq 2$, $\phi_T = \phi\left(\lambda, \lambda^f, \tau\right)$ is strictly increasing in both λ and λ^f. We will proceed by induction on τ. For $\tau = 2$, from (36), (37), (41) and (42), we have ϕ_2 and ϖ_2 is strictly increasing in both λ and λ^f. Suppose for $\tau \geq 2$, ϕ_τ, ϖ_τ is strictly increasing in both λ and λ^f. From conditions (41) and (42), we know $\phi_{\tau+1}$ and $\varpi_{\tau+1}$ are strictly increasing in both λ and λ^f, where we use the fact that $\phi_\tau, \varpi_\tau > 0$. This proves that, for $\tau \geq 2$, $\phi_\tau = \phi\left(\lambda, \lambda^f, \tau\right)$ is strictly increasing in both λ and λ^f. Because of the strict monotonicity, we also have, for $\tau \geq 2$, $\frac{\phi_\tau}{\phi^*_\tau} = \frac{\phi(\lambda, \lambda^f, \tau)}{\phi(1,1,\tau)} < 1$, whenever $\lambda < 1$ and/or $\lambda^f < 1$.

We now prove that, whenever $\lambda < 1$ and/or $\lambda^f < 1$, the ratio $\frac{\phi_\tau}{\phi^*_\tau} = \frac{\phi(\lambda, \lambda^f, \tau)}{\phi^*(\tau)}$ is strictly decreasing in $\tau \geq 1$. We start by noticing, from the proof of Proposition 1, we have, for $\tau \geq 3$,

$$\frac{\phi^*_\tau}{\phi^*_{\tau-1}} + \beta\frac{\phi^*_{\tau-2}}{\phi^*_{\tau-1}} = 1 + \beta + \kappa \quad . \qquad (43)$$

Now we prove that ϕ_τ satisfies an inequality with a similar form as (43):

$$\frac{\phi_\tau}{\phi_{\tau-1}} + \beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \leq 1 + \beta + \kappa\lambda < 1 + \beta + \kappa \qquad \forall \tau \geq 3. \qquad (44)$$

From condition (41), we have, for $\tau \geq 3$,

$$\phi_\tau = (\beta + (1-\beta)\lambda)\,\phi_{\tau-1} + \kappa\lambda\phi_{\tau-1} + \lambda^f \varpi_{\tau-1},$$

$$\frac{\beta}{\beta+(1-\beta)\lambda}\phi_{\tau-1} = \beta\phi_{\tau-2} + \frac{\beta\kappa\lambda}{\beta+(1-\beta)\lambda}\phi_{\tau-2} + \frac{\lambda^f\beta}{\beta+(1-\beta)\lambda}\varpi_{\tau-2}.$$

From the previous two conditions, we have, for $\tau \geq 3$,

$$\phi_\tau + \beta\phi_{\tau-2} = (\beta + (1-\beta)\lambda)\,\phi_{\tau-1} + \kappa\lambda\phi_{\tau-1} + \lambda^f\varpi_{\tau-1} + \frac{\beta}{\beta+(1-\beta)\lambda}\phi_{\tau-1} - \frac{\beta\kappa\lambda}{\beta+(1-\beta)\lambda}\phi_{\tau-2} - \frac{\lambda^f\beta}{\beta+(1-\beta)\lambda}\varpi_{\tau-2}. \tag{45}$$

Note that, for $\tau \geq 3$ and $\lambda, \lambda^f \in (0,1]$, from condition (42), we have

$$\left[(\beta + (1-\beta)\lambda) + \kappa\lambda + \frac{\beta}{\beta+(1-\beta)\lambda}\right]\phi_{\tau-1} \leq (1+\beta+\kappa)\,\phi_{\tau-1},$$

and

$$\begin{aligned}
&\lambda^f\varpi_{\tau-1} - \frac{\beta\kappa\lambda}{\beta+(1-\beta)\lambda}\phi_{\tau-2} - \frac{\lambda^f\beta}{\beta+(1-\beta)\lambda}\varpi_{\tau-2}\\
=&\lambda^f\left(\kappa\beta\lambda\phi_{\tau-2} + \beta\left[\theta + (1-\theta)\lambda^f\right]\varpi_{\tau-2}\right) - \frac{\beta\kappa\lambda}{\beta+(1-\beta)\lambda}\phi_{\tau-2} - \frac{\lambda^f\beta}{\beta+(1-\beta)\lambda}\varpi_{\tau-2}\\
=&\kappa\beta\lambda\left(\lambda^f - \frac{1}{\beta+(1-\beta)\lambda}\right)\phi_{\tau-2} + \beta\lambda^f\left[\theta + (1-\theta)\lambda^f - \frac{1}{\beta+(1-\beta)\lambda}\right]\pi_{\tau+2}\\
\leq&0.
\end{aligned}$$

Together with condition (45), we reach at condition (44).

Now we can prove that $\frac{\phi_\tau}{\phi_\tau^*}$ is strictly decreasing in τ, whenever $\lambda < 1$ and/or $\lambda^f < 1$. We already prove $\frac{\phi_2}{\phi_2^*} < 1 = \frac{\phi_1}{\phi_1^*}$. We proceed by induction on τ. If $\frac{\phi_\tau}{\phi_\tau^*} < \frac{\phi_{\tau-1}}{\phi_{\tau-1}^*}$ for $\tau \geq 2$, we have $\frac{\phi_{\tau-1}}{\phi_\tau} > \frac{\phi_{\tau-1}^*}{\phi_\tau^*}$. From (43) and (44), we have $\frac{\phi_{\tau+1}}{\phi_\tau} < \frac{\phi_{\tau+1}^*}{\phi_\tau^*}$ and thus $\frac{\phi_{\tau+1}}{\phi_{\tau+1}^*} < \frac{\phi_\tau}{\phi_\tau^*}$. This finishes the proof that $\frac{\phi_\tau}{\phi_\tau^*}$ is strictly decreasing in $\tau \geq 1$, whenever $\lambda < 1$ and/or $\lambda^f < 1$.

Now we prove that, whenever $\lambda < 1$ and/or $\lambda^f < 1$, $\frac{\phi_\tau}{\phi_\tau^*}$ converges to 0 as $\tau \to \infty$. Because $\frac{\phi_\tau}{\phi_\tau^*} > 0$ is strictly decreasing in $\tau \geq 1$, there exists $\Gamma \in [0,1)$ such that $\frac{\phi_\tau}{\phi_\tau^*} \to \Gamma$ as $\tau \to \infty$. We next prove by contradiction that $\Gamma = 0$.

Suppose first that $\lambda < 1$. If $\Gamma > 0$, we have $\frac{\phi_\tau}{\phi_\tau^*}\frac{\phi_{\tau-1}^*}{\phi_{\tau-1}} \to 1$ as $\tau \to \infty$. Because $\frac{\phi_\tau^*}{\phi_{\tau-1}^*} \to \tau^*$, we have $\frac{\phi_\tau}{\phi_{\tau-1}} \to \Gamma^*$ and $\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to \frac{1}{\Gamma^*}$ as $\tau \to \infty$. From condition (33), we have $\frac{\phi_\tau}{\phi_{\tau-1}} + \beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to 1 + \beta + \kappa$ as $\tau \to \infty$. However, this is inconsistent with (44) when $\lambda < 1$ and $\kappa > 0$. As a result, $\Gamma = 0$ when $\lambda < 1$.

Suppose next that $\lambda = 1$ but $\lambda_f < 1$. We prove a stronger version of (44):

$$\frac{\phi_\tau}{\phi_{\tau-1}} + \left(1 + \kappa\left(1-\lambda^f\right)\right)\beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \leq 1 + \kappa + \beta \quad \forall\tau \geq 3. \tag{46}$$

From conditions (41) and (42), we have, for $\tau \geq 3$,

$$\phi_\tau = (1+\kappa)\,\phi_{\tau-1} + \lambda^f \varpi_{\tau-1},$$

$$\beta\phi_{\tau-1} = \beta\phi_{\tau-2} + \beta\kappa\phi_{\tau-2} + \beta\lambda^f \varpi_{\tau-2},$$

$$\varpi_{\tau-1} = \kappa\beta\phi_{\tau-2} + \beta\left[\theta + (1-\theta)\,\lambda^f\right]\varpi_{\tau-2}.$$

As a result, for $\tau \geq 3$,

$$\begin{aligned}\phi_\tau + \beta\phi_{\tau-2} &= (1+\kappa+\beta)\,\phi_{\tau-1} + \lambda^f \varpi_{\tau-1} - \beta\kappa\phi_{\tau-2} - \beta\lambda^f\varpi_{\tau-2} \\ &\leq (1+\kappa+\beta)\,\phi_{\tau-1} + \left(\lambda^f - 1\right)\kappa\beta\phi_{\tau-2}.\end{aligned}$$

This proves (46).

Now, if $\Gamma > 0$, similarly, we have $\frac{\phi_\tau}{\phi^*_\tau}\frac{\phi^*_{\tau-1}}{\phi_{\tau-1}} \to 1$ as $\tau \to \infty$. Because $\frac{\phi^*_\tau}{\phi^*_{\tau-1}} \to \Gamma^*$, we have $\frac{\phi_\tau}{\phi_{\tau-1}} \to \Gamma^*$ and $\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to \frac{1}{\Gamma^*}$ as $T \to \infty$. From condition (33), we have $\frac{\phi_\tau}{\phi_{\tau-1}} + \left(1+\kappa\left(1-\lambda^f\right)\right)\beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to 1+\beta+\kappa+\kappa\left(1-\lambda^f\right)\beta\frac{1}{\Gamma^*}$ as $\tau \to \infty$. However, this is inconsistent with equation (44) when $\lambda^f < 1$. As a result, $\Gamma = 0$ when $\lambda = 1$, but $\lambda^f < 1$.

Finally, we prove that, when λ is sufficiently low, $\phi(\lambda, \lambda^f, \tau)$ converges to zero as $\tau \to \infty$. The eigenvalues of the dynamical system (ϕ_τ, ϖ_τ) based on conditions (41) and (42) are

$$m_1 = \frac{\left(\beta + (1-\beta)\,\lambda + \kappa\lambda + \beta\left[(1-\theta)\,\lambda^f + \theta\right]\right) - \sqrt{\left(\beta + (1-\beta+\kappa)\,\lambda - \beta\left[(1-\theta)\,\lambda^f + \theta\right]\right)^2 + 4\beta\lambda^f\lambda\kappa}}{2} > 0$$

$$m_2 = \frac{\left(\beta + (1-\beta)\,\lambda + \kappa\lambda + \beta\left[(1-\theta)\,\lambda^f + \theta\right]\right) + \sqrt{\left(\beta + (1-\beta+\kappa)\,\lambda - \beta\left[(1-\theta)\,\lambda^f + \theta\right]\right)^2 + 4\beta\lambda^f\lambda\kappa}}{2} > m_1$$

Note that $\lim_{\lambda\to 0} m_2 = \beta < 1$. As a result, when λ is sufficiently low, both eigenvalues are below 1, which means that $\phi(\lambda, \lambda^f, \tau)$ converges to zero as $\tau \to \infty$. The proof of part (ii) of Proposition 3 then follows simply from letting $\tau = T$.

Proof of Proposition 4. For $t = T > 0$, from conditions (16), (17) and Assumption 1, we have $\tilde{y}_t = -\tilde{R}_T$ and $\tilde{\pi}_t = -\kappa\tilde{R}_T + \kappa\tilde{\mu}_T$. As a result, $\delta^y_T\left(\lambda, \lambda^f\right) = 1$ and $\delta^\pi_T\left(\lambda, \lambda^f\right) = \kappa$ for all $\lambda, \lambda^f \in (0, 1]$.

For $0 < t \leq T-1$, from condition (34), Assumption 3 and the proof of Proposition 3, we have

$$\begin{aligned}\bar{E}_0[\tilde{y}_t] &= -\bar{E}_0\left[\lambda\phi\left(\lambda, \lambda^f, T-t\right)E\left[\tilde{R}_T|z\right]\right] \\ &= -\lambda\phi\left(\lambda, \lambda^f, T-t\right)\bar{E}_0[\tilde{R}_T]\end{aligned}$$

As a result,

$$\delta^y_t\left(\lambda, \lambda_f\right) = \lambda\phi\left(\lambda, \lambda^f, T-t\right). \tag{47}$$

Similarly, for $0 < t \leq T-1$, from condition (35) and Assumption 3, we have

$$\begin{aligned}\bar{E}_0[\tilde{\pi}_t] &= -\bar{E}_0\left[\kappa\lambda\phi\left(\lambda,\lambda^f,T-t\right)\cdot E\left[\tilde{R}_T|z\right]+\lambda^f\varpi\left(\lambda,\lambda^f,T-t\right)\cdot E\left[\tilde{R}_T|z\right]\right]\\ &= -\left(\kappa\lambda\phi\left(\lambda,\lambda^f,T-t\right)+\lambda^f\varpi\left(\lambda,\lambda^f,T-t\right)\right)\bar{E}_0[\tilde{R}_T].\end{aligned}$$

As a result,

$$\delta_t^{\pi}\left(\lambda,\lambda_f\right)=\kappa\lambda\phi\left(\lambda,\lambda^f,T-t\right)+\lambda^f\varpi\left(\lambda,\lambda^f,T-t\right). \tag{48}$$

Finally, the monotonicity of δ_t^y and δ_t^{π} with respect to λ and λ^f follow directly from Proposition 3.

Proof of Lemma 4. Under Assumption 4, the new information each consumer i receives at period $t \in \{0,...,T-1\}$ consists of the exogenous private signals $s_{i,t}$, price level $\tilde{p}_t$, his own wage and dividend $\tilde{w}_{i,t}$ and dividend $\tilde{e}_{i,t}$. Because $\int s_{i,t}di = z$, $\int \tilde{w}_{i,t}di = \tilde{w}_t = (\epsilon+1)\,\tilde{y}_t$, and $\int \tilde{e}_{i,t}di = \tilde{e}_t = \left[1-\frac{\Omega(\epsilon+1)}{1-\Omega}\right]\tilde{y}_t$, we have $\bar{E}_t\left[\tilde{R}_T\right]$ and $\bar{E}_t\left[\tilde{y}_{t+k}\right]$ $\forall 1\leq k\leq T-t$ are linear functions of z, $\{\tilde{y}_s\}_{s=0}^t$ and $\{\tilde{p}_s\}_{s=0}^t$.[62] From condition (22), we have $\tilde{y}_t$ is a linear function of z, $\{\tilde{y}_s\}_{s=0}^{t-1}$ and $\{\tilde{p}_s\}_{s=0}^t$. Iterating, we have

$$\tilde{y}_t \text{ is a linear function of } z \text{ and } \left\{\tilde{p}_s\right\}_{s=0}^t. \tag{49}$$

Now, we prove by backward induction that, for $t \in \{0,...,T-1\}$,

$$\tilde{p}_t \text{ is a linear function of } z,\ \tilde{\mu}_t, \text{ and } \left\{\tilde{p}_s\right\}_{s=0}^{t-1}. \tag{50}$$

To prove (50), one only need to prove that

$$\tilde{\pi}_t \text{ is a linear function of } z,\ \tilde{\mu}_t, \text{ and } \left\{\tilde{p}_s\right\}_{s=0}^{t-1}. \tag{51}$$

When firms have complete information, $\tilde{\pi}_t$ can be expressed

$$\tilde{\pi}_t = \kappa\tilde{y}_t + \beta E_t\left[\tilde{\pi}_{t+1}\right] + \kappa\tilde{\mu}_t \ \forall t \leq T-1, \tag{52}$$

where E_t denote the rational expectation conditional on all available information in the economy up to period t, $\mathcal{I}_t$[63]. Note that $\tilde{\pi}_T = -\kappa\tilde{R}_T$, $z = E_{T-1}\left[\tilde{R}_T\right]$, and $\tilde{y}_{T-1}$ is a linear function of z and $\{\tilde{p}_s\}_{s=0}^{T-1}$. From condition (6), $\tilde{\pi}_{T-1}$ is a linear function of z, $\tilde{\mu}_{T-1}$ and $\{\tilde{p}_s\}_{s=0}^{T-2}$. Now suppose, for some $0 \leq t \leq T-2$, $\tilde{\pi}_{t+1}$ is a linear function of z, $\tilde{\mu}_{t+1}$ and $\{\tilde{p}_s\}_{s=0}^t$. Together with the fact that $\tilde{y}_t$ is a linear combination of z and $\{\tilde{p}_s\}_{s=0}^t$, we have $\tilde{\pi}_t$ is a linear function of z, $\tilde{\mu}_t$ and $\{\tilde{p}_s\}_{s=0}^{t-1}$. This finishes the proof of (50).

[62] Even if there happen to exist multiple equilibria due to the endogeneity of the information, the above statements are valid for *any* equilibrium, provided at least that we rule out sunspots.

[63] In this case, it contains z and $\{\tilde{\mu}_s\}_{s=0}^t$.

Proof of Lemma 5. To simplify notation, in this proof we let

$$\bar{E}_{t_1,t_2,\cdots,t_k}\left[\cdot\right] \equiv \bar{E}_{t_1}\left[\bar{E}_{t_2}\left[\cdots \bar{E}_{t_k}\left[\cdot\right]\right]\right].$$

Following Lemma 4 and the discussion after it, we know that the information structure assumed in Assumption 4 is equivalent to the following information structure.

Assumption 6 *in each period $t \in \{0, 1, \cdots, T-1\}$, a consumer observes a public signal of the form*

$$\omega_t = z + \eta_t, \tag{53}$$

where $\eta_t \sim N\left(0, \kappa_{\eta,t}^{-1}\right)$ is uncorrelated with z, and a private signal of the form

$$s_{i,t} = z + \epsilon_{i,t}, \tag{54}$$

where $\epsilon_{i,t} \sim N\left(0, \kappa_{\epsilon,t}^{-1}\right)$ is i.i.d. across consumers and uncorrelated with both z and η_t. Moreover, both noises are uncorrelated over time and unpredictable on the basis of past information.

We require $0 < \kappa_{\epsilon,0} < +\infty$. That is, at period 0, each agent receives a nontrivial but not totally informative private signal about z. This is required for agents to have incomplete information about z, and thus also about $\tilde{R}_T$, in period 0. For $1 \leq t \leq T-1$, we let $0 \leq \kappa_{\epsilon,t} < +\infty$. That is, the private information in subsequent periods can be totally uninformative (as in Section 6), but cannot reveal z completely (otherwise agents would attain complete information about $\tilde{R}_T$). Similarly, for $0 \leq t \leq T-1$, we let $0 \leq \kappa_{\eta,t} < +\infty$.

Note that under Assumption 4, for any $t \leq T-1$, z is a summary statistics of $\tilde{R}_T$ based on period t's information set $\mathcal{I}_t$. We have $\bar{E}_t[\tilde{R}_T] = \delta \bar{E}_t\left[z\right]$, where $\delta \equiv \frac{\sigma_z^{-2}}{\sigma_R^{-2}+\sigma_z^{-2}} \in (0,1)$. As a result, for any $t \leq T-1, k \in \{1, \cdots, T-t\}$, and $x = \bar{E}_{t,t_2,\cdots,t_k}[\tilde{R}_T] \in B_t^k$, we have

$$x = \bar{E}_{t,t_2,\cdots,t_k}[\tilde{R}_T] = \delta \bar{E}_{t,t_2,\cdots,t_k}\left[z\right].$$

As a result, to prove Lemma 5, we only need to prove that, under Assumption 6, whenever $t \leq T-2$, $2 \leq k \leq T-t$, and $t = t_1 < t_2 < ... < t_k \leq T-1$, we have

$$Var\left(\bar{E}_{t,t_2,\cdots,t_k}\left[z\right]\right) < Var\left(\bar{E}_t\left[z\right]\right). \tag{55}$$

We first introduce a few new notations and properties that will be useful for the proof. Whenever $t < t_2 \leq T-1$, under the information structure assumed in Assumption 6, we can use $s_{i,t,t_2} = z + \epsilon_{i,t,t_2}$ to denote the sufficient statistic of all the new private signals consumer i receives between period t and t_2 about z. In other words,

$$s_{i,t,t_2} = \frac{\kappa_{\epsilon,t}}{\sum_{s=t}^{t_2}\kappa_{\epsilon,s}} s_{i,t} + \frac{\kappa_{\epsilon,t+1}}{\sum_{s=t}^{t_2}\kappa_{\epsilon,s}} s_{i,t+1} + \cdots \frac{\kappa_{\epsilon,t_2}}{\sum_{s=t}^{t_2}\kappa_{\epsilon,s}} s_{i,t_2}.$$

As a result,

$$\epsilon_{i,t,t_2} = \frac{\kappa_{\epsilon,t}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \epsilon_{i,t} + \frac{\kappa_{\epsilon,t+1}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \epsilon_{i,t+1} + \cdots \frac{\kappa_{\epsilon,t_2}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \epsilon_{i,t_2} \sim N\left(0, \kappa_{\epsilon,t,t_2}^{-1}\right),$$

where $\kappa_{\epsilon,t,t_2} = \sum_{s=t}^{t_2} \kappa_{\epsilon,s}$ denotes the precision of s_{i,t,t_2}.[64] Similarly, we can use $\omega_{i,t,t_2} = z + \eta_{t,t_2}$ to denote the sufficient statistic of all the new public signals between period t and t_1 about z. In other words,

$$\omega_{t,t_2} = \frac{\kappa_{\epsilon,t}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \omega_t + \frac{\kappa_{\epsilon,t+1}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \omega_{t+1} + \cdots \frac{\kappa_{\epsilon,t_2}}{\sum_{s=t}^{t_2} \kappa_{\epsilon,s}} \omega_{t_2}.$$

As a result,

$$\eta_{t,t_2} = \frac{\kappa_{\eta,t}}{\sum_{s=t}^{t_2} \kappa_{\eta,s}} \eta_t + \frac{\kappa_{\eta,t+1}}{\sum_{s=t}^{t_2} \kappa_{\eta,s}} \eta_{t+1} + \cdots + \frac{\kappa_{\eta,t_2}}{\sum_{s=t}^{t_2} \kappa_{\eta,s}} \eta_{t_2} \sim N\left(0, \kappa_{\eta,t,t_2}^{-1}\right),$$

where $\kappa_{\eta,t,t_2} = \sum_{s=t}^{t_2} \kappa_{\eta,s}$ denotes the precision of ω_{t,t_2}.[65] Because $\epsilon_{i,t}$ and η_t are uncorrelated across time, whenever $s < t < t_2 \leq T-1$, we have

$$E_{i,s}\left[s_{i,t,t_2}\right] = E_{i,s}\left[z\right], \tag{56}$$

$$E_{i,s}\left[\omega_{t,t_2}\right] = E_{i,s}\left[z\right]. \tag{57}$$

Now, we provide an explicit formula about each consumer i's belief about z at period $t \leq T-1$, $E_{i,t}[z]$. It depends on his prior about z, $z \sim N\left(0, \kappa_z^{-1}\right)$,[66] private signals he receives between period 0 and t about z, and public signals he receives between period 0 and t about z. As a result,

$$E_{i,t}\left[z\right] = \alpha_t s_{i,0,t} + \beta_t \omega_{0,t}, \tag{58}$$

where $\alpha_t = \frac{\kappa_{\epsilon,0,t}}{\kappa_z + \kappa_{\epsilon,0,t} + \kappa_{\eta,0,t}} \in (0,1)$ is the relative precision of private signals, $\beta_t = \frac{\kappa_{\eta,0,t}}{\kappa_z + \kappa_{\epsilon,0,t} + \kappa_{\eta,0,t}} \in [0,1)$ is the relative precision of public signal.[67] Also, the precision of each consumer i's posterior about z at period t is $\kappa_t = \kappa_z + \kappa_{\epsilon,0,t} + \kappa_{\eta,0,t}$.[68] Aggregating over i, we have

$$\bar{E}_t\left[z\right] = \alpha_t z + \beta_t \omega_{0,t}. \tag{59}$$

From condition (59) and the fact that each consumer i has uncertainty about z at period $t \leq T-1$, the consumer also faces uncertainty about $\bar{E}_t[z]$:

$$Var\left(\bar{E}_t\left[z\right] | I_{i,t}\right) > 0 \quad \forall i \text{ and } t \leq T-1,$$

[64] If $\kappa_{\epsilon,t,t_2} = 0$, we can simply define $s_{i,t,t_2} = z + \epsilon_{i,t,t_2}$ where $\epsilon_{i,t,t_2} \sim N(0, +\infty)$. That is, consumer i effectively receives a totally uninformed private signal between period t and t_2.

[65] If $\kappa_{\eta,t,t_2} = 0$, we can simply define $\omega_{t,t_2} = z + \eta_{t,t_2}$ where $\eta_{t,t_2} \sim N(0, +\infty)$. That is, consumers effectively receive a totally uninformed public signal between period t and t_2.

[66] $0 < \kappa_z^{-1} = \sigma_R^2 + \sigma_z^2 < +\infty$.

[67] We also have $\alpha_t + \beta_t < 1$.

[68] We use $\kappa_{-1} = \kappa_z$ to denote the precision of each consumer i's prior about z.

where $I_{i,t}$ is consumer i's information set at period t. Using the law of total variance, we have

$$Var\left(E_{i,t}\left[\bar{E}_t\left[z\right]\right]\right) < Var\left(\bar{E}_t\left[z\right]\right).$$

Similarly, note that for any random variable X, and any information set I, according to the law of total variance, we have

$$Var\left(E\left[X|I\right]\right) \leq Var\left(X\right).$$

Using the fact that $\bar{E}_t\left[\cdot\right] = E\left[E_{i,t}\left[\cdot\right]|\Omega_t\right]$, where Ω_t is the cross-sectional distribution of consumers' information sets, $\mathcal{I}_{it}$, at time t,

$$Var\left(\bar{E}_t\left[\bar{E}_t\left[z\right]\right]\right) \leq Var\left(E_{i,t}\left[\bar{E}_t\left[z\right]\right]\right) < Var\left(\bar{E}_t\left[z\right]\right). \tag{60}$$

Similarly, for all $h \geq 2$, we have

$$Var\left(\bar{E_t^h}\left[z\right]\right) \leq Var\left(\bar{E_t^2}\left[z\right]\right) < Var\left(\bar{E}_t\left[z\right]\right), \tag{61}$$

where $\bar{E}_t^h\left[z\right] = \int E_{i,t}\left[\bar{E}_t^{h-1}\left[z\right]\right] di$ denote consumers' average h-th order belief of z at period t.

Now, whenever $t < t_2 \leq T-1$, consider the relationship between consumer belief about z at period t and his belief at period t_2. Specifically, consumer i's information about z in period t_2 can be decomposed into three parts: his information about z at period t, new private signals he receives between period $t+1$ and t_2 about z, and new public signals he receives between period $t+1$ and t_2 about z. As a result,

$$E_{i,t_2}\left[z\right] = \left(1-\alpha_{t,t_2}-\beta_{t,t_2}\right)E_{i,t}\left[z\right] + \alpha_{t,t_2}s_{i,t+1,t_2} + \beta_{t,t_2}\omega_{t+1,t_2},$$

where $\alpha_{t,t_2} = \frac{\kappa_{\epsilon,t+1,t_2}}{\kappa_t+\kappa_{\epsilon,t+1,t_2}+\kappa_{\eta,t+1,t_2}} \in [0,1)$ is the relative precision of signal $s_{i,t+1,t_2}$ and $\beta_{t,t_2} = \frac{\kappa_{\eta,t+1,t_2}}{\kappa_t+\kappa_{\epsilon,t+1,t_2}+\kappa_{\eta,t+1,t_2}} \in [0,1)$ is the relative precision of signal η_{t+1,t_2}. Aggregating over i, we have

$$\bar{E}_{t_2}\left[z\right] = \left(1-\alpha_{t,t_2}-\beta_{t,t_2}\right)\bar{E}_t\left[z\right] + \alpha_{t,t_2}z + \beta_{t,t_2}\omega_{t+1,t_2}. \tag{62}$$

Finally, when $s \leq t < t_2 \leq T-1$, we consider $E_{i,t}\left[\omega_{s,t_2}\right]$ and $\bar{E}_t\left[\omega_{s,t_2}\right]$. Note that the public signals between period s and t_2 about z has two components: the public signals between period s and t about z and the public signals between period $t+1$ and t_2 about z. As a result, there exists $\gamma_{s,t,t_2} = \frac{\kappa_{\eta,t+1,t_2}}{\kappa_{\eta,s,t}+\kappa_{\eta,t+1,t_2}} \in [0,1]$ such that[69]

$$\beta_{s-1,t_2}\omega_{s,t_2} = \beta_{s-1,t_2}\left(1-\gamma_{s,t,t_2}\right)\omega_{s,t} + \beta_{s-1,t_2}\gamma_{s,t,t_2}\omega_{t+1,t_2}. \tag{63}$$

[69] If $\kappa_{\eta,s,t} + \kappa_{\eta,t+1,t_2} = \kappa_{\eta,s,t_2} = 0$, condition (63) is automatically satisfied for any γ_{s,t,t_2}, because in that case $\beta_{s-1,t_2} = \frac{\kappa_{\eta,s,t_2}}{\kappa_{s-1}+\kappa_{\epsilon,s,t_2}+\kappa_{\eta,s,t_2}} = 0$.

Using condition (57), we then have

$$\beta_{s-1,t_2}\bar{E}_t\left[\omega_{s,t_2}\right]=\beta_{s-1,t_2}\left(1-\gamma_{s,t,t_2}\right)\omega_{s,t}+\beta_{s-1,t_2}\gamma_{s,t,t_2}\bar{E}_t\left[z\right]. \tag{64}$$

Now we prove that, under Assumption 6, whenever $k\geq 2$ and $t=t_1<t_2<...<t_k\leq T-1$, we have

$$Var\left(\bar{E}_{t,t_2\cdots,t_k}\left[z\right]\right)<Var\left(\bar{E}_t\left[z\right]\right),$$

We proceed by induction on k. When $k=2$, we only need to prove that, whenever $t=t_1<t_2\leq T-1$,

$$Var\left(\bar{E}_{t,t_2}\left[z\right]\right)<Var\left(\bar{E}_t\left[z\right]\right). \tag{65}$$

From conditions (57) and (62), we have

$$E_{i,t}\left[\bar{E}_{t_2}\left[z\right]\right]=\left(1-\alpha_{t,t_2}-\beta_{t,t_2}\right)E_{i,t}\left[\bar{E}_t\left[z\right]\right]+\left(\alpha_{t,t_2}+\beta_{t,t_2}\right)E_{i,t}\left[z\right],$$

$$\bar{E}_{t,t_2}\left[z\right]=\left(1-\alpha_{t,t_2}-\beta_{t,t_2}\right)\bar{E}_t\left[\bar{E}_t\left[z\right]\right]+\left(\alpha_{t,t_2}+\beta_{t,t_2}\right)\bar{E}_t\left[z\right]. \tag{66}$$

Together with (60) and the fact that $0\leq\alpha_{t,t_2}+\beta_{t,t_2}=\frac{\kappa_{\epsilon,t+1,t_2}+\kappa_{\eta,t+1,t_2}}{\kappa_t+\kappa_{\epsilon,t+1,t_2}+\kappa_{\eta,t+1,t_2}}<1$, we have

$$Var\left(\bar{E}_{t,t_2}\left[z\right]\right)\leq\left(\left(1-\alpha_{t,t_2}-\beta_{t,t_2}\right)\sqrt{Var\left(\bar{E}_t\left[\bar{E}_t\left[z\right]\right]\right)}+\left(\alpha_{t,t_2}+\beta_{t,t_2}\right)\sqrt{Var\left(\bar{E}_t\left[z\right]\right)}\right)^2<Var\left(\bar{E}_t\left[z\right]\right), \tag{67}$$

where we used the fact that, for any random variables X,Y and scalars $a,b\geq 0$, the following is true:

$$\begin{aligned}Var(aX+bY) &= a^2Var(X)+2abCov\left(X,Y\right)+b^2Var\left(Y\right)\\ &\leq a^2Var(X)+2ab\sqrt{Var\left(X\right)Var\left(Y\right)}+b^2Var\left(Y\right)\\ &= \left(a\sqrt{Var\left(X\right)}+b\sqrt{Var\left(Y\right)}\right)^2.\end{aligned} \tag{68}$$

This finishes the proof of condition (65).

Now suppose for $k\geq 2$, we have, whenever $1\leq l\leq k$, and $t=t_1<t_2<...<t_l\leq T-1$,

$$Var\left(\bar{E}_{t,t_2,\cdots,t_l}\left[z\right]\right)<Var\left(\bar{E}_t\left[z\right]\right). \tag{69}$$

We prove, whenever $t=t_1<t_2<...<t_{k+1}\leq T-1$, that

$$Var\left(\bar{E}_{t,t_2,\cdots,t_{k+1}}\left[z\right]\right)<Var\left(\bar{E}_t\left[z\right]\right), \tag{70}$$

Using condition (62), we have

$$\bar{E}_{t_{k+1}}\left[z\right]=\left(1-\alpha_{t,t_{k+1}}-\beta_{t,t_{k+1}}\right)\bar{E}_t\left[z\right]+\alpha_{t,t_{k+1}}z+\beta_{t,t_{k+1}}\omega_{t+1,t_{k+1}}.$$

Together with condition (64), we have

$$\begin{aligned}\bar{E}_{t_k,t_{k+1}}[z] &= \left(1-\alpha_{t,t_{k+1}}-\beta_{t,t_{k+1}}\right)\bar{E}_{t_k,t}[z]+\alpha_{t,t_{k+1}}\bar{E}_{t_k}[z]+\beta_{t,t_{k+1}}\bar{E}_{t_k}\left[\omega_{t+1,t_{k+1}}\right]\\ &= \left(1-\alpha_{t,t_{k+1}}-\beta_{t,t_{k+1}}\right)\bar{E}_{t_k,t}[z]+\left(\alpha_{t,t_{k+1}}+\beta_{t,t_{k+1}}\gamma_{t+1,t_k,t_{k+1}}\right)\bar{E}_{t_k}[z]+\beta_{t,t_{k+1}}\left(1-\gamma_{t+1,t_k,t_{k+1}}\right)\omega_{t+1,t_k}.\end{aligned}$$

As a result,

$$\begin{aligned}\bar{E}_{t,t_2,\cdots,t_{k+1}}[z] = {} & \left(1-\alpha_{t,t_{k+1}}-\beta_{t,t_{k+1}}\right)\bar{E}_{t,\cdots,t_k,t}[z]+\left(\alpha_{t,t_{k+1}}+\beta_{t,t_{k+1}}\gamma_{t+1,t_k,t_{k+1}}\right)\bar{E}_{t,\cdots,t_k}[z]\\ & +\beta_{t,t_{k+1}}\left(1-\gamma_{t+1,t_k,t_{k+1}}\right)\bar{E}_{t,\cdots,t_{k-1}}\left[\omega_{t+1,t_k}\right].\end{aligned}$$

Iterating $\beta_{t,t_{k+1}}\bar{E}_{t,\cdots,t_{k-1}}\left[\omega_{t+1,t_k}\right]$ by condition (64), we have

$$\begin{aligned}\bar{E}_{t,t_1,t_2,\cdots,t_{k+1}}[z] = {} & \left(1-\alpha_{t,t_{k+1}}-\beta_{t,t_{k+1}}\right)\bar{E}_{t,\cdots,t_k,t}[z]+\left(\alpha_{t,t_{k+1}}+\beta_{t,t_{k+1}}\gamma_{t+1,t_k,t_{k+1}}\right)\bar{E}_{t,\cdots,t_k}[z]\\ & +\beta_{t,t_{k+1}}\sum_{l=1}^{k-1}\left\{\left[\prod_{s=l+1}^{k}\left(1-\gamma_{t+1,t_s,t_{s+1}}\right)\right]\gamma_{t+1,t_l,t_{l+1}}\bar{E}_{t,\cdots,t_l}[z]\right\}\\ & +\beta_{t,t_{k+1}}\prod_{s=1}^{k}\left(1-\gamma_{t+1,t_s,t_{s+1}}\right)\bar{E}_t[z]\end{aligned} \tag{71}$$

Using (69), we have

$$Var\left(\bar{E}_{t,t_2,\cdots,t_l}[z]\right) < Var\left(\bar{E}_t[z]\right) \quad 1\le l\le k. \tag{72}$$

Now we prove

$$Var\left(\bar{E}_{t,\cdots,t_k,t}[z]\right) < Var\left(\bar{E}_t[z]\right). \tag{73}$$

From condition (59) and the fact that each consumer i has uncertainty about z at period t_k, each consumer also faces uncertainty about $\bar{E}_t[z]$ at period t_k. Similar as (60), we have,

$$Var\left(\bar{E}_{t_k,t}[z]\right) < Var\left(\bar{E}_t[z]\right).$$

Similarly to (61), we have

$$Var\left(\bar{E}_{t,\cdots,t_k,t}[z]\right) \le Var\left(\bar{E}_{t_k,t}[z]\right) < Var\left(\bar{E}_t[z]\right).$$

This proves (73). From conditions (68), (71), (72) and (73), we have

$$Var\left(\bar{E}_{t,t_2,\cdots,t_{k+1}}[z]\right) < Var\left(\bar{E}_t[z]\right).$$

This proves (70), thus also proving (55) and completing the proof of Lemma 5.

Proof of Corollary 3. Whenever $t \leq T-2$, $2 \leq k \leq t-T$ and $x \in B_t^k$, let $\beta_x = \frac{Cov\left(x, \bar{E}_t[\tilde{R}_T]\right)}{Var\left(\bar{E}_t[\tilde{R}_T]\right)}$ denote the coefficient of the projection of x on $\bar{E}_t[\tilde{R}_T]$. Under Assumption 4, $\left|Cov\left(x, \bar{E}_t[\tilde{R}_T]\right)\right| \leq \sqrt{Var\,(x)\,Var\left(\bar{E}_t[\tilde{R}_T]\right)} < Var\left(\bar{E}_t[\tilde{R}_T]\right)$ from Lemma 5. As a result, $\beta_x < 1$. Moreover, we have

$$|\beta_x| = \left|\frac{Cov\left(x, \bar{E}_t[\tilde{R}_T]\right)}{Var\left(\bar{E}_t[\tilde{R}_T]\right)}\right| < 1. \tag{74}$$

Proof of Proposition 5. This is a special case of Proposition 6.

Proof of Lemma 6. Under Assumption 1, aggregate spending in equation (22) can be written as

$$\tilde{y}_t = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{\pi}_{t+k}\right] + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right]. \tag{75}$$

Substituting condition (21) into the above condition and use the fact that $\tilde{\mu}_t$ is i.i.d. over time, we have

$$\begin{aligned}\tilde{y}_t &= -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \kappa\sum_{k=1}^{T-t}\beta^{k-1}\left\{\bar{E}_t\left[\sum_{\tau=0}^{T-t-k}\beta^{\tau}\tilde{y}_{t+k+\tau}\right]\right\} + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right]\\ &= -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \sum_{k=1}^{T-t}\left(1-\beta+k\kappa\right)\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right].\end{aligned}$$

Proof of Proposition 6. For $t < T-1$, under Assumption 4, using condition (20), we can write $\tilde{y}_t$ as a linear function of elements in B_t^k $(1 \leq k \leq T-t)$:

$$\tilde{y}_t = -\sum_{1\leq k\leq T-t \text{ and } x\in B_t^k}\chi_x x \tag{76}$$

where $\chi_x > 0$ for all $x \in B_t^k$ and $1 \leq k \leq T-t$.

Note that, under complete information (Assumption 2), $x = E_t[\tilde{R}_T]$, for all $x \in B_0^k$ and $1 \leq k \leq T$. We have

$$\sum_{1\leq k\leq T-t \text{ and } x\in B_t^k}\chi_x = \phi^*\left(T-t\right) \tag{77}$$

from the proof of Lemma 2.

Let ϕ_t denote the absolute value of coefficient of the projection of $\tilde{y}_t$ on $\bar{E}_t[\tilde{R}_T]$. Using (76), (77) and (74), we know that $\phi_t < \phi_t^* = \phi^*\left(T-t\right)$.[70]

[70]We use the fact that, when $t < T-1$, (76) contains some terms $x \in B_t^k$ with $k \geq 2$. We also use that, from Corollary 3, the coefficient of the projection of such x on $\bar{E}_t[\tilde{R}_T]$ is strictly less than 1.

Proof of Proposition 7. As mentioned in the main text, we maintain the fact that firms have complete information, $\mathcal{I}_{j,t}^{f}=\mathcal{I}_t$. Extending the definition in the main text, for all $\tau,k\geq 0$, we use $\chi_{k,\tau}$ to denote the (absolute value of) influence of k-order average belief of $\tilde{R}_T$ on spending $\tilde{y}_{T-\tau}$.[71] Also, for all $k,\tau\geq 0$, we let $s_{k,\tau}\equiv\sum_{l=1}^{k}\chi_{l,\tau}$ denote the combined (absolute value of) effect of beliefs of order up to k on spending $\tilde{y}_{T-\tau}$.

From condition (22), the effect of k-order average belief on spending, $\chi_{k,\tau}$, can be characterized as

$$\chi_{0,0}=1,$$

$$\chi_{1,\tau}=(1+\tau\kappa)\,\beta^{\tau-1}\quad\forall\tau\geq 1,\tag{78}$$

$$\chi_{k,\tau}=\sum_{l=1}^{\tau-k+1}(1-\beta+l\kappa)\,\beta^{l-1}x_{k-1,\tau-l}\quad\forall k\geq 2\;and\;\tau\geq k,\tag{79}$$

$$\chi_{k,\tau}=0\quad\forall k>\tau.$$

We can then characterize the combined effect of beliefs of order up to k on spending, $s_{k,\tau}$, as

$$s_{0,0}=1,$$

$$s_{0,\tau}=0\quad\forall\tau\geq 1,$$

$$s_{k,\tau}=\beta^{\tau}+\sum_{l=1}^{\tau}(1-\beta+l\kappa)\,\beta^{l-1}s_{k-1,\tau-l}\quad\forall k\geq 1\;and\;\tau\geq 0.\tag{80}$$

Let $d_\tau=s_{\tau,\tau}$ denote the combined effect of beliefs of all different orders on spending. It coincides with the spending response under complete information defined in Proposition 1, $d_\tau=\phi^*(\tau)$. This is because, under Assumption 2, all different orders of beliefs about R_T at period t are $E_t[\tilde{R}_T]$ itself. Following the proof of Proposition 1, we have

$$d_0=1,$$

$$d_1=1+\kappa,$$

$$\frac{d_\tau}{d_{\tau-1}}+\beta\frac{d_{\tau-2}}{d_{\tau-1}}=1+\beta+\kappa\qquad\forall\tau\geq 2.\tag{81}$$

Now we prove $s_{k,\tau}$ satisfies an inequality with a similar form as condition (81):

$$\frac{s_{k,\tau}}{s_{k,\tau-1}}+\beta\frac{s_{k,\tau-2}}{s_{k,\tau-1}}\leq 1+\beta+\kappa\quad\forall\tau\geq 2\;and\;k\geq 1.\tag{82}$$

[71] We have $\chi_{k,\tau}=0\quad\forall k>\tau\geq 0$, that is, when the order of belief k is higher than horizon $\tau=T-t$, k-th order belief does not matter for spending $\tilde{y}_{T-\tau}$. This can be seen from iterating conditions (22) once we impose Assumption 1. Also, as one can see from iterating conditions (22), the influence of k-order belief on consumption and inflation only depends on the distance between t and T, $\tau=T-t$. That is why we can use notations $\chi_{k,\tau}$ and $\psi_{k,\tau}$. Finally, note that we assume firms are completely informed. Higher order beliefs here only involve consumer's belief of consumer's belief of $\cdots$.

From condition (80), we have

$$\beta s_{k,\tau-1} = \beta^{\tau} + \sum_{l=2}^{\tau} \left(1-\beta+(l-1)\kappa\right)\beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 1.$$

As a result, we can write $s_{k,\tau}$ in a recursive form:

$$s_{k,\tau} = \beta s_{k,\tau-1} + (1-\beta)\, s_{k-1,\tau-1} + \kappa \sum_{l=1}^{\tau} \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 1,$$

$$\beta s_{k,\tau-1} = \beta^{2} s_{k,\tau-2} + \beta\,(1-\beta)\, s_{k-1,\tau-2} + \kappa \sum_{l=2}^{\tau} \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 2.$$

Using the previous two conditions, we have, for all $k \geq 1$ and $\tau \geq 2$,

$$s_{k,\tau} + \beta^{2} s_{k,\tau-2} + \beta\,(1-\beta)\, s_{k-1,\tau-2} = 2\beta s_{k,\tau-1} + (1-\beta+\kappa)\, s_{k-1,\tau-1},$$

$$s_{k,\tau} + \beta s_{k,\tau-2} = (1+\beta+\kappa)\, s_{k,\tau-1} + \beta\,(1-\beta)\,\chi_{k,\tau-2} - (1-\beta+\kappa)\,\chi_{k,\tau-1}. \tag{83}$$

To prove (82), we only need to prove:

$$\beta\,(1-\beta)\,\chi_{k,\tau-2} \leq (1-\beta+\kappa)\, x_{k,\tau-1} \quad \forall k \geq 1 \; and \; \tau \geq 2. \tag{84}$$

In fact, we prove the following stronger result:

$$\beta \chi_{k,\tau-2} \leq x_{k,\tau-1} \quad \forall k \geq 1 \; and \; \tau \geq 2. \tag{85}$$

From condition (78), we know that (85) is true for $k=1$ and $\tau \geq 2$. From condition (79), we know that

$$\chi_{k,\tau-1} = \beta^{\tau-1} + \sum_{l=1}^{\tau-k} (1-\beta+l\kappa)\,\beta^{l-1} x_{k-1,\tau-1-l} \quad \forall k \geq 2 \; and \; \tau \geq k+1, \tag{86}$$

$$\beta \chi_{k,\tau-2} = \beta^{\tau-1} + \sum_{l=2}^{\tau-k} \left(1-\beta+(l-1)\kappa\right)\beta^{l-1} x_{k-1,\tau-1-l} \quad \forall k \geq 2 \; and \; \tau \geq k+2.$$

This proves $\beta\chi_{k,\tau-2} \leq x_{k,\tau-1}$ for $k \geq 2$ and $\tau \geq k+2$. Together with the fact that, $\chi_{k,\tau-2} = 0 \; \forall k \geq \tau-1$, we prove (85) and thus (84). This finishes the proof of (82).

Based on (81) and (82), we can then establish Proposition (7). That is, for any given $k \geq 1$ and $\tau \geq k$, the relative contribution of the first k orders, $\frac{s_{k,\tau}}{s_{\tau,\tau}} = \frac{s_{k,\tau}}{d_{\tau}}$, strictly decreases with τ.

First, note that, for any given $k \geq 1, 1 = \frac{s_{k,k}}{d_k} > \frac{s_{k,k+1}}{d_{k+1}}$, because $x_{k+1,k+1} > 0$. Then, we can proceed by induction on $\tau \geq k$, for any fixed $k \geq 1$. If we have $\frac{s_{k,\tau}}{d_{\tau}} > \frac{s_{k,\tau+1}}{d_{\tau+1}}$ for some $\tau \geq k$, we have $\frac{s_{k,\tau}}{s_{k,\tau+1}} > \frac{d_{\tau}}{d_{\tau+1}}$. Using (81) and (82), we have $\frac{s_{k,\tau+2}}{s_{k,\tau+1}} < \frac{d_{\tau+2}}{d_{\tau+1}}$, and thus $\frac{s_{k,\tau+1}}{d_{\tau+1}} > \frac{s_{k,\tau+2}}{d_{\tau+2}}$. This completes the proof that, for any $k \geq 1$ and any $\tau \geq k$, the ratio $\frac{s_{k,\tau}}{s_{\tau,\tau}}$, strictly decreases with the horizon τ.

Finally, we prove that, for any $k \geq 1$,

$$\frac{s_{k,\tau}}{s_{\tau,\tau}} \to 0, \quad as\ \tau \to \infty. \tag{87}$$

In other words, we want to prove the relative contribution of the first k orders of beliefs to aggregate spending converges to zero when the horizon τ goes to infinity. We consider two cases, depending on whether $\kappa = 0$ or $\kappa > 0$.

(i) Suppose $\kappa = 0$. From the proof of Proposition 1, we have $s_{\tau,\tau} = \phi^* (\tau) = 1,\ \forall \tau \geq 0$. From condition (78), $\chi_{1,\tau} = s_{1,\tau} = \beta^{\tau-1}$. This proves (87) for $k = 1$. If there exists $k \geq 2$ such that (87) does not hold, we let $k^* \geq 2$ denote the smallest of such k. Then, (87) holds for $1 \leq k \leq k^* - 1$. Because we already prove that $s_{k^*,\tau} = \frac{s_{k^*,\tau}}{s_{\tau,\tau}} \geq 0$ is decreasing with the horizon τ, there exists $0 < \Gamma < 1$ such that $s_{k^*,\tau} = \frac{s_{k^*,\tau}}{s_{\tau,\tau}} \to \Gamma \quad as\ \tau \to \infty$. Note that $s_{k,\tau} = s_{k-1,\tau} + \chi_{k,\tau},\ \forall k, \tau \geq 1$, we have

$$\chi_{k^*,\tau} = \frac{\chi_{k^*,\tau}}{s_{\tau,\tau}} \to \Gamma \quad as\ \tau \to \infty, \tag{88}$$

and

$$\chi_{k,\tau} = \frac{\chi_{k,\tau}}{s_{\tau,\tau}} \to 0,\ \forall 1 \leq k \leq k^* - 1 \quad as\ \tau \to \infty. \tag{89}$$

Now we prove that, when $\kappa = 0$, for all $k \geq 2$ and $\tau \geq k + 1$,

$$\chi_{k,\tau} = \beta x_{k,\tau-1} + (1-\beta)\, \chi_{k-1,\tau-1}. \tag{90}$$

This comes from conditions (79) and (86). From (88) and (89), we have

$$\chi_{k^*,\tau} \to \Gamma \quad and \quad \beta \chi_{k^*,\tau} + (1-\beta)\, \chi_{k^*-1,\tau} \to \beta\Gamma \quad as\ \tau \to \infty.$$

This contradicts condition (90) when $\Gamma > 0$. As a result, we have proved (87) when $\kappa = 0$.

(ii) Now let $\kappa > 0$. First note that, from condition (80), we have $s_{1,\tau} = (1 + \tau\kappa)\, \beta^{\tau-1} \to 0,$ as $\tau \to +\infty$. From the proof of Proposition 1, we know $s_{\tau,\tau} = d_\tau = \phi^* (\tau) \geq 1$. As a result, (87) is true for $k = 1$.

If there exists $k \geq 2$ such that (87) does not hold, we let $k^* \geq 2$ denote the smallest of such k. Then, (87) holds for $1 \leq k \leq k^* - 1$. Because we already prove that $\frac{s_{k^*,\tau}}{s_{\tau,\tau}} \geq 0$ is decreasing with the horizon τ, there exists $0 < \Gamma < 1$ such that $\frac{s_{k^*,\tau}}{s_{\tau,\tau}} = \frac{s_{k^*,\tau}}{\phi^*(\tau)} \to \Gamma$ as $\tau \to \infty$. As a result, $\frac{s_{k^*,\tau}}{\phi^*(\tau)} \frac{\phi^*(\tau-1)}{s_{k^*,\tau-1}} \to 1$ as $\tau \to \infty$. Because we already prove that, in the proof of 1, $\frac{\phi^*(\tau)}{\phi^*(\tau-1)} \to \Gamma^*$, we have

$$\frac{s_{k^*,\tau}}{s_{k^*,\tau-1}} \to \Gamma^* \quad and \quad \frac{s_{k^*,\tau-2}}{s_{k^*,\tau-1}} \to \frac{1}{\Gamma^*} \quad as\ \tau \to \infty. \tag{91}$$

Note that $s_{k^*,\tau} = s_{k^*-1,\tau} + \chi_{k^*,\tau}$ and $\frac{s_{k^*-1,\tau}}{s_{\tau,\tau}} \to 0\ as\ \tau \to \infty$, we have $\frac{\chi_{k^*,\tau}}{s_{\tau,\tau}} = \frac{\chi_{k^*,\tau}}{\phi^*(\tau)} \to \Gamma\ as\ \tau \to \infty$. As a result

$$\frac{\chi_{k^*,\tau}}{s_{k^*,\tau}} = \frac{\chi_{k^*,\tau}}{\phi^*(\tau)} \frac{\phi^*(\tau)}{s_{k^*,\tau}} \to 1 \quad as\ \tau \to \infty. \tag{92}$$

Now we prove a stronger version of (82)

$$\frac{s_{k,\tau}}{s_{k,\tau-1}}+\beta\frac{s_{k,\tau-2}}{s_{k,\tau-1}}+\kappa\frac{\chi_{k,\tau-1}}{s_{k,\tau-1}}\leq 1+\beta+\kappa \quad \forall\tau\geq 2 \; and \; k\geq 1. \tag{93}$$

This comes from the fact that (85) can be written as

$$\beta\left(1-\beta\right)\chi_{k,\tau-2}+\kappa\chi_{k,\tau-1}\leq\left(1-\beta+\kappa\right)x_{k,\tau-1} \quad \forall\tau\geq 2 \; and \; k\geq 1. \tag{94}$$

Using (33), (91) and (92), we have

$$\frac{s_{k^*,\tau}}{s_{k^*,\tau-1}}+\beta\frac{s_{k^*,\tau-2}}{s_{k^*,\tau-1}}+\kappa\frac{\chi^*_{k,\tau-1}}{s^*_{k,\tau-1}}\to\Gamma^*+\beta\frac{1}{\Gamma^*}+\kappa=1+\beta+2\kappa \quad as \; \tau\to\infty.$$

This contradicts (93) and proves (87).

Proof of Proposition 8. We first prove that, under Assumption (5),

$$Var\left(x\right)\leq\left(1-\epsilon^2\right)^k Var\left(\tilde{R}_T\right). \tag{95}$$

for any $x=\bar{E}_t\bar{E}_{t_2}...\bar{E}_{t_k}[\tilde{R}_T]\in B_t^k$.

To simplify notation, let $y=\bar{E}_{t_2}...\bar{E}_{t_k}[\tilde{R}_T]$ for $k\geq 3$, $y=\bar{E}_{t_2}[\tilde{R}_T]$ for $k=2$, and $y=\tilde{R}_T$ for $k=1$. From Assumption 5, there is at least a mass ϵ of consumers such that

$$Var\left(E_t\left[y\right]|\mathcal{I}_{it}\right)\geq\epsilon Var\left(E_t\left[y\right]\right).$$

As a result,

$$E\left[Var\left(E_t\left[y\right]|\mathcal{I}_{it}\right)|\Omega_t\right]\geq\epsilon^2 Var\left(E_t\left[y\right]\right)$$

where Ω_t is the cross-sectional distribution of consumer i's information set $\mathcal{I}_{it}$ at period t. Using the law of total variance, we have

$$E\left[Var\left(E_t\left[y\right]|\mathcal{I}_{it}\right)|\Omega_t\right]+Var\left(E\left[E_t\left[y\right]|\mathcal{I}_{it}\right]\right)=E\left[Var\left(E_t\left[y\right]|\mathcal{I}_{it}\right)|\Omega_t\right]+Var\left(E\left[y|\mathcal{I}_{it}\right]\right)=Var\left(E_t\left[y\right]\right).$$

As a result, we have[72]

$$\begin{aligned}Var\left(x\right)=Var\left(\bar{E}_t\bar{E}_{t_2}...\bar{E}_{t_k}[\tilde{R}_T]\right)&=Var\left(\bar{E}_t\left[y\right]\right)\leq Var\left(E\left[y|\mathcal{I}_{it}\right]\right)\leq\left(1-\epsilon^2\right)Var\left(E_t\left[y\right]\right)\\&\leq\left(1-\epsilon^2\right)Var\left(y\right)=\left(1-\epsilon^2\right)Var\left(\bar{E}_{t_2}...\bar{E}_{t_k}[\tilde{R}_T]\right).\end{aligned}$$

Iterating the previous condition proves (95).

[72]We use the fact that for any random variable X, and any information set I, $Var\left(E\left[X|I\right]\right)\leq Var\left(X\right)$. We also use the fact that $\bar{E}_t\left[\cdot\right]=E\left[E\left[\cdot\,|\mathcal{I}_{it}\right]|\Omega_t\right]$, where Ω_t is the cross sectional distribution of $\mathcal{I}_{it}$ at time t,

Let ϕ denote the absolute value of the coefficient of the projection of $\tilde{y}_0$ on $\bar{E}_0\left[R_T\right]$. Together with (68) and the fact that $\phi^*(T) = s_{T,T}$[73], we have

$$\left(\frac{\phi}{\phi^*}\right)^2 = \left(\frac{Cov\left(\tilde{y}_0, \bar{E}_0\left[\tilde{R}_T\right]\right)}{\phi^*(T)Var\left(\bar{E}_0\left[\tilde{R}_T\right]\right)}\right)^2 \leq \frac{Var\left(\tilde{y}_0\right)}{\left[\phi^*(T)\right]^2 Var\left(\bar{E}_0[\tilde{R}_T]\right)} \leq \frac{1}{Var\left(\bar{E}_0[\tilde{R}_T]\right)} \Bigg[\sum_{k=1}^{T} \Bigg(\frac{s_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}} \sqrt{Var\left(\tilde{R}_T\right)}$$

$$= \left[\sum_{k=1}^{T}\left(\frac{s_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}}\right)\right]^2 \frac{Var\left(\tilde{R}_T\right)}{Var\left(\bar{E}_0[\tilde{R}_T]\right)}. \tag{96}$$

For any $\vartheta > 0$, there exists $h \in \mathbb{N}_+$ such that $\frac{\left(1-\epsilon^2\right)^{\frac{h}{2}}}{1-\left(1-\epsilon^2\right)^{\frac{1}{2}}} \leq \frac{\vartheta}{2}$. From Proposition 7, there exists $T^* \in \mathbb{N}_+$ such that, for all $T \geq T^*$, $\sum_{k=1}^{h-1} \frac{s_{k,T}}{s_{T,T}} \leq \frac{\vartheta}{2}$. As a result, for all $T \geq T^*$,

$$\sum_{k=1}^{T} \frac{s_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}} \leq \sum_{k=1}^{h-1} \frac{s_{k,T}}{s_{T,T}} + \sum_{k=h}^{T}\left(1-\epsilon^2\right)^{\frac{k}{2}} \leq \frac{\vartheta}{2} + \frac{\left(1-\epsilon^2\right)^{\frac{h}{2}}}{1-\left(1-\epsilon^2\right)^{\frac{1}{2}}} \leq \vartheta.$$

This proves

$$\sum_{k=1}^{T}\left(\frac{s_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}}\right) \to 0 \text{ as } T \to +\infty.$$

Together with (96) and the fact that $Var\left(\bar{E}_0[\tilde{R}_T]\right) \geq \epsilon$, the proof of Proposition (8) is completed.

Proof of Proposition 9. To simplify notation, we use ϕ_τ and ϖ_τ as shortcuts for, respectively, $\phi\left(\lambda, \lambda^f, \tau\right)$ and $\varpi\left(\lambda, \lambda^f, \tau\right)$, where the functions ϕ and ϖ are defined as in the proof of Lemma 3. Similarly, we use ϕ_τ^* to denote $\phi^*\left(\tau\right)$.

From conditions (36) and (37), we have

$$\frac{\partial \phi_1}{\partial \kappa} = \frac{\partial \phi(\lambda, 1, 1)}{\partial \kappa} = 1 > 0 \; and \; \frac{\partial \varpi_1}{\partial \kappa} = \frac{\partial \varpi(\lambda, 1, 1)}{\partial \kappa} = \beta > 0. \tag{97}$$

For any $\tau \geq 2$, when $\lambda^f = 1$, conditions (41) and (42) become

$$\phi_\tau = \left(\beta + \left(1 - \beta + \kappa\right)\lambda\right)\phi_{\tau-1} + \varpi_{\tau-1},$$

$$\varpi_\tau = \kappa\beta\lambda\phi_{\tau-1} + \beta\varpi_{\tau-1}.$$

As a result, for all $\tau \geq 2$, we have

$$\frac{\partial \phi_\tau}{\partial \kappa} = \left(\beta + \left(1-\beta+\kappa\right)\lambda\right)\frac{\partial \phi_{\tau-1}}{\partial \kappa} + \lambda\phi_{\tau-1} + \frac{\partial \varpi_{\tau-1}}{\partial \kappa}, \tag{98}$$

$$\frac{\partial \varpi_\tau}{\partial \kappa} = \kappa\beta\lambda\frac{\partial \phi_{\tau-1}}{\partial \kappa} + \beta\lambda\phi_{\tau-1} + \beta\frac{\partial \varpi_{\tau-1}}{\partial \kappa}. \tag{99}$$

[73] This is pointed out in the proof of Proposition 7.

From conditions (97), (98) and (99), $\frac{\partial\phi_\tau}{\partial\kappa}$ and $\frac{\partial\varpi_\tau}{\partial\kappa}$ are strictly positive for any $\tau \geq 1$ by induction. Moreover, from conditions (36), (37) and (97), we have that $\frac{\partial\phi_2}{\partial\kappa}$ and $\frac{\partial\varpi_2}{\partial\kappa}$ are strictly increasing in λ. Then, from conditions (98), (99) and the fact that ϕ_τ itself is strictly increasing in λ for all $\tau \geq 2$, we have $\frac{\partial\phi_\tau}{\partial\kappa}$ and $\frac{\partial\varpi_\tau}{\partial\kappa}$ are strictly increasing in λ for all $\tau \geq 2$ by induction. The result then follows simply from letting $\tau = T$.

Proof of Proposition 10. To simplify notation, we once again use ϕ_τ and ϖ_τ as shortcuts for, respectively, $\phi\left(\lambda, \lambda^f, \tau\right)$ and $\varpi\left(\lambda, \lambda^f, \tau\right)$.

(i) The Euler condition still holds at the individual level, as condition (10) shows. Imposing Assumption 1 and aggregating equation (10), we have, for all $t \leq T-2$:

$$\tilde{y}_t = \tilde{c}_t = \int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di + \bar{E}_t\left[\tilde{\pi}_{t+1}\right],$$

which is condition (25) in the main text. What remains is to link $\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di$ to $E_t\left[\tilde{y}_{t+1}\right]$ and $\bar{E}_t\left[\tilde{\pi}_{t+1}\right]$ to $E_t\left[\tilde{\pi}_{t+1}\right]$.

From conditions (11) and (16), we have, for all $t \leq T-2$:

$$\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right] di = -\beta^{T-t}\bar{E}_t[\tilde{R}_T] + \sum_{k=2}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{\pi}_{t+k}\right] + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t\left[\tilde{y}_{t+k}\right], \tag{100}$$

$$\beta\tilde{y}_{t+1} = -\beta^{T-t}\bar{E}_{t+1}[\tilde{R}_T] + \sum_{k=2}^{T-t}\beta^{k-1}\bar{E}_{t+1}\left[\tilde{\pi}_{t+k}\right] + \sum_{k=2}^{T-t}(1-\beta)\,\beta^{k-1}\bar{E}_{t+1}\left[\tilde{y}_{t+k}\right]. \tag{101}$$

Under Assumption 3, each consumer has the same information about $\tilde{R}_T$ in all periods $t \leq T-1$. In particular,

$$\bar{E}_t[\tilde{R}_T] = \lambda\delta z \quad \forall t \leq T-1, \tag{102}$$

and, by implication,

$$\bar{E}_t[\bar{E}_{t+1}[\tilde{R}_T]] = \lambda\bar{E}_{t+1}[\tilde{R}_T] \text{ and } \bar{E}_t[\bar{E}^f_{t+1}[\tilde{R}_T]] = \lambda\bar{E}^f_{t+1}[\tilde{R}_T] \quad \forall t \leq T-2. \tag{103}$$

Furthermore, from the proof of Lemma 3, we have that, for all $t \leq T-1$,

$$\tilde{y}_t = -\phi_{T-t}\bar{E}_t[\tilde{R}_T] \qquad \text{and} \qquad \tilde{\pi}_t = -\varpi_{T-t}\bar{E}_t[\tilde{R}_T] + \kappa\tilde{y}_t + \kappa\tilde{\mu}_t.$$

Together with conditions (103) and (103), we have, for all $t \leq T-2$,

$$\bar{E}_t\left[\tilde{y}_{t+1}\right] = -\bar{E}_t\left[\phi_{T-t-1}\bar{E}_{t+1}[\tilde{R}_T]\right] = -\phi_{T-t-1}\lambda\bar{E}_{t+1}[\tilde{R}_T] = \lambda E_t\left[\tilde{y}_{t+1}\right], \tag{104}$$

$$\bar{E}_t\left[\tilde{\pi}_{t+1}\right] = \bar{E}_t\left[-\varpi_{T-t-1}\bar{E}^f_{t+1}[\tilde{R}_T] + \kappa\tilde{y}_{t+1} + \kappa\tilde{\mu}_{t+1}\right] = -\varpi_{T-t-1}\lambda\bar{E}^f_{t+1}[\tilde{R}_T] + \lambda\kappa E_t\left[\tilde{y}_{t+1}\right] = \lambda E_t\left[\tilde{\pi}_{t+1}\right], \tag{105}$$

$$\bar{E}_t\left[\tilde{y}_{t+k}\right]=\bar{E}_{t+1}\left[\tilde{y}_{t+k}\right] \quad \text{and} \qquad \bar{E}_t\left[\tilde{\pi}_{t+k}\right]=\bar{E}_{t+1}\left[\tilde{\pi}_{t+k}\right] \quad \forall k\in\{2,...,T-t\} \tag{106}$$

Together with conditions (100) and (101), we have, for all $t\leq T-2$,

$$\beta\tilde{y}_{t+1}+(1-\beta)\,\bar{E}_t\left[\tilde{y}_{t+1}\right]=\int_0^1 E_{i,t}\left[\tilde{c}_{i,t+1}\right]di.$$

Substituting the previous condition into condition (25), we have, for all $t\leq T-2$,

$$\tilde{y}_t=\beta\tilde{y}_{t+1}+(1-\beta)\,\bar{E}_t\left[\tilde{y}_{t+1}\right]+\bar{E}_t\left[\tilde{\pi}_{t+1}\right].$$

Using conditions (104) and (105), we arrive at the following discounted Euler equation, for all $t\leq T-2$:

$$\tilde{y}_t=(\beta+(1-\beta)\,\lambda)\,E_t\left[\tilde{y}_{t+1}\right]+\lambda E_t\left[\tilde{\pi}_{t+1}\right].$$

(ii) From condition (17), we have, for all $t\leq T-2$:

$$\tilde{\pi}_t=\kappa\sum_{k=1}^{T-t}(\beta\theta)^k\,\bar{E}_t^f\left\{\tilde{y}_{t+k}\right\}+\frac{1-\theta}{\theta}\sum_{k=1}^{T-t}(\beta\theta)^k\,\bar{E}_t^f\left\{\tilde{\pi}_{t+k}\right\}+\kappa\tilde{y}_t+\kappa\tilde{\mu}_t, \tag{107}$$

$$\tilde{\pi}_{t+1}=\kappa\sum_{k=2}^{T-t}(\beta\theta)^{k-1}\,\bar{E}_{t+1}^f\left\{\tilde{y}_{t+k}\right\}+\frac{1-\theta}{\theta}\sum_{k=2}^{T-t}(\beta\theta)^{k-1}\,\bar{E}_{t+1}^f\left\{\tilde{\pi}_{t+k}\right\}+\kappa\tilde{y}_{t+1}+\kappa\tilde{\mu}_{t+1}. \tag{108}$$

Using a similar argument as in part (i), we have that, under Assumption 3, for all $t\leq T-2$,

$$\bar{E}_t^f\left[\tilde{y}_{t+1}\right]=\lambda^f E_t\left[\tilde{y}_{t+1}\right],\quad \bar{E}_t^f\left[\tilde{\pi}_{t+1}\right]=\lambda^f E_t\left[\tilde{\pi}_{t+1}\right],$$

$$\bar{E}_t^f\left[\tilde{\pi}_{t+k}\right]=\bar{E}_{t+1}^f\left[\tilde{\pi}_{t+k}\right] \text{ and } \bar{E}_t^f\left[\tilde{y}_{t+k}\right]=\bar{E}_{t+1}^f\left[\tilde{y}_{t+k}\right] \;\; \forall k\in\{2,...,T-t\}.$$

Together with conditions (34), (107) and (108), we have

$$\begin{aligned}\tilde{\pi}_t&=\kappa\tilde{y}_t+\beta\left[(1-\theta)\,\bar{E}_t^f[\tilde{\pi}_{t+1}]+\theta E_t[\tilde{\pi}_{t+1}]\right]+\beta\theta\kappa\left(\bar{E}_t^f[\tilde{y}_{t+1}]-E_t[\tilde{y}_{t+1}]\right)+\kappa\tilde{\mu}_t\\&=\beta\left[\theta+(1-\theta)\,\lambda^f\right]E_t\left[\tilde{\pi}_{t+1}\right]+\kappa\left\{\tilde{y}_t-\beta\theta\left(1-\lambda^f\right)E_t[\tilde{y}_{t+1}]\right\}+\kappa\tilde{\mu}_t\\&=\beta\left[\theta+(1-\theta)\,\lambda^f\right]E_t\left[\tilde{\pi}_{t+1}\right]+\kappa m_t\tilde{y}_t,\end{aligned}$$

where $m_t\equiv 1-\beta\theta\left(1-\lambda^f\right)\frac{\phi_{T-t-1}}{\phi_{T-t}}$. Finally, to verify that $m_t\in(0,1)$, note that, from condition (41),

$$0<\frac{\phi_{T-t-1}}{\phi_{T-t}}\leq\frac{1}{\beta+(1-\beta+\kappa)\,\lambda}.$$

and therefore

$$1>m_t=1-\beta\theta\left(1-\lambda^f\right)\frac{\phi_{T-t-1}}{\phi_{T-t}}\geq 1-\frac{\beta\theta\left(1-\lambda^f\right)}{\beta+(1-\beta+\kappa)\,\lambda}>0,$$

which completes the proof.

References

Allen, Franklin, Stephen Morris, and Hyun Song Shin. (2006) "Beauty Contests and Iterated Expectations in Asset Markets." *Review of financial Studies*, 19(3): 719–752.

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

Andrade, Philippe, Gaballo Gaetano, Eric Mengus, and Benoit Mojon. (2015) "Forward Guidance and Heterogeneous Beliefs." *Banque De France mimeo*.

Angeletos, George-Marios, and Chen Lian. (2016*a*) "Dampening General Equilibrium: From Micro Elasticities to Macro Effects." *Massachusetts Institute of Technology mimeo*.

Angeletos, George-Marios, and Chen Lian. (2016*b*) "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics*, 2: forthcoming.

Angeletos, George-Marios, and Chen Lian. (2016*c*) "A Theory of Keynesian Multipliers without Nominal Rigidity." *Work in progress*.

Angeletos, George-Marios, and Jennifer La'O. (2010) "Noisy business cycles." In *NBER Macroeconomics Annual 2009, Volume 24*. 319–378. University of Chicago Press.

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

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. (2016) "Quantifying Confidence." *NBER Working Paper No. 20807*.

Angeletos, George-Marios, Luigi Iovino, and Jennifer La'O. (2016) "Real Rigidity, Nominal Rigidity, and the Social Value of Information." *The American Economic Review*, 106(1): 200–227.

Bacchetta, Philippe, and Eric van Wincoop. (2006) "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *The American Economic Review*, 96(3): 552.

Bassetto, Marco. (2015) "Forward Guidance: Communication, Commitment, or Both?" *FRB of Chicago mimeo*.

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.

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

Carlstrom, Charles T, Timothy S Fuerst, and Matthias Paustian. (2012) "How Inflationary is an Extended Period of Low Interest Rates?"

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

Chahrour, Ryan, and Gaetano Gaballo. (2016) "Learning from Prices: Amplification and Business Fluctuations." *Boston College mimeo*.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. (2011) "When Is the Government Spending Multiplier Large?" *Journal of Political Economy*, 119(1): 78–121.

Cochrane, John. (2016*a*) "Do Higher Interest Rates Raise or Lower Inflation?" *Stanford University mimeo*.

Cochrane, John. (2016*b*) "The New-Keynesian Liquidity Trap." *Stanford University mimeo*.

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

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

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. (2012) "The Forward Guidance Puzzle." *FRB of New York mimeo*.

Duffy, John, and Rosemarie Nagel. (1997) "On the Robustness of Behavior in Experimental Beauty Contest Games." *The Economic Journal*, 107(445): 1684–1700.

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

Eggertsson, Gauti B. (2010) "The Paradox of Toil." *FRB of New York Staff Report*, , (433).

Eggertsson, Gauti B, and Paul Krugman. (2012) "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach." *The Quarterly Journal of Economics*, 127(3): 1469–1513.

Farhi, Emmanuel, and Iván Werning. (2012) "Fiscal Multipliers: Liquidity Traps and Currency Unions." *NBER Working Paper No. 18381*.

Farhi, Emmanuel, and Iván Werning. (2016) "The Power of Forward Guidance with Bounded Rationality and Incomplete Markets." *Work in progress*.

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

Gabaix, Xavier. (2016) "A Behavioral New Keynesian Model." *Harvard University mimeo*.

Gaballo, Gaetano. (2016) "Rational Inattention to News: The Perils of Forward Guidance." *American Economic Journal: Macroeconomics*, 8(1): 42–97.

Galí, Jordi. (2008) *Monetary policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications.* Princeton University Press.

Galí, Jordi, J David López-Salido, and Javier Vallés. (2007) "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association*, 5(1): 227–270.

Garcıa-Schmidt, Mariana, and Michael Woodford. (2015) "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *NBER Working Paper No. 21614*.

Hellwig, Christian, and Venky Venkateswaran. (2009) "Setting the Right Prices for the Wrong Reasons." *Journal of Monetary Economics*, 56: 57–77.

Huo, Zhen, and Naoki Takayama. (2015) "Rational Expectations Models with Higher Order Beliefs." *Yale University mimeo*.

Kajii, Atsushi, and Stephen Morris. (1997) "Commonp-Belief: The General Case." *Games and Economic Behavior*, 18(1): 73–82.

Kiley, Michael T. (2016) "Policy Paradoxes in the New Keynesian Model." *Review of Economic Dynamics*, 21: 1–15.

Kocherlakota, Narayana. (2016) "Fragility of Purely Real Macroeconomic Models." *University of Rochester mimeo*.

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

Lorenzoni, Guido. (2009) "A Theory of Demand Shocks." *American Economic Review*, 99(5): 2050–84.

Lucas, Robert E. (1972) "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4(2): 103–124.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. (2016*a*) "The Discounted Euler Equation: A Note." *NBER Working Paper No. 22129*.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. (2016*b*) "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." *The 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." *The 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.

Nimark, Kristoffer. (2008) "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics*, 55(2): 365–382.

Nimark, Kristoffer. (2011) "Dynamic Higher Order Expectations." *Universitat Pompeu Fabra mimeo*.

Samet, Dov. (1998) "Iterated Expectations and Common Priors." *Games and Economic Behavior*, 24(1): 131–141.

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

Singleton, Kenneth J. (1987) "Asset Prices in a Time-series Model with Disparately Informed, Competitive Traders." In *New Approaches to Monetary Economics*. , ed. William A. Barnett and Kenneth J. Singleton, 249–272. Cambridge University Press.

Veldkamp, Laura. (2011) *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Weinstein, Jonathan, and Muhamet Yildiz. (2007) "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements." *Econometrica*, 75(2): 365–400.

Werning, Iván. (2012) "Managing a Liquidity Trap: Monetary and Fiscal Policy." *NBER Working Paper No. 17344*.

Werning, Iván. (2015) "Incomplete Markets and Aggregate Demand." *NBER Working Paper No. 21448*.

Wiederholt, Mirko. (2016) "Empirical Properties of Inflation Expectations and the Zero Lower Bound." *Goethe University Frankfurt mimeo*.

Wieland, Johannes F. (2014) "Are Negative Supply Shocks Expansionary at the Zero Lower Bound?"

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

Woodford, Michael. (2003*b*) *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Woodford, Michael. (2011) "Simple Analytics of the Government Expenditure Multiplier." *American Economic Journal: Macroeconomics*, 3(1): 1–35.

Woodford, Michael. (2012) "Methods of Policy Accommodation at the Interest-rate Lower Bound." *Columbia University mimeo*.

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