

Myopia and Anchoring*

George-Marios Angeletos[†] Zhen Huo[‡]

May 20, 2019

Abstract

We consider a forward-looking economy featuring incomplete information. The latter captures not only imperfect knowledge of the underlying fundamentals but also imperfect reasoning about the behavior of others, in the form of higher-order uncertainty. We develop an observational equivalence between this setting and a representative-agent model featuring two behavioral distortions: myopia, as in models with imperfect foresight; and anchoring of the current outcome to the past outcome, as in models with habit persistence. We further show that the as-if distortions are larger when GE feedback mechanisms, such as the Keynesian income-spending multiplier, are stronger. We use these results to draw useful connections to an emerging literature on bounded rationality and to shed new light on the empirical implications of incomplete information.

*We thank Chris Sims, Alexander Kohlhas and Luigi Iovino for discussing our paper in, respectively, the 2018 NBER Monetary Economics meeting, the 2018 Cambridge/INET conference and the 2018 Hydra workshop. We also acknowledge useful comments from Jaroslav Borovicka, Simon Gilchrist, Michael Golosov, Jennifer La'O, John Leahy, Stephen Morris, Mikkel Plagborg-Møller and seminar participants at the aforementioned conferences, Columbia, NYU, LSE, UCL, Edinburgh, CUHK, the 2018 Duke Macro Jamboree, the 2018 Barcelona GSE Summer Forum EGF Group, the 2018 SED Meeting, and the 2018 NBER Summer Institute. Angeletos acknowledges the support of the National Science Foundation under Grant Number SES-1757198.

[†]MIT and NBER; angelet@mit.edu.

[‡]Yale University; zhen.huo@yale.edu.

1 Introduction

Introducing rational inattention or other forms of incomplete information amounts to accommodating, not only partial knowledge of the underlying shocks, but also an imperfection in how agents reason about the responses of others and the associated general-equilibrium (GE) outcomes, in the form of higher-order uncertainty. In this paper, we establish an observational equivalence between an economy featuring such a friction and a behavioral variant featuring two distortions:

- myopia, or extra discounting of future outcomes; and
- anchoring of current outcomes to past outcomes, or backward-looking behavior.

We further show that these distortions are larger when the GE feedback is stronger, such as when the Keynesian cross is steeper. We use this result and certain extensions of it to recast the informational friction as a form of bounded rationality and to shed new light on its empirical implications.

Framework. We study a dynamic setting in which the optimal action (or best response) in each period depends positively on the expected discounted present values of an exogenous fundamental, denoted by ξ_t , and the average action, denoted by a_t . In the absence of the informational friction, this setting reduces to a representative-agent model, in which a_t obeys the following law of motion:

$$a_t = \varphi \xi_t + \delta \mathbb{E}_t [a_{t+1}], \quad (1)$$

where $\varphi > 0$, $\delta \in (0, 1]$, and $\mathbb{E}_t[\cdot]$ denotes the rational expectation of the representative agent.

Condition (1) nests the Euler condition of the representative consumer, which corresponds to aggregate demand, or the Dynamic IS curve, in the New Keynesian model, as well as the New Keynesian Philips Curve (NKPC), which describes aggregate supply. Alternatively, this condition can be read as an asset-pricing equation, with ξ_t standing for the asset's dividend and a_t for its price.

We depart from these familiar benchmarks by letting agents have imperfect and idiosyncratic knowledge, or “understanding,” of the underlying state of Nature and, by extension, the behavior of others. As in [Morris and Shin \(1998, 2002\)](#), [Woodford \(2003\)](#) and a large follow-up literature, this is accomplished by introducing incomplete information and high-order uncertainty.

The friction could be the product of dispersed private information or, as in the literature on rational inattention ([Sims, 2003](#); [Tirole, 2015](#)), a metaphor for cognitive constraints. What is key for our purposes is that the friction interferes not only with how well the agents tract the underlying fundamentals but also with how they form beliefs, or reason, about the behavior of others. Our contribution rests on making appropriate, simplifying assumptions so as to reveal as transparently as possible its empirical implications and the sense in which it offers a useful model of bounded rationality.

Main results. Under appropriate assumptions, we show that economy under consideration is observationally equivalent to a representative-agent economy in which condition (1) is modified, or distorted, as follows:

$$a_t = \varphi \xi_t + \delta \omega_f \mathbb{E}_t [a_{t+1}] + \omega_b a_{t-1}, \quad (2)$$

for some $\omega_f < 1$ and $\omega_b > 0$. The first distortion ($\omega_f < 1$) represents myopia towards the future, the second ($\omega_b > 0$) anchors current outcomes to past outcomes. The one dulls the forward-looking behavior, the other adds a backward-looking element akin to habit or adjustment costs.

We further show that both of the aforementioned distortions increase with the relative importance of GE effects (e.g., the slope of the Keynesian cross). As explained below, this offers useful insight into how the friction works, how it relates to other forms of bounded rationality such as Level-k Thinking, and how its empirical footprint may differ from that of habit and adjustment costs.

Underlying insights and marginal contribution. Our observational-equivalence result encapsulates two basic insights, which have previously appeared in the literature, albeit in different forms. The one insight is that, even in the absence of learning, higher-order uncertainty causes forward-looking agents to act as if they discount future GE considerations (Angeletos and Lian, 2018). The other insight is that learning adds momentum in beliefs, and more so in higher-order beliefs than in first-order beliefs (Woodford, 2003; Morris and Shin, 2006; Nimark, 2008).

Relative to these earlier works, the marginal *theoretical* contribution of our paper contains the following elements: (i) the extension of the first insight to a stationary environment with recurring shocks and learning; (ii) the proof of equilibrium existence and uniqueness; (iii) the analytical characterization of the equilibrium; and (iv) the observational-equivalence result.

The later in turn serves as an instrument for the paper's *applied* contribution, which itself consists of the following elements: (i) a bridge to the DSGE literature; (ii) the resolution of a certain disconnect between micro and macro; (iii) the dependence of the as-if distortions on GE mechanisms and policy; (iv) the recasting of the informational friction as a form of bounded rationality; and (v) an empirical application in the context of inflation. We expand on these applied aspects of our contribution below.

DSGE and macroeconomic dynamics. Sims (2003), Mankiw and Reis (2002), Woodford (2003), and a large follow-up literature have argued that informational frictions may substitute for the more dubious forms of sluggish adjustment employed in the DSGE literature. Our observational equivalence offers a particularly sharp illustration of this idea: the backward-looking element seen in condition (2) is akin to that introduced by habit persistence in consumption, by adjustment costs to investment, or by the so-called hybrid version of the NKPC. In addition, our result paves the way to an empirical evaluation of this idea, which is discussed further below.

Macro vs micro. Higher-order uncertainty and the associated imperfect reasoning about the behavior of others influences actual behavior only when agents respond to common shocks, not when

they respond to idiosyncratic shocks. As long as GE effects are strong enough, the documented distortions may therefore loom at the macroeconomic level even if they appear to be absent in the microeconomic level. This basic insight helps rationalize why the macroeconomic estimates of the habit persistence in consumption and the adjustment costs in investment are much larger than their microeconomic counterparts (Havranek, Rusnak, and Sokolova, 2017; Groth and Khan, 2010; Zorn, 2018), or perhaps even why the momentum in asset prices is more pronounced at the level of the entire stock market than at the level of individual stocks (Jung and Shiller, 2005).

GE and policy. As already noted, both the backward-looking element and the extra discounting of the future increase with the relative importance of the GE consideration. This underscores that a key feature of the accommodated friction is to distort how agents reason about the behavior of others and, thereby, about the GE consequences of any change in the environment. In effect, agents behave as if they fail to take into account the full GE interaction.

This in turn offers not only a bridge to bounded rationality, which we discuss next, but also a few practical insights into how macroeconomic dynamics may depend on market structures or policy. First, the effective myopia and habit in the Dynamic IS curve is shown to increase with the Keynesian income-spending multiplier, which, in richer models, may depend on the severity of liquidity constraints and the design of fiscal policy. Second, the observed inertia in inflation, or the effective past-price indexation parameter in the so-called hybrid version of the NKPC, is predicted to increase with the frequency of price adjustments, adding a new twist to the paradox of flexibility. And third, all these distortions are shown to decrease with the responsiveness of monetary policy.

Bounded rationality. How does the approach taken here compare to plausible alternatives that touch on similar issues by relaxing rational expectations? And how robust are our insights to such relaxations? We address this question by studying four variants of our framework.

The first variant replaces incomplete information with Level-k Thinking, as in Garcia-Schmidt and Woodford (2018) and Farhi and Werning (2017). This concept directly captures the idea that agents fail to take into account the full extent of their GE interaction. It therefore has the same flavor as our approach, but is observably distinct: it maps to $\omega_f < 1$ and $\omega_b = 0$. That is, it produces myopia but not anchoring or momentum.¹ The same is true of the second variant, which replaces incomplete information with the form of “cognitive discounting” assumed in Gabaix (2017): this, too, maps to $\omega_f < 1$ and $\omega_b = 0$. By contrast, our approach naturally produces both $\omega_f < 1$ and $\omega_b > 0$, offering a better account of both the macroeconomic time series and the available evidence on expectations.

The remaining two variants aim at blending incomplete information with belief mis-specification rather than pitting the one against the other. In one variant, agents fail to anticipate the fact that GE

¹There is additional, conceptual difference. By its very definition, Level-k Thinking is based on the premise that agents actively engage in higher-order reasoning, albeit up to a finite order. This is not true for our approach by virtue of the rational-expectations fixed point: agents need to be good statisticians but not good game theorists.

effects gain force as others adjust their beliefs and behavior with the passage of time. In another, agents have a mis-specified prior about the persistence of the fundamental and the precision of their own signals, making room for “overconfidence” and for the kind of over-reaction documented in [Bordalo et al. \(2018\)](#) and [Kohlhas and Walther \(2018a\)](#). Both of these variants preserve our observational-equivalence result, mapping to different combinations of $\omega_f < 1$ and $\omega_b > 0$. In so doing, they not only help illustrate the likely robustness of our insights to plausible, structured relaxations of rational expectations, but also offer guidance on how to preserve the tractability of our approach in richer settings and how to map the theory to the data.

Application to Inflation. Our contribution is completed with an empirical implementation in the context of inflation. Applying our observational-equivalence result to an incomplete-information extension of the New Keynesian Philips Curve (NKPC) yields a version of the Hybrid NKPC similar to that estimated in [Gali and Gertler \(1999\)](#) and [Gali, Gertler, and Lopez-Salido \(2005\)](#), but also ties the actual inflation dynamics to the inertia of the expectations of inflation. We are thus able to show that the theory rationalizes the existing estimates of the Hybrid NKPC, not just for *some* level of the assumed friction, but rather for a level that is broadly consistent with the expectations evidence documented in [Coibion and Gorodnichenko \(2015\)](#). This offers, not only a quantitative evaluation of the assumed friction in the specific context, but also an illustration of how our results facilitate a “sufficient statistics” approach to the mapping between the theory and the data.”²

Layout. The rest of the paper is organized as follows. Section 2 introduces our framework and explains the eclectic perspective we take on the assumed friction. Section 3 develops the observational-equivalence result. Section 4 expands on our result’s applicability and added value. Section 5 allows for mis-specification, or bounded rationality. Section 6 contains our empirical exercise in the context of inflation. Section 7 concludes. The Appendices contain proofs and a few additional results.

2 Framework

In this section we set up our framework and illustrate its applicability. We also motivate the approach we take in the rest of the paper by discussing the essence of the friction we are after and the complexity we aim at bypassing.

²This part of our contribution is closely related to [Nimark \(2008\)](#), which is the first to study an incomplete-information extension of the NKPC. Relative to that paper, the assumptions made here allow for a tighter connection between the theory and the existing estimates of the Hybrid NKPC. Most importantly, we map the empirical moment documented in [Coibion and Gorodnichenko \(2015\)](#) to the theoretical object of interest, namely the impulse response of inflation to the real marginal cost. This mapping is essential for the quantification of the impact of the friction on the actual inflation dynamics, but is absent from both of the aforementioned papers, as well as from other important contributions such as [Mankiw and Reis \(2002\)](#), [Reis \(2006\)](#), [Kiley \(2007\)](#), [Melosi \(2016\)](#), and [Matejka \(2016\)](#). A few other works evaluate or estimate the standard NKPC after replacing the representative agent’s expectation of inflation with the average forecast in surveys. This approach is inconsistent with the foundations laid out here. Still, our results allow for a simple mapping from the relevant survey evidence to the object of interest.

2.1 Set up

Time is discrete, indexed by $t \in \{0, 1, \dots\}$, and there is a continuum of players, indexed by $i \in [0, 1]$. In each period t , each agent i chooses an action $a_{it} \in \mathbb{R}$. We denote the corresponding average action by a_t . We next specify the best response of player i in period t as follows:

$$a_{it} = \mathbb{E}_{it} [\varphi \xi_t + \beta a_{it+1} + \gamma a_{t+1}] \quad (3)$$

where ξ_t is the underlying exogenous fundamental, $\mathbb{E}_{it}[\cdot]$ denotes the player's subjective expectation in period t , and (φ, β, γ) are parameters, with $\varphi > 0$, $\beta, \gamma \in [0, 1)$, and $\beta + \gamma < 1$.³

As evident in condition (3), γ regulates the dynamic strategic complementarity, or how much an agent has to reason about the future actions of others when deciding what to do today. In applications, this captures the relative importance of GE considerations. By contrast, β regulates the relative importance of PE considerations, or how much an agent has to reason about his own future behavior.

When all agents share the same expectations and this fact is common knowledge, the economy admits a representative agent. That is, $a_{i,t} = a_t$, $\mathbb{E}_{i,t} = \mathbb{E}_{t,t}$, and condition (3) reduces to the following:

$$a_t = \mathbb{E}_t[\varphi \xi_t + \delta a_{t+1}], \quad (4)$$

where $\delta \equiv \beta + \gamma$. This may correspond to the Euler equation of the representative consumer in the New Keynesian model or an elementary asset-pricing equation.

As anticipated in the Introduction, we depart from these familiar benchmarks by relaxing their common-knowledge foundations. The specifics of this departure are spelled out shortly. For now, let us only require that beliefs satisfy the Law of Iterated Expectations at the *individual* level: for all i , t and $\tau \geq t$, $\mathbb{E}_{it} [\mathbb{E}_{i\tau} [\cdot]] = \mathbb{E}_{it} [\cdot]$. That is, agents may question the knowledge or the rationality of others, but none expects *herself* to hold biased beliefs in the future.

Using this property and iterating on condition (3) yields the following extensive form:

$$a_{i,t} = \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t} [\varphi \xi_{t+k}] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t} [a_{t+k+1}]. \quad (5)$$

This condition makes clear that a player's optimal behavior at any given point of time depends on

³The best responses assumed here are the same as those in [Angeletos and Lian \(2018\)](#). But whereas that paper considers a non-stationary setting where ξ_t is fixed at zero in all $t \neq T$, for some fixed $T \geq 1$, we consider a stationary setting in which ξ_t varies in all t and, in addition, there is gradual learning over time. These features are essential for our observational-equivalence result and our applied contributions. Our framework also resembles the beauty contests considered by [Morris and Shin \(2002\)](#), [Woodford \(2003\)](#), [Angeletos and Pavan \(2007\)](#), [Bergemann and Morris \(2013\)](#), and [Huo and Pedroni \(2017\)](#). Because behavior is *not* forward-looking in these settings, the relevant higher-order beliefs are those regarding the *concurrent* beliefs of others. By contrast, the relevant higher-order beliefs in our setting are those regarding the *future* beliefs of others, as in [Allen, Morris, and Shin \(2006\)](#), [Morris and Shin \(2006\)](#) and [Nimark \(2008, 2017\)](#).

her subjective expectations of the entire future paths of the fundamental and of the average action. Aggregating it also yields the following equilibrium restriction:

$$a_t = \varphi \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t [\xi_{t+k}] + \gamma \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t [a_{t+k+1}], \quad (6)$$

where $\bar{\mathbb{E}}_t[\cdot]$ denotes the average expectation in the cross-section of the population.

Condition (6) stylizes the kind of fixed-point relation between the actual outcomes and expectations thereof that is embedded in the forward-looking equations of macroeconomic models. To help fix ideas, we next sketch how condition (6) nests certain incomplete-information extensions of the Dynamic IS curve and the New Keynesian Philips Curve. This also illustrates why, in applications, the dynamic strategic complementarity considered here corresponds to GE mechanisms such as the modern version of the Keynesian cross. After this illustration, we then discuss the essence of the friction we are after and the complexity we aim at bypassing.

2.2 Two Examples: Dynamic IS and NKPC

The New Keynesian model boils down to two forward-looking equations, the Dynamic IS curve and the New Keynesian Philips Curve (NKPC), along with a specification of monetary policy. The familiar, representative-agent versions of these equations are given by, respectively,

$$c_t = \mathbb{E}_t[-\varsigma r_t + c_{t+1}] \quad \text{and} \quad \pi_t = \mathbb{E}_t[\kappa \mathfrak{m}c_t + \beta \pi_{t+1}],$$

where c_t is aggregate consumption, r_t is the real interest rate, π_t is inflation, $\mathfrak{m}c_t$ is the real marginal cost, $\varsigma > 0$ is the elasticity of intertemporal substitution, $\kappa \equiv \frac{(1-\beta\theta)(1-\theta)}{\theta}$ is the slope of the Philips curve, $\theta \in (0, 1)$ is the Calvo parameter, $\beta \in (0, 1)$ is the subjective discount factor, and \mathbb{E}_t is the expectation of the representative agent.⁴ The first equation describes how aggregate spending responds to interest rates, the second how inflation responds to the real marginal cost or the output gap.

Relaxing the common-knowledge foundations of the New Keynesian model along the lines of [Angeletos and Lian \(2018\)](#) yields the following incomplete-information extensions of these equations:

$$c_t = -\varsigma \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t [r_{t+k}] + (1 - \beta) \sum_{k=1}^{\infty} \beta^{k-1} \bar{\mathbb{E}}_t [c_{t+k}], \quad (7)$$

$$\pi_t = \kappa \sum_{k=0}^{\infty} (\beta\theta)^k \bar{\mathbb{E}}_t [\mathfrak{m}c_{t+k}] + \beta(1 - \theta) \sum_{k=0}^{\infty} (\beta\theta)^k \bar{\mathbb{E}}_t [\pi_{t+k+1}], \quad (8)$$

⁴Whenever we refer to the New Keynesian model, we suppress the micro-foundations and work directly with the log-linearized equilibrium conditions. All variables are thus expressed as log-deviations of from the flexible-price steady state.

where $\bar{\mathbb{E}}_t$ denotes the average expectation of the consumers in (7) and that of the firms in (8). The first equation is nested in condition (6) by letting $a_t = c_t$, $\xi_t = r_t$, $\varphi = -\varsigma$, $\beta = \beta$, and $\gamma = 1 - \beta$; the second by letting $a_t = \pi_t$, $\xi_t = \mathfrak{m}c_t$, $\varphi = \kappa$, $\beta = \beta\theta$ and $\gamma = \beta(1 - \theta)$. The detailed derivations can be found in Appendix B. Here, we briefly discuss the economics behind conditions (7) and (8).

To understand condition (7), recall that the Permanent Income Hypothesis gives consumption as a function of the present discounted value of income. Extending this elementary result to accommodate variation in the real interest rate and heterogeneity in beliefs, and using the fact that aggregate income equals aggregate spending in equilibrium, yields condition (7). Finally, to understand why γ is given by $1 - \beta$, recall that in the textbook version of the Permanent Income Hypothesis the quantity $1 - \beta$ corresponds to the marginal propensity to consume out of the present discounted value of income. Since the latter captures the behavior of others, we have that the degree of strategic complementarity is herein related to how much consumption responds to income, or the slope of the Keynesian cross.⁵

To understand this condition, recall that a firm's optimal reset price is given by the discounted present value of the nominal marginal cost. Aggregating across firms and using the identity that ties inflation to the average rest price yields condition (8). When all firms share the same, rational expectations, this condition reduces to the familiar, textbook version of the NKPC. Away for that benchmark, condition (8) reveals the precise manner in which expectations of future inflation (the behavior of the firms) feed into current inflation. Note in particular that $\gamma = 1 - \theta$, which means that the effective degree of strategic complementarity increases with the frequency of price adjustment. This is because the feedback from the expectations of future inflation to current inflation increases when a higher fraction of firms are able to adjust their prices today on the basis of such expectations.

2.3 Higher-Order Beliefs: The Wanted Essence and the Unwanted Complexity

Consider condition (4), which stylizes the representative-agent or complete-information benchmarks we depart from. Iterating this forward yields the following solution for the equilibrium outcome:

$$a_t = \varphi \sum_{k=0}^{\infty} (\beta + \gamma)^k \mathbb{E}_t[\xi_{t+k}]. \quad (9)$$

where $\mathbb{E}_t[\cdot]$ stands for the representative agent's rational expectation. The following properties are then evident. First, the outcome is pinned down by the first-order beliefs of the fundamental. And second, the distinction between PE and GE effects, or the decomposition of the sum $\beta + \gamma$, is inconsequential.

⁵It is easy to extend condition (7) to a perpetual-youth, overlapping-generations model as in [Del Negro, Giannoni, and Patterson \(2015\)](#). Under appropriate assumptions, this boils down to replacing $\gamma = 1 - \beta$ with $\gamma = 1 - \beta\chi$, where χ is the survival probability. The latter can be a proxy for the length of planning horizons, either in the literal sense of expected lifespans or in the sense described in [Woodford \(2018\)](#). Alternatively, as in [Farhi and Werning \(2017\)](#), a lower χ , and hence a higher γ , may represent a higher probability of binding liquidity constraints. These observations further corroborate the interpretation of γ in this context as a proxy for the strength of the GE feedback between income and spending.

Once we relax the complete-information assumption, these properties cease to hold, even if we maintain rational expectations. To illustrate, consider two incomplete-information economies, called A and B . These economies share the same information structure and the same sum of PE and GE effects but a different composition of it: economy A features $\beta = \alpha$ and $\gamma = 0$, whereas economy B features $\beta = 0$ and $\gamma = \alpha$, for the same $\alpha \in (0, 1)$.

In economy A , condition (6) becomes $a_t = \varphi \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\xi_{t+k}]$, meaning that only the first-order beliefs of the fundamental matter, as in the representative-agent benchmark. In economy B , instead, condition (6) reduces to $a_t = \varphi \bar{\mathbb{E}}_t[\xi_t] + \gamma \bar{\mathbb{E}}_t[a_{t+1}]$. Recursive iteration of this condition allows the replacement of $\bar{\mathbb{E}}_t[a_{t+1}]$ with higher-order beliefs of the fundamental, leading to the following solution:

$$a_t = \varphi \sum_{h=0}^{\infty} \gamma^h \bar{\mathbb{F}}_t^{h+1}[\xi_{t+h}], \quad (10)$$

where, for any random variable X , $\bar{\mathbb{F}}_t^1[X] \equiv \bar{\mathbb{E}}_t[X]$ denotes the average first-order forecast of X and, for all $h \geq 2$, $\bar{\mathbb{F}}_t^h[X] \equiv \bar{\mathbb{E}}_t[\bar{\mathbb{F}}_{t+1}^{h-1}[X]]$ denotes the corresponding h -th order forecast.

This example illustrates the following point, which summarizes our eclectic perspective on the modeling role of higher-order uncertainty.

Fact. *Higher-order beliefs reflect GE considerations, or how the typical agent reasons about the behavior of others. Their deviation from first-order beliefs is the hallmark of the form of “imperfect reasoning” accommodated by incomplete information, and is the force that makes the distinction between PE and GE consequential in our setting.*

This point extends to the more general class of economies we study in this paper, in which both β and γ are positive. The only twist is that, in this case, the relevant set of higher-order beliefs is significantly richer than that seen in condition (10). Indeed, let $\zeta_t \equiv \sum_{\tau=0}^{\infty} \beta^\tau \xi_{t+\tau}$ and consider the following set of forward-looking, higher-order beliefs:

$$\bar{\mathbb{E}}_{t_1}[\bar{\mathbb{E}}_{t_2}[\cdots[\bar{\mathbb{E}}_{t_h}[\zeta_{t+k}]\cdots]],$$

for any $t \geq 0$, $k \geq 2$, $h \in \{2, \dots, k\}$, and $\{t_1, t_2, \dots, t_h\}$ such that $t = t_1 < t_2 < \dots < t_h = t + k$. It is easy to show that, when $\beta > 0$, the period- t outcome depends on all of these higher-order beliefs.

To further appreciate the added complexity relative to the $\beta = 0$ case, which is the case studied numerically in prior works such as [Nimark \(2008, 2017\)](#) and [Melosi \(2016\)](#), note that, for any t and any $k \geq 2$, there are now $k - 1$ types of second-order beliefs, plus $(k - 1) \times (k - 2)/2$ types of third-order beliefs, plus $(k - 1) \times (k - 2) \times (k - 3)/6$ types of fourth-order beliefs, and so on. For instance, when $k = 10$ (thinking about the outcome 10 periods later), there are 210 beliefs of the fourth order that are relevant when $\beta > 0$ compared to only one such belief when $\beta = 0$.

An integral part of our contribution is the bypassing of this immense complexity. The assumptions that permit this are spelled out in the beginning of the next section. This comes at the cost of generality, in particular the exclusion of endogenous information. The benefits will become evident as we develop our results. For now, we wish to emphasize the following, complementary point.

By bypassing the complexity of higher-order beliefs and characterizing directly the fixed point between actual and expected outcomes, the approach taken in the rest of our paper also epitomizes the following basic idea. Even though the analyst may find it useful to study the structure of higher-order beliefs as a means for a deeper understanding of the equilibrium and its empirical properties, the economic agents themselves do not need to engage in higher-order reasoning. They only need to have in mind the correct statistical model of the economy. This contrasts solution concepts that are defined on the premise the agents actively engage in higher-order reasoning, such as Level- k Thinking. In short, our approach requires agents to be good statisticians but not good game theorists.

3 The Equivalence Result

This section contains the core of our contribution. We introduce the assumptions that let us bypass the complexity of higher-order beliefs and solve directly the rational-expectations fixed point, proceed to develop our observation-equivalence result, and discuss the insights encapsulated in it.

3.1 Specification

We henceforth make two assumptions. First, we let the fundamental ξ_t follow an AR(1) process:

$$\xi_t = \rho\xi_{t-1} + \eta_t = \frac{1}{1 - \rho L}\eta_t, \quad (11)$$

where $\eta_t \sim \mathcal{N}(0, 1)$ is the period- t innovation, L is the lag operator, and $\rho \in (0, 1)$ parameterizes the persistence of the fundamental. Second, we assume that player i receives a new private signal in each period t , given by

$$x_{it} = \xi_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \sigma^2) \quad (12)$$

where $\sigma \geq 0$ parameterizes the informational friction (the level of noise). The player's information in period t is the history of signals up to that period.

As anticipated in the previous section, these assumptions aim at minimizing complexity without sacrificing essence. They are, of course, restrictive. In particular, they rule out the possibility that information is endogenous, either because agents extract information from market signals whose informativeness depends on the decisions of others, or because they choose what source of information to pay attention to. These issues are the subject of a large literature—but not of our paper.

Borrowing from the literature on rational inattention, we also invite a flexible interpretation of our setting as one where fundamentals and outcomes are observable but cognitive limitations makes agents act as *if* they observe the entire state of Nature with idiosyncratic noise.⁶ But instead of endogenizing the structure of this noise, we fix it in a way that best serves our purposes.

3.2 Solving the Rational-Expectations Fixed Point

Consider first the frictionless benchmark ($\sigma = 0$), in which case the outcome is pinned down by first-order beliefs, as in condition (9). Thanks to the AR(1) specification assumed above, $\mathbb{E}_t[\xi_{t+k}] = \rho^k \xi_t$, for all $t, k \geq 0$. We thus reach the following result, which states that the complete-information outcome follows the same AR(1) process as the fundamental, rescaled by the factor $\frac{\varphi}{1-\rho\delta}$.

Proposition 1. *In the frictionless benchmark ($\sigma = 0$), the equilibrium outcome is given by*

$$a_t = a_t^* \equiv \frac{\varphi}{1-\rho\delta} \xi_t = \frac{\varphi}{1-\rho\delta} \frac{1}{1-\rho L} \eta_t. \quad (13)$$

Consider next the case in which information is incomplete ($\sigma > 0$). As already explained, the outcome is then a function of an infinite number of higher-order beliefs. Despite the simplifying assumptions made here, the dynamic structure of these beliefs is quite complex. To illustrate, consider again the special case in which $\beta = 0$, which reduces the dimensionality of the relevant higher-order beliefs and gives the outcome as in condition (10). Using the Kalman filter, we can readily show that the first-order belief $\bar{\mathbb{E}}_t[\xi_t]$ follows an AR(2) process:

$$\bar{\mathbb{E}}_t[\xi_t] = \left(1 - \frac{\lambda}{\rho}\right) \left(\frac{1}{1-\lambda L}\right) \xi_t = \left(1 - \frac{\lambda}{\rho}\right) \left(\frac{1}{1-\lambda L}\right) \left(\frac{1}{1-\rho L}\right) \eta_t, \quad (14)$$

where $\lambda = \rho(1-G)$ and G is the Kalman gain. This implies that the second-order belief of the type $\bar{\mathbb{E}}_t[\bar{\mathbb{E}}_{t+1}[\xi_{t+1}]]$ follows an ARMA(3,1) and, by induction, for any $h \geq 1$, the h -th order belief of the type $\bar{\mathbb{E}}_t[\bar{\mathbb{E}}_{t+1}[\dots\bar{\mathbb{E}}_{t+h}[\xi_{t+h}]]]$ follows an ARMA($h+1, h-1$). In short, beliefs of higher order exhibit increasingly complex dynamics and the state space needed to track the entire belief hierarchy is infinite, even when $\beta = 0$.

Yet, as anticipated in the beginning of this section, this complexity is *not* inherited by the rational-expectations fixed point. The methods of [Huo and Takayama \(2018\)](#) guarantee that, insofar as the fundamental and the signals follow finite ARMA processes, the fixed point we are interested in is also a finite ARMA process. Under the assumptions made here, we can show that, regardless of β , γ and σ , the fixed point is merely an AR(2) process, whose exact form is characterized below.

⁶Indeed, note that the history of the outcome up to, and including, period t is measurable in $\xi^t \equiv (\xi_0, \dots, \xi_t)$, which defines the aggregate state of Nature in period. The corresponding Harsanyi type of agent i is $x_i^t \equiv (x_{i0}, \dots, x_{it})$. It follows that the assumed signals represent signals not only of the fundamental but also of the outcome.

Proposition 2. *The equilibrium exists, is unique and is such that the aggregate outcome obeys the following law of motion:*

$$a_t = \left(1 - \frac{\vartheta}{\rho}\right) \left(\frac{1}{1 - \vartheta L}\right) a_t^*, \quad (15)$$

where a_t^* is the frictionless counterpart, obtained in Proposition 1, and where ϑ is a scalar that satisfies $\vartheta \in (0, \rho)$ and that is given by the reciprocal of the largest root of the following cubic:

$$C(z) \equiv -z^3 + \left(\rho + \frac{1}{\rho} + \frac{1}{\rho\sigma^2} + \beta\right) z^2 - \left(1 + \beta \left(\rho + \frac{1}{\rho}\right) + \frac{\beta + \gamma}{\rho\sigma^2}\right) z + \beta,$$

Condition (15) expresses the incomplete-information dynamics as a simple transformation of the complete-information counterpart. This transformation is indexed by the scalar ϑ , which plays a dual role: relative to the frictionless benchmark (which is herein nested by $\vartheta = 0$), a higher ϑ means both a smaller impact effect, captured by the factor $1 - \frac{\vartheta}{\rho}$ in condition (15), and a more sluggish build up over time, captured by the lag term ϑL .

To understand the structure of the result, consider momentarily the special case in which $\gamma = 0$. By shutting down the strategic complementarity, this case isolates the role of first-order uncertainty. Using condition (6) along with $\gamma = 0$ (and hence $\delta \equiv \beta + \gamma = \beta$) and the fact that $\bar{\mathbb{E}}_t[\xi_{t+k}] = \rho^k \bar{\mathbb{E}}_t[\xi_t]$ for all $k \geq 0$, we infer that the aggregate outcome is given by

$$a_t = \varphi \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\xi_{t+k}] = \frac{\varphi}{1 - \delta\rho} \bar{\mathbb{E}}_t[\xi_t]. \quad (16)$$

This is the same as the complete-information outcome, modulo the replacement of ξ_t , the actual fundamental, with $\bar{\mathbb{E}}_t[\xi_t]$, the average first-order forecast of it. And since the latter follows the AR(2) process given in condition (14), we infer that Proposition 2 holds with $\vartheta = \lambda$ when $\gamma = 0$.

What happens when $\gamma > 0$? Higher-order beliefs then become relevant. As already noted, such beliefs follow ARMA processes of ever increasing order. And yet, the equilibrium continues to follow an AR(2) process, as in the case with $\gamma = 0$. The only twist is that the AR(2) process has different coefficients. In particular, ϑ is strictly higher than λ , which, as mentioned above, means that the equilibrium outcome exhibits less amplitude and more persistence than the first-order beliefs. This is the empirical footprint of the kind of imperfect GE reasoning accommodated in our analysis.

3.3 The Equivalence Result

Let us momentarily put aside the economy under consideration and, instead, consider a variant, representative-agent economy in which the aggregate Euler condition (4) is modified as follows:

$$a_t = \varphi \xi_t + \delta \omega_f \mathbb{E}_t [a_{t+1}] + \omega_b a_{t-1} \quad (17)$$

for some $\omega_f < 1$ and $\omega_b > 0$. The original representative-agent economy is nested with $\omega_f = 1$ and $\omega_b = 0$. Relative to this benchmark, a lower ω_f represents a higher discounting of the future, or less forward-looking behavior; a higher ω_b represents a greater anchoring of the current outcome to the past outcome, or more backward-looking behavior. We refer to this variant economy as the “behavioral economy.”

It is easy to verify that the equilibrium outcome of this economy is given by an AR(2) process, whose coefficients (ζ_0, ζ_1) are functions of (ω_f, ω_b) and (φ, δ, ρ) . In comparison, the equilibrium outcome in our incomplete-information economy is an AR(2) process with coefficients determined as in Proposition 2. Matching the coefficients of the two AR(2) processes, and characterizing the mapping from the latter to the former, we reach the following result.

Proposition 3 (Observational Equivalence). *Fix $(\varphi, \beta, \gamma, \rho)$. For any $\sigma > 0$ in the incomplete-information economy, there exists a unique pair (ω_f, ω_b) in the behavioral economy, with $\omega_f < 1$ and $\omega_b > 0$, such that the two economies generate the same joint dynamics for the fundamental and the aggregate outcome. Furthermore, a higher σ maps to a lower ω_f and a higher ω_b .*

This proposition, which is the main result of our paper, allows one to recast the informational friction as the combination of two behavioral distortions: extra discounting of the future, or myopia, in the form of $\omega_f < 1$; and backward-looking behavior, or anchoring of the current outcome to past outcome, in the form of $\omega_b > 0$.

As anticipated in the Introduction, this result embeds two broader insights. The one is that higher-order uncertainty induces agents to act as if they are myopic towards the future. The other is that learning induces momentum in higher-order beliefs and thereby in the outcome. The first insight manifests as $\omega_f < 1$, the second as $\omega_b > 0$.

[Angeletos and Lian \(2018\)](#) develop the first insight in a non-stationary economy in which higher-order uncertainty pertains only to the value of the fundamental in a single future period. Our result extends this insight to a stationary setting with recurring shocks and learning, thus also blending it with the second insight. [Woodford \(2003\)](#) highlight the second insight in a setting that, like ours, features recurring shocks and learning but, unlike ours, abstracts from forward-looking behavior and therefore does not yield the documented form of myopia. Furthermore, whereas [Woodford \(2003\)](#) and most of the follow-up works are based almost exclusively on numerical simulations, our analysis allows for an exact characterization of the equilibrium and of its comparative statics.

This in turn allows to complement our observational-equivalence result with the following one, which circles back to our earlier discussion of how the distinction between PE and GE channels

becomes consequential once the assumed form of “imperfect reasoning” is taken into account.

Proposition 4 (GE). *Fix the sum $\beta + \gamma$, or the sum of PE and GE effects. A stronger GE component (higher γ) maps to both greater myopia (lower ω_f) and greater anchoring (higher ω_b).*

We expand on the applicability and the usefulness of these results in the next few sections. We close this section with a comment on the robustness of our results.

3.4 Robustness

The AR(2) solution offered in Proposition 2 and, by extension, the exact equivalence obtained in Propositions 3 depend on the assumptions made about the process of ξ_t and the available signals. Without these or appropriate substitute assumptions, the elegance is lost. Nevertheless, the insights about myopia and anchoring survive more generally. This is illustrated in four ways.

- In Appendix ?, we show how to generalize these insights in a setting that allows for an arbitrary process for ξ_t and arbitrary learning dynamics, provided that we orthogonalize the information about the innovations occurring at different points of time.⁷
- In the next two sections we show how extensions of Propositions 3 can be obtained in settings featuring multiple forward-looking equations, a combination of private and public signals, and certain departures from rational expectations.
- In Section ? and Appendix ? we show how Proposition 3 may offer a good approximation of the more complicated dynamics obtained in extensions of our setting that accommodate endogenous signals.

These observations also explain why we prefer the perspective offered by Proposition 3 to the specific solution obtained in Proposition 2: the latter is fragile, the perspective is robust.

4 Applications and Extensions

In this section, we discuss the application of our results to the New Keynesian context, explain the resolution offered to a certain disconnect between micro and macro, and explore two extensions, one to multivariate settings and another to the combination of private and public signals.

⁷Although the proposed orthogonalization is somewhat exotic, it allows for an explicit characterization of the impulse response functions of the outcome to the innovations, adding further clarity to what’s going on behind our main result.

4.1 Connection to DSGE

In the end of Section 2 we sketched how our framework nests incomplete-information extensions of the Dynamic IS curve and the NKPC. We also discussed how γ relates to the slope of the Keynesian cross, or the income-spending multiplier, in the first context and to the frequency of price adjustment in the second. The following results are now immediate.

Corollary 1. *Applying our result to condition (7) yields the following incomplete-information version of the Dynamic IS curve:*

$$c_t = -\varsigma r_t + \omega_f \mathbb{E}_t[c_{t+1}] + \omega_b c_{t-1} \quad (18)$$

In this context, the distortions increase with the slope of the Keynesian cross.

Corollary 2. *Applying our result to condition (8) yields the following incomplete-information version of the the following version of the NKPC:*

$$\pi_t = \kappa m c_t + \omega_f \mathbb{E}_t[\pi_{t+1}] + \omega_b \pi_{t-1} \quad (19)$$

In this context, the distortions increase with the frequency of price adjustment.

Condition (18) looks like the Euler condition of representative consumer who exhibits myopia along with habit persistence. Condition (19) looks like the so-called Hybrid NKPC. Our results thus offer an informational micro-foundation of the more ad hoc backward-looking elements employed in the DSGE literature (Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007).

Of course, this idea is not entirely new: a large literature has already pushed the idea that informational frictions can be the source of sluggish adjustment.⁸ Our analysis adds to this line of work, not only by offering the sharpest, up to date, illustration of the aforementioned idea, but also by bringing in the following additional insights.

1. It highlights that the backward-looking elements featured in the DSGE literature may be endogenous, not only to the level of the informational friction, but also to GE mechanisms—and thereby also to market structures and policies that regulate the strength of such GE mechanisms. For instance, a fiscal-policy reform that alleviates liquidity constraints and reduces the income-spending multiplier may also reduce the as-if habit and myopia in the consumption dynamics.
2. It helps reduce a discomfoting gap between the macroeconomic and the microeconomics estimates of these elements, a point we discuss next.

⁸See, inter alia, Sims (2003); Woodford (2003); Mankiw and Reis (2002, 2007); Mackowiak and Wiederholt (2009, 2015); Nimark (2008).

3. It blends these backward-looking DSGE elements with a form of imperfect foresight, thus also building a bridge to the literature on bounded rationality discussed in the next section.
4. It facilitates the quantitative evaluation we conduct in Section 6.

4.2 Multivariate Systems [to be added soon]

4.3 Micro- vs Macro-level Distortions

As mentioned in the Introduction, the macroeconomic estimates of the habit in consumption and of the adjustment costs in investment are much larger than the corresponding microeconomic estimates; see [Havranek, Rusnak, and Sokolova \(2017\)](#) for a meta-analysis of multiple studies in the context of consumption habit, and [Groth and Khan \(2010\)](#) and [Zorn \(2018\)](#) for investment. Our results also offer a simple resolution to this disconnect.

To illustrate, consider the application to consumption studied above and allow for idiosyncratic income or preference shocks. Suppose further that each consumer has perfect knowledge of her idiosyncratic shocks, while maintaining the informational friction regarding the real interest rate (the aggregate fundamental) and aggregate spending (the aggregate outcome). In this context, Corollary ?? continues to hold: the dynamics of aggregate consumption exhibit habit-like behavior. At the same time, the response of individual consumption to idiosyncratic shocks exhibit no such behavior. It follows that an econometrician may estimate a positive habit at the macro level (i.e., in the response of aggregate outcomes to aggregate shocks) along with a zero habit at the micro level (i.e., in the response of individual outcomes to idiosyncratic shocks).

In the case just described, the absence of habit-like behavior at the micro level hinges on the assumption that agents observe perfectly their idiosyncratic shocks. Relaxing this assumption—for example, letting agents be rationally inattentive to both aggregate and idiosyncratic shocks—allows the micro responses to display a similar form of anchoring as the macro responses. Yet, the distortion is likely to remain more pronounced at the macro level than at the micro one for two reasons, the one highlighted in [Mackowiak and Wiederholt \(2009\)](#) and the one highlighted here.

Insofar as the friction is the product of costly information acquisition or rational inattention, it is natural to expect that the typical agent will collect relative more information about, or allocate relatively more cognitive capacity to, idiosyncratic shocks, simply because such shocks are more volatile and there is higher return in reducing uncertainty about them. This is the mechanism articulated in [Mackowiak and Wiederholt \(2009\)](#) and boils down to having less first-order uncertainty about idiosyncratic than aggregate shocks. But even if the first-order uncertainty about the two kind of shocks were the same, the distortion at the macro level would remain larger insofar as there are positive GE feedback effects, such as the Keynesian income-spending multiplier or the dynamic strategic com-

plementarity in price-setting decisions of the firms. In short, the mechanism identified in our paper and the one identified in the aforementioned work complement each other towards generating more pronounced distortions at the macro level than at the micro level.⁹

4.4 Inertia vs Excess Sensitivity [to be completed soon]

Business-cycle theory often tries to obtain both inertia vis-a-vis certain shocks and excess sensitivity vis-a-vis other shocks.

Take inflation as an example. On the one hand, it displays a muted and sluggish response to innovations in measured output gaps or real marginal costs. On the other hand, it displays significant and highly transient residual volatility. How can one make sense of these features?

The DSGE literature has tried to address the first issue by assuming that NKPC is very flat and by adding an ad-hoc backward looking component; and the second by introducing mysterious cost-push shocks. But if the NKPC is very flat, inflation should not move much with the cost-push shock either; and if the aforementioned backward-looking component is a true structural feature, it should induce inertial response to that shock too. In this sense, the modeling approach taken in the DSGE literature appears to fix the one problem only by exaggerating the other.

An extension of our main result offers an appealing alternative that avoids this trade off. The extension allows the agents to observe noisy public signals in addition to the noisy private signals considered in the main analysis. Such public signals can be interpreted either literally or as a modeling device for introducing correlated correlated mistakes in the agents' reasoning about the behavior of others.

By its very nature, such noise has an immediate and transient effect on beliefs and the equilibrium outcome. It is therefore as if the myopia and anchoring apply *only* to the response of the economy to the innovations in the fundamental. What is more, as the GE feedback increases, the impact of noise also increases, while the response to the fundamental gets more muted. In this sense, "excess sensitivity" may co-exist with "inertia."

5 Bounded Rationality [to be updated soon]

As already noted, the idea that incomplete information can rationalize a certain kind of myopia was first put forward in [Angeletos and Lian \(2018\)](#). But whereas that paper focused on a non-stationary environment featuring a single, once-and-for-all anticipated change in the value of the fundamental at some predetermined future date (a particular type of "MIT shock"), our analysis considers a stationary

⁹We verify all these intuitions in Appendix D with a variant that lets both aggregate and idiosyncratic shocks be observed with noise. The quantitative potential of this particular idea, however, is left open for future research.

setting with recurring shocks. This step is crucial for the development and the applicability of our observational-equivalence result. Furthermore, by accommodating learning over time, our analysis blends the myopia with the backward-looking element sought after by the DSGE literature.

The last point also helps distinguish our contribution from those of [Gabaix \(2017\)](#) and [Farhi and Werning \(2017\)](#). These works depart from rational expectations in a manner that helps capture a similar form of imperfect foresight as ours. The former achieves this by assuming that the perceived law of motion of all the relevant economic variables exhibit less amplitude and less persistence than the true one (an assumption called “cognitive discounting”), the latter by letting agents have limited depth of reasoning in the sense of Level-k Thinking.¹⁰ These works do not, however, provide a theory of momentum in beliefs and behavior. In terms of our observational-equivalence result, they accommodate $\omega_f < 1$ but restrict $\omega_b = 0$. It follows that an elementary testable difference between incomplete information and these alternatives is whether ω_b is positive or zero.

From this perspective, the approach taken here seems to be empirically superior. First, the macroeconomic data demands $\omega_b > 0$, which is precisely the reason why the DSGE literature departed from baseline, forward-looking macroeconomic models by adding habit persistence in consumption, adjustment costs to investment, etc. Second, and perhaps more tellingly, the available evidence on expectations also demands $\omega_b > 0$: as shown in [Coibion and Gorodnichenko \(2012, 2015\)](#), [Coibion, Gorodnichenko, and Kumar \(2015\)](#) and [Vellekoop and Wiederholt \(2017\)](#), the average forecast errors of both professional forecasters and firm managers exhibit positive serial autocorrelation, in line with the learning dynamics induced by incomplete information.

It may be possible to reconcile these facts with the aforementioned forms of bounded rationality by letting their otherwise arbitrary “default point” (e.g., the level-0 behavior) be an increasing function of the past outcome; but this begs the question of why this is true. Alternatively, one may try to augment them with some kind of non-Bayesian learning; but it is unclear at this point how this can be done.¹¹ By contrast, the approach taken here and in the related literature on informational frictions readily captures the relevant facts. What is more, the exercise conducted in the next section suggests that this success is, not only in qualitative terms, but also in quantitative terms.

Finally, [Coibion et al. \(2018\)](#) provides additional, and more direct, support for the approach taken here. The authors conduct an innovative survey where firms are asked, not only their own beliefs about inflation, but also their beliefs about the beliefs of other firms. In addition, the participants were asked to play a beauty-contest game like that considered in the experimental literature on Level-k

¹⁰This follows the lead of [Garcia-Schmidt and Woodford \(2018\)](#), whose solution concept (“reflective equilibrium”) is essentially the same as Level-k Thinking. See also [Iovino and Sergeyev \(2017\)](#) for another, topical application.

¹¹For instance, the experimental literature has allowed the depth of reasoning to increase with the rounds of the experiment in order to capture the gradual adjustment in beliefs and actions, and has interpreted this as learning how to play the rational expectations equilibrium. This may make sense when the agents face a completely new situation (as in experiments) but is not directly amendable to the kind of stationary settings we are interested in.

Thinking. The solicited beliefs display properties consistent with incomplete information and Bayesian learning. By contrast, they appear to be unrelated to the depth of thinking (k level) inferred by the participants' play at the beauty-contest game.

Notwithstanding these points, we view our approach and the aforementioned behavioral alternatives as close cousins: they represent plausible, and related, departures from the full-information rational-expectations benchmark. We also think that a fruitful direction for future research is one that combines incomplete information with bounded rationality. We explore two such extensions in the end of Section 7. The one merges our approach with Level- k Thinking.¹² The other relaxes the assumption that agents can perfectly anticipate that others will learn in the future. Both of these extensions preserve the essence of our results, but also intensify the documented myopia. Another interesting direction for future research may be one that augments our work with the kind of belief extrapolation studied in [Bordalo et al. \(2018\)](#) and [Kohlhas and Walther \(2018a,b\)](#).

6 Application to the NKPC

As already noted, the application of our observational-equivalence result to the incomplete-information extension of the NKPC seen in condition () yields the following hybrid version of the NKPC:

$$\pi_t = \kappa \mathfrak{m}c_t + \omega_f \beta \mathbb{E}_t[\pi_{t+1}] + \omega_b \pi_{t-1} \quad (20)$$

where π_t denotes inflation, $\mathfrak{m}c_t$ denotes the real marginal cost, $\kappa > 0$ is the slope of the standard NKPC, $\beta \in (0, 1)$ is the subjective discount factors, and $\omega_f < 0$ and $\omega_b > 0$ are functions of σ , the level of noise, and ϑ , the Calvo parameter. In this section, we use this result to show the theory can match *jointly* existing estimates of the Hybrid NKPC ([Gali and Gertler, 1999](#); [Gali, Gertler, and Lopez-Salido, 2005](#)) and independent evidence on inflation expectations ([Coibion and Gorodnichenko, 2015](#)). This offers a quantitative evaluation of the assumed friction in the particular context, and more broadly an illustration of how our approach facilitates a clean mapping between theory and data.

Matching Estimates of the Hybrid NKPC. The Hybrid NKPC estimated in [Gali and Gertler \(1999\)](#) and [Gali, Gertler, and Lopez-Salido \(2005\)](#), is similar to the one seen in (35). There is, however, a difference. An unrestricted version of the Hybrid NKPC, such as that estimated in those papers, allows ω_f and ω_b to be free. By contrast, even when if we can choose the level of noise at will, these coefficients are restricted in our setting: a higher ω_b can be obtained only if the noise is larger, which in turns requires ω_f to be smaller. This restriction is described below.

Proposition 5. *The dynamics of inflation produced by the Hybrid NKPC can be replicated by an*

¹²We thank Alexander Kohlhas for suggesting us to explore such a variant.

incomplete-information extension of the standard NKPC for some $\sigma > 0$ if and only if $\omega_b > 0$ and

$$\omega_f = 1 - \frac{1}{\beta\rho^2}\omega_b. \quad (21)$$

where β is the subjective discount factor and ρ is the persistence of the real marginal cost. Furthermore, for any pair (ω_b, ω_f) that satisfies the above restriction and any value θ of the Calvo parameter, there exists a unique level of noise σ such that the two economies are observationally equivalent.

A quick test of our theory is therefore whether the existing estimates of the Hybrid NKPC happen to satisfy the restriction in Proposition 5. This proposition gives the locus of the pairs (ω_f, ω_b) that are compatible with our theory for some level of noise. To construct this locus, and to be able to identify σ by inverting the provided estimates of (ω_f, ω_b) , we need to specify δ , θ , and ρ . We set $\delta = 0.99$, $\theta = 0.6$, and $\rho = 0.95$. The value of θ corresponds to a modest degree of price stickiness, broadly in line with textbook calibrations of the New Keynesian model and with the micro data. The value of ρ is obtained by estimating an AR(1) process on the labor share, a standard empirical proxy for the real marginal cost. The locus implied under this parameterization of our model is then represented by the solid red line in Figure 1.

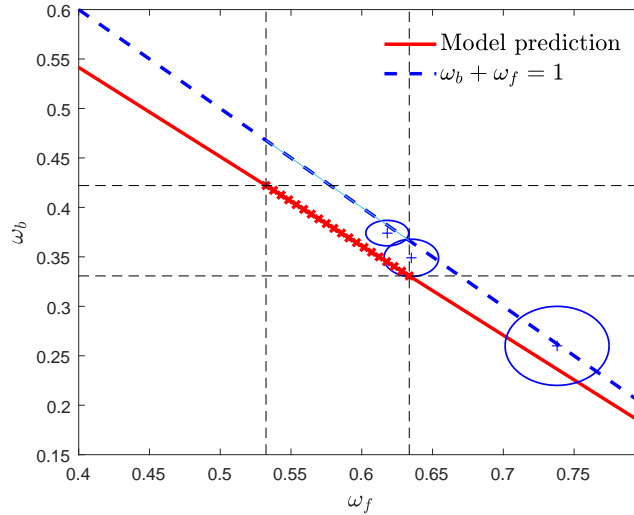


Figure 1: Testing the Theory

Gali, Gertler, and Lopez-Salido (2005) provide three baseline estimates of (ω_f, ω_b) . These estimates and their confidence regions are represented by the blue crosses and the surrounding disks in Figure 1. A priori, there is no reason to expect that the estimates obtained in Gali, Gertler, and Lopez-Salido (2005) should fall on, or close to, the locus implied by our theory. And yet, as evident in the figure, that's the case. In other words, our model matches the existing estimates on the Hybrid NKPC and allows one to rationalize them as the product of informational frictions.

Mavroeidis, Plagborg-Møller, and Stock (2014) review the extensive literature on the empirical literature of the NKPC and questions the robustness of the estimates provided by Gali, Gertler, and Lopez-Salido (2005). This debate is beyond the scope of our paper. In any event, the exercise conducted next bypasses the estimation of the Hybrid NKPC on macroeconomic data and instead infers it by calibrating our theory to survey data on expectations.

Matching Survey Evidence on Informational Frictions. Although the theory passes the test of matching existing estimates of the Hybrid NKPC, it is not clear at this point whether this success hinges on an empirically implausible magnitude for the informational friction. We now address this question, and impose the theory to an additional test, by examining whether the informational friction required in order to rationalize the existing estimates of ω_f and ω_b is consistent with survey evidence on expectations.

To this goal, we utilize the findings of Coibion and Gorodnichenko (2015). That paper uses data on inflation forecasts from the Survey of Professional Forecasters to measure a key moment that can help gauge the magnitude of the informational friction. The basic idea is that the friction should manifest itself in the predictability of the average forecast errors. In particular, Coibion and Gorodnichenko (2015) run the following regression:

$$\pi_{t+k} - \bar{\mathbb{E}}_t[\pi_{t+k}] = K (\bar{\mathbb{E}}_t[\pi_{t+k}] - \bar{\mathbb{E}}_{t-1}[\pi_{t+k}]) + v_{t+k,t} \quad (22)$$

With complete information, K is zero, because the current forecast correction is independent of past information. By contrast, when information is incomplete, average forecasts adjust sluggishly towards the truth, implying that past innovations in forecasts predict future forecast corrections, that is, $K > 0$. Furthermore, K is larger the larger the noise and the slower the speed of learning.

Coibion and Gorodnichenko (2015) illustrate this logic with an example in which inflation follows an exogenous AR(1) process and K is a direct transformation of the level of noise. Clearly, this not the case here. Because actual inflation and forecasts of inflation are jointly determined in equilibrium, the regression coefficient K implied by our theory is more complicated than that in their example and is indeed endogenous to the GE interaction among the firms. Nevertheless, we can use the theory to characterize K as a function of σ and of (δ, θ, ρ) . With the latter fixed in the way described earlier, this gives us a mapping from the 90% confidence interval of K provided in Coibion and Gorodnichenko (2015) to an interval for σ in our model. That is, we have a confidence interval for the informational friction itself. For any σ in this interval, we can then compute the pair (ω_f, ω_b) predicted by our theory.

We can thus map the evidence reported in Coibion and Gorodnichenko (2015) to a segment of the (ω_f, ω_b) locus we obtained earlier on. This segment is identified by the red crosses in Figure 1 and gives the pairs of (ω_f, ω_b) that are consistent with the confidence interval for K provided in Coibion

and [Gorodnichenko \(2015\)](#). It is then evident from the figure that our model can pass jointly the test of matching that evidence and the test of matching the existing estimates of the Hybrid NKPC.¹³

One may push further the empirical evaluation of the theory by testing its predictions against data on higher-order beliefs. We are not aware of any such data in the US context studied above. However, the evidence provided recently by [Coibion et al. \(2018\)](#) seems reassuring: in a survey of firms in New Zealand, higher-order expectations of inflation display patterns consistent with those at the core of incomplete-information models.

Quantitative Bite. The quantitative implications of the exercise conducted above are further illustrated in Figure 2. This figure compares the impulse response function of inflation under three scenarios. The solid black line corresponds to frictionless benchmark, with perfect information. The dashed blue line corresponds to the frictional case, with an informational friction that matches the baseline estimation of [Coibion and Gorodnichenko \(2015\)](#). The dotted red line is explained later.

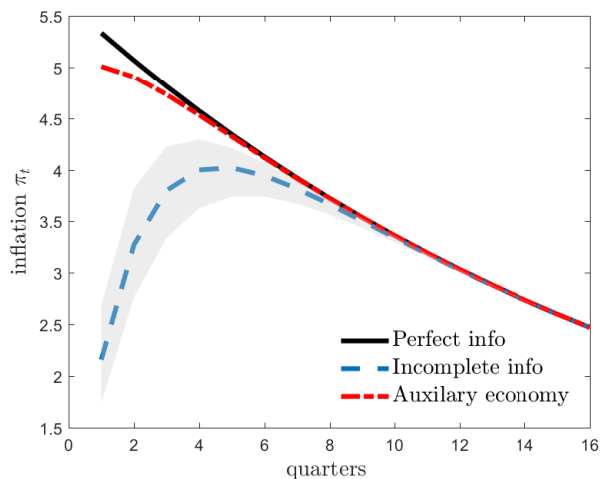


Figure 2: Impulse Response Function of Inflation

As evident by comparing the dashed blue line to the solid black one, the quantitative bite of the informational friction is significant: the impact effect on inflation is about 60% lower than its complete-information counterpart, and the peak of the inflation response is attained 5 quarters after impact rather than on impact. This is suggestive of how informational frictions may help reconcile quantitative macroeconomic models, which can account for the business cycle only by assuming significant sluggishness in the inflation dynamics, with realistic menu-cost models, which appear to be unable to produce such sluggishness.¹⁴

¹³The statement is true for two of the three estimates provided in [Gali, Gertler, and Lopez-Salido \(2005\)](#). These happen to be, not only those that our theory rationalizes, but also those that the authors prefer for other, econometric reasons.

¹⁴See, for example, [Golosov and Lucas Jr \(2007\)](#), [Midrigan \(2011\)](#), [Alvarez and Lippi \(2014\)](#), and [Nakamura and Steinsson](#)

Let us now explain the dotted red line in the figure. Using condition (33), the incomplete-information inflation dynamics can be decomposed into two components: the belief of the present discounted value of real marginal costs, $\varphi \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\mathbf{mc}_{t+k}]$; and the belief of the present discounted value of inflation, $\gamma \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\pi_{t+k+1}]$. The same decomposition can also be applied when agents have perfect information:

$$\pi_t^* = \varphi \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t[\mathbf{mc}_{t+k} | \mathbf{mc}_t] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t[\pi_{t+k+1}^* | \mathbf{mc}_t], \quad (23)$$

A natural question is which component contributes more to the anchoring of inflation as we move from the complete to incomplete information.

To answer this question, we define the following auxiliary variable:

$$\tilde{\pi}_t = \varphi \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t[\mathbf{mc}_{t+k}] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t[\pi_{t+k+1}^* | \mathbf{mc}_t]. \quad (24)$$

The difference between π_t^* and $\tilde{\pi}_t$ measures the importance of beliefs about real marginal costs, and the difference between $\tilde{\pi}_t$ and π_t measures the importance of beliefs about inflation. The dotted red line in Figure 2 corresponds to $\tilde{\pi}_t$. Clearly, most of the difference between complete and incomplete information is due the anchoring of beliefs about future inflation. These beliefs are tied with higher-order beliefs, which display less responsiveness and more inertia than the first-order beliefs.

7 Conclusion

We showed how the accommodation of incomplete information, higher-order uncertainty and learning in forward-looking, general-equilibrium models is akin to the introduction of two behavioral distortions: myopia, or extra discounting of the future; and backward-looking behavior, or anchoring of current outcomes to past outcomes. We formalized this perspective with the help of an observational-equivalence result, which rested on strong assumptions, but also elaborated on the robustness of the underlying insights and the offered perspective.

Our observational-equivalence result was instrumental, not only for the formalization of the above perspective, but also for the following additional, applied purposes:

1. It illustrated how incomplete information can, not only offer a substitute for the more ad hoc backward-looking features of the DSGE literature, but also help resolve the gap between the

(2013). Although different specifications can rationalize a degree of price rigidity either much smaller than or almost as large as the one predicted by the standard NKPC, this literature has not produced the kind of hump-shaped inflation dynamics that the DSGE literature has captured with the Hybrid NKPC.

macroeconomic and microeconomic estimates of such features.

2. It blend these backward-looking features with a form of imperfect foresight.
3. It highlighted how both of these elements may be endogenous to GE mechanisms and thereby also to market structures and policies that regulate them.
4. It facilitated the empirical evaluation of our theory in the context of inflation dynamics.
5. It let us relate our approach and the existing literature on informational frictions to an emerging literature on bounded rationality.

Although our paper was focused on macroeconomic applications, our results and insights are relevant more broadly. We conclude the paper by illustrating this in the context of asset pricing.

In Appendix C, we take a setting with overlapping generations of traders and dispersed private information, as in [Singleton \(1987\)](#). Under appropriate assumptions, this setting gives rise to the following equilibrium asset-pricing condition:

$$p_t = \mathbb{E}_t[d_{t+1}] + \omega_f \delta \mathbb{E}_t[p_{t+1}] + \omega_b p_{t-1},$$

where p_t is the asset's price, d_t is its dividend, $\mathbb{E}_t[\cdot]$ is the full-information, rational-expectations operator, δ is the standard discount factor, and (ω_f, ω_b) are our familiar coefficients.

The above offers a sharp illustration of how incomplete information can rationalize momentum and predictability in asset prices ($\omega_b > 0$), as well as excessive discounting of news about future fundamentals ($\omega_f < 1$). The latter in turn hints to a possible fragility of the predictions of the literature that emphasizes long-term risks to the accommodation of higher-order uncertainty or related forms of bounded rationality. Finally, our insight about the dependence of the as-if distortions on strategic complementarity and GE feedbacks adds a new angle to the Samuelson dictum ([Jung and Shiller, 2005](#)): to the extent that the positive feedback in asset trading is stronger at the stock-market level than at the individual-stock level because of fire-sale externalities and liquidity black holes, our results may help explain why asset-price anomalies are more pronounced at the former level than at the latter.

References

- Abel, Andrew B. and Olivier J. Blanchard. 1983. "An Intertemporal Model of Saving and Investment." *Econometrica* 51 (3):675–692.
- 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.
- Alvarez, Fernando and Francesco Lippi. 2014. "Price Setting with Menu Cost for Multiproduct Firms." *Econometrica* 82 (1):89–135.
- Angeletos, George-Marios and Chen Lian. 2018. "Forward Guidance without Common Knowledge." *American Economic Review* 108 (9):2477–2512.
- Angeletos, George-Marios and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4):1103–1142.
- Bachmann, Rüdiger, Ricardo J. Caballero, and Eduardo M. R. A. Engel. 2013. "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model." *American Economic Journal: Macroeconomics* 5 (4):29–67.
- Bergemann, Dirk and Stephen Morris. 2013. "Robust Predictions in Games with Incomplete Information." *Econometrica* 81 (4):1251–1308.
- Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. "Over-reaction in Macroeconomic Expectations." *NBER Working Paper No. 24932* .
- Caballero, Ricardo J. and Eduardo M. R. A. Engel. 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach." *Econometrica* 67 (4):783–826.
- Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1):1–45.
- Coibion, Olivier and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1):116–159.
- . 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–78.
- Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2015. "How Do Firms Form Their Expectations? New Survey Evidence." *NBER Working Paper Series* .

- Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert. 2018. "Do You Know That I Know That You Know...? Higher-Order Beliefs in Survey Data." *NBER Working Paper No. 24987* .
- Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." *FRB of New York mimeo* .
- Farhi, Emmanuel and Iván Werning. 2017. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *NBER Working Paper No. 23281* .
- Gabaix, Xavier. 2017. "A Behavioral New Keynesian Model." *Harvard mimeo* .
- Galí, Jordi. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition*. Princeton University Press.
- Gali, Jordi and Mark Gertler. 1999. "Inflation dynamics: A structural econometric analysis." *Journal of Monetary Economics* 44 (2):195–222.
- Gali, Jordi, Mark Gertler, and J David Lopez-Salido. 2005. "Robustness of the estimates of the hybrid New Keynesian Phillips curve." *Journal of Monetary Economics* 52 (6):1107–1118.
- Garcia-Schmidt, Mariana and Michael Woodford. 2018. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review* forthcoming.
- Golosov, Mikhail and Robert E Lucas Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2):171–199.
- Groth, Charlota and Hashmat Khan. 2010. "Investment Adjustment Costs: An Empirical Assessment." *Journal of Money, Credit and Banking* 42 (8):1469–1494.
- Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit Formation in Consumption: A Meta-Analysis." *European Economic Review* 95:142–167.
- Hayashi, Fumio. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation." *Econometrica* 50 (1):213–224.
- Huo, Zen and Naoki Takayama. 2018. "Rational Expectations Models with Higher Order Beliefs." *miméo, Yale University*.
- Huo, Zhen and Marcelo Zouain Pedroni. 2017. "Infinite Higher Order Beliefs and First Order Beliefs: An Equivalence Result Infinite Higher Order Beliefs and First Order Beliefs: An Equivalence Result." *miméo, Yale University/University of Amsterdam*.

- lovino, Luigi and Dmitriy Sergeyev. 2017. "Quantitative Easing without Rational Expectations." *Work in progress* .
- Jung, Jeeman and Robert J. Shiller. 2005. "Samuelson's Dictum and the Stock Market." *Economic Inquiry* 43 (2):221–228.
- Kasa, Kenneth, Todd B. Walker, and Charles H. Whiteman. 2014. "Heterogeneous Beliefs and Tests of Present Value Models." *The Review of Economic Studies* 81 (3):1137–1163.
- Khan, Aubhik and Julia K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76 (2):395–436.
- Kiley, Michael T. 2007. "A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting." *Journal of Money, Credit and Banking* 39 (s1):101–125.
- Kohlhas, Alexandre and Ansgar Walther. 2018a. "Asymmetric Attention." *IIES miméo* .
- . 2018b. "Sparse Expectations: A Unified Explanation of Forecast Data." *IIES miméo* .
- Mackowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803.
- . 2015. "Business Cycle Dynamics under Rational Inattention." *Review of Economic Studies* 82 (4):1502–1532.
- Mankiw, N. Gregory and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4):1295–1328.
- Mankiw, N Gregory and Ricardo Reis. 2007. "Sticky information in general equilibrium." *Journal of the European Economic Association* 5 (2-3):603–613.
- Matejka, Filip. 2016. "Rationally Inattentive Seller: Sales and Discrete Pricing." *Review of Economic Studies* 83 (3):1125–1155.
- Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock. 2014. "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve." *Journal of Economic Literature* 52 (1):124–88.
- Melosi, Leonardo. 2016. "Signalling effects of monetary policy." *The Review of Economic Studies* 84 (2):853–884.
- Midrigan, Virgiliu. 2011. "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations." *Econometrica* 79 (4):1139–1180.

- Morris, Stephen and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-fulfilling Currency Attacks." *American Economic Review* :587–597.
- . 2002. "Social Value of Public Information." *American Economic Review* 92 (5):1521–1534.
- . 2006. "Inertia of Forward-looking Expectations." *The American Economic Review* :152–157.
- Nakamura, Emi and Jón Steinsson. 2013. "Price Rigidity: Microeconomic Evidence and Macroeconomic Implications." *Annual Review of Economics* 5 (1):133–163.
- Nimark, Kristoffer. 2008. "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics* 55 (2):365–382.
- . 2017. "Dynamic Higher Order Expectations." *Cornell University mimeo* .
- Reis, Ricardo. 2006. "Inattentive producers." *The Review of Economic Studies* 73 (3):793–821.
- 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*, edited by William A. Barnett and Kenneth J. Singleton. Cambridge University Press, 249–272.
- Smets, Frank and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3):586–606.
- Tirole, Jean. 2015. "Cognitive Games and Cognitive Traps." *Toulouse School of Economics mimeo* .
- Vellekoop, Nathanael and Mirko Wiederholt. 2017. "Inflation Expectations and Choices of Households." *Goethe University Frankfurt mimeo* .
- Vives, Xavier and Liyan Yang. 2017. "Costly Interpretation of Asset Prices." *mimeo, IESE/University of Toronto*.
- Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* .
- . 2018. "Monetary Policy Analysis When Planning Horizons Are Finite." In *NBER Macroeconomics Annual 2018, Volume 33*. University of Chicago Press.
- Zorn, Peter. 2018. "Investment under Rational Inattention: Evidence from US Sectoral Data." *mimeo, University of Munich*.

Appendix A: Proofs

Proof of Proposition 1. Follows directly from the analysis in the main text.

Proof of Proposition 2. Suppose that the agent's equilibrium policy function is given by

$$a_{it} = h(L)x_{it}$$

for some lag polynomial $h(L)$. The aggregate outcome can then be expressed as follows:

$$a_t = h(L)\xi_t = \frac{h(L)}{1 - \rho L}\eta_t.$$

In the sequel, we verify that the above guess is correct and characterize $h(L)$.

First, we look for the fundamental representation of the signals. Define $\tau_\eta = \sigma_\eta^{-2}$ and $\tau_u = \sigma^{-2}$ as the reciprocals of the variances of, respectively, the innovation in the fundamental and the noise in the signal. (In the main text, we have normalized $\sigma_\eta = 1$.) The signal process can be rewritten as

$$x_{it} = \mathbf{M}(L) \begin{bmatrix} \hat{\eta}_t \\ \hat{u}_{it} \end{bmatrix}, \quad \text{with} \quad \mathbf{M}(L) = \begin{bmatrix} \tau_\eta^{-\frac{1}{2}} & 1 \\ \frac{1}{1-\rho L} & \tau_u^{-\frac{1}{2}} \end{bmatrix}.$$

Let $B(L)$ denote the fundamental representation of the signal process. By definition, $B(L)$ needs to be an invertible process and it needs to satisfy the following requirement

$$B(L)B(L^{-1}) = \mathbf{M}(L)\mathbf{M}'(L^{-1}) = \frac{\tau_\eta^{-1} + \tau_u^{-1}(1 - \rho L)(L - \rho)}{(1 - \rho L)(L - \rho)}, \quad (25)$$

which leads to

$$B(L) = \tau_u^{-\frac{1}{2}} \sqrt{\frac{\rho(1 - \lambda L)}{\lambda(1 - \rho L)}},$$

where λ is the inside root of the numerator in equation (25)

$$\lambda = \frac{1}{2} \left[\rho + \frac{1}{\rho} \left(1 + \frac{\tau_u}{\tau_\eta} \right) - \sqrt{\left(\rho + \frac{1}{\rho} \left(1 + \frac{\tau_u}{\tau_\eta} \right) \right)^2 - 4} \right].$$

Next, we characterize the beliefs of ξ_t , $a_{i,t+1}$, and a_{t+1} , that is, the beliefs that show up in the best-response condition of the agent. The forecast of a random variable

$$f_t = \mathbf{A}(L) \begin{bmatrix} \hat{\eta}_t \\ \hat{u}_{it} \end{bmatrix}$$

can be obtained by using the Wiener-Hopf prediction formula:

$$\mathbb{E}_{it}[f_t] = [\mathbf{A}(L)\mathbf{M}'(L^{-1})B(L^{-1})^{-1}]_+ B(L)^{-1}x_{it}.$$

Consider the forecast of the fundamental. Note that

$$\xi_t = \begin{bmatrix} \tau_\eta^{-\frac{1}{2}} \frac{1}{1-\rho L} & 0 \end{bmatrix} \begin{bmatrix} \hat{\eta}_t \\ \hat{u}_{it} \end{bmatrix},$$

from which it follows that

$$\mathbb{E}_{it}[\xi_t] = G_1(L)x_{it}, \quad G_1(L) \equiv \frac{\lambda \tau_u}{\rho \tau_\eta} \frac{1}{1-\rho\lambda} \frac{1}{1-\lambda L}.$$

Consider the forecast of the future own and average actions. Using the guess that $a_{it+1} = h(L)x_{i,t+1}$ and $a_{t+1} = h(L)\xi_{t+1}$, we have

$$a_{t+1} = \begin{bmatrix} \tau_\eta^{-\frac{1}{2}} \frac{h(L)}{L(1-\rho L)} & 0 \end{bmatrix} \begin{bmatrix} \hat{\eta}_t \\ \hat{u}_{it} \end{bmatrix}, \quad a_{it+1} - a_{t+1} = \begin{bmatrix} 0 & \tau_u^{-\frac{1}{2}} h(L) \end{bmatrix} \begin{bmatrix} \hat{\eta}_t \\ \hat{u}_{it} \end{bmatrix},$$

and the forecasts are

$$\mathbb{E}_{it}[a_{t+1}] = G_2(L)x_{it}, \quad G_2(L) \equiv \frac{\lambda \tau_u}{\rho \tau_\eta} \left(\frac{h(L)}{(1-\lambda L)(L-\lambda)} - \frac{h(\lambda)(1-\rho L)}{(1-\rho\lambda)(L-\lambda)(1-\lambda L)} \right),$$

$$\mathbb{E}_{it}[a_{it+1} - a_{t+1}] = G_3(L)x_{it}, \quad G_3(L) \equiv \frac{\lambda}{\rho} \left(\frac{h(L)(L-\rho)}{L(L-\lambda)} - \frac{h(\lambda)(\lambda-\rho)}{\lambda(L-\lambda)} - \frac{\rho h(0)}{\lambda L} \right) \frac{1-\rho L}{1-\lambda L}$$

Now, turn to the fixed point problem that characterizes the equilibrium:

$$a_{it} = \mathbb{E}_{it}[\varphi \xi_t + \beta a_{it+1} + \gamma a_{t+1}]$$

Using our guess, we can replace the left-hand side with $h(L)x_{it}$. Using the results derived above, on the other hand, we can replace the right-hand side with $[G_1(L) + (\beta + \gamma)G_2(L) + \beta G_3(L)]x_{it}$. It follows that our guess is correct if and only if

$$h(L) = G_1(L) + (\beta + \gamma)G_2(L) + \beta G_3(L)$$

Equivalently, we need to find an analytic function $h(z)$ that solves

$$\begin{aligned} h(z) &= \varphi \frac{\lambda \tau_u}{\rho \tau_\eta} \frac{1}{1 - \rho\lambda} \frac{1}{1 - \lambda z} + \\ &+ (\beta + \gamma) \frac{\lambda \tau_u}{\rho \tau_\eta} \left(\frac{h(z)}{(1 - \lambda z)(z - \lambda)} - \frac{h(\lambda)(1 - \rho z)}{(1 - \rho\lambda)(z - \lambda)(1 - \lambda z)} \right) \\ &+ \beta \frac{\lambda}{\rho} \left(\frac{h(z)(z - \rho)}{z(z - \lambda)} - \frac{h(\lambda)(\lambda - \rho)}{\lambda(z - \lambda)} - \frac{\rho h(0)}{\lambda z} \right) \frac{1 - \rho z}{1 - \lambda z}, \end{aligned}$$

which can be transformed as

$$C(z)h(z) = d(z; h(\lambda), h(0))$$

where

$$\begin{aligned} C(z) &\equiv z(1 - \lambda z)(z - \lambda) - \frac{\lambda}{\rho} \left\{ \beta(z - \rho)(1 - \rho z) + (\beta + \gamma) \frac{\tau_u}{\tau_\eta} z \right\} \\ d(z; h(\lambda), h(0)) &\equiv \varphi \frac{\lambda \tau_u}{\rho \tau_\eta} \frac{1}{1 - \rho\lambda} z(z - \lambda) - \frac{1}{\rho} \left(\frac{\tau_u \lambda (\beta + \gamma)}{\tau_\eta} + \beta(\lambda - \rho) \right) z(1 - \rho z)h(\lambda) \\ &\quad - \beta(z - \lambda)(1 - \rho z)h(0) \end{aligned}$$

Note that $C(z)$ is a cubic equation and therefore contains with three roots. We will verify later that there are two inside roots and one outside root. To make sure that $h(z)$ is an analytic function, we choose $h(0)$ and $h(\lambda)$ so that the two roots of $d(z; h(\lambda), h(0))$ are the same as the two inside roots of $C(z)$. This pins down the constants $\{h(0), h(\lambda)\}$, and therefore the policy function $h(L)$

$$h(L) = \left(1 - \frac{\vartheta}{\rho}\right) \frac{\varphi}{1 - \rho\delta} \frac{1}{1 - \vartheta L},$$

where ϑ^{-1} is the root of $C(z)$ outside the unit circle.

Now we verify that $C(z)$ has two inside roots and one outside root. Note that $C(z)$ can be rewritten as

$$C(z) = \lambda \left\{ -z^3 + \left(\rho + \frac{1}{\rho} + \frac{1}{\rho} \frac{\tau_u}{\tau_\eta} + \beta \right) z^2 - \left(1 + \beta \left(\rho + \frac{1}{\rho} \right) + \frac{\beta + \gamma}{\rho} \frac{\tau_u}{\tau_\eta} \right) z + \beta \right\}.$$

With the assumption that $\beta > 0$, $\gamma > 0$, and $\beta + \gamma < 1$, it is straightforward to verify that the following

properties hold:

$$\begin{aligned}
C(0) &= \beta > 0 \\
C(\lambda) &= -\lambda\gamma \frac{1}{\rho} \frac{\tau_u}{\tau_\eta} < 0 \\
C(1) &= \frac{\tau_u(1-\beta-\gamma)}{\tau_\eta\rho} + (1-\beta) \left(\frac{1}{\rho} + \rho - 2 \right) > 0
\end{aligned}$$

Therefore, the three roots are all real, two of them are between 0 and 1, and the third one ϑ^{-1} is larger than 1.

Finally, to show that ϑ is less than ρ , it is sufficient to show that

$$C\left(\frac{1}{\rho}\right) = \frac{\tau_u(1-\rho\beta-\rho\gamma)}{\tau_\eta\rho^3} > 0.$$

Since $C(\vartheta^{-1}) = 0$, it has to be that ϑ^{-1} is larger than ρ^{-1} , or $\vartheta < \rho$.

Proof of Proposition 3. The equilibrium outcome in the hybrid economy is given by the following AR(2) process:

$$a_t = \frac{\zeta_0}{1-\zeta_1 L} \xi_t$$

where

$$\zeta_1 = \frac{1}{2\omega_f\delta} \left(1 - \sqrt{1-4\delta\omega_f\omega_b} \right) \quad \text{and} \quad \zeta_0 = \frac{\varphi\zeta_1}{\omega_b - \rho\omega_f\delta\zeta_1} \quad (26)$$

and $\delta \equiv \beta + \gamma$. The solution to the incomplete-information economy is

$$a_t = \left(1 - \frac{\vartheta}{\rho} \right) \left(\frac{\varphi}{1-\rho\delta} \right) \left(\frac{1}{1-\vartheta L} \xi_t \right),$$

To match the hybrid model, we need

$$\zeta_1 = \vartheta \quad \text{and} \quad \zeta_0 = \left(1 - \frac{\vartheta}{\rho} \right) \frac{\varphi}{1-\rho\delta}. \quad (27)$$

Combining (26) and (27), and solving for the coefficients of ω_f and ω_b , we infer that the two economies generate the same dynamics if and only if the following two conditions hold:

$$\omega_f = \frac{\delta\rho^2 - \vartheta}{\delta(\rho^2 - \vartheta^2)} \quad (28)$$

$$\omega_b = \frac{\vartheta(1-\delta\vartheta)\rho^2}{\rho^2 - \vartheta^2} \quad (29)$$

Since $\delta \equiv \beta + \gamma$ and since ϑ is a function of the primitive parameters $(\sigma, \rho, \beta, \gamma)$, the above two conditions give the coefficients ω_f and ω_b as functions of the primitive parameters, too.

It is immediate to check that $\omega_f < 1$ and $\omega_b > 0$ if $\vartheta \in (0, \rho)$, which in turn is necessarily true for any $\sigma > 0$; and that $\omega_f = 1$ and $\omega_b = 0$ if $\vartheta = \rho$, which in turn is the case if and only if $\sigma = 0$. The proof of the comparative statics in terms of σ is contained in the proof of Proposition 4.

Proof of Proposition 4. We first show that ω_f is decreasing in ϑ and ω_b is increasing in ϑ . This can be verified as follows

$$\begin{aligned}\frac{\partial \omega_f}{\partial \vartheta} &= \frac{-\delta(\rho^2 + \vartheta^2) + 2\delta^2 \rho^2 \vartheta}{(\delta(\rho^2 - \vartheta^2))^2} < \frac{-\delta(\rho^2 + \vartheta) + 2\delta \rho \vartheta}{(\delta(\rho^2 - \vartheta^2))^2} = \frac{-\delta(\rho - \vartheta)^2}{(\delta(\rho^2 - \vartheta^2))^2} < 0 \\ \frac{\partial \omega_b}{\partial \vartheta} &= \frac{\rho^2(\rho^2 + \vartheta^2 - 2\delta\vartheta\rho^2)}{(\rho^2 - \vartheta^2)^2} > \frac{\rho^2(\rho^2 + \vartheta^2 - 2\vartheta\rho)}{(\rho^2 - \vartheta^2)^2} = \left(\frac{\rho}{\rho + \vartheta}\right)^2 > 0\end{aligned}$$

Now it is sufficient to show that ϑ is increasing in γ . Note that

$$C\left(\frac{1}{\rho}\right) = \frac{\tau_u(1 - \rho\beta - \rho\gamma)}{\tau_\eta \rho^3} > 0 \quad \text{and} \quad C\left(\frac{1}{\lambda}\right) = -\frac{\tau_u \gamma \beta}{\tau_\eta \rho \lambda^2} < 0$$

By the continuity of $C(z)$, it must be the case that $C(z)$ admits a root between $\frac{1}{\rho}$ and $\frac{1}{\lambda}$. Recall from the proof of Proposition 2, ϑ^{-1} is the only outside root, and it follows that $\lambda < \vartheta < \rho$. It also implies that $C(z)$ is decreasing in z in the neighborhood of $z = \vartheta^{-1}$, a property that we use in the sequel to characterize comparative statics of ϑ .

Next, using the definition of $C(z)$, namely

$$C(z) \equiv -z^3 + \left(\rho + \frac{1}{\rho} + \frac{1}{\rho} \frac{\tau_u}{\tau_\eta} + \beta\right) z^2 - \left(1 + \beta \left(\rho + \frac{1}{\rho}\right) + \frac{\beta + \gamma}{\rho} \frac{\tau_u}{\tau_\eta}\right) z + \beta,$$

taking its derivative with respect to γ , and evaluating that derivative at $z = \vartheta^{-1}$, we get

$$\frac{\partial C(\vartheta^{-1})}{\partial \gamma} = -\frac{\tau_u}{\rho \tau_\eta} < 0$$

Combining this with the earlier observation that $\frac{\partial C(\vartheta^{-1})}{\partial z} < 0$, and using the Implicit Function Theorem, we infer that ϑ is an increasing function of γ .

Similarly, taking derivative with respect to τ_u , we have

$$\frac{\partial C(\vartheta^{-1})}{\partial \tau_u} = \frac{1}{\rho \tau_\eta} \vartheta^{-1} (\vartheta^{-1} - \beta - \gamma) > \frac{1}{\rho \tau_\eta} \vartheta^{-1} (1 - \beta - \gamma) > 0.$$

Since $\tau_u = \sigma^2$, we conclude that ϑ is also increasing in σ .

Proof of Proposition 5. The hybrid and the incomplete-information economies generate the same dynamics if and only if conditions (28) and (29) hold. Using (29), we can rewrite (28) as follows:

$$\omega_f = \Omega(\omega_b; \delta, \rho) \equiv 1 - \frac{1}{\delta \rho^2} \omega_b. \quad (30)$$

Furthermore, any (β, γ, ρ) , the equilibrium of the incomplete-information economy gives an invertible mapping from $\sigma \in (0, \infty)$ to $\vartheta \in (0, \rho)$, whereas condition (29) gives an invertible mapping from $\vartheta \in (0, \rho)$ to $\omega_b \in (0, \infty)$. It follows that there exists a $\sigma \in (0, \infty)$ such that the equilibrium dynamics of the incomplete-information economy replicates that of the hybrid economy if and only if the pair (ω_b, ω_f) satisfies condition (30) along with $\omega_b \in (0, \infty)$. Finally, the level of the informational friction that achieves this replication is obtained by inverting condition (29) to obtain ϑ , and thereby also σ , as an implicit function of ω_b .

Proof of Proposition 9. First, let us prove $g_k < \widehat{g}_k$. Recall that $\{g_k\}$ is given by

$$g_k = \sum_{h=0}^{\infty} \gamma^h \lambda_k \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h}$$

Clearly,

$$0 < g_k < \sum_{h=0}^{\infty} \gamma^h \lambda_k \rho_{k+h} = \widehat{g}_k,$$

which proves the first property. If $\lim_{k \rightarrow \infty} \lambda_k = 1$ and $\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}$ exists for all k , then it follows that

$$\lim_{k \rightarrow \infty} \frac{\widehat{g}_k}{g_k} = \frac{\lim_{k \rightarrow \infty} \sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}{\lim_{k \rightarrow \infty} \sum_{h=0}^{\infty} \gamma^h \rho_{k+h}} = 1.$$

Next, let us prove that $\frac{g_{k+1}}{g_k} > \frac{\widehat{g}_{k+1}}{\widehat{g}_k}$. By definition,

$$\begin{aligned} \frac{\widehat{g}_{k+1}}{\widehat{g}_k} &= \frac{\lambda_{k+1} \sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\lambda_k \sum_{h=0}^{\infty} \gamma^h \rho_{k+h}} \\ \frac{g_{k+1}}{g_k} &= \frac{\lambda_{k+1} \sum_{h=0}^{\infty} \gamma^h \lambda_{k+2} \dots \lambda_{k+h+1} \rho_{k+h+1}}{\lambda_k \sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h}} \end{aligned}$$

Since $\{\lambda_k\}$ is strictly increasing and $\rho_k > 0$, we have

$$\frac{g_{k+1}}{g_k} \Big/ \frac{\widehat{g}_{k+1}}{\widehat{g}_k} > \frac{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h}} \Big/ \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}$$

It remains to show that the term on the right-hand side is greater than 1. To proceed, we start with

the following observation. If $\theta_1 \geq \theta_2 > 0$, and $\frac{y_2}{y_1+y_2} \geq \frac{x_2}{x_1+x_2}$, then

$$\frac{x_1\theta_1 + x_2\theta_2}{x_1 + x_2} \geq \frac{y_1\theta_1 + y_2\theta_2}{y_1 + y_2}$$

Note that

$$\frac{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \dots \lambda_{k+h} \rho_{k+h}} = \frac{\rho_{k+1}}{\rho_k} \frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \gamma^2 \lambda_{k+1} \lambda_{k+2} \frac{\rho_{k+3}}{\rho_{k+1}} + \dots}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \gamma^2 \lambda_{k+1} \lambda_{k+2} \frac{\rho_{k+2}}{\rho_k} + \dots}$$

and

$$\frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}} = \frac{\rho_{k+1}}{\rho_k} \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \gamma^2 \frac{\rho_{k+3}}{\rho_{k+1}} + \dots}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \gamma^2 \frac{\rho_{k+2}}{\rho_k} + \dots}$$

Based on the observation, we will show that

$$\frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \gamma^2 \lambda_{k+1} \lambda_{k+2} \frac{\rho_{k+3}}{\rho_{k+1}} + \dots}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \gamma^2 \lambda_{k+1} \lambda_{k+2} \frac{\rho_{k+2}}{\rho_k} + \dots} \geq \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \gamma^2 \frac{\rho_{k+3}}{\rho_{k+1}} + \dots}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \gamma^2 \frac{\rho_{k+2}}{\rho_k} + \dots}$$

by induction. We first establish the following

$$\frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k}} \geq \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}}}{1 + \gamma \frac{\rho_{k+1}}{\rho_k}}$$

This inequality is obtained by labeling $\theta_1 = 1, \theta_2 = \frac{\rho_k \rho_{k+2}}{\rho_{k+1}^2}, x_1 = y_1 = 1, x_2 = \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k},$ and $y_2 = \gamma \frac{\rho_{k+1}}{\rho_k}.$ By assumption, $\frac{\rho_k \rho_{k+2}}{\rho_{k+1}^2} \leq 1.$ Meanwhile,

$$\frac{x_2}{x_1 + x_2} = \frac{\gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k}} \leq \frac{\gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k}}{\lambda_{k+1} + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k}} = \frac{y_2}{y_1 + y_2}$$

Now suppose that

$$\frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n}}{\rho_{k+1}}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k}} \geq \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \frac{\rho_{k+n}}{\rho_{k+1}}}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k}}$$

We want to show

$$\begin{aligned} & \frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n}}{\rho_{k+1}} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n+1}}{\rho_{k+1}}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}} \\ & \geq \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \frac{\rho_{k+n}}{\rho_{k+1}} + \gamma^n \frac{\rho_{k+n+1}}{\rho_{k+1}}}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \frac{\rho_{k+n}}{\rho_k}} \end{aligned}$$

Let $\theta_1 = \frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \frac{\rho_{k+n}}{\rho_{k+1}}}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k}}$, $\theta_2 = \frac{\rho_k \rho_{k+n+1}}{\rho_{k+1} \rho_{k+n}}$, $x_1 = 1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k}$,
 $x_2 = \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}$, $y_1 = 1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k}$, $y_2 = \gamma^n \frac{\rho_{k+n}}{\rho_k}$. We have

$$\begin{aligned} & \frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n}}{\rho_{k+1}} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n+1}}{\rho_{k+1}}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}} \\ &= \frac{x_1 \frac{1 + \gamma \lambda_{k+1} \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n}}{\rho_{k+1}}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k}} + x_2 \theta_2}{x_1 + x_2} \\ &\geq \frac{x_1 \theta_1 + x_2 \theta_2}{x_1 + x_2} \end{aligned}$$

and

$$\frac{1 + \gamma \frac{\rho_{k+2}}{\rho_{k+1}} + \dots + \gamma^{n-1} \frac{\rho_{k+n}}{\rho_{k+1}} + \gamma^n \frac{\rho_{k+n+1}}{\rho_{k+1}}}{1 + \gamma \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \frac{\rho_{k+n}}{\rho_k}} = \frac{y_1 \theta_1 + y_2 \theta_2}{y_1 + y_2}$$

It remains to show that $\theta_1 \geq \theta_2$ and $\frac{x_2}{x_1 + x_2} \leq \frac{y_2}{y_1 + y_2}$. Note that

$$\frac{\theta_1}{\theta_2} = \frac{1 + \gamma \frac{\rho_{k+1}}{\rho_k} \frac{\rho_{k+2} \rho_k}{\rho_{k+1}^2} + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k} \frac{\rho_{k+n} \rho_k}{\rho_{k+1} \rho_{k+n-1}}}{\theta_2 + \gamma \frac{\rho_{k+1}}{\rho_k} \theta_2 + \dots + \gamma^{n-1} \frac{\rho_{k+n-1}}{\rho_k} \theta_2}$$

By assumption, $\theta_2 < 1$ and $\theta_2 \leq \frac{\rho_k \rho_{k+i+1}}{\rho_{k+1} \rho_{k+i}}$ when $i \leq n$, which leads to $\theta_1 \geq \theta_2$. Also note that

$$\begin{aligned} & \frac{x_2}{x_1 + x_2} \\ &= \frac{\gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}}{1 + \gamma \lambda_{k+1} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n-1} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}} \\ &\leq \frac{\gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}}{\lambda_{k+1} \dots \lambda_{k+n} + \gamma \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+1}}{\rho_k} + \dots + \gamma^{n-1} \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n-1}}{\rho_k} + \gamma^n \lambda_{k+1} \dots \lambda_{k+n} \frac{\rho_{k+n}}{\rho_k}} \\ &= \frac{y_2}{y_1 + y_2} \end{aligned}$$

This completes the proof that $\frac{g_{k+1}}{g_k} > \frac{\widehat{g}_{k+1}}{\widehat{g}_k}$.

Appendix B1: Dynamic IS Curve [to be completed]

Appendix B2: NKPC

In this appendix we show how to nest an incomplete-information extension of the NKPC to our framework. Apart for the introduction of incomplete information, the micro-foundations are the same as in familiar textbook treatments of the NKPC (e.g., Galí, 2008). There is a continuum of firms, each producing a differentiated commodity. Firms set prices optimally, but can adjust them only infrequently. Each period, a firm has the option to reset its price with probability $1-\theta$, where $\theta \in (0, 1)$; otherwise, it is stuck at the previous-period price. Technology is linear, so that the real marginal cost of a firm is invariant to its production level.

The optimal reset price solves the following problem:

$$P_{it}^* = \arg \max_{P_{it}} \sum_{k=0}^{\infty} (\delta\theta)^k \mathbb{E}_{it} \left\{ Q_{t|t+k} \left(P_{it} Y_{it+k|t} - P_{t+k} \text{mc}_{t+k} Y_{i,t+k|t} \right) \right\}$$

subject to the demand equation, $Y_{it+k} = \left(\frac{P_{it}}{P_{t+k}} \right)^{-\epsilon} Y_{t+k}$, where $Q_{t|t+k}$ is the stochastic discount factor between t and $t+k$, Y_{t+k} and P_{t+k} are, respectively, aggregate income and the aggregate price level in period $t+k$, P_{it} is the firm's price, as set in period t , $Y_{i,t+k|t}$ is the firm's quantity in period $t+k$, conditional on not having changed the price since t , and mc_{t+k} is the real marginal cost in period $t+k$.

Taking the first-order condition and log-linearizing around a steady state with no shocks and zero inflation, we get the following, familiar, characterization of the optimal rest price:

$$p_{it}^* = (1 - \delta\theta) \sum_{k=0}^{\infty} (\delta\theta)^k \mathbb{E}_{it} [\text{mc}_{t+k} + p_{t+k}]. \quad (31)$$

We next make the simplifying assumption that the firms observe that past price level but do not extract information from it. Following Vives and Yang (2017), this assumption can represent a form of bounded rationality. It can also be motivated on empirical grounds: in the data, inflation contains little statistical information about real marginal costs and output gaps—it's dominated by the residual, or what the DSGE literature interprets as "markup shocks." This means that, even if we were to allow firms to extract information from past inflation, this would make little quantitative difference, provided that we accommodate this empirically relevant source of noise. Furthermore, as we show later on, our observational-equivalence result remains an excellent approximation of the true equilibrium in extension that allow for such endogenous information.

With this simplifying assumption, we can restate condition (31) as

$$p_{it}^* - p_{t-1} = (1 - \delta\theta) \sum_{k=0}^{\infty} (\delta\theta)^k \mathbb{E}_{it}[\mathfrak{m}c_{t+k} + \pi_{t+k}], \quad (32)$$

Since only a fraction $1 - \theta$ of the firms adjust their prices each period, the price level in period t is given by $p_t = (1 - \theta) \int p_{it}^* di + \theta p_{t-1}$. By the same token, inflation is given by

$$\pi_t \equiv p_t - p_{t-1} = (1 - \theta) \int (p_{it}^* - p_{t-1}).$$

Combining this with condition (32) and rearranging, we arrive at the following expression:

$$\pi_t = \frac{(1 - \delta\theta)(1 - \theta)}{\theta} \sum_{k=0}^{\infty} (\delta\theta)^k \bar{\mathbb{E}}_t[\mathfrak{m}c_{t+k}] + \delta(1 - \theta) \sum_{k=0}^{\infty} (\delta\theta)^k \bar{\mathbb{E}}_t[\pi_{t+k+1}]. \quad (33)$$

When information is complete, we can replace $\bar{\mathbb{E}}_t[\cdot]$ with $\mathbb{E}_t[\cdot]$, the expectation of the representative agent. We can then use the Law of Iterated Expectations to reduce condition (33) to the following:

$$\pi_t = \kappa \mathfrak{m}c_t + \delta \mathbb{E}_t[\pi_{t+1}], \quad (34)$$

where $\kappa \equiv \frac{(1 - \delta\theta)(1 - \theta)}{\theta}$. This is the standard NKPC.

When instead information is incomplete, the Law of Iterated Expectations does not apply at the aggregate level, because average forecast errors can be auto-correlated, and therefore condition (33) cannot be reduced to condition (34). Furthermore, the approach, taken in some papers, of replacing $\mathbb{E}_t[\pi_{t+1}]$ in the standard NKPC with the average forecast in surveys is invalid under the micro-foundations laid out here.¹⁵ Instead, the applicable version of the NKPC is the one given in condition (33), which relates current inflation to forecasts of the entire future.

The incomplete-information version of the NKPC obtained above has the advantage of being robust to a rich class of specifications of the stochastic process of the real marginal cost and of the available signals about it. But it is too cumbersome to take to the data, precisely because it requires that the econometrician have data on subjective expectations of the entire future paths of the real marginal cost and of inflation. This is where our observational-equivalence result comes to rescue.

Condition (33) is readily nested the analysis of Sections 2 and 3 by mapping $\mathfrak{m}c_t$ and π_t to ξ_t and a_t , respectively, and by letting $\varphi = \kappa$, $\beta = \delta\theta$, and $\gamma = \delta(1 - \theta)$. The following is then an immediate application of Propositions 3 and 4, provided of course that we maintain the assumptions introduced in Section 3.

¹⁵Mavroeidis, Plagborg-Møller, and Stock (2014) also question the validity this approach and effectively invite the exercise conducted here.

Proposition 6. (i) There exist $\omega_f < 1$ and $\omega_b > 0$ such that, when information is incomplete, the equilibrium process for inflation solves the following equation:

$$\pi_t = \kappa mc_t + \omega_f \delta \mathbb{E}_t[\pi_{t+1}] + \omega_b \pi_{t-1} \quad (35)$$

(ii) For any given level of noise, increasing the degree of price flexibility (i.e., reducing θ) results to a lower ω_f and a higher ω_b .

Part (i) establishes that, when information is incomplete, it is as if inflation is governed by a variant of the NKPC that introduces myopia, in the form of $\omega_f < 1$, along with a backward-looking component, in the form of $\omega_b > 0$. This is similar to the Hybrid NKPC considered in, inter alia, [Gali and Gertler \(1999\)](#), [Christiano, Eichenbaum, and Evans \(2005\)](#) and [Smets and Wouters \(2007\)](#). It bypasses the complexity of condition (33) and facilitates the empirical exercises conducted below.

Part (ii) adds the following interesting lesson. When information is complete, higher price flexibility contributes merely to a steeper NKPC, that is, to a higher κ in condition (34). This is generally bad for the empirical fit of the New Keynesian model, which in turn explains why the literature has tried hard to justify a degree of price stickiness at the aggregate level that is higher than the one that appears to present at the micro-economic level under the lens of menu-cost models. But once information is incomplete, a moderate degree of price flexibility can be good in the sense that it contributes to more sluggishness in the inflation dynamics by reinforcing the role of higher-order uncertainty. This point helps explain why our quantitative implementation reconciles salient features of the inflation dynamics with a relatively modest degree of price stickiness.

Appendix C1: Investment

A long tradition in macroeconomics that goes back to [Hayashi \(1982\)](#) and [Abel and Blanchard \(1983\)](#) has studied representative-agent models in which the firms face a cost in adjusting their capital stock. In this literature, the adjustment cost is specified as follows:

$$\text{Cost}_t = \Phi \left(\frac{I_t}{K_{t-1}} \right) \quad (36)$$

where I_t denotes the rate of investment, K_{t-1} denotes the capital stock inherited from the previous period, and Φ is a convex function. This specification gives the level of investment as a decreasing function of Tobin's Q. It also generates aggregate investment responses that are broadly in line with those predicted by more realistic, heterogeneous-agent models that account for the dynamics of investment at the firm or plant level ([Caballero and Engel, 1999](#); [Bachmann, Caballero, and Engel,](#)

2013; Khan and Thomas, 2008).¹⁶

By contrast, the DSGE literature that follows [Christiano, Eichenbaum, and Evans \(2005\)](#) and [Smets and Wouters \(2007\)](#) assumes that the firms face a cost in adjusting, not their capital stock, but rather their rate of investment. That is, this literature specifies the adjustment cost as follows:

$$\text{Cost}_t = \Psi \left(\frac{I_t}{I_{t-1}} \right) \quad (37)$$

As with the Hybrid NKPC, this specification was adopted because it allows the theory to generate sluggish aggregate investment responses to monetary and other shocks. But it has no obvious analogue in the literature that accounts for the dynamics of investment at the firm or plant level.

In the sequel, we set up a model of aggregate investment with two key features: first, the adjustment cost takes the form seen in condition (36); and second, the investments of different firms are strategic complements because of an aggregate demand externality. We then augment this model with incomplete information and show that it becomes observationally equivalent to a model in which the adjustment cost takes the form seen in condition (37). This illustrates how incomplete information can merge the gap between the different strands of the literature and help reconcile the dominant DSGE practice with the relevant microeconomic evidence on investment.

Let us fill in the details. We consider an AK model with costs to adjusting the capital stock. There is a continuum of monopolistic competitive firms, indexed by i and producing different varieties of intermediate investment goods. The final investment good is a CES aggregator of intermediate investment goods. Letting X_{it} denote the investment good produced by firm i , we have that the aggregate investment is given by

$$I_t = \left[\int X_{it}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

And letting Q_{it} denote the price faced by firm i , we have that the investment price index is given by

$$Q_t = \left[\int Q_{it}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

A representative final goods producer has perfect information and purchases investment goods to maximize its discounted profit

$$\max_{\{K_t, I_t\}} \sum_{t=0}^{\infty} \chi^t \mathbb{E}_0 \left[\exp(\xi_t) AK_t - Q_t I_t - \Phi \left(\frac{I_t}{K_t} \right) K_t \right],$$

¹⁶These works differ on the importance they attribute to heterogeneity, lumpiness, and non-linearities, but appear to share the prediction that the impulse response of aggregate investment is peaked on impact. They therefore do not provide a micro-foundation of the kind of sluggish investment dynamics featured in the DSGE literature.

subject to

$$K_{t+1} = K_t + I_t.$$

Here, the fundamental shock, ξ_t , is an exogenous productivity shock to the final goods production, and $\Phi\left(\frac{I_t}{K_t}\right)K_t$ represents the quadratic capital-adjustment cost. The following functional form is assumed:

$$\Phi\left(\frac{I_t}{K_t}\right) = \frac{1}{2}\psi\left(\frac{I_t}{K_t}\right)^2.$$

Let $Z_t \equiv \frac{I_t}{K_t}$ denote the investment-to-capital ratio. On a balanced growth path, this ratio and the price for the investment goods remain constant, i.e., $Z_t = Z$ and $Q_t = Q$. The log-linearized version of the final goods producer's optimal condition around the balanced growth path can be written as

$$Qq_t + \psi Z z_t = \chi \mathbb{E}_t \left[A\xi_{t+1} + Qq_{t+1} + \psi Z(1 + Z)z_{t+1} \right]. \quad (38)$$

When the producers of the intermediate investment goods choose their production scale, they may not observe the underlying fundamental ξ_t perfectly. As a result, they have to make their decision based on their expectations about fundamentals and others' decisions. Letting

$$\max_{X_{it}} \mathbb{E}_{it} [Q_{it}X_{it} - cX_{it}],$$

subject to

$$Q_{it} = \left(\frac{X_{it}}{I_t}\right)^{-\frac{1}{\sigma}} Q_t.$$

Define $Z_{it} \equiv \frac{X_{it}}{K_t}$ as the firm-specific investment-to-capital ratio, and the log-linearized version of the optimal choice of X_{it} is

$$z_{it} = \mathbb{E}_{it} [z_t + \sigma q_t].$$

In steady state, the price Q simply equals the markup over marginal cost c ,

$$Q = \frac{\sigma}{\sigma - 1}c,$$

and the investment-to-capital ratio Z solves the quadratic equation

$$Q + \psi Z = \chi \left(A + Q + \psi Z + \psi Z^2 - \frac{1}{2}\psi Z^2 \right).$$

Frictionless Benchmark. If all intermediate firms observe ξ_t perfectly, then we have

$$z_{it} = z_t + \sigma q_t$$

Aggregation implies that $z_{it} = z_t$ and $q_t = 0$. It follows that z_t obeys the following Euler condition:

$$z_t = \varphi \xi_t + \delta \mathbb{E}_t [z_{t+1}]$$

where

$$\varphi = \frac{\rho \chi A}{\psi Z} \quad \text{and} \quad \delta = \chi(1 + Z).$$

Incomplete Information. Suppose now that firms receive a noisy signal about the fundamental ξ_t as in Section 2. Here, we make the same simplifying assumption as in the NKPC application. We assume that firms observe current z_t , but preclude them from extracting information from it. Together with the pricing equation (38), the aggregate investment dynamics follow

$$z_t = \frac{\rho \chi A}{\psi Z} \sum_{k=0}^{\infty} \chi^k \mathbb{E}_t [\xi_{t+k}] + \chi Z \sum_{k=0}^{\infty} \chi^k \mathbb{E}_t [z_{t+k+1}]$$

The investment dynamics can be understood as the solution to the dynamic beauty contest studied in Section 2 by letting

$$\varphi = \frac{\rho \chi A}{\psi Z}, \quad \beta = \chi, \quad \text{and} \quad \gamma = \chi Z.$$

The following is then immediate.

Proposition 7. *When information is incomplete, there exist $\omega_f < 1$ and $\omega_b > 0$ such that the equilibrium process for investment solves the following equation:*

$$z_t = \varphi \xi_t + \omega_f \delta \mathbb{E}_t [z_{t+1}] + \omega_b z_{t-1}$$

Finally, it is straightforward to show that the above equation is of the same type as the one that governs investment in a complete-information model where the adjustment cost is in terms of the investment rate, namely a model in which the final good producer's problem is modified as follows:

$$\max_{\{K_t, I_t\}} \sum_{t=0}^{\infty} \chi^t \mathbb{E}_0 \left[\exp(\xi_t) A K_t - Q_t I_t - \Psi \left(\frac{I_t}{\tilde{I}_{t-1}} \right) I_t \right]$$

where \tilde{I}_t is the aggregate investment.

Appendix C2: Asset Prices

Consider a log-linearized version of the standard asset-pricing condition in an infinite horizon, representative-agent model:

$$p_t = \mathbb{E}_t[d_{t+1}] + \delta \mathbb{E}_t[p_{t+1}],$$

where p_t is the price of the asset in period t , d_{t+1} is its dividend in the next period, \mathbb{E}_t is the expectation of the representative agent, and δ is his discount factor. Iterating the above condition gives the equilibrium price as the expected present discounted value of the future dividends.

By assuming a representative agent, the above condition conceals the importance of higher-order beliefs. A number of works have sought to unearth that role by considering variants with heterogeneously informed, short-term traders, in the tradition of Singleton (1987); see, for example, Allen, Morris, and Shin (2006), Kasa, Walker, and Whiteman (2014), and Nimark (2017). We can capture these works in our setting by modifying the equilibrium pricing condition as follows:

$$p_t = \bar{\mathbb{E}}_t[d_{t+1}] + \delta \bar{\mathbb{E}}_t[p_{t+1}] + \epsilon_t,$$

where $\bar{\mathbb{E}}_t$ is the *average* expectation of the traders in period t and ϵ_t is an i.i.d shock interpreted as the price effect of noisy traders. The key idea embedded in the above condition is that, as long as the traders have different information and there are limits to arbitrage, asset markets are likely to behave like (dynamic) beauty contests.

Let us now assume that the dividend is given by $d_{t+1} = \xi_t + u_{t+1}$, where ξ_t follows an AR(1) process and u_{t+1} is i.i.d. over time, and that the information of the typical trader can be represented by a series of private signals as in condition (12).¹⁷ Applying our results, and using the fact that $\xi_t = \mathbb{E}_t[d_{t+1}]$, we then have that the component of the equilibrium asset price that is driven by ξ_t obeys the following law of motion, for some $\omega_f < 1$ and $\omega_b > 0$:

$$p_t = \mathbb{E}_t[d_{t+1}] + \omega_f \delta \mathbb{E}_t[p_{t+1}] + \omega_b p_{t-1},$$

where $\mathbb{E}_t[\cdot]$ is the fully-information, rational expectations. We thus have that asset prices can display both myopia, in the form of $\omega_f < 1$, and momentum, or predictability, in the form of $\omega_b > 0$.

Kasa, Walker, and Whiteman (2014) have already emphasized how incomplete information and higher-order uncertainty can help explain momentum and predictability in asset prices. Our result

¹⁷Here, we are abstracting from the complications of the endogenous revelation of information and we think of the signals in (12) as convenient proxies for all the information of the typical trader. One can also interpret this as a setting in which the dividend is observable (and hence so is the price, which is measurable in the dividend) and the assumed signals are the representation of a form of rational inattention. Last but not least, we have verified that the solution with endogenous information can be approximated very well by the solution obtained with exogenous information.

offers a sharp illustration of this insight and blends it with the insight regarding myopia. In the present context, the latter insight seems to challenge the asset-price literature that emphasizes long-run risks: news about the long-run fundamentals may be heavily discounted when there is higher-order uncertainty. Finally, our result suggests that both kinds of distortions are likely to be greater at the level of the entire stock market than at the level of the stock of a particular firm insofar as financial frictions and GE effects cause the trades to be strategic complements at the macro level even if they are strategic substitutes at the micro level, which in turn may help rationalize Samuelson’s dictum (Jung and Shiller, 2005).

We leave the exploration of these—admittedly speculative—ideas open for future research. We conclude this appendix by illustrating how our observational-equivalence result, which relies on assuming away the endogenous revelation of information through the equilibrium price, can be seen as an approximation of the dynamics that obtain when this assumption is relaxed.

Allowing learning from prices adds more realism, but typically rules out an analytic characterization of the equilibrium.¹⁸ Suppose, in particular, that the traders in our setting can perfectly observe the current price as well as the last-period dividend. In this case, the equilibrium pricing dynamics does not admit a finite state-space representation. To illustrate, set $\delta = 0.98$, $\rho = 0.95$, $\sigma_u = 2$, and $\sigma_\epsilon = \sigma_\nu = 5$, and approximate the equilibrium dynamics with an MA(100) process. The solid blue in Figure 3 gives the resulting IRF of the equilibrium price to an innovation in ξ_t . The dashed red line is obtained by taking our hybrid economy, which assumes away the learning from either the price or the past dividend, and recalibrating the level of the idiosyncratic noise so that the implied IRF is close as possible to the one obtained in the economy in which such learning is allowed. As evident in the figure, the hybrid economy does a very good job in replicating the dynamics of the latter economy.

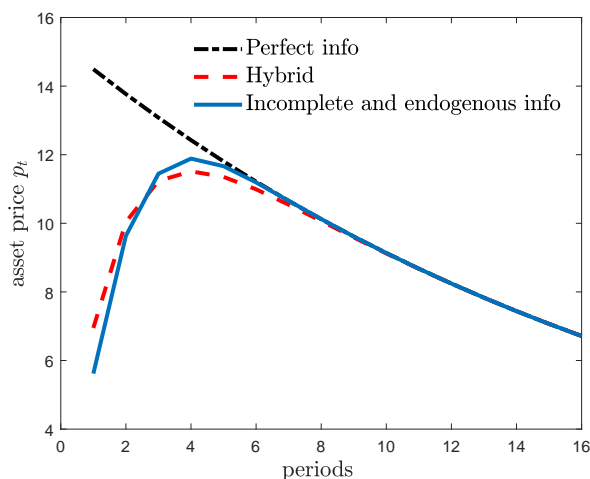


Figure 3: Impulse Response Function of Asset Price

¹⁸See Nimark (2017) and Huo and Takayama (2018) for a more detailed discussion.

We have verified that this similarity extends to a wide range of values for the parameters of the assumed setting. This similarity may, of course, be broken by assuming a more complex stochastic process for the fundamental and a more convoluted learning dynamics. However, the analysis of Section 7 together with the example presented here illustrate why our analysis can be thought of as a convenient proxy of settings with endogenous information aggregation.

Appendix D: Variant with Idiosyncratic Shocks

In this appendix, we extend the analysis to a setting that features both aggregate and idiosyncratic shocks. This serves to illustrate how our theory offers a natural explanation of why significant levels of as-if myopia and anchoring can be present at the macroeconomic level (i.e., in the response to aggregate shocks) even if they are absent at the microeconomic level (i.e., in the response to idiosyncratic shocks), which complements the discussion in Section 4 .

To accommodate idiosyncratic shocks, we extend the model so that the optimal behavior of agent i obeys the following equation:

$$a_{it} = \mathbb{E}_{it}[\varphi\xi_{it} + \beta a_{it+1} + \gamma a_{t+1}]$$

where

$$\xi_{it} = \xi_t + \zeta_{it}$$

and where ζ_{it} is a purely idiosyncratic shock. We let the latter follow a similar AR(1) process as the aggregate shock: $\zeta_{it} = \rho\zeta_{it-1} + \epsilon_{it}$, where ϵ_{it} is i.i.d. across both i and t .¹⁹

We then specify the information structure as follows. First, we let each agent observe the same signal x_{it} about the aggregate shock ξ_t as in our baseline model. Second, we let each agent observe the following signal about the idiosyncratic shock ζ_{it} :

$$z_{it} = \zeta_{it} + v_{it},$$

where v_{it} is independent of ζ_{it} , of ξ_t , and of x_{it} .

Because the signals are independent, the updating of the beliefs about the idiosyncratic and the aggregate shocks are also independent. Let $1 - \frac{\lambda}{\rho}$ be the Kalman gain in the forecasts of the aggregate fundamental, that is,

$$\mathbb{E}_{it}[\xi_t] = \lambda\mathbb{E}_{it-1}[\xi_t] + \left(1 - \frac{\lambda}{\rho}\right)x_{it}$$

¹⁹The restriction that the two kinds of shocks have the same persistence is only for expositional simplicity.

Next, let $1 - \frac{\hat{\lambda}}{\rho}$ be the Kalman gain in the forecasts of the idiosyncratic fundamental, that is,

$$\mathbb{E}_{it}[\zeta_{it}] = \hat{\lambda}\mathbb{E}_{it-1}[\zeta_{it}] + \left(1 - \frac{\hat{\lambda}}{\rho}\right) z_{it}$$

It is straightforward to extend the results of Section 3.2 to the current specification. It can thus be shown that the equilibrium action is given by the following:

$$a_{it} = \left(1 - \frac{\hat{\lambda}}{\rho}\right) \frac{\varphi}{1 - \rho\beta} \frac{1}{1 - \hat{\lambda}L} \zeta_{it} + \left(1 - \frac{\vartheta}{\rho}\right) \frac{\varphi}{1 - \rho\delta} \frac{1}{1 - \vartheta L} \xi_t + u_{it}$$

where ϑ is determined in the same manner as in our baseline model and where u_{it} is a residual that is orthogonal to both ζ_{it} and ξ_t and that captures the combined effect of all the idiosyncratic noises in the information of agent i . Finally, it is straightforward to check that $\vartheta = \lambda$ when $\gamma = 0$; $\vartheta > \lambda$ when $\gamma > 0$; and the gap between ϑ and λ increases with the strength of the GE effect, as measured with γ .

In comparison, the full-information equilibrium action is given by

$$a_{it}^* = \frac{\varphi}{1 - \rho\beta} \zeta_{it} + \frac{\kappa}{1 - \rho\delta} \xi_t.$$

It follows that, relative to the full-information benchmark, the distortions of the micro- and the macro-level IRFs are given by, respectively,

$$\left(1 - \frac{\hat{\lambda}}{\rho}\right) \frac{1}{1 - \hat{\lambda}L} \quad \text{and} \quad \left(1 - \frac{\vartheta}{\rho}\right) \frac{1}{1 - \vartheta L}.$$

The macro-level distortions is therefore higher than its micro-level counterpart if and only if $\vartheta > \hat{\lambda}$.

Following [Mackowiak and Wiederholt \(2009\)](#), it is natural to assume that $\hat{\lambda}$ is lower than λ , because the typical agent is likely to allocate more attention to idiosyncratic shocks than to aggregate shocks. This guarantees a lower distortion at the micro level than at the macro level even if we abstract from GE interactions (which amounts to setting $\gamma = 0$, or abstracting from role higher-order uncertainty). But once such interactions are taken into account, we have that ϑ remains higher than $\hat{\lambda}$ even if $\hat{\lambda} = \lambda$. In short, the macro-level response can display a bigger distortion, not only because of the mechanism identified in the aforesaid paper, but also because of the role of higher-order uncertainty identified here.

Appendix E: Robustness of Main Insights

In the main text we claimed that, although our observational-equivalence result depends on stringent assumptions about the process of the fundamental and the available signals, it encapsulates a few broader insights, which in turn justify the perspective put forward in our paper. In this section, we substantiate this claim by clarifying these insights and by elaborating on their robustness.

Setup. We henceforth let the fundamental ξ_t follow a flexible, possibly infinite-order, MA process:

$$\xi_t = \sum_{k=0}^{\infty} \rho_k \eta_{t-k}, \quad (39)$$

where the sequence $\{\rho_k\}_{k=0}^{\infty}$ is non-negative and square summable. Clearly, the AR(1) process assumed earlier on is nested as a special case where $\rho_k = \rho^k$ for all $k \geq 0$. The present specification allows for richer, possibly hump-shaped, dynamics in the fundamental, as well as for “news shocks,” that is, for innovations that shift the fundamental only after a delay.

Next, for every i and t , we let the incremental information received by agent i in period t be given by the series $\{x_{i,t,t-k}\}_{k=0}^{\infty}$, where

$$x_{i,t,t-k} = \eta_{t-k} + \epsilon_{i,t,t-k} \quad \forall k,$$

where $\epsilon_{i,t,t-k} \sim \mathcal{N}(0, (\tau_k)^{-2})$ is i.i.d. across i and t , uncorrelated across k , and orthogonal to the past, current, and future innovations in the fundamental, and where the sequence $\{\tau_k\}_{k=0}^{\infty}$ is non-negative and non-decreasing. In plain words, whereas our baseline specification has the agents observe a signal about the concurrent fundamental in each period, the new specification lets them observe a series of signals about the entire history of the underlying past and current innovations.

This specification is similar to our baseline in that it allows for more information to be accumulated as time passes. It differs, however, in two respects. First, it “orthogonalizes” the information structure in the sense that, for every t , every k , and every $k' \neq k$, the signals received at or prior to date t about the shock η_{t-k} are independent of the signals received about the shock $\eta_{t-k'}$. Second, it allows for more flexible learning dynamics in the sense that the precision τ_k does not have to be flat in k : the quality of the incremental information received in any given period about a past shock may either increase or decrease with the lag since the shock has occurred.

The first property is essential for tractability. The pertinent literature has struggled to solve for, or accurately approximate, the complex fixed point between the equilibrium dynamics and the Kalman filtering that obtains in dynamic models with incomplete information, especially in the presence of endogenous signals; see, for example, [Nimark \(2017\)](#). By adopting the aforementioned orthogonalization, we cut the Gordian knot and facilitate a closed-form solution of the entire dynamic structure

of the higher-order beliefs and of the equilibrium outcome.²⁰ The second property then permits us, not only to accommodate a more flexible learning dynamics, but also to disentangle the speed of learning from level of noise—a disentangling that was not possible in Section 3 because a single parameter, σ , controlled both objects at once.

Dynamics of Higher-Order Beliefs. The information regarding η_{t-k} that an agent has accumulated up to, and including, period t can be represented by a sufficient statistic, given by

$$\tilde{x}_{i,t}^k = \sum_{j=0}^k \frac{\tau_j}{\pi_k} x_{i,t-j,t-k}$$

where $\pi_k \equiv \sum_{j=0}^k \tau_j$. That is, the sufficient statistic is constructed by taking a weighted average of all the available signals, with the weight of each signal being proportional to its precision; and the precision of the statistic is the sum of the precisions of the signals. Letting $\lambda_k \equiv \frac{\pi_k}{\sigma_\eta^2 + \pi_k}$, we have that $\mathbb{E}_{it}[\eta_{t-k}] = \lambda_k \tilde{x}_{i,t}^k$, which in turn implies $\bar{\mathbb{E}}_t[\eta_{t-k}] = \lambda_k \eta_{t-k}$ and therefore

$$\bar{\mathbb{E}}_t[\xi_t] = \bar{\mathbb{E}}_t \left[\sum_{k=0}^{\infty} \rho_k \eta_{t-k} \right] = \sum_{k=0}^{\infty} f_{1,k} \eta_{t-k}, \quad \text{with} \quad f_{1,k} = \lambda_k \rho_k. \quad (40)$$

The sequence $\mathbf{F}_1 \equiv \{f_{1,k}\}_{k=0}^{\infty} = \{\lambda_k \rho_k\}_{k=0}^{\infty}$ identifies the IRF of the average first-order forecast to an innovation. By comparison, the IRF of the fundamental itself is given by the sequence $\{\rho_k\}_{k=0}^{\infty}$. It follows that the relation of the two IRFs is pinned down by the sequence $\{\lambda_k\}_{k=0}^{\infty}$, which describes the dynamics of learning. In particular, the smaller λ_0 is (i.e., the less precise the initial information is), the larger the initial gap between the two IRFs (i.e., a larger the initial forecast error). And the slower λ_k increases with k (i.e., the slower the learning over time), the longer it takes for that gap (and the average forecast) to disappear.

These properties are intuitive and are shared by the specification studied in the rest of the paper. In the information structure specified in Section 3, the initial precision is tied with the subsequent speed of learning. By contrast, the present specification disentangles the two. As shown next, it also allows for a simple characterization of the IRFs of the higher-order beliefs, which is what we are after.

Consider first the forward-looking higher-order beliefs. Applying condition (40) to period $t + 1$

²⁰Such an orthogonalization may not square well with rational inattention or endogenous learning: in these contexts, the available signals may naturally confound information about current and past innovations, or even about entirely different kinds of fundamentals. The approach taken here is therefore, not a panacea, but rather a sharp instrument for understanding the specific friction we are after in this paper, namely the inertia of first- and higher-order beliefs. The possible confusion of different shocks is a conceptual distinct matter, outside the scope of this paper.

and taking the period- t average expectation, we get

$$\bar{\mathbb{F}}_t^2 [\xi_{t+1}] \equiv \bar{\mathbb{E}}_t [\bar{\mathbb{E}}_{t+1} [\xi_{t+1}]] = \bar{\mathbb{E}}_t \left[\sum_{k=0}^{\infty} \lambda_k \rho_k \eta_{t+1-k} \right] = \sum_{k=0}^{\infty} \lambda_k \lambda_{k+1} \rho_{k+1} \eta_{t-k}$$

Notice here, agents in period t understand that in period $t+1$ the average forecast will be improved, and this is why λ_{k+1} shows up in the expression. By induction, for all $h \geq 2$, the h -th order, forward-looking belief is given by

$$\bar{\mathbb{F}}_t^h [\xi_{t+h-1}] = \sum_{k=0}^{\infty} f_{h,k} \eta_{t-k}, \quad \text{with} \quad f_{h,k} = \lambda_k \lambda_{k+1} \dots \lambda_{k+h-1} \rho_{k+h-1}. \quad (41)$$

The increasing components in the product $\lambda_k \lambda_{k+1} \dots \lambda_{k+h-1}$ seen above capture the anticipation of learning. We revisit this point at the end of this section.

The set of sequences $\mathbf{F}_h = \{f_{h,k}\}_{k=0}^{\infty}$, for $h \geq 2$, provides a complete characterization of the IRFs of the relevant, forward-looking, higher-order beliefs. Note that $\frac{\partial \mathbb{E}[\xi_{t+h} | \eta_{t-k}]}{\partial \eta_{t-k}} = \rho_{k+h-1}$. It follows that the ratio $\frac{f_{h,k}}{\rho_{k+h-1}}$ measures the effect of an innovation on the relevant h -th order belief relative to its effect on the fundamental. When information is complete, this ratio is identically 1 for all k and h . When, instead, information is incomplete, this ratio is given by

$$\frac{f_{h,k}}{\rho_{k+h-1}} = \lambda_k \lambda_{k+1} \dots \lambda_{k+h-1}.$$

The following result is thus immediate.

Proposition 8. Consider the ratio $\frac{f_{h,k}}{\rho_{k+h-1}}$, which measures the effect at lag k of an innovation on the h -th order forward-looking belief relative to its effect on the fundamental.

- (i) For all k and all h , this ratio is strictly between 0 and 1.
- (ii) For any k , this is decreasing in h .
- (iii) For any h , this ratio is increasing in k .
- (iv) As $k \rightarrow \infty$, this ratio converges to 1 for any $h \geq 2$ if and only if it converges for $h = 1$, and this in turn is true if and only if $\lambda_k \rightarrow 1$.

These properties shed light on the dynamic structure of higher-order beliefs. Part (i) states that, for any belief order h and any lag k , the impact of a shock on the h -th order belief is lower than that on the fundamental itself. Part (ii) states that higher-order beliefs move less than lower-order beliefs both on impact and at any lag. Part (iii) states that the gap between the belief of any order and the fundamental decreases as the lag increases; this captures the effect of learning. Part (iv) states that, regardless of h , the gap vanishes in the limit as $k \rightarrow \infty$ if and only if $\lambda_k \rightarrow 1$, that is, if and only if the learning is bounded away from zero.

Myopia and Anchoring. To see how these properties drive the equilibrium behavior, we henceforth restrict $\beta = 0$ and normalize $\varphi = 1$. As noted earlier, the law of motion for the equilibrium outcome is then given by $a_t = \mathbb{E}_t[\xi_t] + \gamma \mathbb{E}_t[a_{t+1}]$, which in turn implies that $a_t = \sum_{h=1}^{\infty} \gamma^{h-1} \mathbb{F}_t^h[\xi_{t+h-1}]$. From the preceding characterization of the higher-order beliefs $\mathbb{F}_t^h[\xi_{t+h-1}]$, it follows that

$$a_t = \sum_{k=0}^{\infty} g_k \eta_{t-k}, \quad \text{with} \quad g_k = \sum_{h=1}^{\infty} \gamma^{h-1} f_{h,k} = \left\{ \sum_{h=1}^{\infty} \gamma^{h-1} \lambda_k \lambda_{k+1} \dots \lambda_{k+h-1} \rho_{k+h-1} \right\}. \quad (42)$$

This makes clear how the IRF of the equilibrium outcome is connected to the IRFs of the first- and higher-order beliefs. Importantly, the higher γ is, the more the dynamics of the equilibrium outcome tracks the dynamics higher-order beliefs relative to the dynamics of lower-order beliefs. On the other hand, when the growth rate of the IRF of the fundamental $\frac{\rho_{k+1}}{\rho_k}$ is higher, it also increases the relative importance of higher-order beliefs.²¹

We are now ready to explain our result regarding myopia. For this purpose, it is best to abstract from learning and focus on how the mere presence of higher-order uncertainty affects the beliefs about the future. In the absence of learning, $\lambda_k = \lambda$ for all k and for some $\lambda \in (0, 1)$. The aforementioned formula for the IRF coefficients then reduces to the following:

$$g_k = \left\{ \sum_{h=1}^{\infty} (\gamma \lambda)^{h-1} \rho_{k+h-1} \right\} \lambda.$$

Clearly, this is the same IRF as that of a complete-information, representative-economy economy in which the equilibrium dynamics satisfy

$$a_t = \xi'_t + \gamma' \mathbb{E}_t[a_{t+1}], \quad (43)$$

where $\xi'_t \equiv \lambda \xi_t$ and $\gamma' \equiv \gamma \lambda$. It is therefore as if the fundamental is less volatile and, in addition, the agents are less forward-looking. The first effect stems from first-order uncertainty: it is present simply because the forecast of the fundamental move less than one-to-one with the true fundamental. The second effect originates in higher-order uncertainty: it is present because the forecasts of the actions of others move *even* less than the forecast of the fundamental.

This is the crux of the forward-looking component of our observational-equivalence result (that is, the one regarding myopia). Note in particular that the extra discounting of the future remains present

²¹The last point is particularly clear if we set $\rho_k = \rho^k$ (meaning that ξ_t follows an AR(1) process). In this case, the initial response is given by

$$g_0 = \sum_{h=1}^{\infty} (\gamma \rho)^{h-1} \lambda_0 \lambda_1 \dots \lambda_{h-1},$$

from which it is evident that the importance of higher-order beliefs increases with both γ and ρ . This further illustrates the point made in Section 3.3 regarding the role of the persistence of the fundamental.

even if when if control for the impact of the informational friction on first-order beliefs. Indeed, replacing ξ'_t with ξ_t in the above shuts down the effect of first-order uncertainty. And yet, the extra discounting survives, reflecting the role of higher-order uncertainty. This complements the related points we make in Section ?? and 4.3.

So far, we shed light on the source of myopia, while shutting down the role of learning. We next elaborate on the robustness of the above insights to the presence of learning and, most importantly, on how the presence of learning and its interaction with higher-order uncertainty drive the backward-looking component of our observational-equivalence result.

To this goal, and as a benchmark for comparison, we consider a variant economy in which all agents share the same subjective belief about ξ_t , this belief happens to coincide with the average first-order belief in the original economy, and these facts are common knowledge. The equilibrium outcome in this economy is proportional to the subjective belief of ξ_t and is given by

$$a_t = \sum_{k=0}^{\infty} \hat{g}_k \eta_{t-k}, \quad \text{with} \quad \hat{g}_k = \sum_{h=1}^{\infty} \gamma^{h-1} \lambda_k \rho_{k+h-1}.$$

This resembles the complete-information benchmark in that the outcome is pinned down by the first-order belief of ξ_t , but allows this belief to adjust sluggishly to the underlying innovations in ξ_t .

By construction, the variant economy preserves the effects of learning on first-order beliefs but shuts down the interaction of learning with higher-order uncertainty. It follows that the comparison of this economy with the original economy reveals the role of this interaction.

Proposition 9. *Let $\{g_k\}$ and $\{\hat{g}_k\}$ denote the Impulse Response Function of the equilibrium outcome in the two economies described above.*

(i) $0 < g_k < \hat{g}_k$ for all $k \geq 0$

(ii) If $\frac{\rho_k}{\rho_{k-1}} \geq \frac{\rho_{k+1}}{\rho_k}$ and $\rho_k > 0$ for all $k > 0$, then $\frac{g_{k+1}}{g_k} > \frac{\hat{g}_{k+1}}{\hat{g}_k}$ for all $k \geq 0$

Consider property (i), in particular the property that $g_k < \hat{g}_k$. This property means that our economy exhibits a uniformly smaller dynamic response for the equilibrium outcome than the aforementioned economy, in which higher-order uncertainty is shut down. But note that the two economies share the following law of motion:

$$a_t = \varphi \bar{\mathbb{E}}_t[\xi_t] + \gamma \bar{\mathbb{E}}_t[a_{t+1}]. \quad (44)$$

Furthermore, the two economies share the same dynamic response for $\bar{\mathbb{E}}_t[\xi_t]$. It follows that the response for a_t in our economy is smaller than that of the variant economy because, and only because, the response of $\bar{\mathbb{E}}_t[a_{t+1}]$ is also smaller in our economy. This verifies that the precise role of higher-order uncertainty is to arrest the response of the expectations of the future outcome (the future actions

of others) beyond and above how much the first-order uncertainty (the unobservability of ξ_t) arrests the response of the expectations of the future fundamental.

A complementary way of seeing this point is to note that g_k satisfies the following recursion:

$$g_k = f_{1,k} + \lambda_k \gamma g_{k+1}. \quad (45)$$

The first term in the right-hand side of this recursion corresponds to the average expectation of the future fundamental. The second term corresponds the average expectation of the future outcome (the actions of others). The role of first-order uncertainty is captured by the fact that $f_{1,k}$ is lower than ρ_k . The role of higher-order uncertainty is captured by the presence of λ_k in the second term: it is as if the discount factor γ has been replaced by a discount factor equal to $\lambda_k \gamma$, which is strictly less than γ . This represents a generalization of the form of myopia seen in condition (43). There, learning was shut down, so that that λ_k and the extra discounting of the future were invariant in the horizon k . Here, the additional discounting varies with the horizon because of the anticipation of future learning (namely, the knowledge that λ_k will increase with k).

Consider next property (ii), namely the property that

$$\frac{g_{k+1}}{g_k} > \frac{\widehat{g}_{k+1}}{\widehat{g}_k}$$

This property helps explain the backward-looking component of our observational-equivalence result (that is, the one regarding anchoring).

To start with, consider the variant economy, in which higher-order uncertainty is shut down. The impact of a shock $k + 1$ periods from now relative to its impact k periods from now is given by

$$\frac{\widehat{g}_{k+1}}{\widehat{g}_k} = \frac{\lambda_{k+1}}{\lambda_k} \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}} > \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}.$$

The inequality captures the effect of learning on first-order beliefs. Had information being perfect, we would have had $\frac{\widehat{g}_{k+1}}{\widehat{g}_k} = \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}$; now, we instead have $\frac{\widehat{g}_{k+1}}{\widehat{g}_k} > \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}$. This means that, in the variant economy, the impact of the shock on the equilibrium outcome can build force over time because, and only because, learning allows for a gradual build up in first-order beliefs.²²

Consider now our economy, in which higher-order uncertainty is present. We now have

$$\frac{g_{k+1}}{g_k} > \frac{\widehat{g}_{k+1}}{\widehat{g}_k}$$

²²This is easiest to see when $\rho_k = 1$ (i.e., the fundamental follows a random walk), for then \widehat{g}_{k+1} is necessarily higher than \widehat{g}_k for all k . In the AR(1) case where $\rho_k = \rho^k$ with $\rho < 1$, \widehat{g}_{k+1} can be either higher or lower than \widehat{g}_k , depending on the balance between two opposing forces: the build-up effect of learning and the mean-reversion in the fundamental.

This means that higher-order uncertainty amplifies the build-up effect of learning: as time passes, the impact of the shock on the equilibrium outcome builds force more rapidly in our economy than in the variant economy. But since the impact is always lower in our economy,²³ this means that the IRF of the equilibrium outcome is likely to display a more pronounced hump shape in our economy than in the variant economy. Indeed, the following is a directly corollary of the above property.

Corollary 3. *Let the variant economy display a hump-shaped response: $\{\hat{g}_k\}$ is single peaked at $k = k^b$ for some $k^b \geq 1$. Then, the equilibrium outcome also displays a hump-shaped response: $\{g_k\}$ is also single peaked at $k = k^g$. Furthermore, the peak of the equilibrium response is after the peak of the variant economy: $k^g \geq k^b$ necessarily, and $k^g > k^b$ for an open set of $\{\lambda_k\}$ sequences.*

To interpret this result, think of k as a continuous variable and, similarly, think of λ_k , \hat{g}_k , and g_k as differentiable functions of k . If \hat{g}_k is hump-shaped with a peak at $k = k_b > 0$, it must be that \hat{g}_k is weakly increasing prior to k_b and locally flat at k_b . But since we have proved that the growth rate of g_k is strictly higher than that of \hat{g}_k , this means that g_k attains its maximum at a point k_g that is strictly above k^b . In the result stated above, the logic is the same. The only twist is that, because k is discrete, we must either relax $k_g > k_b$ to $k_g \geq k_b$ or put restrictions on $\{\lambda_k\}$ so as to guarantee that $k_g \geq k_b + 1$.

Summing up, learning by itself contributes towards a gradual build up of the impact of any given shock on the equilibrium outcome; but its interaction with higher-order uncertainty makes this build up even more pronounced. It is precisely these properties that are encapsulated in the backward-looking component of our observational equivalence result: the coefficient ω_b , which captures the endogenous build up in the equilibrium dynamics, is positive because of learning and it is higher the higher the importance of higher-order uncertainty.

Two Forms of Bounded Rationality. We now shed light on two additional points, which were anticipated earlier on: the role played by the anticipation that others will learn in the future; and the possible interaction of incomplete information with Level-k Thinking.

To illustrate the first point, we consider a behavioral variant where agents fail to anticipate that others will learn in the future. To simplify, we also set $\beta = 0$. Recall from equation (41), when agents are rational, the forward higher-order beliefs are

$$\bar{\mathbb{F}}_t^h [\xi_{t+h-1}] = \sum_{k=0}^{\infty} \lambda_k \lambda_{k+1} \dots \lambda_{k+h-1} \rho_{k+h-1} \eta_{t-k}.$$

In the variant economy, by shutting down the anticipation of learning, the nature of higher-order beliefs changes, as $\mathbb{E}_{it} [\bar{\mathbb{E}}_{t+k} [\xi_{t+q}]] = \mathbb{E}_{it} [\bar{\mathbb{E}}_t [\xi_{t+q}]]$ for $k, q \geq 0$, and the counterpart of $\bar{\mathbb{F}}_t^h [\xi_{t+h-1}]$

²³Recall, this is by property (i) of Proposition 9.

becomes

$$\bar{\mathbb{E}}_t^h [\xi_{t+h-1}] \equiv \bar{\mathbb{E}}_t [\bar{\mathbb{E}}_t [\dots \bar{\mathbb{E}}_t [\xi_{t+h-1}] \dots]] = \sum_{k=0}^{\infty} \lambda_k^h \rho_{t+h-1} \eta_{t-k}.$$

Learning implies $\lambda_{k+1} > \lambda_k$, and the anticipation of learning implies $\lambda_k \lambda_{k+1} \dots \lambda_{k+h-1} > \lambda_k^h$. As a result, higher-order beliefs in the behavioral variant under consideration vary *less* than those under rational expectations. By the same token, the aggregate outcome in this economy, which is given

$$a_t = \sum_{h=1}^{\infty} \gamma^{h-1} \bar{\mathbb{E}}_t^h [\xi_{t+h-1}],$$

behaves as if the myopia and anchoring are stronger than in the rational-expectations counterpart. In line with these observations, it can be shown that, if we go back to our baseline specification and impose that agents fail to anticipate that others will learn in the future, Proposition 3 continues to hold with the following modification: ω_f is smaller and ω_b is higher.

To illustrate the second point, we consider a variant that lets agents have limited depth of reasoning in the sense of Level- k Thinking. With level-0 thinking, agents believe that the aggregate outcome is fixed at zero for all t , but still form rational beliefs about the fundamental. Therefore, $a_{it}^0 = \mathbb{E}_{it}[\xi_t]$, and the implied aggregate outcome for level-0 thinking is $a_t^0 = \bar{\mathbb{E}}_t[\xi_t]$.

With level-1 thinking, agent i 's action changes to

$$a_{it}^1 = \mathbb{E}_{it}[\xi_t] + \gamma \mathbb{E}_{it}[a_{t+1}^0] = \mathbb{E}_{it}[\xi_t] + \gamma \mathbb{E}_{it} [\bar{\mathbb{E}}_{t+1}[\xi_{t+1}]],$$

where the second-order higher-order belief shows up. By induction, the level- k outcome is given by

$$a_t^k = \sum_{h=1}^{k+1} \gamma^{h-1} \bar{\mathbb{F}}_t^h [\xi_{t+h-1}].$$

In a nutshell, Level- k Thinking truncates the hierarchy of beliefs at a finite order.

Compared with the rational-expectations economy that has been the focus of our analysis, the GE feedback effects in both of the aforementioned two variants are attenuated, and the resulting as-if myopia is strengthened. Furthermore, by selecting the depth of thinking, we can make sure that the second variant produces a similar degree of myopia as the first one.²⁴ That said, the source of the additional myopia is different. In the first, the relevant forward-looking higher-order beliefs have been replaced by myopic counterparts, which move less. In the second, the right, forward-looking higher-order beliefs are still at work, but they have been truncated at a finite point.

²⁴This follows directly from the fact that impact of effect of an innovation in the first variant is bounded between those of the level-0 and the level- ∞ outcome in the second variant.