Myopia and Anchoring*

George-Marios Angeletos† **Zhen Huo**‡

November 8, 2018

Abstract

We consider a stationary setting featuring forward-looking behavior, strategic complementarity, and incomplete information. We obtain an observational equivalence result that recasts the aggregate dynamics of this setting as that of a representative-agent model featuring two distortions: myopia, as in models with imperfect foresight; and anchoring of the current outcome to the past outcome, as in models with habit persistence and adjustment costs. We further show that the as-if distortions are larger when the general-equilibrium feedback, or the strategic complementarity, is stronger. These results offer a fresh perspective on the observable implications of informational frictions; build a useful bridge to the DSGE literature; and help reduce an uncomfortable gap between the prevailing structural interpretations of the macroeconomic time series and the related microeconomic evidence. Finally, an empirical evaluation is offered in the context of inflation, wherein it is shown how our results can rationalize existing estimates of the Hybrid NKPC while also matching survey evidence on expectations.

*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-Moller 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

Forward-looking behavior and equilibrium feedbacks are central to our understanding of business cycles, asset-price fluctuations, industry dynamics, and more. In this paper, we open a new window into the interplay of these features with a realistic friction in the amount of knowledge agents have about the fundamentals and about one another's beliefs and actions.

We first establish an observational equivalence between a rational-expectations setting featuring such an informational 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 the as-if distortions are larger when the general-equilibrium (GE) feedback, or the strategic complementarity, is stronger, and elaborate on the underlying theoretical principles. We finally build useful connections to multiple strands of the literature, address a disturbing disconnect between the prevailing structural interpretations of the macroeconomic time series and the related microeconomic evidence, and offer an empirical evaluation in the context of inflation.

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:

$$a_t = \varphi \xi_t + \delta \mathbb{E}_t \left[a_{t+1} \right], \tag{1}$$

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. These examples indicate the broader applicability of the results we develop in this paper.[1]

We depart from the representative-agent benchmark by allowing information to be incomplete (i.e., noisy and heterogeneous). This can be the product of either dispersed private information (Lucas, 1972; Morris and Shin, 2002) or rational inattention (Sims, 2003). Either way, the key is that we accommodate, not only *first-order* uncertainty (imperfect knowledge of the underlying fundamental), but also *higher-order* uncertainty (uncertainty about the beliefs and actions of others). That is, we let agents face realistic doubts about the awareness, attentiveness, or responsiveness of *others*.

[1] Applications not explored here may include dynamic Bertrand competition and markets with network externalities.

Main result. Under appropriate assumptions, the incomplete-information economy is shown to be observationally equivalent to a variant, complete-information, representative-agent economy in which condition (1) is modified as follows:

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

for some $\omega_f < 1$ and $\omega_b > 0$. The first modification ($\omega_f < 1$) represents myopia towards the future, the second ($\omega_b > 0$) anchors the current outcome to the past outcome. The one dulls the forward-looking behavior, the other adds a backward-looking element. Furthermore, both distortions are shown to increase with the strength of the GE feedback, or the degree of the strategic complementarity.[2]

The exact form of this result relies on strong assumptions about the stochastic process for ξ_t and the information structure. It nevertheless stylizes a few more general insights, whose robustness we document in Section 7 under a flexible specification of the stochasticity and the information.

One key insight, which builds on Angeletos and Lian (2018) and Morris and Shin (2006), explains the myopia. In our setting, behavior depends on expectations of the *future* actions of others—e.g., current inflation depends on expectations of future inflation, current spending depends on expectations of future spending, etc. In equilibrium, such expectations are pinned down by the current beliefs of the future beliefs of others. Because such higher-order beliefs tend to vary less than the corresponding first-order beliefs in response to any innovation in ξ_t, it is *as if* the agents discount the future outcomes more heavily than in the frictionless, representative-agent benchmark. Furthermore, because the dependence of equilibrium behavior on higher-order beliefs increases with the GE feedback, or the strategic complementarity, this kind of myopia also increases with it.[3]

Another key insight, which builds on Woodford (2003) and Nimark (2008), regards the role played by learning. Learning induces extra persistence in the beliefs of the fundamental and of the future outcomes. Furthermore, this persistence is stronger in the latter beliefs than in the former, reflecting the smaller dependence of higher-order beliefs on recent information. This explains why the current outcome appears to be anchored to the past outcome at a rate that, like our form of myopia, increases with the strength of the GE feedback, or the strategic complementarity.

We add to th'e literature, not only by blending these insights and offering a few more, but also by operationalizing them in terms of our observational-equivalence result, which is new. Also new is the

[2] Such GE feedbacks are often "hidden" behind the kind of representative-agent equilibrium conditions herein stylized by condition (1). They include the Keynesian income-spending multiplier in the context of the Dynamic IS curve, the dynamic strategic complementarity in the firms' price-setting decisions in the context of the NKPC, and the positive feedback from expectations of future trades to current trades in the context of asset pricing.

[3] To be precise, our observational-equivalence result combines the aforementioned effect with an additional effect, which is that the first-order beliefs of the fundamental also move less than in the frictionless benchmark. As explained in Section 7, this effect is not strictly needed: higher-order uncertainty is alone sufficient for the documented form of myopia. However, the two effects may naturally come together in applications and only complement each other.

analysis in Section 7. This allows for a flexible specification of the stochasticity and the higher-order belief dynamics, while also cutting the Gordian knot of complex signal-extraction problems, so as to develop a sharp understanding of the principles that underly our observational-equivalence result.

Together, these results offer a fresh perspective on the implications of informational frictions. They allow us to draw useful connections to the literature. And they facilitate our empirical exercise in the context of inflation. We elaborate on these applied aspects of our contribution below.

DSGE. Baseline macroeconomic models emphasize forward-looking behavior but have hard time capturing a salient feature of the aggregate time series: inflation, consumption and investment alike appear to respond sluggishly to the underlying innovations, as if there is a strong backward-looking component in their laws of motion. To address this challenge and provide a successful structural interpretation of the data, the DSGE literature has sacrificed on the micro-foundations.

For instance, to match the inflation dynamics, the literature has followed Gali and Gertler (1999) and Christiano, Eichenbaum, and Evans (2005) in replacing the standard NKPC with the so-called Hybrid NKPC, a backward-looking variant that finds no support in the related menu-cost literature (Alvarez and Lippi, 2014; Golosov and Lucas Jr, 2007; Nakamura and Steinsson, 2013). Similarly, to match the aggregate investment dynamics, the literature has employed a form of adjustment cost that is incompatible with both standard Q theory (Hayashi, 1982) and the literature that studies investment at the micro level (Bachmann, Caballero, and Engel, 2013; Caballero and Engel, 1999).

Our observational-equivalence result offers the sharpest possible illustration of how informational frictions can provide a plausible micro-foundation for the DSGE add-ons: accommodating incomplete information is akin to adding habit persistence in consumption, adjustment costs in investment, and a backward-looking element in the NKPC. At the same time, our result indicates that the as-if distortions may be endogenous to GE feedback mechanisms and thereby also to policies that regulate the such mechanisms.[4]

Micro vs Macro. The GE feedback is active, and higher-order beliefs are relevant, only when agents respond to aggregate shocks. It follows that the documented forms of myopia and anchoring may loom large at the macro level even if they appear to be small in micro data. This provides a simple, unified explanation to why the macroeconomic estimates of both the habit persistence in consumption and the adjustment costs in investment are much higher their microeconomic counterparts (Havranek, Rusnak, and Sokolova, 2017; Groth and Khan, 2010; Zorn, 2018); why the persistence of inflation is higher in the aggregate time series than in disaggregated data (Altissimo et al., 2010); and perhaps even why the momentum in asset prices is more pronounced at the stock-market level than at the individual-stock level (Jung and Shiller, 2005).

[4]Eearlier works such as Sims (2003), Mankiw and Reis (2002, 2007), Mackowiak and Wiederholt (2009, 2015), and Nimark (2008) share the idea that informational frictions can substitute for the DSGE add-ons, but do not contain our observational-equivalence result and our insights about GE effects and the micro-to-macro gap, which we discuss next.

Bounded Rationality. The form of myopia, or imperfect foresight, documented here is rationalized by an informational friction, as in Angeletos and Lian (2018). Similar forms of imperfect foresight are obtained in Garcıa-Schmidt and Woodford (2018) and Farhi and Werning (2017) by replacing rational expectations with Level-k Thinking, and in Gabaix (2017) by introducing a belief bias called "cognitive discounting." These works, however, do not produce the kinds of backward-looking behavior and belief momentum that characterize our approach. In terms of condition (2), they effectively let $\omega_f < 1$ but restrict $\omega_b = 0$. By contrast, the data appear to demand both $\omega_f < 1$ and $\omega_b > 0$.[5] This favors of our approach over the aforementioned alternatives, but also invites us to study two variants that combine incomplete information with bounded rationality.

Application to Inflation. Our contribution is completed with an emprical exercise in the context of inflation. We revisit the micro-foundations of the NKPC, adding incomplete information. The general formulation of the incomplete-information NKPC turns out to be too cumbersome to estimate. By imposing additional structure, our observational-equivalence result bypasses this obstacle and facilitates a simple and sharp mapping to the data.

In particular, we first show that our results help rationalize existing estimates of the Hybrid NKPC, such as those found in Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2005). We next show that this is achieved with an informational friction that *also* matches the evidence on inflation expectations provided by Coibion and Gorodnichenko (2015).

Here, it is useful to clarify the following point. By measuring the predictability of the average forecast error in surveys, Coibion and Gorodnichenko (2015) provided a key empirical moment that helps gauge the level of the informational friction. Yet, by treating inflation as an exogenous stochastic process, this work could not possibly quantify the equilibrium impact of this friction on the actual inflation dynamics. Our paper fills this gap by solving the fixed point between inflation expectations and actual inflation, by connecting the aforementioned empirical moment to its theoretical counterpart, and by evaluating the implied bite on the equilibrium outcomes.

Layout. The rest of the paper is organized as follows. Section 2 expands on the related literature. Section 3 introduces our framework. Section 4 develops the observational-equivalence result. Section 5 illustrates the applicability of this result and discussed its value-added. Section 6 contains our quantitative exercise in the context of inflation. Section 7 illustrates the robustness of the insights underlying our observational-equivalence result. Section 8 concludes. The Appendices contain proofs and a few additional results.

[5] While the literature on the forward-guidance puzzle (Del Negro, Giannoni, and Patterson, 2015; McKay, Nakamura, and Steinsson, 2016; Angeletos and Lian, 2018; Farhi and Werning, 2017) has focused on getting extra discounting, the DSGE and SVARs literatures have long pointed out the need of a strong backward-looking component. Furthermore, the available evidence on expectations (e.g., Coibion and Gorodnichenko, 2012, 2015; Coibion, Gorodnichenko, and Kumar, 2015) suggests, not only that the average forecasts of future outcomes respond little on impact to aggregate shocks, but also that they adjust more and more with the passage of time. The first property maps to $\omega_f < 1$, the second to $\omega_b > 0$. Finally, Coibion et al. (2018) offers additional supporting evidence by directly soliciting higher-order beliefs in a survey.

2 Related Literature

As already noted, our paper builds heavily on the existing literature on incomplete information and higher-order beliefs, most notably on Morris and Shin (2002, 2006), Woodford (2003), Nimark (2008), and Angeletos and Lian (2018); see Angeletos and Lian (2016) for a review and additional references. Our main contributions vis-a-vis this literature are: (i) the observational-equivalence result and its applications; (ii) the bridges built to three other strands of the literature, on DSGE, on micro-to-macro, and on bounded rationality; (iii) the empirical exercise in the context of inflation; and (iv) the robustness analysis in Section 7. Our paper also adds to the agenda of Sims (2003, 2010) by shedding new light on how rational inattention, as a source of incomplete information, interacts with GE mechanisms in forward-looking models.

On the methodological front, our paper builds on a literature that bypasses the higher-order beliefs and attacks directly the rational-expectations fixed point; see, e.g., Kasa, Walker, and Whiteman (2014), Huo and Takayama (2018) and the references therein. In particular, we utilize the methods of Huo and Takayama (2018) in order to solve explicitly for the rational-expectations equilibrium under our baseline specification. We then translate the obtained equilibrium in the form of our observational-equivalence result and study its comparative statics with respect to the strength of GE feedback and other parameters. What is new here is the second step, the various applied lessons that derive from it, and the analysis in Section 7, which follows a different methodological route.

On the applied front, our paper is most closely related to Nimark (2008). This paper is the first to study the interaction of sticky prices and incomplete information and shares the basic idea of attributing the backward-looking component of the Hybrid NKPC to the inertia of higher-order beliefs. It does not, however, contain either our analytical results or the tight connection we draw between the estimates of the Hybrid NKPC (Gali and Gertler, 1999; Gali, Gertler, and Lopez-Salido, 2005) and the evidence on inflation expectations (Coibion and Gorodnichenko, 2012, 2015).[6]

Mackowiak and Wiederholt (2009) argue that rational inattention can reconcile price rigidity at the macro level with price flexibility at the micro level.[7] Although this message sounds similar to ours regarding the gap between micro and macro, the mechanism is different. In that work, the gap is explained by greater first-order uncertainty about aggregate shocks than idiosyncratic shocks. In our paper, instead, the gap is attributed to GE effects and higher-order uncertainty; it is therefore present *even if* the first-order uncertainty about, or the attention to, the two kinds of shocks is the same. The two mechanisms are nevertheless complementary and naturally come together in applications.

[6]These points distinguish more broadly our contribution from a larger literature, including Mankiw and Reis (2002), Reis (2006), Kiley (2007), Melosi (2016), and Matejka (2016), that studies price-setting in the presence of informational frictions. A few other works estimate the standard NKPC after replacing the representative agent's expectation of inflation with the average forecast in surveys. Mavroeidis, Plagborg-Møller, and Stock (2014) note that this approach lacks solid theoretical foundations. It is indeed inconsistent with the micro-foundations laid out here.

[7]See Zorn (2018) and Carroll et al. (2018) for related points in the context of, respectively, investment and consumption.

Bordalo et al. (2018) and Kohlhas and Walther (2018a,b) argue that the expectations data paint a more varied picture than the one contained in our paper and the existing literature on informational frictions: while average forecasts tend to under-react to aggregate shocks, individual forecasts tend to over-react to idiosyncratic news, suggesting either a particular departure from rational expectations (Bordalo et al., 2018) or a more elaborate informational friction (Kohlhas and Walther, 2018a,b). The interplay of these ideas with the mechanisms we study is an important question, beyond the scope of this paper. Nevertheless, the most relevant fact for our purposes is the under-reaction of average forecasts to aggregate shocks, which is consistent with our theory.

The relation of our paper to the literature on Level-k Thinking has already been commented on and is further discussed in the last part of Section 5. Related is also the literature on adaptive learning (Sargent, 1993; Evans and Honkapohja, 2012; Marcet and Nicolini, 2003). This literature allows for the anchoring of current outcomes to past outcomes; see, in particular, Carvalho et al. (2017) for an application in the context of inflation. The anchoring found in our paper has three distinct qualities: it is consistent with rational expectations; it is tied to the strength of the GE feedback; and it is directly comparable to that found in the DSGE literature.

3 The Abstract Framework

In this section we set up our framework, review the frictionless, complete-information benchmark we depart from, and illustrate the interaction of forward-looking behavior and higher-order beliefs.

Set up. Time is discrete, indexed by $t \in \{0, 1, ...\}$, 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:

$$a_{it} = \mathbb{E}_{it} \left[\varphi \xi_t + \beta a_{it+1} + \gamma a_{t+1} \right] \tag{3}$$

where ξ_t is the exogenous fundamental, $\mathbb{E}_{it}[\cdot]$ is the rational-expectation operator conditional on the period-t information of player i, and (φ, β, γ) are parameters, with $\varphi > 0$, $\beta, \gamma \in [0, 1)$, and $\beta + \gamma < 1$.

Condition (3) specifies best responses in recursive form. To see more clearly how current behavior depends on expectations of the entire future paths of the fundamental and of the average action, we iterate this condition forward and reach the following extensive-form representation:

$$a_{i,t} = \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t} \left[\varphi \xi_{t+k} \right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t} \left[a_{t+k+1} \right], \tag{4}$$

Aggregating the above condition across agents, we then also obtain the following equilibrium restriction between outcomes and expectations:

$$a_t = \varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[\xi_{t+k}\right] + \gamma \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[a_{t+k+1}\right], \tag{5}$$

where $\overline{\mathbb{E}}_t[.]$ denotes the average expectation in the cross-section of the population.[8]

The last condition is useful, not only because it facilitates the characterization of the aggregate outcome as a function of first- and higher-order beliefs (see below), but also because it helps nest applications in which a direct analogue to the individual-level best-response condition (3) is unavailable, as in cases where a_t corresponds to the price determined in a Walrasian market. Finally, either of these conditions makes clear that the scalars β and γ parameterize two distinct aspects of forward-looking behavior. On the one hand, β determines the extent to which an individual discounts the future values of either the exogenous fundamental or the endogenous outcome. On the other hand, γ regulates the extent to which an individual conditions his current behavior on her expectations of the future actions of *others*. In applications, this kind of dynamic strategic complementarity corresponds to GE effects such as the positive feedback from expectations of future inflation to current inflation, or that from expectations of future aggregate spending to current aggregate spending.

Frictionless, Representative-Agent Benchmark. Suppose, momentarily, that information is complete, by which we mean that all agents share the same information and therefore face no uncertainty about one another's beliefs. In this case, we can replace $\overline{\mathbb{E}}_t[\cdot]$ in condition (5) with the expectation of a representative agent, that is, the expectation conditional on the common information. Regardless of how noisy that common information might be, we can then use the Law of Iterated Expectations to reduce condition (5) to the following, representative-agent, Euler-like condition:

$$a_t = \mathbb{E}_t[\varphi \xi_t + \delta a_{t+1}], \tag{6}$$

where $\mathbb{E}_t[\cdot]$ denotes the expectation of the representative agent and $\delta \equiv \beta + \gamma \in (0, 1)$.

It is then immediate to see that the complete-information version of our framework nests the two building blocks of the New Keynesian model: the NKPC is nested with a_t standing for inflation and ξ_t for the real marginal cost or the output gap; and the Dynamic IS Cure (that is, the Euler condition

[8] The 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), Angeletos and La'O (2010), 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). The implications of this subtle difference are discussed as we proceed; they include the dependence of the documented form of myopia on the persistence of the fundamental and on the anticipation of the future learning of others.

of the representative consumer) is nested with a_t standing for consumption and ξ_t for the real interest rate. Alternatively, condition (6) can represent an asset-pricing equation with a_t standing for the asset price and ξ_t for the next-period dividend.

By iterating condition (6), we can obtain the complete-information outcome as follows:

$$a_t = \varphi \sum_{k=0}^{\infty} \delta^k \mathbb{E}_t[\xi_{t+k}]. \tag{7}$$

This stylizes how, in the aforementioned applications and beyond, outcomes are pinned down by first-order beliefs of fundamentals. As for extensions of these applications that add incomplete information, we will later show how such extensions can indeed be nested in condition (5).

Incomplete Information and Higher-Order Beliefs. Once information is incomplete, condition (7) ceases to hold and, instead, the aggregate outcome hinges on a certain kind of forward-looking, higher-order beliefs. Later on, appropriate assumptions about the information structure will permit an explicit characterization of these beliefs. For now, we explain how the aggregate outcome can be expressed as a function of these beliefs *regardless* of the information structure.

To illustrate, let $\beta = 0 < \gamma$,[9] which means that the equilibrium outcome satisfies

$$a_t = \varphi \overline{\mathbb{E}}_t \left[\xi_t\right] + \gamma \overline{\mathbb{E}}_t \left[a_{t+1}\right]. \tag{8}$$

Iterating the above condition once gives

$$a_t = \varphi \overline{\mathbb{E}}_t \left[\xi_t\right] + \gamma \varphi \overline{\mathbb{E}}_t \left[\overline{\mathbb{E}}_{t+1} \left[\xi_{t+1}\right]\right] + \gamma^2 \overline{\mathbb{E}}_t \left[\overline{\mathbb{E}}_{t+1} \left[a_{t+2}\right]\right],$$

from which it is evident that the equilibrium outcome depends on a particular kind of forward-looking, second-order beliefs, namely the current beliefs of the next-period beliefs of the next-period fundamental and the next-period outcome (see second and third term in the above condition, respectively). By iterating condition (8) again and again, we can ultimately express the equilibrium outcome as a function of an infinite hierarchy of beliefs about the current and future values of the fundamental:

$$a_t = \varphi \sum_{h=0}^{\infty} \gamma^h \overline{\mathbb{F}}_t^{h+1} \left[\xi_{t+h}\right] \tag{9}$$

where, for any random variable X, $\overline{\mathbb{F}}_t^1 [X] \equiv \overline{\mathbb{E}}_t [X]$ and $\overline{\mathbb{F}}_t^h [X] \equiv \overline{\mathbb{E}}_t \left[\overline{\mathbb{F}}_{t+1}^{h-1} [X]\right]$ for all $h \geq 2$.

Consider next the more general case in which both $\gamma > 0$ and $\beta > 0$. In this case, which is relevant for the applications of interest, the class of higher-order beliefs that drive the equilibrium outcome is

[9] This case is too narrow for the applications to inflation and aggregate demand, which we study later on; but it is relatively simple and nests the games studied in Allen, Morris, and Shin (2006), Morris and Shin (2006) and Nimark (2017).

much richer than the one described above. To see this, let $\chi_t \equiv \varphi \sum_{k=0}^{\infty} \beta^k \xi_{t+k}$ and rewrite condition (8) as follows:

$$a_t = \overline{\mathbb{E}}_t\left[\chi_t\right] + \gamma \sum_{k=1}^{\infty} \beta^{k-1} \overline{\mathbb{E}}_t\left[a_{t+k}\right]$$

Applying this condition to period $t+k$, for any $k \geq 1$, and taking the expectations as of period t, we obtain the following representation of the period-t beliefs of the future outcomes:

$$\overline{\mathbb{E}}_t[a_{t+k}] = \overline{\mathbb{E}}_t\left[\overline{\mathbb{E}}_{t+k}\left[\chi_{t+k}\right]\right] + \gamma \sum_{j=1}^{\infty} \beta^{j-1} \overline{\mathbb{E}}_t\left[\overline{\mathbb{E}}_{t+k}\left[a_{t+k+j}\right]\right]$$

Combining and rearranging, we reach the following characterization of the period-t outcome:

$$a_t = \overline{\mathbb{E}}_t\left[\chi_t\right] + \gamma \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}_t\left[\overline{\mathbb{E}}_{t+k}\left[\chi_{t+k}\right]\right] + \gamma^2 \left\{ \sum_{k=1}^{\infty} \beta^{k-1} \sum_{j=1}^{\infty} \beta^{j-1} \overline{\mathbb{E}}_t\left[\overline{\mathbb{E}}_{t+k}\left[a_{t+k+j}\right]\right] \right\}$$

The relevant second-order beliefs are therefore those regarding the beliefs of others, not only in the next period, but also in *all* future periods, namely $\overline{\mathbb{E}}_t[\overline{\mathbb{E}}_{t+k}[\chi_{t+k}]]$ for every $k \geq 1$. Intuitively, because $\beta > 0$ means that an individual's decision depends on her own first-order belief of the entire future path of the both the exogenous fundamental and the endogenous outcome, $\beta > 0$ also implies that the second-order beliefs of the entire future path of the fundamental become relevant.

As we iterate this argument again and again, the set of higher-order beliefs that emerge gets richer and richer. In particular, fix a t and pick any $k \geq 2$, any $h \in \{2, ..., k\}$, and any $\{t_1, t_2, ...t_h\}$ such that $t = t_1 < t_2 < ... < t_h = t + k$. Then, the period-$t$ outcome depends on all of the following types of forward-looking higher-order beliefs:

$$\overline{\mathbb{E}}_{t_1}[\overline{\mathbb{E}}_{t_2}[\cdots[\overline{\mathbb{E}}_{t_h}[\chi_{t+k}]\cdots]].$$

For any t and any $k \geq 2$, there are $k-1$ 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.

What's next. The above derivations indicate the potential complexity of dynamic, incomplete-information models. An integral part of our contribution is the bypassing of this complexity and the development of sharp analytical results. This is achieved by following two different, but complementary, paths. In Sections 4-6, we employ a somewhat rigid specification of the fundamental process and the information structure so as to facilitate our observation-equivalence result and its various applications. In Section 7, we then use a more flexible specification but also cut the Gordian knot of complex signal-extraction problems so as to shed further light on the underlying principles and their robustness; this sacrifices the elegance of our observational-equivalence result but not its essence.

4 The Equivalence Result

In this section we develop our observation-equivalence result and discuss the broader insights that are encapsulated in it.

4.1 Specification

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

$$\xi_t = \rho \xi_{t-1} + \eta_t = \frac{1}{1-\rho L}\eta_t, \tag{10}$$

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

$$x_{it} = \xi_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \sigma^2) \tag{11}$$

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.

These assumptions are restrictive. They rule out, inter alia, endogenous signals such as prices. The main justification for them is that they guarantee the exact validity of our observational-equivalence result. This result may nevertheless serve as a good proxy of the equilibrium even in extensions that allow for endogenous signals; we illustrate this in Appendix C. Furthermore, the presence of idiosyncratic noise in the available information can be motivated as the byproduct of rational inattention, subject to the caveat that we do not study the problem of finding the optimal signal.

All in all, we view our baseline specification as a useful and empirically plausible perturbation of the complete-information benchmark. This benchmark, which is herein nested by setting $\sigma = 0$, imposes, not only that every agent knows the current value of ξ_t, but also that she is confident that every other agent shares the same beliefs with him about the entire future path of both ξ_t and a_t. By contrast, setting $\sigma > 0$ lets agents face, not only uncertainty about the underlying aggregate shocks, but also doubts about the attentiveness, awareness, and responsiveness of others.

4.2 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 (7). 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*

$$a_t = a_t^* \equiv \frac{\varphi}{1-\rho\delta}\xi_t = \frac{\varphi}{1-\rho\delta}\frac{1}{1-\rho L}\eta_t. \tag{12}$$

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 (9). Using the Kalman filter, we can readily show that the first-order belief $\overline{\mathbb{E}}_t[\xi_t]$ follows an AR(2) process:

$$\overline{\mathbb{E}}_t[\xi_t] = \left(1-\frac{\lambda}{\rho}\right)\left(\frac{1}{1-\lambda L}\right)\xi_t = \left(1-\frac{\lambda}{\rho}\right)\left(\frac{1}{1-\lambda L}\right)\left(\frac{1}{1-\rho L}\right)\eta_t, \tag{13}$$

where $\lambda = \rho(1-G)$ and G is the Kalman gain. This implies that the second-order belief of the type $\overline{\mathbb{E}}_t[\overline{\mathbb{E}}_{t+1}[\xi_{t+1}]]$ follows an ARMA(3,1) and, by induction, for any $h \geq 1$, the h-th order belief of the type $\overline{\mathbb{E}}_t[\overline{\mathbb{E}}_{t+1}[...\overline{\mathbb{E}}_{t+h}[\xi_{t+h}]]$ 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.

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

$$a_t = \left(1-\frac{\vartheta}{\rho}\right)\left(\frac{1}{1-\vartheta L}\right)a_t^*, \tag{14}$$

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 (14) 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 (14), and a more sluggish build up over time, captured by the lag term ϑL.

To develop some intuition for 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 (5) along with $\gamma = 0$ (and hence $\delta \equiv \beta + \gamma = \beta$) and the fact that $\overline{\mathbb{E}}_t[\xi_{t+k}] = \rho^k \overline{\mathbb{E}}_t[\xi_t]$ for all $k \geq 0$, we infer that the aggregate outcome is given by

$$a_t = \varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t[\xi_{t+k}] = \frac{\varphi}{1 - \delta\rho} \overline{\mathbb{E}}_t[\xi_t]. \tag{15}$$

This is the same as the complete-information outcome, modulo the replacement of ξ_t, the actual fundamental, with $\overline{\mathbb{E}}_t[\xi_t]$, the average first-order forecast of it. And since the latter follows the AR(2) process given in condition (13), 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 $\gamma = 0$, albeit with different coefficients.

Suppose that one had guessed a priori that the equilibrium follows an AR(2) process, but did not know the coefficients of this process. For any such guess, one could have computed all the expectations showing up in the right-hand side of condition (5), which we repeat here:

$$a_t = \varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[\xi_{t+k}\right] + \gamma \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[a_{t+k+1}\right].$$

For the guess to be correct, all these expectations would have to sum up to an AR(2) process, with the same coefficients as the guess. Proposition 2 effectively establishes that this is true if and only if the guess satisfies the following three restrictions: (i) the first root of the AR(2) process is the same as the persistence of the fundamental or, equivalently, of the complete-information outcome; (ii) its volatility is tied to its second root in the manner seen in condition (14); and (iii) its second root is given by the inverse of the highest root of the provided cubic. The proof provided in the Appendix derives these restrictions via a different method, which permits one to construct the equilibrium without having to guess it and also establishes its uniqueness.

But let us put aside the technical asmipects of Proposition 2 and, instead, focus on its economic content. When $\gamma > 0$, ϑ is strictly higher than λ.[10] That is, the equilibrium dynamics exhibits less amplitude and more persistence, not only relative to the complete-information counterpart, but also relative to first-order beliefs. This because the equilibrium depends on higher-order beliefs, which themselves display less amplitude and more persistence than first-order beliefs.

[10] This fact may not be obvious from looking at Proposition 2, but follows from the property that ϑ is increasing in γ, which is established as a part of the proof of Proposition 4.

In Section 7, we make this logic clear, and elaborate on its robustness, by working with a more flexible specification of the fundamental process and the information structure. The bottom line is that the more stringent specification assumed here guarantees that the equilibrium inherits the key qualitative properties of higher beliefs without, however, inheriting their complexity. This in turn facilitates our observational-equivalence result, which we present next.

4.3 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 (6) is modified as follows:

$$a_t = \varphi\xi_t + \delta\omega_f \mathbb{E}_t\left[a_{t+1}\right] + \omega_b a_{t-1} \tag{16}$$

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.

Condition (16) nests the Euler condition of a representative consumer who exhibits habit; a variant of the Q theory that has the representative firm face a cost for adjusting its rate of investment rather than a cost for adjusting its capital stock; and the so-called Hybrid NKPC. With the latter example in mind, we henceforth refer to the economy described above as the "hybrid 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 hybrid 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 $\omega_b > 0$. We compliment this result with the following, which studies the comparative statics of the as-if distortions with respect to the GE effect, or the strategic complementarity.

Proposition 4 (**GE**)**.** *A stronger GE feedback (higher γ) maps to both greater myopia (lower ω_f) and greater anchoring (higher ω_b) in the hybrid model.*

We expand on the applicability and the usefulness of these results in Sections 5 and 6. Before that, we next explain the broader principles that underly them. We thereby also explain why we prefer the perspective developed above over the characterization provided in Proposition 2: whereas the latter depends critically on the details of the assumed specification for fundamental process and the information structure, the insights encapsulated by Propositions 3 and 4 are more general.

4.4 Underlying Principles

To understand what drives $\omega_f < 1$, it suffices to abstract from learning and, instead, focus on the effects of first- and higher-order uncertainty. As evident from conditions (4) and (5), the optimal behavior in our setting is pinned down by two forward-looking objects: the expected present discounted value of the exogenous fundamental, and the expected present discounted value of the endogenous outcome. When an innovation occurs in the fundamental, both of these objects move, but less so under incomplete information than under complete information. First-order uncertainty arrests the movement of the former. Higher-order uncertainty arrests the movement of the latter, indeed at an even greater degree than that characterizing the former. Both effects cause the economy to respond less to news about the future that in the frictionless benchmark, explaining the documented form of myopia. And because the relevance of the second effect (i.e., that regarding higher-order beliefs) increases with the strength of the GE feedback, this myopia also increases with it.

To illustrate this point, consider the following exercise, which draws from Angeletos and Lian (2018). First, let $\beta = 0$, which reduces the dimensionality of the relevant higher-order beliefs and gives the aggregate outcome as in condition (9). Next, fix a $t \geq 0$ and a $H \geq 2$ and assume that, as of period t, it is commonly known that, with probability one, $\xi_\tau = 0$ for all $\tau \neq t + H$; what is uncertain is only the value of ξ_{t+H}, that is, the value of the fundamental at horizon H. Under these assumptions, condition (9) reduces to

$$a_t = \varphi\gamma^H \overline{\mathbb{E}}_t \left[\overline{\mathbb{E}}_{t+1} \left[...\overline{\mathbb{E}}_{t+H} \left[\xi_{t+H}\right] ...\right]\right]. \tag{17}$$

Note that a longer horizon maps to beliefs of higher order, for it involves a longer chain of iterating on the underlying forward-looking, best-response condition. Finally, let the common prior about ξ_{t+H} be a Normal distribution with mean 0 and variance σ_ξ^2; model the period-t information about ξ_{t+H} as a collection of private signals of the form

$$x_{i,t} = \xi_{t+H} + \epsilon_{i,t},$$

where $\epsilon_{i,t}$ is i.i.d. across i, drawn from an Normal distribution with mean 0 and variance σ_x^2; and assume that there is no learning after period t.

In this context, we now ask the following question: how does a_t respond to the news about ξ_{t+H}? By condition (17), the answer to this question boils down to studying the response of the relevant H-order belief. Let $\lambda \equiv \frac{\sigma_\xi^2}{\sigma_\xi^2+\sigma_x^2} \in (0,1)$ and note that λ is a decreasing function of the level of noise in the available information: a lower λ means a larger informational friction. By the assumption that there is no learning after t, we have that the first-order belief in period $t+H$ is given by $\mathbb{E}_{i,t+H}[\xi_{t+H}] = \lambda x_{i,t}$ for the typical agent. The corresponding average is therefore given by

$$\overline{\mathbb{E}}_{t+H}\left[\xi_{t+H}\right] = \lambda \xi_{t+H}$$

It then follows that the second-order belief in the previous period is given by

$$\overline{\mathbb{E}}_{t+H-1}\left[\overline{\mathbb{E}}_{t+H}\left[\xi_{t+H}\right]\right] = \overline{\mathbb{E}}_{t+H-1}\left[\lambda \xi_{t+H}\right] = \lambda^2 \xi_{t+H}$$

Iterating, we infer that the kind of H-order average belief that drives the period-t outcome is given by

$$\overline{\mathbb{E}}_t\left[\overline{\mathbb{E}}_{t+1}\left[...\overline{\mathbb{E}}_{t+H}\left[\xi_{t+H}\right]...\right]\right] = \lambda^{H+1}\xi_{t+H}.$$

Finally, combining the above with condition (17), we conclude that the period-t outcome can be expressed as follows:

$$a_t = \varphi(\gamma\lambda)^H \overline{\mathbb{E}}_t[\xi_{t+H}] \qquad \text{with} \qquad \overline{\mathbb{E}}_t[\xi_{t+H}] = \lambda \xi_{t+H}.$$

This makes clear the two channels via which the informational friction influences the response of the equilibrium outcome to the news about ξ_{t+H}. On the one hand, first-order uncertainty reduces the response of $\overline{\mathbb{E}}_t[\xi_{t+H}]$ to innovations in ξ_{t+H}. On the other hand, higher-order uncertainty reduces the effective discount factor applied to the news about the future. Both effects are stronger when the informational friction is larger (i.e.., when λ is smaller). But whereas the first effect is invariant to the horizon H, the second effect interacts with it: news about the distant future are effectively discounted more heavily than news about the immediate future.

The exercise conducted above has departed from the specification studied in the rest of this section in two key respects: it has replaced the stationary process for the fundamental with a once-and-for all "news shock" about the value of fundamental at a fixed horizon; and it has ruled out learning. It has nevertheless offered a succinct illustration of the form of as-if myopia produced by incomplete information. It has also clarified that this form of myopia does not depend on the presence of leaning. From this perspective, the key contribution of Propositions 3 and 4 is to show how the insight of this special exercise applies to a stationary setting featuring recurrent and persistent shocks as well as learning. The analysis in Section 7, on the other hand, elaborates on the robustness of this insight.

Let us now turn to the second key feature of our observational-equivalence result, the as-if anchoring captured by $\omega_b > 0$. Unlike the documented form of myopia, this crucially depends on the presence learning. To start with, consider how learning influences the dynamics of first-order beliefs. As evident from condition (13), the first-order belief of the current fundamental depends, not only on the fundamental itself, but also on the corresponding previous-period first-order beliefs. That is, the past first-order belief emerges as a requisite state variable for tracking the current first-order belief. A similar property is true for higher-order beliefs, except that the dimensionality of the requisite state variables increases without bound as the belief order increases.[11] Intuitively, because higher-order beliefs are more strongly anchored to the common prior than first-order beliefs, they depend less to recent information and have "longer memory." As long as $\gamma > 0$, the equilibrium depends positively on both first- and higher-order beliefs. It follows that, in general, the equilibrium outcome, too, has longer memory than either the fundamental or the first-order beliefs.

This explains the essence of the documented from of anchoring: $\omega_b > 0$ reflects the additional persistence, or memory, of the first- and higher-order beliefs relative to that of the fundamental. But whereas the basic logic has already been put forward in previous works (e.g., Woodford, 2003), the result in Proposition 3 offers a particularly succinct illustration of it thanks to the "magic" afforded by our baseline specification: whereas the logic itself allows the current outcome to depends on the entire infinitely long state vector required for capturing the infinitely long memory of the entire belief hierarchy, the assumed specification guarantees that, for any level of noise and any degree of complementarity, the past aggregate outcome is, conveniently, a sufficient statistic for the impact of this infinitely long state vector on the current equilibrium outcome.[12] That is, the "magic" is that condition (16) holds with a_{t-1} in place of the entire history of past first- and higher-order beliefs.[13] Finally, the result that ω_b increases with γ reflects the property that a higher γ increases the dependence of the outcome on beliefs of higher order, which themselves have longer memory.

While the "magic" behind our observational-equivalence is fragile, the underlying insights are not. This is made clear in Section 7 with the help of a flexible specification that, not only illustrates the robustness of these insights, but also disentangles the level of first- and higher-order uncertainty from the speed of learning. By contrast, such a disentangling is not possible under the more rigid specification considered here, because σ, a single parameter, regulates *both* the speed of learning and the level of first- and higher-order uncertainty.

[11] This is evident in the property, noted before, that the beliefs of h-th order follow ARMA$(h+1, h-1)$ processes.

[12] Note that a_{t-1} is a sufficient statistic only in the eyes of the analyst who observes the times series of the fundamental and the outcome: the agents inside the model do not observe a_{t-1}.

[13] The same magic was present in Woodford (2003). Relative to that paper, however, ours accommodates forward-looking behavior and offers a sharp analytical characterization of the equilibrium instead of relying on numerical simulations.

The analysis in Section 7 also reveals that learning shapes the equilibrium behavior, not only in the manner described above (i.e., by adding persistence, or memory, in beliefs and outcomes), but also in another, more subtle, manner: through the anticipation that other agents will learn in the future. This anticipatory effect matters only because agents are forward-looking and is therefore absent in static beauty contests, such as those studied by Morris and Shin (2002) and Woodford (2003).

The forward-looking nature of the problem under consideration also explains why ϑ, the equilibrium persistence of the aggregate outcome, is increasing in ρ, the exogenous persistence of the fundamental. Holding σ constant, an increase in ρ raises the horizon of the "news component" of any given innovation in the fundamental: the higher ρ is, the more information any such innovation contains about, not only about the fundamental, but also about the equilibrium outcome further into the future. Recall that forecasting the equilibrium outcome further and further into the future involves beliefs of higher and higher order. It follows that an increase in ρ raises the relative importance of higher-order uncertainty, in the same way as an increase in the degree of strategic complementarity.

All these insights can not easily be understood by inspecting the AR(2) solution obtained in Proposition 2. This underscores how our contribution hinges on the combination of the observational-equivalence result presented above and the more elaborate analysis offered in Section 7. For the applied purposes of the next two sections, however, one can sidestep that analysis, trust the summary of insights provided above, and rely on the observational-equivalence result alone.

4.5 Testable Restrictions

Although Proposition 3 guarantees that an incomplete-information economy can always be mapped to a hybrid economy, the converse is not true: a hybrid economy can be replicated by an incomplete-information economy only when ω_f and ω_b satisfy a certain restriction.

Proposition 5. *The equilibrium dynamics of a hybrid economy can be replicated by that of an incomplete-information economy for some $\sigma > 0$ if and only if $\omega_b > 0$ and*

$$\omega_f = 1 - \frac{1}{\delta \rho^2} \omega_b. \tag{18}$$

Furthermore, for any pair (ω_b, ω_f) that satisfies the above restriction, there exists a unique $\sigma > 0$ such that the two economies are observationally equivalent.

This result offers a simple test for our theory. Suppose that one uses a times series of ξ_t and a_t to estimate ρ, the persistence of the fundamental, and the pair (ω_f, ω_b), which governs the law of motion (16) of the outcome. Suppose further that one knows β and γ, and hence also δ, from independent sources. One can then test whether condition (18) is satisfied. If it does, then and only then the data is compatible with our theory.

Additional testable predictions, or overidentifying restrictions, can be obtained by looking at the forecasts of future outcomes. Let $\epsilon_t^k \equiv a_{t+k} - \overline{\mathbb{E}}_t\left[a_{t+k}\right]$ be the realized average k-period ahead forecast error. As long as information is incomplete, ϵ_t^k is serially correlated. And because the magnitude of the serial correlation depends on σ, this provides us with an additional restriction that can be used to identify σ and/or to test the model. We put these ideas at work in Section 6.

5 Implications and Discussion

The DSGE literature that follows Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) has added three distinct backward-looking elements to the key forward-looking equations of baseline macroeconomic models: habit persistence in consumption, adjustment costs to investment, and automatic past-price indexation. These modifications are crucial for this literature's capacity to offer a successful structural interpretation of the macroeconomic times series,[14] but lack independent empirical support. They are also inconsistent with the models employed in strands of the literature that aim at understanding the microeconomic data.[15] All in all, these kinds of backward-looking elements are considered as crude proxies for other, unspecified mechanisms.

Prior work has already pushed the idea that informational frictions can be such a mechanism (Sims, 2003; Woodford, 2003; Mankiw and Reis, 2002, 2007; Mackowiak and Wiederholt, 2009, 2015; Nimark, 2008). Our analysis adds to this line of work in four ways:

1. It offers the sharpest, up to date, illustration of the aforementioned idea.

2. 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.

3. It helps reduce a discomforting gap between the macroeconomic and the microeconomics estimates of these elements, a point we discuss next.

4. It blends these backward-looking DSGE elements with a form of imperfect foresight.

[14] Not only do they allow the theory to match the sluggishness in the dynamic responses of consumption, investment, and inflation to a variety of identified shocks, but also help fix the comovement properties of the New Keynesian model.

[15] For instance, although the inflation dynamics implied by the standard NKPC are broadly consistent—qualitatively if not quantitatively—with menu-cost models (Alvarez and Lippi, 2014; Golosov and Lucas Jr, 2007; Nakamura and Steinsson, 2013), such models do not produce the kind of backward-looking behavior implied by the Hybrid NKPC. Similarly, the form of investment adjustment cost assumed in the DSGE literature is at odds with the literature that studies investment at the plant or firm level (Bachmann, Caballero, and Engel, 2013; Bloom et al., 2018; Caballero and Engel, 1999).

5. It clarifies how an emerging literature on bounded rationality relates to our approach and the related literature on informational frictions.

6. It facilitates the quantitative evaluation we conduct in Section 6.

In the sequel, we elaborate on each one of these aspects of our contribution, starting with a sketch of how our results can be applied to aggregate demand in the New Keynesian model.

5.1 Applications

In the textbook New Keynesian model, which abstracts from investment, aggregate demand is given by the Euler condition of the representative consumer:[16]

$$c_t = -r_t + \mathbb{E}_t\left[c_{t+1}\right], \tag{19}$$

where c_t is aggregate consumption in period t, r_t is the real interest rate between period t and $t+1$, and $\mathbb{E}_t$ is the rational expectation conditional on period-t information.

As shown in Angeletos and Lian (2018), an incomplete-information extension of condition (19) is given by the following:

$$c_t = -\sum_{k=0}^{\infty} \theta^k \overline{\mathbb{E}}_t[r_{t+k}] + (1-\theta)\sum_{k=1}^{\infty} \theta^{k-1}\overline{\mathbb{E}}_t[c_{t+k}]. \tag{20}$$

where $\overline{\mathbb{E}}_t$ stand for the average expectation of the consumers, and $\theta \in (0,1)$ for their subjective discount factor. To understand where this condition comes from, consider the textbook version of Permanent Income Hypothesis. This gives consumption as a function of the expected present discounted value of income. Extending this so as to accommodate variation in the real interest rate and heterogeneity in information, and using the fact that aggregate income equals aggregate consumption in equilibrium, results to condition (20).

This condition recasts the aggregate-demand block of the New Keynesian model as a dynamic game among the consumers. This game is nested in our abstract framework by mapping r_t and c_t to ξ_t and a_t, respectively, and by letting

$$\varphi = -1, \qquad \beta = \theta, \qquad \text{and} \qquad \gamma = 1-\theta.$$

The following result is then an immediate corollary of Propositions 3 and 4, provided of course that we maintain the assumptions introduced in the beginning of Section 4.[17]

[16] Throughout, we work with the log-linearized model: all variables are in log-deviations from steady state.

[17] Online Appendix C of Angeletos and Lian (2018) develops a "discounted" Euler equation that resembles condition

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

$$c_t = -r_t + \omega_f \mathbb{E}_t[c_{t+1}] + \omega_b c_{t-1} \tag{21}$$

Furthermore, a lower θ, which represents a stronger income-spending multiplier, results to both a lower ω_f and a higher ω_b.

It is therefore as if the economy is populated by a representative agent whose consumption exhibits habit persistence, of the kind assumed in Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). There are, though, two subtle differences. First, whereas the true habit model imposes $\omega_f + \omega_b = 1$, our model implies $\omega_f + \omega_b < \rho < 1$. This means that the overall movement in consumption is smaller, reflecting the myopia produced by incomplete information. Second, and perhaps most importantly, the coefficients ω_f and ω_b depends critically on θ, because this parameter governs the strength of the relevant GE feedback, namely the Keynesian income-spending multiplier.

It is useful *not* to interpret θ literally, as the subjective discount factor. For instance, we can readily extend the analysis to a perpetual-youth, overlapping-generations model along the lines of Del Negro, Giannoni, and Patterson (2015) and, under appropriate assumptions, replicate the results reported above with θ replaced by $\chi\theta$, where χ is the survival probability. The latter can in turn be thought of as a measure of the length of planning horizons, either in the sense of expected lifespans or in the sense described in Woodford (2018). Alternatively, as in Farhi and Werning (2017), $1 - \chi$ can serve as a proxy for the probability of binding liquidity constraints. For our purposes, the key observation is that, even if such features happen to be irrelevant under complete information due to offsetting PE and GE effects,[18] they can be crucial under incomplete information because they determine the strength of the relevant GE effect and the consequent importance of higher-order uncertainty.

This in turn builds a bridge to a growing theoretical and empirical literature that studies the determinants of aggregate demand and of the Keynesian multiplier in settings that allow for incomplete markets and rich heterogeneity.[19] In the light of our results, the interaction of these features with informational frictions can perhaps rationalize both significant myopia vis-a-vis the future and a strong, habit-like, backward-looking force in the aggregate consumption dynamics. The dependence of these distortions on policies that regulate the strength of the Keynesian multiplier, such as fiscal reforms that alleviate liquidity constraints, is another implication of our analysis that warrants further investigation.

(20), but also differs from it in two crucial respects. First, it imposes $\omega_b = 0$, ruling out the habit-like element. Second, it only describes the particular path of consumption triggered by a once-and-for-all shift in the expectations of the real interest rate that will prevail at a single, and fixed, future date. By contrast, our result describes the entire stochastic process of consumption in a stationary setting with recurrent shocks. The same points distinguish the discounted NKPC found in that paper from the version of the Hybrid NKPC we develop later on, in Section 6.

[18] In the present context, this offsetting is evident the property that the sum $\gamma + \beta$ is invariant to θ and χ.

[19] E.g., Auclert (2017), Kaplan and Violante (2014), Kaplan, Moll, and Violante (2016), and Werning (2015).

In the above, we focused on consumption. In Appendix B, we turn to investment. We take a model that features a conventional form of adjustment costs to capital, as in Hayashi (1982) and Abel and Blanchard (1983), and show how the introduction of incomplete information to this model can make investment behave *as if* the adjustment cost takes the more exotic form assumed in the DSGE literature. And in Section 6, we show how the Hybrid NKPC can be obtained by augmenting the standard NKPC with incomplete information. Together, these applications explain how our analysis connects to each of the three building blocks of the modern macroeconomic framework.[20]

5.2 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 1 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

[20] These applications treat each block in isolation of each other. Accommodating their interaction may break the exact observational equivalence, but, as in the case of richer information structures, need not upset the key insights.

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 complementarity 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.[21]

5.3 Imperfect Foresight and Bounded Rationality

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 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.[22] 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

[21]We 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.

[22]This follows the lead of García-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.

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.[23] 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 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.[24] 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

In this section, we study the application of our theory to the aggregate-supply block of the New Keynesian model, that is, the NKPC. This application is shown to 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). Furthermore, the implied quantitative bite of the informational friction on the inflation dynamics is non-trivial.

[23] For 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.

[24] We thank Alexander Kohlhas for suggesting us to explore such a variant.

Setup and Mapping to Abstract 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:

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

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 Ψ_{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:

$$p_{it}^{*} = (1-\delta\theta) \sum_{k=0}^{\infty} (\delta\theta)^{k} \mathbb{E}_{it}[\psi_{t+k} + p_{t+k}]. \tag{22}$$

We next make the simplifying assumption that the firms observe that past price level but do not extract information from it.[25] This permits us to restate condition (22) as

$$p_{it}^{*} - p_{t-1} = (1-\delta\theta) \sum_{k=0}^{\infty} (\delta\theta)^{k} \mathbb{E}_{it}[\psi_{t+k} + \pi_{t+k}], \tag{23}$$

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

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

[25] 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. In any event, this assumption is only for convenience: as shown in Appendix C, our observational-equivalence result can be a good approximation of the true equilibrium in settings that allow agents to extract information from current or past outcomes.

Combining this with condition (23) and rearranging, we arrive at the following expression:

$$\pi_t = \frac{(1-\delta\theta)(1-\theta)}{\theta}\sum_{k=0}^{\infty}(\delta\theta)^k\overline{\mathbb{E}}_t\left[\psi_{t+k}\right] + \delta(1-\theta)\sum_{k=0}^{\infty}(\delta\theta)^k\overline{\mathbb{E}}_t\left[\pi_{t+k+1}\right]. \quad (24)$$

When information is complete, we can replace $\overline{\mathbb{E}}_t[\cdot]$ with $\mathbb{E}_t[\cdot]$, the expectation of the representative agent. We can then use the Law of Iterated Expectations to reduce condition (24) to the following:

$$\pi_t = \kappa\psi_t + \delta\mathbb{E}_t[\pi_{t+1}], \quad (25)$$

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 (24) cannot be reduced to condition (25). 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.[26] Instead, the applicable version of the NKPC is the one given in condition (24), 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 (24) is readily nested the analysis of Sections 3 and 4 by mapping ψ_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 4.

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:*

$$\pi_t = \kappa\psi_t + \omega_f\delta\mathbb{E}_t[\pi_{t+1}] + \omega_b\pi_{t-1} \quad (26)$$

(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 compo-

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

nent, 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 (24) 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 (25). This is generally bad for the empirically 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 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.

Testing the Theory. The Hybrid NKPC estimated in Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2005), is similar to the one seen in (26). There are, however, two differences. First, our theory restricts the pair (ω_f, ω_b) in the way described in Proposition 5, whereas unrestricted estimations of the Hybrid NKPC allow these parameters to be free. And second, our theory ties the pair (ω_f, ω_b) to the dynamics of inflation forecasts. We now use these restrictions to test our theory.

Matching Estimates of the Hybrid NKPC. Gali, Gertler, and Lopez-Salido (2005) synthesize the literature on the estimation of the Hybrid NKPC and provide a few estimates of the pair (ω_f, ω_b). A quick test of our theory is whether these estimates 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.

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

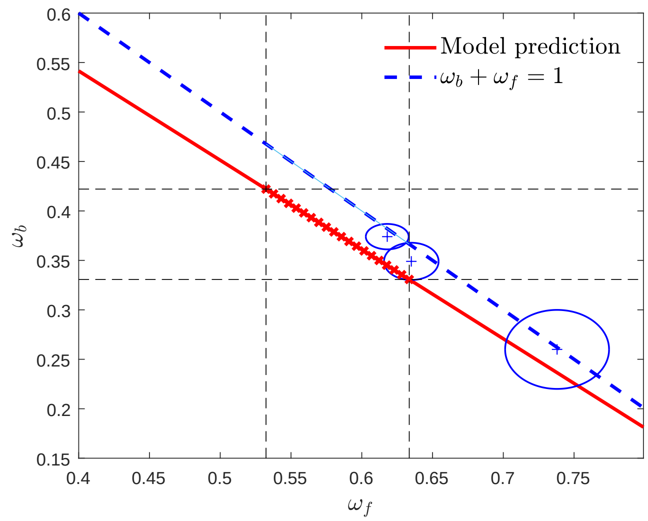

Figure 1: Testing the Theory

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:

$$\pi_{t+k} - \overline{\mathbb{E}}_t[\pi_{t+k}] = K\left(\overline{\mathbb{E}}_t[\pi_{t+k}] - \overline{\mathbb{E}}_{t-1}[\pi_{t+k}]\right) + v_{t+k,t} \tag{27}$$

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 is 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.[27]

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.

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.[28]

Let us now explain the dotted red line in the figure. Using condition (24), the incomplete-

[27] The 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.

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

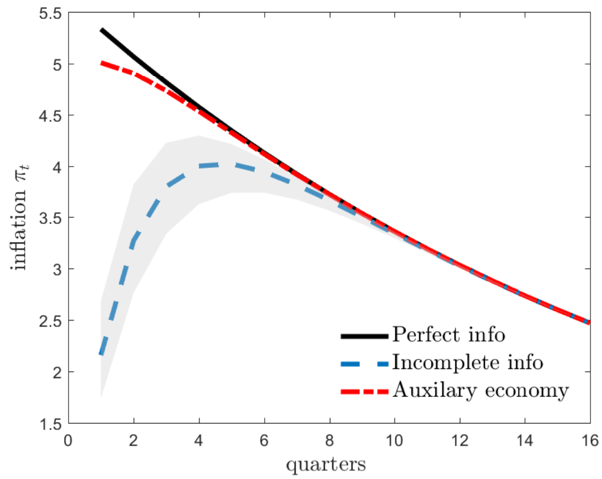

Figure 2: Impulse Response Function of Inflation

information inflation dynamics can be decomposed into two components: the belief of the present discounted value of real marginal costs, $\varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t[\psi_{t+k}]$; and the belief of of the present discounted value of inflation, $\gamma \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t[\pi_{t+k+1}]$. The same decomposition can also be applied when agents have perfect information:

$$\pi_t^* = \varphi \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\psi_{t+k} | \psi_t\right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\pi_{t+k+1}^* | \psi_t\right], \tag{28}$$

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:

$$\widetilde{\pi}_t = \varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[\psi_{t+k}\right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\pi_{t+k+1}^* | \psi_t\right]. \tag{29}$$

The difference between π_t^* and $\widetilde{\pi}_t$ measures the importance of beliefs about real marginal costs, and the difference between $\widetilde{\pi}_t$ and π_t measures the importance of beliefs about inflation. The dotted red line in Figure 2 corresponds to $\widetilde{\pi}_t$. 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 Robustness: Incomplete Information as Myopia and Anchoring

Earlier on 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:

$$\xi_t = \sum_{k=0}^{\infty} \rho_k \eta_{t-k}, \tag{30}$$

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.[29] 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 4 because a single parameter, σ, controled 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

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

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 $\mathbb{E}_{it}[\eta_{t-k}] = \lambda_k \widetilde{x}_{i,t}^{k}$, which in turn implies $\overline{\mathbb{E}}_t[\eta_{t-k}] = \lambda_k \eta_{t-k}$ and therefore

$$\overline{\mathbb{E}}_t\left[\xi_t\right] = \overline{\mathbb{E}}_t\left[\sum_{k=0}^{\infty} \rho_k \eta_{t-k}\right] = \sum_{k=0}^{\infty} f_{1,k} \eta_{t-k}, \qquad \text{with} \qquad f_{1,k} = \lambda_k \rho_k. \tag{31}$$

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 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 4, 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 (31) to period $t+1$

[29] Such 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

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

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

$$\overline{\mathbb{F}}_t^h\left[\xi_{t+h-1}\right] = \sum_{k=0}^{\infty} f_{h,k} \eta_{t-k}, \qquad \text{with} \qquad f_{h,k} = \lambda_k \lambda_{k+1} ... \lambda_{k+h-1} \rho_{k+h-1}. \tag{32}$$

The increasing components in the product $\lambda_k \lambda_{k+1} ... \lambda_{k+h-1}$ 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

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

The following result is thus immediate.

Proposition 7. *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 $k \to \infty$, 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 $\lambda_k \to 1$.

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 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 $k \to \infty$ if and only if $\lambda_k \to 1$, 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 $a_t = \overline{\mathbb{E}}_t[\xi_t] + \gamma\overline{\mathbb{E}}_t[a_{t+1}]$, which in turn implies that $a_t = \sum_{h=1}^{\infty} \gamma^{h-1}\overline{\mathbb{F}}_t^h [\xi_{t+h-1}]$. From the preceding characterization of the higher-order beliefs $\overline{\mathbb{F}}_t^h [\xi_{t+h-1}]$, it follows that

$$a_t = \sum_{k=0}^{\infty} g_k \eta_{t-k}, \quad \text{with} \quad g_k = \sum_{h=1}^{\infty} \gamma^{h-1} f_{h,k} = \left\{ \sum_{h=1}^{\infty} \gamma^{h-1} \lambda_k \lambda_{k+1} ... \lambda_{k+h-1} \rho_{k+h-1} \right\}. \tag{33}$$

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.[30]

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

$$a_t = \xi_t' + \gamma' \mathbb{E}_t[a_{t+1}], \tag{34}$$

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

[30] The 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 4.3 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 4.4 and 5.2.

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

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

This resembles the complete-information benchmark in that the outcome is pined 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 8. *Let $\{g_k\}$ and $\{\widehat{g}_k\}$ denote the Impulse Response Function of the equilibrium outcome in the two economies described above.*

(i) $0 < g_k < \widehat{g}_k$ 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 $\frac{g_{k+1}}{g_k} > \frac{\widehat{g}_{k+1}}{\widehat{g}_k}$ for all $k \geq 0$

Consider property (i), in particular the property that $g_k < \widehat{g}_k$. 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:

$$a_t = \varphi \overline{\mathbb{E}}_t[\xi_t] + \gamma \overline{\mathbb{E}}_t[a_{t+1}]. \tag{35}$$

Furthermore, the two economies share the same dynamic response for $\overline{\mathbb{E}}_t[\xi_t]$. 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 $\overline{\mathbb{E}}_t[a_{t+1}]$ 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:

$$g_k = f_{1,k} + \lambda_k \gamma g_{k+1}. \tag{36}$$

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 (34). 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.[31]

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

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

[31] This 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,[32] 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 2. *Let the variant economy display a hump-shaped response: $\{\widehat{g}_k\}$ 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, $\widehat{g}_k$, and g_k as differentiable functions of k. If $\widehat{g}_k$ is hump-shaped with a peak at $k = k_b > 0$, it must be that $\widehat{g}_k$ 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 $\widehat{g}_k$, 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 (32), when agents are rational, the forward higher-order beliefs are

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

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

[32] Recall, this is by property (i) of Proposition 8.

becomes

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

Learning implies $\lambda_{k+1} > \lambda_k$, and the anticipation of learning implies $\lambda_k \lambda_{k+1} \ldots \lambda_{k+h-1} > \lambda_k^h$. 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

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

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 $a_t^0 = \overline{\mathbb{E}}_t[\xi_t]$.

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

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

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

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

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.[33] 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.

[33] This 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.

8 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 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 ($\omega_b > 0$), in line with Kasa, Walker, and Whiteman (2014). But it also highlights the possibility of excessive discounting of news about future fundamentals ($\omega_f < 1$). This 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. 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. 5.1, 8

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. 8, 9, 8

Altissimo, Filippo, Laurent Bilke, Andrew Levin, Thomas Matha, and Benoit Mojon. 2010. "Sectoral and Aggregate Inflation Dynamics in the Euro Area." *Journal of the European Economic Association* 4 (2-3):585–593. 1

Alvarez, Fernando and Francesco Lippi. 2014. "Price Setting with Menu Cost for Multiproduct Firms." *Econometrica* 82 (1):89–135. 1, 15, 28

Angeletos, George-Marios and Jennifer La'O. 2010. "Noisy Business Cycles." In *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, 319–378. 8

Angeletos, George-Marios and Chen Lian. 2016. "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics* 2:1065–1240. 2

———. 2018. "Forward Guidance without Common Knowledge." *American Economic Review* 108 (9):2477–2512. 1, 5, 2, 8, 4.4, 5.1, 17, 5.3

Angeletos, George-Marios and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4):1103–1142. 8

Auclert, Adrien. 2017. "Monetary Policy and the Redistribution Channel." *NBER Working Paper No. 23451* . 19

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. 1, 15, 8

Bergemann, Dirk and Stephen Morris. 2013. "Robust Predictions in Games with Incomplete Information." *Econometrica* 81 (4):1251–1308. 8

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* forthcoming. 15

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. "Over-reaction in Macroeconomic Expectations." *NBER Working Paper No. 24932* . 2, 5.3

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. 1, 15, 8

Carroll, Christopher D., Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White. 2018. "Sticky Expectations and Consumption Dynamics." Working Paper 24377, National Bureau of Economic Research. 7

Carvalho, Carlos, Stefano Eusepi, Emanuel Moench, and Bruce Preston. 2017. "Anchored Inflation Expectations." . 2

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. 1, 5, 5.1, 6, 8

Coibion, Olivier and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1):116–159. 5, 2, 5.3

———. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–78. 1, 5, 2, 5.3, 6, 6, 6

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2015. "How Do Firms Form Their Expectations? New Survey Evidence." *NBER Working Paper Series* . 5, 5.3

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* . 5, 5.3, 6

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." *FRB of New York mimeo* . 5, 5.1

Evans, George W and Seppo Honkapohja. 2012. *Learning and expectations in macroeconomics*. Princeton University Press. 2

Farhi, Emmanuel and Iván Werning. 2017. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *NBER Working Paper No. 23281* . 1, 5, 5.1, 5.3

Gabaix, Xavier. 2017. "A Behavioral New Keynesian Model." *Harvard mimeo* . 1, 5.3

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. 6

Gali, Jordi and Mark Gertler. 1999. "Inflation dynamics: A structural econometric analysis." *Journal of Monetary Economics* 44 (2):195–222. 1, 2, 6, 6

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. 1, 2, 6, 6, 6, 27

Garcıa-Schmidt, Mariana and Michael Woodford. 2018. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review* forthcoming. 1, 22

Golosov, Mikhail and Robert E Lucas Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2):171–199. 1, 15, 28

Groth, Charlota and Hashmat Khan. 2010. "Investment Adjustment Costs: An Empirical Assessment." *Journal of Money, Credit and Banking* 42 (8):1469–1494. 1, 5.2

Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit Formation in Consumption: A Meta-Analysis." *European Economic Review* 95:142–167. 1, 5.2

Hayashi, Fumio. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation." *Econometrica* 50 (1):213–224. 1, 5.1, 8

Huo, Zen and Naoki Takayama. 2018. "Rational Expectations Models with Higher Order Beliefs." miméo, Yale University. 2, 4.2, 36

Huo, Zhen and Marcelo Zouain Pedroni. 2017. "Infinite Higher Order Beliefs and First Order Beliefs: An Equivalnce Result Infinite Higher Order Beliefs and First Order Beliefs: An Equivalnce Result." miméo, Yale University/University of Amsterdam. 8

Iovino, Luigi and Dmitriy Sergeyev. 2017. "Quantitative Easing without Rational Expectations." *Work in progress* . 22

Jung, Jeeman and Robert J. Shiller. 2005. "Samuelson's Dictum and the Stock Market." *Economic Inquiry* 43 (2):221–228. 1, 8, 8

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2016. "Monetary Policy According to HANK." *NBER Working Paper No. 21897* . 19

Kaplan, Greg and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82 (4):1199–1239. 19

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. 2, 8, 8

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. 8

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. 6

Kohlhas, Alexandre and Ansgar Walther. 2018a. "Asymmetric Attention." *IIES miméo* . 2, 5.3

———. 2018b. "Sparse Expectations: A Unified Explanation of Forecast Data." *IIES miméo* . 2, 5.3

Lucas, Robert E. Jr. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (2):103–124. 1

Mackowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803. 4, 2, 5, 5.2, 8

———. 2015. "Business Cycle Dynamics under Rational Inattention." *Review of Economic Studies* 82 (4):1502–1532. 4, 5

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. 4, 6, 5

Mankiw, N Gregory and Ricardo Reis. 2007. "Sticky information in general equilibrium." *Journal of the European Economic Association* 5 (2-3):603–613. 4, 5

Marcet, Albert and Juan P Nicolini. 2003. "Recurrent hyperinflations and learning." *American Economic Review* 93 (5):1476–1498. 2

Matejka, Filip. 2016. "Rationally Inattentive Seller: Sales and Discrete Pricing." *Review of Economic Studies* 83 (3):1125–1155. 6

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. 6, 26, 6

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The Power of Forward Guidance Revisited." *American Economic Review* 106 (10):3133–3158. 5

Melosi, Leonardo. 2016. "Signalling effects of monetary policy." *The Review of Economic Studies* 84 (2):853–884. 6

Midrigan, Virgiliu. 2011. "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations." *Econometrica* 79 (4):1139–1180. 28

Morris, Stephen and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review* 92 (5):1521–1534. 1, 2, 8, 4.4

———. 2006. "Inertia of Forward-looking Expectations." *The American Economic Review* :152–157. 1, 2, 8, 9

Nakamura, Emi and Jón Steinsson. 2013. "Price Rigidity: Microeconomic Evidence and Macroeconomic Implications." *Annual Review of Economics* 5 (1):133–163. 1, 15, 28

Nimark, Kristoffer. 2008. "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics* 55 (2):365–382. 1, 4, 2, 8, 5

———. 2017. "Dynamic Higher Order Expectations." *Cornell Univeristy mimeo* . 8, 9, 7, 8, 36

Reis, Ricardo. 2006. "Inattentive producers." *The Review of Economic Studies* 73 (3):793–821. 6

Sargent, Thomas J. 1993. *Bounded rationality in macroeconomics*. Oxford University Press. 2

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3):665–690. 1, 4, 2, 5

———. 2010. "Rational Inattention and Monetary Economics." *Handbook of Monetary Economics* 3:155–181. 2

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. 8, 8

Smets, Frank and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3):586–606. 5, 5.1, 6, 8

Vellekoop, Nathanael and Mirko Wiederholt. 2017. "Inflation Expectations and Choices of Households." *Goethe University Frankfurt mimeo* . 5.3

Vives, Xavier and Liyan Yang. 2017. "Costly Interpretation of Asset Prices." mimeo, IESE/University of Toronto. 25

Werning, Iván. 2015. "Incomplete Markets and Aggregate Demand." *NBER Working Paper No. 21448* . 19

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* . 1, 2, 8, 4.4, 13, 5

———. 2018. "Monetary Policy Analysis When Planning Horizons Are Finite." In *NBER Macroeconomics Annual 2018, Volume 33*. University of Chicago Press. 5.1

Zorn, Peter. 2018. "Investment under Rational Inattention: Evidence from US Sectoral Data." miméo, University of Munich. 1, 7, 5.2

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

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

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

$$B(L)B(L^{-1}) = \mathbf{M}(L)\mathbf{M}'(L^{-1}) = \frac{\tau_\eta^{-1} + \tau_u^{-1}(1-\rho L)(L-\rho)}{(1-\rho L)(L-\rho)}, \tag{37}$$

which leads to

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

where λ is the inside root of the numerator in equation (37)

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

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

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

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

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

Consider the forecast of the fundamental. Note that

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

from which it follows that

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

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

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

and the forecasts are

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

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

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 $\left[G_1(L) + (\beta+\gamma)G_2(L) + \beta G_3(L)\right]x_{it}$. 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

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

which can be transformed as

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

where

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

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

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

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

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

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

$$\zeta_1 = \frac{1}{2\omega_f\delta}\left(1-\sqrt{1-4\delta\omega_f\omega_b}\right) \quad \text{and} \quad \zeta_0 = \frac{\varphi\zeta_1}{\omega_b-\rho\omega_f\delta\zeta_1} \tag{38}$$

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

$$\zeta_1 = \vartheta \quad \text{and} \quad \zeta_0 = \left(1-\frac{\vartheta}{\rho}\right)\frac{\varphi}{1-\rho\delta}. \tag{39}$$

Combining (38) and (39), 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:

$$\omega_f = \frac{\delta\rho^2-\vartheta}{\delta(\rho^2-\vartheta^2)} \tag{40}$$
$$\omega_b = \frac{\vartheta(1-\delta\vartheta)\rho^2}{\rho^2-\vartheta^2} \tag{41}$$

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 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

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

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

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

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 (40) and (41) hold. Using (41), we can rewrite (40) as follows:

$$\omega_f = \Omega(\omega_b; \delta, \rho) \equiv 1 - \frac{1}{\delta\rho^2}\omega_b. \tag{42}$$

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 (41) 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 (42) along with $\omega_b \in (0, \infty)$. Finally, the level of the informational friction that achieves this replication is obtained by inverting condition (41) to obtain ϑ, and thereby also σ, as an implicit function of ω_b.

Proof of Proposition 8. First, let us prove $g_k < \widehat{g}_k$. Recall that $\{g_k\}$ is given by

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

Clearly,

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

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

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

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

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

$$\frac{g_{k+1}}{g_k} = \frac{\lambda_{k+1}}{\lambda_k} \frac{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+2} \ldots \lambda_{k+h+1} \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \lambda_{k+1} \ldots \lambda_{k+h} \rho_{k+h}}$$

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

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

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

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

and

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

Based on the observation, we will show that

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

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

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

We want to show

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

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

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

and

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

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

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

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

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

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

Appendix B: 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:

$$\text{Cost}_t = \Phi\left(\frac{I_t}{K_{t-1}}\right) \tag{43}$$

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).[34]

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:

$$\text{Cost}_t = \Psi\left(\frac{I_t}{I_{t-1}}\right) \tag{44}$$

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 (43); 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 (44). 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

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

[34] These 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.

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

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

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

$$Q q_t + \psi Z z_t = \chi \mathbb{E}_t \Big[A \xi_{t+1} + Q q_{t+1} + \psi Z (1 + Z) z_{t+1} \Big]. \tag{45}$$

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

$$\max_{X_{it}} \ \mathbb{E}_{it} \left[Q_{it} X_{it} - c X_{it} \right],$$

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

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

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:

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

where

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

Incomplete Information. Suppose now that firms receive a noisy signal about the fundamental ξ_t as in Section 3. 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 (45), the aggregate investment dynamics follow

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

The investment dynamics can be understood as the solution to the dynamic beauty contest studied in Section 3 by letting

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

The following is then immediate.

Proposition 9. *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 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:

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

where $\widetilde{I}_t$ is the aggregate investment.

Appendix C: 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:

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

where $\overline{\mathbb{E}}_t$ 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 (11).[35] 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

[35] Here, we are abstracting from the complications of the endogenous revelation of information and we think of the signals in (11) 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.

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 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.[36] 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.

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

[36] See Nimark (2017) and Huo and Takayama (2018) for a more detailed discussion.

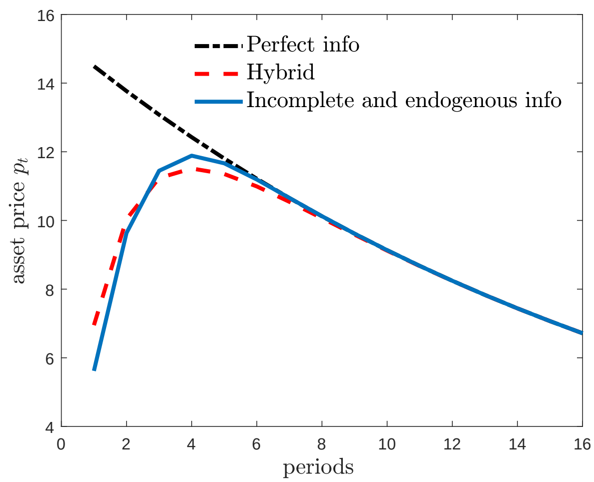

Figure 3: Impulse Response Function of Asset Price

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 5 .

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.[37]

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

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

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,

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

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 4.2 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,

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

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 aforecited paper, but also because of the role of higher-order uncertainty identified here.