Myopia and Anchoring

George-Marios Angeletos*  Zhen Huo†

April 11, 2018

Abstract

We consider a stationary setting featuring forward-looking behavior, slow learning, and higher-order uncertainty. 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 in the sense of extra discounting of the future; and anchoring of the current outcome to the past outcome, as in models that feature consumption habit, adjustment costs to investment, or past-price indexation. This builds a bridge to the DSGE literature; it also distinguishes our approach from an emerging literature on bounded rationality that captures the first feature (myopia) but misses the second feature (anchoring). We further show that the as-if distortions are larger when the general-equilibrium interaction is stronger; this property reflects the role of higher-order uncertainty and helps reduce the gap between macroeconomic and microeconomic estimates of adjustment frictions. We finally illustrate the quantitative potential of our theory in the context of inflation by showing how it can help rationalize existing estimates of the Hybrid NKPC while also matching survey evidence on expectations.

*MIT and NBER; angelet@mit.edu.
†Yale University; zhen.huo@yale.edu.
1 Introduction

To account for salient features of the data and aid quantitative policy evaluation, macroeconomics often sacrifices on the micro-foundations. For example, consider the response of inflation to identified monetary shocks, or to innovations in the output gap. To match the magnitude of this response, the DSGE literature has assumed a much higher degree of price rigidity than what seems palatable given the available microeconomic evidence. And to match its sluggishness, the literature has utilized ad-hoc backward-looking mechanisms that have no obvious counterpart in the microeconomic data.\(^1\) Similarly, the sluggish response of aggregate consumption has been explained by a degree of habit that is much higher than that estimated in microeconomic data.\(^2\)

In this paper, we offer a potential resolution to this disconnect between microeconomics and macroeconomics. This resolution is based on the accommodation of incomplete information, slow learning, and higher-order uncertainty. The key insight is that these features cause the economy to respond to shocks in a similar manner as the combination of two behavioral distortions:

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

Furthermore, these distortions are larger the stronger general-equilibrium feedback loops in the economy are, reflecting the role of higher-order uncertainty.

This insight is formalized with an exact observational-equivalence result under appropriate assumptions. Its empirical potential is illustrated in the context of inflation by showing how our theory can help rationalize both existing estimates of the Hybrid NKPC and survey evidence on expectations. Additional lessons include an application to consumption and the comparison of the friction considered here to related forms of bounded rationality.

**Framework.** Our starting point is a dynamic beauty contest, namely a game with a continuum of players, best responses that depend on the expectations of the future actions of others, and strategic complementarity. These features help capture the role of forward-looking behavior and general-equilibrium (GE) effects in the New Keynesian model, among other contexts.

Denote with \(\xi_t\) the exogenous, payoff-relevant fundamental and let it follow a persistent Gaussian process. Next, denote with \(a_t\) the equilibrium outcome (e.g., inflation). When information is complete, our setting reduces to a representative-agent model and \(a_t\) obeys the following law of motion:

\[
a_t = \varphi \xi_t + \delta E_t[\xi_{t+1}],
\]

\(^1\)We have in mind the hybrid version of the New Keynesian Philips Curve estimated in Gali and Gertler (1999) or, relatedly, the indexation to the past price level introduced in Christiano, Eichenbaum, and Evans (2005).
\(^2\)See the meta-analysis of micro and macro estimates of habit in Havranek, Rusnak, and Sokolova (2017).
where $\varphi > 0$ and $\delta \in (0, 1]$ are known scalars and $E_t[\cdot]$ is the rational expectations of the representative agent. Note that condition (1) nests the two key equations of the New Keynesian model: the NKPC is nested by interpreting $\xi_t$ as the output gap, or the real marginal cost, and $a_t$ as inflation; the Dynamic IS curve is nested by interpreting $\xi_t$ as the real interest rate and $a_t$ as aggregate spending.

We depart from this benchmark by allowing information to be incomplete. By this we mean, not only imperfect observability of $\xi_t$, but also the accommodation of higher order uncertainty: we allow the agents to be doubtful about what others know, how attentive others are, and how others will respond to any news about the current state and the future prospects of the economy.

Such uncertainty can be the product either of the geographic segmentation of information (Lucas, 1972) or of bounded rationality in the form of rational inattention (Sims, 2003). Alternatively, it can formalize the difficulty in contemplating the equilibrium implications of exogenous shocks (Tirole, 2015). One way or another, the key is that the agents face uncertainty, not only about the path of the underlying fundamental, but also about the beliefs and the responses of others. Furthermore, because behavior is forward-looking, the relevant higher-order beliefs are forward-looking, too: the equilibrium outcome depends on the current beliefs of the future beliefs of others.

**Results.** Our main result is an observational equivalence between the incomplete-information economy described above and a variant, complete-information, representative-agent economy in which condition (1) is modified as follows:

$$a_t = \varphi \xi_t + \delta \omega_f E_t[a_{t+1}] + \omega_b a_{t-1}$$

for some scalars $\omega_f < 1$ and $\omega_b > 0$. The first modification ($\omega_f < 1$) represents myopia towards the future; the second ($\omega_b > 0$) represents anchoring to the past.

Although our observational-equivalence result relies on strong assumptions about the information structure, it encapsulates two general insights. The first is that the absence of common knowledge arrests the response of the current beliefs of the future beliefs of others to any news about the future fundamentals; this explains why agents behave as if they discount the future more heavily than what it is rational. The second insight is that the dynamics of learning causes extra persistence in the equilibrium outcome relative to that in the underlying fundamental; this explains why the current outcome appears to be anchored to the past outcome, conditional on the fundamental.

Each one of these insights has previously appeared in isolation of the other; see Angeletos and Lian (2016a) for the first and Woodford (2003) for the second. Our contribution is, not only to blend the two insights, but also to operationalize them in terms of the observational-equivalence result presented above. This builds a bridge to the DSGE literature (Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007) and offers a sharp illustration of how informational frictions help address the empirical challenges mentioned in the beginning of the Introduction.
Our result also clarifies how the belief friction accommodated in our paper—and more broadly in the literature on higher-order uncertainty—compares to the alternatives put forward in Gabaix (2016) and Farhi and Werning (2017). These works produce a similar form of myopia as ours by assuming that the subjective beliefs of the future actions of others do not adjust as much as their rational-expectations counterparts. They do not, however, produce our second feature, the anchoring of current outcomes to past outcomes. In terms of condition (2), they let $\omega_f < 1$ but restrict $\omega_b = 0$. But note that both the macroeconomic times series and the available evidence on expectations (Coibion and Gorodnichenko, 2012, 2015) favor theories that deliver $\omega_b > 0$, so as to capture the sluggish dynamics of outcomes and beliefs. Our approach therefore appears to, not only preserve the methodological benefits of the rational-expectations solution concept, but also offer a more successful account of salient features of the data.

Holding constant the level of noise in the economy, the myopia and the anchoring that obtain at the aggregate level intensify with the strength of the underlying general-equilibrium (GE) interactions. This is because such interactions regulate the importance of higher-order beliefs. When such interactions are absent, or when the agents contemplate how to respond to purely idiosyncratic shocks, there is no need to predict the behavior of others, so higher-order beliefs are irrelevant. When, instead, GE feedback effects are important, higher-order beliefs are also important. And because higher-order beliefs respond with both less amplitude and more sluggishness than lower-order beliefs, both the as-if myopia and the as-if anchoring are stronger.

Decision-theoretic frictions such as adjustment costs, habit, and sparsity can arrest the response of individual choices to idiosyncratic shocks, but do not necessarily produce a markedly different picture when considering the response of aggregate outcomes to aggregate shocks. In contrast, our approach predicts that the as-if distortions ought to be more salient at the macro level due to the aforementioned role played by GE effects and higher-order beliefs. This helps explain why the macroeconomic estimates of the habit persistence in consumption are much higher than their microeconomic counterparts (Havranek, Rusnak, and Sokolova, 2017), or why the persistence of inflation is much higher in the aggregate time series than in disaggregated data (Altissimo et al., 2010).

A similar point applies to rational inattention (Sims, 2003) insofar as this is treated purely as a theory of imperfect adjustment at the individual level, which is a perspective often encountered in the literature, as opposed to a theory of frictional coordination at the aggregate level, which is our preferred perspective. To understand what we mean, note that rational inattention can be disentangled

---

3These benefits include immunity to Lucas’ critique and the bypassing of the conundrum of whether and how the agents adjust their subjective beliefs in the light of experiences (actual outcomes) that systematically reject their beliefs.

4In fact, as illustrated by Caplin and Spulber (1987), certain forms of adjustment costs can be neutral at the aggregate level even if they induce large inertia at the individual level.

5A complementary but different point is made in Mackowiak and Wiederholt (2009), by arguing that rational agents find it optimal to pay little attention to aggregate shocks relative to idiosyncratic shocks because the former are relatively smaller and less relevant for individual payoffs.
from higher-order uncertainty either by abstracting from GE interactions or by letting all agents observe the same signal of the underlying state. But if GE effects are at work and the signal observed by one agent is, conditionally on the state, independent from the signals observed by other agents, then rational inattention gives rise to higher-order uncertainty and leads to the effects documented here.

**Application to Inflation.** The main application of our theory concerns the dynamics of inflation. We take the supply block of the New Keynesian model, which is summarized in the standard NKPC, and introduce incomplete information. We assume a modest degree of price stickiness, that is, one in line with the microeconomic evidence and textbook parameterizations, not the much higher one typically assumed in the DSGE literature. We next show that our theory helps rationalize existing estimates of the Hybrid NKPC, such as those found in Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2005). What is more, this is achieved with a level of informational friction that is consistent with estimates of that friction provided by Coibion and Gorodnichenko (2015) on the basis of surveys of inflation forecasts. This indicates how realistic informational frictions can 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 are unable to produce such sluggishness.6

**Application to Consumption.** Our second application shifts the focus to the demand block of the New Keynesian model: we show how incomplete information can arrest the response of aggregate consumption to monetary policy, can offer a micro-foundation of the form of consumption habit assumed in the DSGE literature, and can help reconcile the micro and macro estimates of such habit. We further show that the as-if distortions are largest when the Keynesian income-spending multiplier is stronger, a point that suggests an important interaction between our mechanism and financial frictions.

**Layout.** Section 2 expands on the relation of our paper to the literature. Section 3 introduces our framework. Section 4 characterizes the equilibrium. Section 5 develops the observational-equivalence result. Section 6 isolates the role of higher-order beliefs and elaborates on the distinct implications that our theory has at the macro and the micro level. Section 7 contains our main application, the one regarding the NKPC. Section 8 turns to the application to consumption. Section 9 expands on the logic behind our observational-equivalence result and on the roles played by higher-order uncertainty and learning. Section 10 concludes.

---

6See, for example, Golosov and Lucas (2007), Midrigan (2011), Alvarez and Lippi (2014), and Nakamura and Steinsson (2013). Different “details” such as the number of products sold by a firm and the so-called selection effect can rationalize a degree of price rigidity either much smaller than or almost as large as the one predicted by the standard NKPC. Yet, the menu-cost literature has not offered an explanation of the pronounced hump-shaped inflation dynamics that the DSGE literature has captured with ad hoc past-price indexation and the hybrid NKPC. Although our framework does not nest menu-cost models, its contributes towards filling that gap.
2 Related Literature

Our paper builds heavily on existing insights from the growing literature on informational frictions. Our main contribution is the observational-equivalence result, the recasting of the informational friction in terms of myopia and anchoring, and the mapping to the DSGE literature. An additional contribution is the empirical exercise in the context of inflation.

Sims (2003) emphasizes that rational inattention can generate sluggish response to shocks, but abstracts from GE effects and higher-order uncertainty and does not address why such sluggishness appears to be more pronounced at the macroeconomic level than the microeconomic one. Morris and Shin (2002) and Woodford (2003) emphasize that higher-order beliefs move less than first-order beliefs, but abstract from forward-looking behavior and do not explain how incomplete information rationalizes a form of myopia vis-a-vis the future. Conversely, Angeletos and Lian (2016a) provide the latter insight, but abstract from learning and do not study the dynamic adjustment in beliefs; they also consider a non-stationary environment that rules out recurring shocks. Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006) and Nimark (2017) share our emphasis on forward-looking beliefs, but do not share our analytical results and our empirical application.

Our paper also contains a modest methodological contribution. The literature has struggled with the complexity of the fixed point between the Kalman filter and the equilibrium dynamics. On the one hand, our baseline analysis addresses this problem with the assistance of the methods of Huo and Takayama (2015). On the other hand, the analysis is Section 9 cuts the Gordian knot by adapting an appropriate “orthogonalization” of the information structure. This allows for a closed-form characterization of the dynamics of the belief hierarchy and thereby for a sharp understanding of the interaction between higher-order uncertainty, learning, and forward-looking behavior.

Turning to the applied front, our paper is most closely related to Nimark (2008) and Coibion and Gorodnichenko (2012, 2015). Nimark (2008) is the first to combine Calvo-like sticky prices with incomplete information. It also shares our emphasis on forward-looking beliefs. It does not, however, contain either our quantitative evaluation or our mapping between incomplete information and the Hybrid NKPC. Coibion and Gorodnichenko (2012, 2015) argue that survey evidence on inflation forecasts are consistent with models featuring informational frictions. However, by treating inflation as an exogenous process, that work could not address whether such frictions explain the observed inflation dynamics. Our paper fills this gap and builds a bridge to the evidence on the Hybrid NKPC.

Related are also Mackowiak and Wiederholt (2009, 2015), Mankiw and Reis (2002, 2006), and Chung, Herbst, and Kiley (2015). These papers share with ours the broader theme that informational frictions can help explain salient features of the macroeconomic time series, but do not share either our empirical application to the Hybrid NKPC or our insights regarding the interaction of higher-order uncertainty with GE effects and forward-looking behavior.
Finally, our application to consumption connects to Carroll et al. (2018). This paper also argues that informational frictions help resolve the gap between the micro and the macro estimates of habit. However, it attributes the gap solely to an asymmetry in how much consumers know about idiosyncratic versus aggregate shocks. We instead argue that, even in the absence of such an asymmetry, the as-if habit can be higher at the macro level due to GE effects and the role of higher-order uncertainty.

3 Model

Time is discrete, indexed by \( t \in \{0, 1, \ldots\} \), 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 specify the best response of player \( i \) in period \( t \) as follows:

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

where \( \xi_t \) is the exogenous payoff-relevant fundamental, \( \mathbb{E}_{it}[] \) is the expectation operator conditional on the period–\( t \) information of player \( i \), and \( \varphi > 0 \) and \( \beta, \gamma \in [0, 1] \) are fixed parameters.

Our framework is similar to the one used in Section 5 of Angeletos and Lian (2016a). But whereas that paper focuses on the more narrow question of how \( a_0 \) responds to news about \( \xi_T \) for some \( T \geq 1 \) holding \( \xi_t = 0 \) for all \( t \neq T \), our paper accommodates recurrent shocks and proceeds to characterize how learning and higher-order uncertainty affect the equilibrium dynamics. Our framework also resembles the beauty-contest games studied in Morris and Shin (2002), Angeletos and Pavan (2007), and Woodford (2003), except that behavior is not forward looking in these papers. By contrast, forward-looking behavior is at the core of the applications we are concerned with in this paper.

The extent to which behavior is forward-looking is parameterized by the scalars \( \beta \) and \( \gamma \). In particular, \( \beta \) identifies the extent to which the optimal action of an agent depends on her expectations of her own future action, whereas \( \gamma \) controls the extent to which the optimal action of an agent depends on her expectations of the future actions of others. In applications, this kind of dynamic strategic complementarity captures GE effects such as the feedback from expectations of future inflation to current inflation (this feedback lies behind the NKPC), or the feedback from expectations of future aggregate spending to current aggregate spending (this is the modern version of the Keynesian cross).

To see more clearly how the current outcome depends on expectations of the entire future, iterate condition (3) and aggregate across agents to obtain the following expression:

\[
a_t = \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t [\varphi_t \xi_{t+k}] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t [a_{t+k+1}],
\]

where \( \mathbb{E}_t[.] \) denotes the average expectation in the cross-section of the population. It is then evident
that the aggregate outcome today depends, not only on the beliefs of the future path of the exogenous fundamental, but also on the beliefs of the future path of the endogenous outcome itself. Furthermore, $\beta$ controls the rate at which the future is discounted, and $\gamma$ controls the strength of the GE feedback from the expectations of the future values of the aggregate outcome to its current value.

**Interpretation.** When all agents share the same information, we can replace $\mathbb{E}_t[\cdot]$ with the expectation of the representative agent, that is, the expectation conditional on the common information set. We can then use the Law of Iterated Expectations to reduce condition (4) to the following:

$$a_t = \mathbb{E}_t[\varphi \xi_t + \delta a_{t+1}], \quad (5)$$

where $\mathbb{E}_t[\cdot]$ denotes the expectation of the representative agent and $\delta \equiv \beta + \gamma$. It is then immediate to see that the complete-information version of our framework nests the two building blocks of the NKPC is nested with $a_t$ standing for inflation and $\xi_t$ for the real marginal cost or the output gap; and the Dynamic IS Cure (that is, the Euler condition of the representative consumer) is nested with $a_t$ standing for consumption and $\xi_t$ for the real rate of return. Alternatively, condition (5) can represent a risk-neutral asset-pricing equation with $a_t$ standing for the asset price and $\xi_t$ for the next-period dividend. A similar point applies to the Q-theory of investment. We will study these applications, and their incomplete-information variants, in Sections 7 and 8.

**Shocks and Information.** To complete the model, we need to specify the stochastic process for the exogenous fundamental and of the information structure. For the bulk of our analysis (i.e., with the exception of Section 9), we make the following two assumptions. First, we let the fundamental $\xi_t$ follow an AR(1) process:

$$\xi_t = \rho \xi_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, 1) \quad (6)$$

where $\rho \in (0, 1)$ parameterizes the persistence of the fundamental. (Note that the volatility is normalized to 1.) Second, we assume that player $i$ receives a new signal each period, given by

$$x_{it} = \xi_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \sigma^2) \quad (7)$$

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. The main justification is that they facilitate a sharp characterization of the incomplete-information outcomes and lead to our observational-equivalence result. Yet, the insights that obtain from our analysis are not unduly sensitive to these assumptions. This will become evident in Section 9, where we study the structure and the dynamics of the underlying higher-order beliefs under a more flexible information structure. The “beauty” of our observation-equivalence result is then lost, but the key ideas survive.
Beyond FIRE (full-information, rational expectations). Throughout, we use Rational Expectations Equilibrium as our solution concept. But whereas a standard practice is to combine this solution concept with complete, or “full,” information, which herein means setting $\sigma = 0$, we allow for $\sigma > 0$ and study how this affects the predictions of the theory.

It is important to recognize that $\sigma = 0$ imposes, not only that every agent knows $\xi_t$, but also that she is perfectly confident that every other agent knows $\xi_t$, that everyone knows that everyone knows, and so on, ad infinitum. Such common knowledge together with the rational-expectations solution concept implies that the agents can reach a common belief about one another’s behavior and thereby about the future path of the endogenous outcome. In a nutshell, setting $\sigma = 0$ prevents the agents from having any doubts about the awareness and the responsiveness of others. Conversely, letting $\sigma > 0$ helps accommodate such doubts. That’s the essence of higher-order uncertainty.\(^7\)

Last but not least, letting $\sigma > 0$ can be thought of as a proxy for rational inattention in the sense of Sims (2003) and for costly contemplation in the sense of Tirole (2015). We wholeheartedly embrace this interpretation, but do not attempt to endogenize the level of attention, or the depth of cognition: we treat the information structure as exogenous.

4 Equilibrium and Higher-Order Beliefs

In this section we characterize the equilibrium and elaborate on the role of higher-order beliefs.

Let us start with the case in which information is complete, in the sense that that all agents share the same information and therefore face no uncertainty about one another’s beliefs. In this case, the aggregate outcome satisfies condition (5). Iterating this condition forward gives

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

where, recall, $\delta \equiv \beta + \gamma$. This stylizes how outcomes are determined in any unique-equilibrium model in which expectations are rational and agents share the same information: outcomes are pinned down by first-order beliefs of the underlying fundamentals.

To sharpen the characterization of the complete-information benchmark, let the information set of the representative agent be the history of $\xi_t$; this is nested in (7) by letting $\sigma = 0$. In this case, we have $\mathbb{E}_t[\xi_{t+k}] = \rho^k \xi_t$ for all $t, \tau$. We thus reach the following result, which states that the aggregate outcome follows the same AR(1) process as the fundamental, rescaled by the factor $\frac{\varphi}{1-\rho \delta}$.

\(^7\)Like the bulk of the applied literature, our model confounds first- and higher-order uncertainty. Although the two are likely to come together in practice, they play distinct roles in the inner workings of the theory and have distinct observable implications. This will become clear as we proceed.
Proposition 1. In the frictionless benchmark \((\sigma = 0)\), the equilibrium outcome is given by

\[
a_t = \frac{\varphi}{1 - \rho^\delta} \xi_t = \Phi^*(L) \eta_t, \quad \text{with} \quad \Phi^*(L) \equiv \frac{\varphi}{1 - \rho^\delta} \frac{1}{1 - \rho^L},
\]

where \(L\) henceforth denotes the lag operator.

Consider now the case in which information is incomplete, in the sense that the agents have differential information and cannot reach a common belief about the future path of either the exogenous fundamental, \(\xi_t\), or the endogenous outcome, \(a_t\). As already explained, this is captured in our setting by specifying the signals as in (7) and letting \(\sigma > 0\). But before characterizing the equilibrium under this particular information structure, let us explain how the equilibrium depends on higher-order beliefs for arbitrary information structures. This helps understand the underpinnings of our observational-equivalence and the robustness of our insights.

As noted before, by iterating condition (4) forward and aggregating across \(i\), we get

\[
a_t = \varphi \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t [\xi_{t+k}] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t [a_{t+k+1}] \tag{9}
\]

This underscores that the equilibrium outcome in period \(t\) depends, not only on the average forecast of the fundamental, but also on the average forecast of the current and future values of the outcome itself. To understand the equilibrium dynamics of \(a_t\), we have must therefore also understand the equilibrium dynamics of this kind of forecasts—and this is where higher-order beliefs come into play.

To illustrate, let \(\beta = 0 < \gamma\). This case is too narrow for the applications of interest, but it useful because of its relative simplicity; it is also the case studied in Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), and Nimark (2017). In this case, the equilibrium outcome satisfies

\[
a_t = \varphi \mathbb{E}_t [\xi_t] + \gamma \mathbb{E}_t [a_{t+1}] \tag{10}
\]

Iterating this condition once gives

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

from which it is evident that the equilibrium outcome depends on on a particular kind of second-order beliefs, namely the future beliefs of the future fundamental and the future outcome. By iterating condition (10) again and again, we can ultimately express the equilibrium outcome as a function of the hierarchy of beliefs about the current and future values of the fundamental:

\[
a_t = \varphi \sum_{h=0}^{\infty} \gamma^h \mathbb{E}_t^{h+1} [\xi_{t+h}] \tag{11}
\]
where, for any random variable $X$, $\mathbb{F}^h_t [X]$ is defined recursively by

$$
\mathbb{F}^1_t [X] \equiv \mathbb{E}_t [X] \quad \text{and} \quad \mathbb{F}^h_t [X] \equiv \mathbb{E}_t \left[ \mathbb{F}^{h-1}_{t+1} [X] \right] \quad \forall h \geq 2.
$$

Condition (11) is similar to related results in Morris and Shin (2002) and Woodford (2003), except for the following subtle difference. In these papers, the relevant higher-order beliefs are the beliefs about the current beliefs of others. Here, they are the beliefs about the future beliefs of others. This difference reflects the importance of forward-looking behavior in our setting.

Consider next the more general case, in which both $\gamma > 0$ and $\beta > 0$. This case is relevant for the applications studied in Section 7 and 8. In this case, the class of higher-order beliefs that drive the equilibrium outcome is much richer than the one described above. To see this, let $\varphi^2 = 1$ (this is completely innocuous) and rewrite condition (10) as follows:

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

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:

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

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

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

It is then evident that the relevant second-order beliefs are, not only those corresponding to the next period, but also those corresponding to all future periods, namely $\mathbb{E}_t [\mathbb{E}_{t+k} [\xi_{t+k}]]$ for every $k \geq 1$.

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:

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

Note that, for any $t$ and any $k \geq 2$, there are $k - 1$ types of second-order beliefs about $\xi_{t+k}$, 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. In short, there is severe curse of dimensionality if one tries to understand the joint
dynamics of all the relevant higher-order beliefs.

This in turn explains why it is impossible to obtain a tractable solution for arbitrary specifications of the process for \( \xi_t \) and of the dynamics of learning. A highly tractable solution, however, becomes possible under the particular specification we introduced earlier on.

**Proposition 2.** (i) The equilibrium exists and is unique.

(ii) There exists a scalar \( \vartheta \in (0, \rho) \) such that the aggregate outcome is given by

\[
a_t = \Phi(L; \rho, \vartheta) \eta_t, \quad \text{with} \quad \Phi(L; \rho, \vartheta) \equiv \left( 1 - \frac{\vartheta}{\rho} \right) \left( \frac{1}{1 - \vartheta \rho} \right) \Phi^*(L),
\]

and where \( \Phi^*(L) \) is the frictionless counterpart obtained in Proposition 1.

(iii) The scalar \( \vartheta \) is a function of \( (\sigma, \rho, \beta, \gamma) \) and is given by the reciprocal of the largest root of the following cubic:

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

Part (i) establishes existence and uniqueness of equilibrium. Part (ii) obtains the equilibrium dynamics as a transformation of the frictionless counterpart: relative to that benchmark, the incomplete-information case features a smaller impact effect, captured by the term \( 1 - \frac{\vartheta}{\rho} \in (0, 1) \) in condition (12), and additional persistence, captured by the term \( \vartheta L \). Part (iii) completes the characterization of the equilibrium by delivering \( \vartheta \) as the solution to a simple cubic. The latter, and hence also the value of \( \vartheta \), depends on the level of noise \( (\sigma) \), the persistence of the underlying fundamental \( (\rho) \), and the parameters that control how much the agents care about the future \( (\beta, \gamma) \).

To develop some intuition for the result, consider momentarily the case in which \( \gamma = 0 \). By shutting down the GE effect, this case also shuts down the role of higher-order uncertainty. The aggregate outcome is then given by

\[
a_t = \frac{\varphi}{1 - \delta \rho} \mathbb{E}_t[\xi_t],
\]

which is the same as the complete-information outcome, modulo the replacement of \( \xi_t \), the actual fundamental, with \( \mathbb{E}_t[\xi_t] \), the average first-order forecast of it.\(^8\) Furthermore, using the Kalman filter, one can show that the average first-order forecast follows an AR(2) process given by

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

where \( \lambda = \rho(1 - G) \), with \( G \) being the Kalman gain. Combining (13) and (14) proves, in effect, that Proposition 2 holds with \( \vartheta = \lambda \) when \( \gamma = 0 \).

---

\(^8\)To see this, note that, when \( \gamma = 0 \), condition (3) reduces to \( a_{it} = \varphi \mathbb{E}_t[\xi_t] + \delta \mathbb{E}_t[a_{it+1}] \), with \( \delta = \beta \). Iterating this forward gives \( a_{it} = \varphi \mathbb{E}_t[\sum_{k} \delta^k \xi_{t+k}] = \frac{\varphi}{1 - \delta \rho} \mathbb{E}_t[\xi_t] \). Aggregating gives the result.
What happens when $\gamma > 0$, that is, when we switch on the GE effect? Proposition 2 then applies with $\vartheta > \lambda$. That is, the equilibrium dynamics exhibits less amplitude and more persistence, not only relative to the complete-information counterpart, but also relative to the first-order forecast of the fundamental. This reflects the role of higher-order uncertainty.

We elaborate on the precise mechanics of higher-order uncertainty in Section 9. The basic logic is that the dynamics of higher-order beliefs display less amplitude and more persistence than the dynamics of first-order beliefs. Because the equilibrium outcome is driven in part by first-order beliefs and in part by higher-order beliefs, we then also have that the equilibrium outcome displays less amplitude and more persistence than the first-order beliefs. Proposition 2 is a sharp illustration of this logic, made possible thanks to the particular signal structure we have assumed.

Unlike the first-order beliefs, the second- and higher-order beliefs do not follow AR(2) processes. There is therefore a certain “magic” behind the property that the equilibrium outcome follows an AR(2) even though the higher-order beliefs do not. This magic, and the proof of Proposition 2, builds on a method developed in Huo and Takayama (2015). We look for a function $\Phi$ in the space of analytic functions. The key is the use of the Wiener filter in characterizing the forecasts of the behavior of others. Insofar as the fundamental and the signals follow ARMA processes, this strategy permits the analyst to bypass the infinite state space that is necessary for tracking the hierarchy of beliefs and, instead, to identify a finite state space that is sufficient for tracking the equilibrium dynamics. With the assumed specification, it can be shown the solution for $a_t$ is an AR(2) process, regardless of the magnitude of $\gamma$ and of the associated importance of higher-order beliefs. Furthermore, an analytic solution for the coefficient $\vartheta$ can be obtained, as described in part (iii) of the proposition.

Note that $\vartheta$ plays a dual role in the impulse response of $a_t$: a higher $\vartheta$ means both a smaller impact effect, captured by the factor $1 - \frac{\vartheta}{\rho}$ in condition (12), and a more sluggish build up over time, captured by the lag term $\vartheta L$. These properties anticipate the observational equivalence result presented in the next section. With this point in mind, we next study the determinants of $\vartheta$.

**Proposition 3.** (i) $\vartheta$ is continuously increasing in $\sigma$, with $\vartheta \to 0$ as $\sigma \to 0$ and $\vartheta \to \rho$ as $\sigma \to \infty$.

(ii) $\vartheta$ is increasing in $\gamma$.

(iii) $\vartheta$ is increasing in $\beta$ if $\gamma > 0$, and is invariant to $\beta$ if $\gamma = 0$.

Part (i) verifies that a larger informational friction implies both a smaller initial response and a more sluggish build up over time. Part (ii) establishes that both of these effects are intensified when the GE feedback is stronger, in the sense of a larger $\gamma$. This follows from two key points made earlier: that a higher $\gamma$ raises the relative importance of higher-order beliefs; and that the dynamic response of the latter displays less amplitude and more persistence than that of the first-order beliefs. Part (iii) adds

---

9The fact that $\vartheta > \lambda$ is not obvious from looking at Proposition 2, but follows from the property that $\vartheta$ is increasing in $\gamma$, which is proved in Proposition 3.
that, insofar the GE effect is non-zero, \( \vartheta \) is strictly increasing in \( \beta \) as well. To understand this finding, recall that \( \beta \) measures the extent to which the agent is forward-looking vis-a-vis her own future action. But as long as \( \gamma > 0 \), the agent’s own future actions are sensitive to her future expectations of the future actions of others. It follows that a higher \( \beta \) indirectly increases the dependance of the current outcome to forward-looking higher-order beliefs.

It is worth emphasizing that, by design of our baseline model, the results obtained above confound the roles of first- and higher-order uncertainty. In particular, if we shut down the role of higher-order uncertainty by setting \( \gamma = 0 \) (no GE feedback), these results to hold, subject to the restriction that the value of \( \vartheta \) that obtains with \( \gamma = 0 \) is smaller than the one that obtains with \( \gamma > 0 \). From this perspective, higher-order uncertainty amplifies the equilibrium effects of first-order uncertainty. That said, the extensions considered in Sections 6 and 9 allow us to elaborate on why higher-order uncertainty is, by itself, sufficient for the qualitative properties we have established here. These extensions also corroborate the interpretation we put forward in the next section, in terms of myopia and anchoring.

5 Equivalence Result

Let us momentarily put aside our model and, instead, consider a variant, representative-agent economy in which the aggregate Euler condition (5) is modified as follows:

\[
a_t = \varphi \xi_t + \delta \omega_f E_t [a_{t+1}] + \omega_b a_{t-1}
\]

(15)

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 \( \omega_f \) represents a higher discounting of the future, or less forward-looking behavior; a higher \( \omega_b \) represents a greater anchoring of the current outcome to the past outcome, or more backward-looking behavior.

Condition (15) nests the following examples: the Euler condition of a representative consumer who exhibits habit in consumption; a variant of the Q-theory of investment that lets the firms face a cost to adjusting their rate of investment rather than to adjusting their capital stock; and the so-called Hybrid NKPC, a variant of the NKPC that pegs current inflation to past inflation. 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 \( a_t = \frac{\zeta_0}{1 - \delta \rho} \xi_t \), for some coefficients \( (\zeta_0, \zeta_1) \) that are functions of \( (\omega_f, \omega_b) \) and \( (\varphi, \delta, \rho) \). In comparison, the equilibrium outcome of our incomplete-information economy is given by \( a_t = \left( 1 - \frac{\vartheta}{\rho} \right) \left( \frac{\varphi}{1 - \delta \rho} \right) \left( \frac{1}{1 - \delta \rho} \right) \xi_t \). It follows that the two economies generate the same joint process for \( \xi_t \) and \( a_t \) if and only if \( \zeta_0 = \left( 1 - \frac{\vartheta}{\rho} \right) \left( \frac{\varphi}{1 - \delta \rho} \right) \) and \( \zeta_1 = \vartheta \). Using this observation along with the characterization of the coefficients \( \vartheta, \zeta_0, \) and \( \zeta_1 \) (which can be found in the Appendix), we reach the following result.
Proposition 4. Fix $(\varphi, \beta, \gamma, \rho)$ and let $\delta \equiv \beta + \gamma$.

(i) For any $\sigma > 0$ in the incomplete-information economy, there exists a unique pair $(\omega_f, \omega_b)$ in the hybrid economy, with $\omega_f < 1$ and $\omega_b > 0$, such that the two economies are observationally equivalent in the sense of generating the same joint dynamics for the fundamental and the aggregate outcome.

(ii) A greater informational friction (higher $\sigma$) and/or a stronger GE feedback (higher $\gamma$) maps to both greater myopia (lower $\omega_f$) and greater anchoring (higher $\omega_b$) in the hybrid model.

Part (i) allows us to recast the informational friction as the combination of two behavioral distortions: extra discounting of the future, or myopia, in the form of $\omega_f < 1$; backward-looking behavior, or anchoring of the current outcome to past outcome, in the form of $\omega_b > 0$. Part (ii), in turn, establishes that both of these as-if distortions get intensified as we increase either the severity of the informational friction (measured by $\sigma$) or the strength of the underlying GE effects (parameterized by $\gamma$).

This result draws a link between two strands of the literature: the theoretical one on incomplete information and higher-order uncertainty, which builds on Morris and Shin (2002, 1998, 2003) and Woodford (2003) and is reviewed in Angeletos and Lian (2016b); and the quantitative one on DSGE models, which follows Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007).

As noted in the Introduction, the latter literature has taken the key forward-looking equations of textbook macroeconomic models—the Euler condition of the representative consumer, the Q theory of investment, and the NKPC—and has modified them in an arguably ad hoc manner, and equation-by-equation, by adding habit in consumption, adjustment costs to the rate of investment, and automatic indexation of the prices of the firms that do not have the option to reset their prices. These modifications lack supporting microeconomic evidence, but are necessary in DSGE models for two related purposes: the pronounced hump shapes in the dynamic response of consumption, investment, and inflation to identified shocks; and to accommodate the strong comovement of the macroeconomic quantities over the business cycle. Our result illustrates how incomplete information can offer a unified substitute to, or micro-foundation of, these modifications.

This result helps also clarify how our contribution relates to Gabaix (2016) and Farhi and Werning (2017). These works depart from rational expectations in a manner that helps capture a similar form of myopia as the one captured here and in the related contribution by Angeletos and Lian (2016a). They do not, however, provide a theory of why the current outcomes may be anchored to past outcomes. That is, they accommodate $\omega_f < 1$ but restrict $\omega_b = 0$. But the facts appear to demand $\omega_b > 0$ at least as much as they demand $\omega_f < 1$, which is why the DSGE literature was lead to include the aforementioned bells and whistles in the first place. It follows that, in comparison to the forms of bounded foresight put forward in the aforecited works, the alternative of allowing for rational inattention and higher-order uncertainty offers, not only the methodological benefit of reconciling bounded foresight with rational expectations, but also a more promising account of the macroeconomic data.
But let us put aside these connections to the literature and let us expand on the empirical content of our result. Although Proposition 4 guarantees that the incomplete-information economy can always be mapped to a hybrid economy, the converse is not true: a hybrid economy is observationally equivalent to and incomplete-information economy only when \( \omega_f \) and \( \omega_b \) happen to lie together on a particular line in the \((\omega_f, \omega_b)\) space.

**Proposition 5.** There exists a function \( \Omega \) such that 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 the following restriction is satisfied:

\[
\omega_f = \Omega(\omega_b; \delta, \rho)
\]

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

Apart from clarifying the mapping between the two models, this result offers a simple test for our theory. Suppose, in particular, that one estimates condition (15) on the data and gets some estimates from \( \omega_f \) and \( \omega_b \). Suppose further that one knows \( \delta \) and \( \rho \) from independent sources. One can then test the hypothesis that condition (16) is satisfied. If it does, then and only then the data is compatible with our theory. And provided that this is true, the estimates of \( \omega_f \) and \( \omega_b \) can be used to identify \( \sigma \).

Additional testable predictions, or overidentifying restrictions, can be obtained by looking at the forecasts of future outcomes. Let \( \epsilon_t^k \equiv a_{t+k} - \bar{E}_t [a_{t+k}] \) be the realized average \( k \)-period ahead forecast error. In the hybrid model, \( \epsilon_t^k \) is serially uncorrelated, because the economy is populated by a rational representative agent. In our model, instead, \( \epsilon_t^k \) is serially correlated, because the economy is populated by a large number of differentially informed agents, whose forecast errors are uncorrelated at the individual level but not at the aggregate level. Importantly, because the serial correlation of the average forecast error in our model depends on \( \sigma \), this provides us with an additional restriction that can be used to identify \( \sigma \) and to test the model. We will put these ideas at work in Section 7.

### 6 GE vs PE, and Macro vs Micro Responses

In this section, we extend the analysis to a setting that features both aggregate and idiosyncratic shocks. This serves two purposes. First, it helps isolate the role of higher-order beliefs and recast our mechanism as one that regards exclusively GE effects. Second, it illustrates 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).
6.1 Adding Idiosyncratic Shocks and Disentangling GE from PE

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

\[ a_{it} = E_{it} [\varphi \xi_{it} + \beta a_{it+1} + \gamma a_{t+1}] \] (17)

where

\[ \xi_{it} = \xi_t + \xi_{it} \]

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

Regardless of the information structure, the equilibrium action of agent \(i\) can be expressed as

\[ a_{it} = \sum_{k=0}^{\infty} \beta^k E_{it} [\xi_{it+k}] + \gamma \sum_{k=0}^{\infty} \beta^k E_{it} [a_{t+k+1}] = PE_{it} + GE_{it} \]

where

\[ PE_{it} \equiv \sum_{k=0}^{\infty} \beta^k E_{it} [\xi_{it+k}] \quad \text{and} \quad GE_{it} \equiv \gamma \sum_{k=0}^{\infty} \beta^k E_{it} [a_{t+k+1}] \]

The first term captures the effect of any shock on the optimal behavior of agent \(i\) holding constant her beliefs of the actions of others. This effect represents a direct or partial-equilibrium effect and is pinned down by the first-order beliefs of agent \(i\) about the current value and the future path of her own fundamental. The second term captures the additional effect that obtains through the adjustment of agent \(i\)'s beliefs of the future actions of others. This effect represents an indirect or general-equilibrium effect and is pinned down by the agent's second- and higher-order beliefs, namely by her beliefs of the beliefs of others about their own future fundamentals, her beliefs of the second-order beliefs of others, and so on. By the same token, the aggregate effect of a shock can be split into a PE and a GE component by using the identity

\[ a_t = PE_t + GE_t, \]

where \(PE_t\) and \(GE_t\) are the cross-sectional averages of, respectively, \(PE_{it}\) and \(GE_{it}\).

Clearly, the GE effect of a shock is non-zero only insofar as it affects a non-zero mass of agents and \(\gamma\) is positive (there is a GE interaction to start with). Furthermore, the GE effect is tied exclusively to higher-order beliefs: predicting the aggregate outcome is the same as predicting the actions of others, which in turn is the same as predicting the beliefs of others. Conversely, the PE effect isolates the role of the first-order beliefs, namely the beliefs of the agents about their own fundamentals.

\(^{10}\) The restriction that the two kinds of shocks have the same persistence is only for expositional simplicity.
These properties are true even in our baseline specification. However, because that specification rules out idiosyncratic shocks, it equates the fundamental of one agent with the fundamental of every other agent. This precludes typical agent from facing uncertainty about the beliefs of others unless she also faces uncertainty about her own (and common) fundamental. That is, the baseline specification confounds the first-order uncertainty of an agent about her own fundamental with her higher-order uncertainty about the beliefs of others. We next consider a variant that helps separate the two forms of uncertainty.

6.2 Isolating Higher-Order Uncertainty

To isolate the role of higher-order uncertainty, we let the agents observe perfectly their own fundamental. Specifically, the information received by agent $i$ in period $t$ is given by the pair

$$z_{it} = \{\xi_{it}, x_{it}\},$$

where $\xi_{it}$ is the agent’s own fundamental and $x_{it}$ is a signal of the aggregate fundamental, specified as in our baseline model. We also focus on the limit as $\frac{V(\xi_{it})}{V(t)} \rightarrow \infty$ (large idiosyncratic shocks).

In this variant, we can effectively separate the forecasts of one’s own fundamental from the forecasts of the fundamentals of others: relative to the signal $x_{it}$, the own fundamental $\xi_{it}$ contains negligible information about the current and the future values of $\xi_t$; and symmetrically, relative to $\xi_{it}$, $x_{it}$ contains negligible information about the current and future values of $\xi_{it}$. It follows that

$$\text{PE}_{it} \equiv \sum_{k=0}^{\infty} \beta^k E_{it} [\xi_{it+k}] = \frac{1}{1 - \rho\beta} \xi_{it},$$

which is the same PE effect as the one under full information. This confirms that the specification under consideration shuts down the effects of the informational friction that operate via first-order beliefs and PE responses; what remains active is only the mechanism that regards higher-order beliefs and GE feedback loops. The next result completes the analysis of the variant under consideration by characterizing the GE effect.

**Proposition 6.** Consider the variant described above. The individual and the aggregate outcomes are given by, respectively,

$$a_{it} = \frac{1}{1 - \rho\beta} \xi_{it} + \overline{GE}_t + u_{it} \quad \text{and} \quad a_t = \frac{1}{1 - \rho\beta} \xi_t + \overline{GE}_t,$$

where $u_{it}$ is idiosyncratic noise, orthogonal to both $\xi_t$ and $\xi_{it}$, and $\overline{GE}_t$ follows the same law of motion as the equilibrium outcome in the baseline model.
The response to idiosyncratic shocks is therefore the same as in the full-information benchmark. By contrast, the response to aggregate shocks is distorted because, and only because, the agents face uncertainty about the beliefs and the actions of others. Furthermore, this is now cleanly decomposed into two parts: a PE effect, which is captured by $\frac{1}{1-\rho^3} \xi_t$; and a GE effect, which is denoted by $\overline{GE}_t$.

The latter follows the same law of motion as the equilibrium outcome in the baseline model.

This result refines the main lessons of our paper: the myopia and the anchoring documented in our earlier analysis are herein recast as mechanisms that regard exclusively the response of higher-order beliefs and the GE effects of the aggregate shocks. In other words, although first- and higher-order uncertainty complement each other and are likely to go hand-in-hand in practice, higher-order uncertainty alone suffices for the kinds of myopia and anchoring we have documented.

### 6.3 Micro- vs Macro-level Distortions

The preceding result also offers a sharp illustration of how our approach helps resolve the disconnect between micro and macro estimates of habit, adjustment costs, etc. As already explained, in the case considered above the informational friction distorts the response of the aggregate outcome to aggregate shocks without distorting at all the response of individual outcomes to idiosyncratic shocks. It follows that, if an econometrician misinterprets the anchoring effect of the informational friction as the product of habit, she will 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).

Of course, the complete absence of a distortion at the micro level hinges on the assumption that each agent observes perfectly her own fundamentals. Relaxing this assumption 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 the following reasons.

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, and pay relatively more attention to, idiosyncratic shocks, simply because such shocks are more volatile and there is higher return in reducing uncertainty about them. This is the mechanism articulated in Mackowiak and Wiederholt (2009) and boils down to having less first-order uncertainty about idiosyncratic than aggregate shocks, a property that contributes towards a lower distortion at the micro level. 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 because of the role played by higher-order uncertainty. In short, the mechanism identified in our paper and the one identified in the aforementioned paper complement each other towards generating more pronounced distortions at the macro level than at the micro level.\footnote{We verify all these intuitions in Appendix B with a variant that lets both kinds of shocks be observed with noise.}
7 Myopia and Inertia in Inflation

In this section, we consider an incomplete-information version of the aggregate-supply block of the New Keynesian model. The question of interest is how inflation responds to innovations in the output gap, or the real marginal cost. When information is complete, this response is governed by the standard version of the New Keynesian Phillips Curve (NKPC). In this case, the predicted response is too large and too fast compared to what seems to be true in the data. When instead information is incomplete, it is as if the response of inflation is governed by the Hybrid NKPC, a variant that introduces myopia and inertia. Importantly, our theory is able 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).

7.1 Setup and Theoretical Results

There is a continuum of firms, each producing a differentiated commodity under a linear technology. Firms set prices optimally, but can adjust them only infrequently. Each period, a firm has the option to reset its price with probability $1 - \theta$; with remaining probability, it is stuck at the previous-period price. Technology is linear, so that the real marginal cost faced by each firm is invariant to its production level. The real marginal cost corresponds to the aggregate fundamental in our earlier analysis; the outcome is inflation.

Consider a firm that has the option to adjust its price in period $t$. The optimal reset price is given by the solution to the following problem:

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

subject to the demand function

$$Y_{i,t+k} = \left( \frac{P_t}{P_{t+k}} \right)^{-\epsilon} Y_{t+k}$$

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 $t + k$, $P_t$ 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 the firm not having changed its price since period $t$, and $\Psi_{t+k}$ 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 reach the following characterization of the optimal rest price:

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

(18)
with the understanding that all the variables are henceforth expressed in terms of log-deviations from the steady state. This condition means that the optimal reset price is a weighted average of the firm’s belief of the current and future nominal marginal costs. The only difference from the textbook New Keynesian model is that the expectation operator is allowed to differ across firms, reflecting the heterogeneity of their information and the associated higher-order uncertainty.

Since only a fraction \(1 - \theta\) of the firms adjust their prices each period, the following accounting identify holds:

\[
p_t = (1 - \theta) \int p^*_t + \theta p_{t-1}
\]

Combining this with condition (18) gives the nominal price level today as a function of both the past price level and the current average beliefs of the future price levels. Because this contains both a backward-looking and a forward-looking component, it is not directly nested in our earlier analysis. However, such nesting becomes possible if (i) we transform the problem from one stated in terms of the price level to one stated in terms of the inflation rate and (ii) we make the simplifying assumption that the firms observe that current price level but preclude them from extracting information from it.\(^{12}\)

This permits us to restate condition (18) as

\[
p^*_t - p_{t-1} = (1 - \delta \theta) \sum_{k=0}^{\infty} (\delta \theta)^k \mathbb{E}_t[\psi_{t+k} + \pi_{t+k}],
\]

By (19), on the other hand, we have

\[
\pi_t \equiv p_t - p_{t-1} = (1 - \theta) \int (p^*_t - p_{t-1})
\]

Combining and rearranging, we arrive at the following expression for inflation:

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

When information is complete, we can replace \(\mathbb{E}_t[\cdot]\) with \(\mathbb{E}_t[\cdot]\), the expectation operator conditional the common information set. We can then use the Law of Iterated Expectations to reduce condition (21) to the following:

\[
\pi_t = \kappa \psi_t + \delta \mathbb{E}_t[\pi_{t+1}],
\]

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

\(^{12}\)This assumption can be justified on the grounds that the observation of inflation contains little information about the underlying output gap because most of the short-run variation in inflation is due to orthogonal cost-push shocks. Alternatively, as in Vives and Yang (2017), the failure to extract information from the realized outcomes can be interpreted as a form of bounded rationality. In any event, this assumption facilitates an exact nesting but does not drive our findings.
When, instead, information is incomplete, the above condition no more holds. Instead, inflation must be understood as a solution to a dynamic beauty contest of the type we have studied in the rest of the paper. In particular, the current setting is nested in our earlier analysis by mapping $\psi_t$ and $\pi_t$ to, respectively, $\xi_t$ and $a_t$; and by letting

$$\varphi = \kappa \quad \beta = \delta \theta \quad \text{and} \quad \gamma = \delta (1 - \theta).$$

It is worth noting that the GE feedback, as measured by $\gamma$, is larger when prices are more flexible. This means that higher-price flexibility reinforces the role of higher-order uncertainty, a point we revisit shortly.

To apply our analytical results, we finally assume that $\psi_t$, the real marginal cost, follows an AR(1) process and that the information structure take the form introduced in Section 3. The following is then an immediate application of Proposition 4.

**Proposition 7.** (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:

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

(ii) For any given level of noise, increasing the degree of price flexibility (i.e., reducing $\theta$) results to a lower $\omega_f$ and a higher $\omega_b$.

Part (i) establishes that, when information is incomplete, it is as if inflation is governed by a variant of the NKPC that introduces myopia, in the form of $\omega_f < 1$, along with a backward-looking component, in the form of $\omega_b > 0$. This is similar to the Hybrid NKPC considered in Gali and Gertler (1999), Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007).

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 $\kappa$ in condition (22). 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 seems to present at the micro-economic level, especially if one reads the evidence through the lenses 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. We corroborate this point in the sequel by showing how our framework can reconcile salient features of the inflation dynamics with a modest degree of price stickiness.
7.2 Testing our Theory

The Hybrid NKPC estimated in Gali and Gertler (1999), Gali, Gertler, and Lopez-Salido (2005), and elsewhere is the same as the one seen in (23). There are, however, two differences. First, our theory restricts the pair \((\omega_f, \omega_b)\) 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 \((\omega_f, \omega_b)\) to the stochastic properties of inflation forecasts. We now use these restrictions to test our theory.

**Matching Existing Estimates of the Hybrid NKPC.** In a special issue of the Journal of Monetary Economics devoted to the estimation of the NKPC, Gali, Gertler, and Lopez-Salido (2005) discuss various estimation issues, synthesize the related literature, and offer a few baseline estimates of the pair \((\omega_f, \omega_b)\). A quick test of our theory is then provided by checking whether these estimates happen to satisfy the restriction seen in Proposition 5.

This proposition gives the locus of the pairs \((\omega_f, \omega_b)\) that are compatible with our theory for some level of noise. To construct this locus, we only need to specify \(\delta, \theta, \) and \(\rho\). We set \(\delta = 0.99, \theta = 0.6, \) and \(\rho = 0.95\). The value of \(\theta\) 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 \(\rho\), on the other hand, is obtained by estimating an AR(1) process on the labor share, which is the 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.

Consider now the estimates of \((\omega_f, \omega_b)\) obtained in Gali, Gertler, and Lopez-Salido (2005). That paper provides three baseline estimates. These estimates and the associated 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 precisely the case. In other words, our model matches the existing estimates on the Hybrid NKPC and allows one to rationalize those estimates as the product of informational frictions.

**Matching Survey Evidence on Informational Frictions.** Although our model passes the aforementioned test, it is not clear at this point whether this success hinges on an empirically implausible specification of the informational friction. We now address this question, and impose our model to an additional test, by examining whether the level of noise required by our model in order to rationalize the existing estimates of \(\omega_f\) and \(\omega_b\) is consistent with more direct, survey-based, evidence about the level of the informational friction.

To this goal, we draw a mapping between our model and the survey evidence on expectations reported Coibion and Gorodnichenko (2015). That paper provides an estimate on the magnitude of the information friction based on the inflation forecasts data from the Survey of Professional Forecasters. The basic idea is that the magnitude of the informational friction should manifest itself in the
predictability of the average forecast errors. In particular, Coibion and Gorodnichenko (2015) run the following regression:

\[
\pi_{t+k} - \mathbb{E}_t[\pi_{t+k}] = K \left( \mathbb{E}_t[\pi_{t+k}] - \mathbb{E}_{t-1}[\pi_{t+k}] \right) + \nu_{t,k,t} \tag{24}
\]

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) verify the aforementioned logic in a model in which inflation is assumed to follow an exogenous AR(1) process, and proceed to show how \( K \) is negatively related to the Kalman gain (and thereby positively related to the level of noise). Their exact characterization does not apply in our setting because inflation is endogenous and follows a different process that the one assumed in that paper. Yet, the logic is robust.

Specifically, although the regression coefficient \( K \) implied by our theory is more complicated that the one in the aforecited paper, we can characterize this coefficient as function of the level of noise \( \sigma \) along with the parameters \((\delta, \theta, \rho)\). Having fixed the latter in the way described earlier, this gives a mapping from the 90% confidence interval of \( K \) provided in Coibion and Gorodnichenko (2015) to an interval for \( \sigma \) in our model. For any \( \sigma \) in this interval, we can then compute the pair \((\omega_f, \omega_b)\) predicted by our theory.

We can thus map the evidence reported in Coibion and Gorodnichenko (2015) to a segment of the \((\omega_f, \omega_b)\) locus we obtained earlier on. This segment is identified by the red crosses in Figure 2 and gives the pairs of \((\omega_f, \omega_b)\) 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 both the test of matching that evidence and the test of matching the existing estimates of the Hybrid NKPC.\footnote{To be precise, the above statement is true for two of the three baseline estimates provided in Gali, Gertler, and Lopez-Salido (2005). But these happen to be the estimates associated with the smallest $\omega_y$ and the largest $\omega_b$. This is good news for the quantitative significance of our theory: once our theory is disciplined by the survey evidence, it rationalizes significant levels of both myopia and anchoring.}

The quantitative implications of our theory for the dynamics of inflation are further illustrated in Figure 3. This figure compares the impulse response function of inflation under two scenarios. The solid black solid line corresponds to frictionless benchmark, with perfect information. The red solid line corresponds to the frictional case, with an informational friction that matches the baseline estimation of Coibion and Gorodnichenko (2015). As evident in the figure, the latter case displays both dampened amplitude and significant sluggishness: 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.

### 8 Myopia and Habit Persistence in Consumption

We now shift the focus from the supply block of the New Keynesian model to its demand block. In particular, we show how informational frictions can arrest the response of aggregate consumption to news about monetary policy, can offer a micro-foundation of the form of consumption habit assumed in the DSGE literature, and can help reconcile the micro and macro estimates of such habit.

To this goal, we consider a perpetual-youth, overlapping-generations version of the New Keynesian model, along the lines of Del Negro, Giannoni, and Patterson (2015) and Piergallini (2007). In each period, each consumer remains alive with probability $\theta \in (0, 1]$; with the remaining probability,
he dies and gets replaced by a new consumer. As noted in Farhi and Werning (2017), the probability of death is such a model can be thought of as a proxy for the probability of binding liquidity constraints. The case of an unconstrained, infinitely-lived, representative agent is nested by setting \( \theta = 1 \). For our purposes, letting \( \theta < 1 \) is useful because it permits us to parameterize the GE effect that runs inside the demand block of the New Keynesian model, namely the income-spending multiplier.

Consider a household \( i \) born in period \( t \). Taking into account the mortality risk, expected lifetime utility is given by

\[
\sum_{t=\tau}^{\infty} (\chi \theta)^{t-\tau} U(C_{it}^\tau, N_{it}^\tau),
\]

where \( C_{it}^\tau \) denotes consumption, \( N_{it}^\tau \) denotes labor supply, \( U(C, N) = \log C - \frac{1}{1+\epsilon} N^\epsilon, \epsilon > 0 \) is the inverse of the Frisch elasticity of labor supply, and \( \chi \in (0,1) \) is the subjective discount factor. We allow the households to trade a riskless bond, whose real return is denoted by \( R_t \), and give them access to actuarily fair annuities, so that the effective return to saving is \( \frac{R_t}{\tau} \) conditional on survival and zero otherwise. We nevertheless preclude them from trading more sophisticated assets, such as GDP futures, so that we can bypass the complications of endogenous information aggregation. We finally assume that firms profits are taxed by the government and distributed in a lump sum manner to all the households, regardless of age. This guarantees that old and young households have the same exposure to any variation in the present value of the aggregate profits in the economy, a property that distinguishes the exercise conducted here from those in the aforecited papers and that ultimately guarantees that the perpetual-youth structure affects the aggregate dynamics only when it interacts with incomplete information.

We henceforth work with the log-linearized solution around a steady state in which there are no shocks, \( \chi R_t = 1 \), and \( C_t = Y_t = Y^* \), where \( Y^* \) represents the natural rate of output.\(^{14}\) Using lower-case variables to represent log-deviations from the steady state (e.g., \( r_t \equiv \log R_t - \log \chi \)), we

---

\(^{14}\)To simplify the exposition, we suppress the production side of the economy and the determination of the flexible-price outcomes. These “details” can be filled as in Angeletos and Lian (2016a) and are immaterial for the purposes of the result we describe here.
can express aggregate consumption in period $t$ as follows:

$$c_t = (1 - \chi \theta) \sum_{k=0}^{\infty} (\chi \theta)^k \mathbb{E}_t[y_{t+k}] - \chi \theta \sum_{k=0}^{\infty} (\chi \theta)^k \mathbb{E}_t[r_{t+k}],$$

where $\mathbb{E}_t$ is the average expectation. This is essentially the Permanent Income Hypothesis (see the first term), adapted to allow for variation in the real interest rate (see the second term).

Using the fact that $y_{t+k} = c_{t+k}$ for all $t$ and $k$; adding the simplifying assumption that the consumers observe the current income but do not extract information from it; and finally solving for $c_t$, we obtain the following equilibrium restriction on aggregate consumption:

$$c_t = \frac{1}{1 - \chi \theta} \sum_{k=0}^{\infty} (\chi \theta)^k \mathbb{E}_t[r_{t+k}] + (1 - \chi \theta) \sum_{k=1}^{\infty} (\chi \theta)^k \mathbb{E}_t[c_{t+k}],$$

(25)

This is a modern, forward-looking version of the Keynesian cross: current aggregate consumption depends on expectations of future aggregate consumption, because the latter pins down income.

We henceforth treat $r_t$ as an exogenous AR(1) process and investigate how aggregate consumption responds to innovations in $r_t$. The answer to this question is straightforward when information is complete. In this benchmark, the average expectation, $\mathbb{E}_t$, can be replaced by $\mathbb{E}_t$, the expectation conditional on the common information set. Using the Law of Iterated Expectations, it is then easy to show that, this benchmark, condition (25) can be restated in recursive form as follows:

$$c_t = -r_t + \mathbb{E}_t[c_{t+1}].$$

(26)

That is, regardless of $\theta$, aggregate consumption is determined in the same fashion as in the textbook, infinite-horizon, representative-agent model.

When instead information is incomplete, the Law of Iterated Expectations does not hold at the aggregate level and, as a result, condition (25) cannot be reduced to condition (26). Instead, it is best to read condition (25) as a dynamic beauty contest among the consumers. This game is nested in our framework by letting $a_t = c_t$, $\xi_t = -r_t$, $\varphi = 1$, $\beta = \chi \theta$, and $\gamma = 1 - \chi \theta$. The following result is then immediate, provided, of course, that the information structure is specified as in our abstract analysis.

---

15 This assumption simplifies the exposition but is not essential. For instance, it can be replaced by adding transitory idiosyncratic income shocks and letting the consumers observe their own income but not the aggregate own.

16 One can think of this as studying the aggregate-demand effects of a monetary policy that targets a specific process for the real interest rate. Alternatively, one can assume that prices are infinitely rigid, in which case $r_t$ coincides with the nominal rate (the policy instrument), and directly interpret the randomness in $r_t$ as exogenous shocks to monetary policy.

17 As noted earlier, the irrelevance of $\theta$ under complete information hinges on the assumption that firms profits are taxed and distributed in a lump sum manner. Otherwise, $\theta$ distorts the aggregate-level Euler condition in a similar manner as in Blanchard (1985); see Del Negro, Giannoni, and Patterson (2015). By shutting this distortion down, we zero in on the interaction of $\theta$ with incomplete information.
**Proposition 8.** When information is incomplete, there exist scalars $\omega_f < 1$ and $\omega_h > 0$ such that the equilibrium process for consumption solves the following equation:

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

(27)

Furthermore, a lower $\theta$ (short horizon) maps to a lower $\omega_f$ and a higher $\omega_h$.

It is therefore as if the economy is populated by a representative consumer who is myopic vis-a-vis the future movements in the real interest rate and her consumption exhibits habit persistence, as in the works of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007).\textsuperscript{18} Furthermore, the survival probability $\theta$ emerges as a key co-determinant, along with the level of noise, of the as-if myopia and habit. This is because a lower $\theta$ raises the magnitude of the relevant GE effect (this is evident from the property that $\gamma = 1 - \chi(\theta)$, thus also increasing the bite of higher-order uncertainty.

The reason that Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) replaced condition (26) with condition (27) is quite simple. In the data, aggregate consumption responds little to monetary and other shocks on impact, but builds up force over time. Such sluggish dynamics is inconsistent with condition (26). Replacing the latter with condition (27) fixes the problem and improves the empirical performance of the model.

The existing interpretation of condition (27), however, runs at a problem. If this condition is the product of true habit persistence in preferences, one would expect estimates of the habit parameter $\omega_h$ to be the comparable at aggregate and individual data. Yet, as the meta-analysis by Havranek, Rusnak, and Sokolova (2017) shows, the available microeconomic estimates of habit persistence tend to be much lower than their macroeconomic counterparts.

As anticipated in Section 6, our theory offers a simple resolution to this puzzle. When a consumer decides how to respond to a change in monetary policy or other aggregate shocks, she has to forecast the impact of that shock on her future income. In general equilibrium, her income is determined by the spending decisions of all the other consumers. It follows that the response of aggregate consumption can be understood as the solution to a dynamic beauty contest. Because higher-order beliefs adjust less and more slowly than first-order beliefs, expectations of income may adjust with less amplitude and more sluggishness than what predicted by the standard, complete-information, New Keynesian model, explaining in turn the sluggish dynamics of aggregate consumption. Finally, because this mechanism is not active in the context of the response of individual choices to idiosyncratic shocks (in such a context, higher-order uncertainty is irrelevant), our theory can help explain the aforementioned gap between the microeconomic and macroeconomic estimates of habit.

The quantitative evaluation of this insight is left for future work. Building on the point made

---

\textsuperscript{18}There is a subtle difference. The form of habit assumed in those papers imposes $\omega_f + \omega_h = 1$. In our setting, instead, $\omega_f + \omega_h < \rho < 1$. 
earlier about the relation between the survival probability and the bite of higher-order uncertainty, we only conjecture that the quantitative effects are likely to be largest if our mechanism is combined with frictions that boost the Keynesian income-spending multiplier, such as short horizons and liquidity constraints. Finally, although we have focused on consumption as the sole determinant of aggregate demand, similar points apply to investment as well: Appendix C develops an example in which informational frictions help reconcile the more exotic kind of adjustment costs to investment assumed in the DSGE literature with those that were featured in the original formulation of the Q theory of investment and that are more consistent with the microeconomic evidence on investment dynamics.

9 Incomplete Information as Myopia and Anchoring: Robustness

Although the observational-equivalence result presented in Section 5 depends on strong assumptions about the process of the fundamental and the available signals, it encapsulates two broader insights. The first is that the absence of common knowledge arrests the response of the current beliefs of the future beliefs of others to any news about the future fundamentals; this explains why agents behave as if they discount the future more heavily than what it is rational even if we control for the effect of the informational friction on first-order beliefs. The second insight is that the dynamics of learning causes extra persistence in the equilibrium outcome relative to that in the underlying fundamental; this explains why the current outcome appears to be anchored to the past outcome, conditional on the fundamental.

In this section, we elaborate on the robustness of these insights. To this goal, we modify the information structure as follows. For every $i$ and $t$, the incremental information received by agent $i$ in period $t$ is given by the series $\{x_{i,t,t-k}\}_{k=0}^{\infty}$, where

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

and where $\epsilon_{i,t,t-k} \sim N(0, (\tau_k)^{-2})$ is i.i.d. across $i$ and $t$, uncorrelated across $k$, and orthogonal to the past, current, and future innovations in the fundamental. That is, whereas our baseline specification has the agents observe a signal about $\xi_t$ in each period, the new specification lets them observe a series of signals about the entire history of the underlying 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 $\eta_{t-k}$ 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 $\tau_k$ 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
increases or decrease with the lag since the shock has occurred.

The first property is essential for tractability. The pertinent literature has struggled to solve the complex fixed point between the equilibrium dynamics and the Kalman filtering that obtains in dynamic models with incomplete information. By adapting 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. 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 is not possible in our baseline because $\sigma$ controls both objects at once.

By the familiar argument, the information regarding $\eta_{t-k}$ that an agent has accumulated up to, and including, period $t$ can be represented by a sufficient statistic, given by

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

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 $\lambda_k \equiv \frac{\pi_k}{\pi^k_\eta + \pi_k}$, we have that $\mathbb{E}_t[\eta_{t-k}] = \lambda_k \tilde{x}^k_{t,t}$, which in turn implies $\mathbb{E}_t[\eta_{t-k}] = \lambda_k \eta_{t-k}$ and therefore

$$\mathbb{E}_t[\xi_t] = \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \rho^k \eta_{t-k} \right] = \sum_{k=0}^{\infty} \lambda_k \rho^k \eta_{t-k}. \quad (28)$$

We conclude that the IRF of the (average) first-order belief to any given innovation is given by the sequence $F_1 = \{f_{1,k}\}_{k=0}^{\infty}$, where

$$f_{1,k} \equiv \partial \mathbb{E}_t \left[ \frac{\mathbb{E}_t[\xi_t]}{\eta_{t-k}} \right] = \lambda_k \rho^k.$$

By comparison, the IRF of the fundamental itself is given by the sequence $\{\rho^k\}_{k=0}^{\infty}$. It follows that the IRF of the first-order belief relative to that of the fundamental is pinned down by the sequence $\{\lambda_k\}_{k=0}^{\infty}$, which describes the dynamics of learning. In particular, the smaller $\lambda_0$ 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 $\lambda_k$ 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. Indeed, that specification can be captured here by restricting $\{\lambda_k\}_{k=0}^{\infty}$ so that $\lambda_k = 1 - G^k$ for all $k \geq 1$, where $G \in (0,1)$ is a scalar that corresponds to the Kalman gain and is inversely related to $\sigma$. This makes clear that our baseline specification ties the initial precision of the information about
any given innovation 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 (28) to period $t+1$ and taking the period-$t$ average expectation, we get

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

By induction, for all $h \geq 2$, the $h$-th order, forward-looking belief is given by

$$F^h_t [\xi_{t+h}] = \sum_{k=0}^{\infty} \lambda_k \lambda_{k+1} \cdots \lambda_{k+h-1} \rho^{k+h-1} \eta_{t-k}.$$ 

It follows that the corresponding IRF is given by the sequence $F_h = \{f_{h,k}\}_{k=0}^{\infty}$, where

$$f_{h,k} = \frac{\partial \mathbb{E}_t \left[ F^h_t [\xi_{t+h}] \right]}{\partial \eta_{t-k}} = \lambda_k \lambda_{k+1} \cdots \lambda_{k+h-1} \rho^{k+h-1}$$

(29)

Note next that $\rho^{k+h-1} = \frac{\partial \mathbb{E}_t [\xi_{t+h} \mid \eta_{t-k}]}{\partial \eta_{t-k}}$. It follows that the ratio $f_{h,k} / \rho^{k+h-1}$ measures the effect of an innovation on the $h$-th order forward-looking belief relative to its effect on the fundamental. When information is complete, we have that the aforementioned ratio is identically 1 for all $k$ and $h$. When, instead, information is incomplete, we have that

$$\frac{\partial}{\partial \eta_{t-k}} \mathbb{E}_t \left[ F^h_t [\xi_{t+h}] \mid \eta_{t-k} \right] = \frac{f_{h,k}}{\rho^{k+h-1}} = \lambda_k \lambda_{k+1} \cdots \lambda_{k+h-1}.$$

The following result is thus immediate.

**Proposition 9.** 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 $\lambda_k \to 1$.

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

These properties shed light on the dynamic structure of higher-order beliefs. To see how these properties in turn 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

\[
a_t = \sum_{h=0}^{\infty} g_h \eta_{t-h}
\]

where, for all \( k \),

\[
g_k \equiv \frac{\partial \mathbb{E}[a_t|\eta_{t-k}]}{\partial \eta_{t-k}} = \sum_{h=1}^{\infty} \gamma^{h-1} f_{h,k} = \left\{ \sum_{h=1}^{\infty} (\rho \gamma)^{h-1} \lambda_k \lambda_{k+1} \ldots \lambda_{k+h-1} \right\} \rho^k.
\]

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 \( \gamma \) is, the more the dynamics of the equilibrium outcome tracks the dynamics higher-order beliefs relative to the dynamics of lower-order beliefs.

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:

\[
g_k = \left\{ \sum_{h=1}^{\infty} (\rho \gamma \lambda)^{h-1} \right\} \rho^k \lambda.
\]

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

\[
a_t = \xi'_t + \gamma' \mathbb{E}_t[a_{t+1}],
\]

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 even if when control for the impact of the informational friction on first-order beliefs. Indeed, replacing $\xi^*_t$ with $\xi_t$ 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 6.

So far, we shed light on the forward-looking component (myopia) of our observational-equivalence result 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 $\xi_t$, 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 $\xi_t$ and is given by

$$a_t = \sum_{k=0}^{\infty} \hat{g}_k \eta_{t-k}, \quad \text{with} \quad \hat{g}_k = \frac{1}{1 - \gamma \rho} f_{1,k}.$$  

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

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

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

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

$$a_t = \phi E_t[\xi_t] + \gamma E_t[a_{t+1}].$$  

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

\[
g_k = f_{1,k} + \lambda_k \gamma g_{k+1}.
\] (33)

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 \(\rho^k\). The role of higher-order uncertainty is captured by the presence of \(\lambda_k\) in the second term: it is as if the discount factor \(\gamma\) has been replaced by a discount factor equal to \(\lambda_k \gamma\), which is strictly less than \(\gamma\). This represents a generalization of the form of myopia seen in condition (31). There, learning was shut down, so that that \(\lambda_k\) 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 \(\lambda_k\) will increase with \(k\)).

Consider next property (ii), namely the property that

\[
\frac{g_{k+1}}{g_k} > \frac{\hat{g}_{k+1}}{\hat{g}_k} > \rho
\]

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. In this economy, the impact of a shock \(k + 1\) periods from now relative to its impact \(k\) periods from now is given by

\[
\frac{\hat{g}_{k+1}}{\hat{g}_k} = \frac{f_{1,k+1}}{f_{1,k}} > \rho.
\]

The inequality captures the effect of learning on first-order beliefs. Had information being perfect, we would have had \(\frac{\hat{g}_{k+1}}{\hat{g}_k} = \rho\;\text{now, we instead have}\;\frac{\hat{g}_{k+1}}{\hat{g}_k} > \rho\). 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.\(^{19}\)

\(^{19}\)This is easiest to see when \(\rho = 1\) (i.e., the fundamental follows a random walk), for then \(\hat{g}_{k+1}\) is necessarily higher than \(\hat{g}_k\) for all \(k\). When instead \(\rho < 1\), \(\hat{g}_{k+1}\) can be either higher or lower than \(\hat{g}_k\), depending on the balance between two opposing forces: the build-up effect of learning and the mean-reversion in the fundamental.
Consider now our economy, in which higher-order uncertainty is present. We now have
\[
\frac{g_{k+1}}{g_k} > \frac{\hat{g}_{k+1}}{\hat{g}_k}
\]
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,\(^\text{20}\) 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 1.** Suppose that the first-order belief displays a hump-shaped response, namely \(\{f_{1,k}\}\) is single peaked at \(k = k^b\) for some \(k^b \geq 1\). Then, the equilibrium outcome also displays a hump-shaped response, namely \(\{g_k\}\) is also single peaked at \(k = k^g\). Furthermore, the peak of the equilibrium response is after the peak of the first-order belief: \(k^g \geq k^b\) necessarily, and \(k^g > k^b\) for an open set of \(\{\lambda_k\}\) sequences.

To interpret this result, think momentarily of \(k\) as a continuous variable and, similarly, think of \(\lambda_k, g_k, \) and \(f_{1,k}\) as differentiable functions of \(k\). If \(f_{1,k}\) is hump-shaped with a peak at \(k = k^b > 0\), it must be that \(b_k\) 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 \(b_k\), 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 \(\omega_b\), 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.

### 10 Conclusion

In this paper we showed how the accommodation of incomplete information, higher-order uncertainty and learning in forward-looking models is akin to the introduction of two kinds of behavioral frictions: myopia vis-a-vis the future; and anchoring of current outcomes to past outcomes.

\(^{20}\)Recall, this is by property (i) of Proposition 10.
We offered a stark illustration of this point, in terms of an observational-equivalence result, under appropriate assumptions about the information structure. These assumptions are restrictive, but the insights apply more broadly.

The observation-equivalence result was useful, not only for illustrating the above point, but also for three additional reasons. First, it let us build a sharp connection between incomplete information and a set of more ad hoc adjustment frictions assumed in the DSGE literature. Second, it helped explain how incomplete information can help resolve the gap between the macroeconomic and microeconomic estimates of such frictions. Last but not least, it facilitated the application and the quantitative evaluation of our theory in the context of inflation dynamics.

This application brought together the theory with two kinds of evidence: one regarding the actual inflation dynamics; and another regarding inflation expectations. We thus show how our approach can, not only rationalize evidence that have traditionally been associated with the hybrid version of the New Keynesian Philips Curve, but also achieve this while also matching independent evidence on inflation expectations.

Another application shifted the focus to the demand block of the New Keynesian model, namely to the dynamic relation between aggregate consumption and the real interest rate. This permitted us to illustrate how information frictions can help, not only arrest the response of aggregate consumption to monetary policy, but also reconcile the relatively high degree of consumption habit assumed in the DSGE literature with one estimated in microeconomic evidence.

When exploring these two applications of our theory to the New Keynesian model, we studied each block of the model (namely the NKPC and the Euler condition for consumption) in isolation of the other. We thus emphasized the GE effects and the higher-order beliefs that operate within each block of the model (respectively, the dynamic complementary in the price-setting behavior of the firms and the dynamic income-spending multiplier), but abstracted from their interaction. The extension of our analysis in this direction and the quantitative evaluation of a fully-fledged DSGE setting augmented with informational frictions are important open questions for future research.
References


Huo, Zhen and Naoki Takayama. 2015. “Rational Expectations Models with Higher Order Beliefs.” *Yale mimeo*. 2, 4


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 \( \tau_{u} = \sigma_{u}^{-2} \).\(^{21}\)

The signal process can be rewritten as

\[ x_{it} = M(L) \begin{bmatrix} \tilde{\eta}_{it} \\ \tilde{u}_{it} \end{bmatrix} \]

where

\[ M(L) = \begin{bmatrix} \tau_{\eta}^{-\frac{1}{2}} & \frac{1}{1 - \rho L} & \tau_{u}^{-\frac{1}{2}} \end{bmatrix}. \]

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

\[ B(L)B(L^{-1}) = M(L)M'(L^{-1}) = \frac{\tau_{\eta}^{-1} + \tau_{u}^{-1}(1 - \rho L)(L - \rho)}{(1 - \rho L)(L - \rho)}, \quad (34) \]

which leads to

\[ B(L) = \tau_{u}^{-\frac{1}{2}} \sqrt{\frac{\rho}{\lambda}} \frac{1 - \lambda L}{1 - \rho L}, \]

where \( \lambda \) is the inside root of the numerator in equation (34)

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

\(^{21}\)In the main text, we have normalized \( \sigma_{\eta} = 1 \).
Next, we characterize the beliefs of $\xi_t$, $a_{i,t+1}$, and $a_{t+1}$, that is, the beliefs that show up in the best-response condition of the agent. The forecast of a random variable

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

can be obtained by using the Wiener-Hopf prediction formula

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

We start with the forecast of the fundamental. Note that

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

from which it follows that

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

Using the guess that $a_{it+1} = h(L)x_{i,t+1}$ and $a_{t+1} = h(L)\xi_{t+1}$, we have

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

and the forecasts are

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

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

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

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

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

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

41
Equivalently, we need to find an analytic function \( h(z) \) that solves

\[
\begin{align*}
    h(z) &= \varphi \frac{\lambda \tau u}{\rho \tau \eta} \frac{1}{1 - \rho \lambda} \frac{1}{1 - \lambda z} + \\
    &\quad + (\beta + \gamma) \frac{\lambda \tau u}{\rho \tau \eta} \left( \frac{h(z)}{(1 - \lambda z)(z - \lambda)} - \frac{h(\lambda)(1 - \rho z)}{(1 - \rho \lambda)(z - \lambda)(1 - \lambda z)} \right) \\
    &\quad + \beta \frac{\lambda}{\rho} \left( \frac{h(z)(z - \rho)}{z(z - \lambda)} - \frac{h(\lambda)(\lambda - \rho)}{\lambda(z - \lambda)} - \frac{\rho h(0)}{\lambda - z} \right) \frac{1 - \rho z}{1 - \lambda z},
\end{align*}
\]

which can be transformed as

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

where

\[
C(z) \equiv z(1 - \lambda z)(z - \lambda) - \frac{\lambda}{\rho} \left\{ \beta(z - \rho)(1 - \rho z) + (\beta + \gamma) \frac{\tau u}{\tau \eta} z \right\}
\]

\[
d(z; h(\lambda), h(0)) \equiv \varphi \frac{\lambda \tau u}{\rho \tau \eta} \frac{1}{1 - \rho \lambda} z(z - \lambda) - \frac{1}{\rho} \left( \frac{\tau u \lambda(\beta + \gamma)}{\tau \eta} + \beta(\lambda - \rho) \right) z(1 - \rho z) h(\lambda) - \beta(z - \lambda)(1 - \rho z) h(0)
\]

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 \( h(\lambda) \) to make sure 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 \( h(L) \)

\[
h(L) = \left( 1 - \frac{\vartheta}{\rho} \right) \frac{\varphi}{1 - \rho \vartheta} \frac{1}{1 - \vartheta L}.
\]

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

Finally, we show that \( C(z) \) has two inside roots and one outside root. Note that \( C(z) \) can be rewritten as

\[
C(z) = \lambda \left\{ - z^3 + \left( \rho + \frac{1}{\rho} + \frac{\lambda \tau u}{\rho \tau \eta} + \beta \right) z^2 - \left( 1 + \beta \left( \rho + \frac{1}{\rho} \right) + \frac{\beta + \gamma \tau u}{\rho \tau \eta} \right) z + \beta \right\}.
\]

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

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

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

**Proof of Proposition 3**

Note that

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

By the continuity of \( C(z) \), it must therefore be that \( C(z) \) admits a root between \( \frac{1}{\rho} \) and \( \frac{1}{\lambda} \). But since this root has to be higher than 1 and since the only such root is given by the reciprocal of \( \vartheta \), this proves 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 \( \vartheta \).

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

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

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

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

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

Similarly, we have

\[
\frac{\partial C(\vartheta^{-1})}{\partial \beta} = (\vartheta^{-1} - \lambda)(\vartheta^{-1} - \lambda^{-1}) < 0 \quad \text{and} \quad \frac{\partial C(\vartheta^{-1})}{\partial \gamma} = -\frac{\tau_u}{\rho \tau_\eta} < 0
\]

which proves that \( \vartheta \) increases with both \( \beta \) and \( \gamma \).

When \( \gamma = 0 \), the three roots of \( C(z) \) become \( \beta, \lambda \) and \( \lambda^{-1} \). For this case, the outside root is independent of \( \beta \) and \( \gamma \).
Proof of Proposition 4

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

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

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

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

To match the hybrid model, we need

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

Combining (35) and (36), and solving for the coefficients of \( \omega_f \) and \( \omega_b \), we infer that the two economies generate the same dynamics if and only if the following two conditions hold:

\[ \omega_f = \frac{\delta \rho^2 - \vartheta}{\delta (\rho^2 - \vartheta^2)} \]  
\[ \omega_b = \frac{\vartheta (1 - \delta \vartheta) \rho^2}{\rho^2 - \vartheta^2} \]  

Since \( \delta \equiv \beta + \gamma \) and since \( \vartheta \) is a function of the primitive parameters \( (\sigma, \rho, \beta, \gamma) \), the above two conditions give the coefficients \( \omega_f \) and \( \omega_b \) 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 \( \sigma = 0 \). This completes the proof of part (i). Part (ii) follows from Proposition 3 together with the fact that a higher \( \vartheta \) maps to a lower \( \omega_f \) and a higher \( \omega_b \).

Proof of Proposition 5

As already noted, the hybrid and the incomplete-information economies generate the same dynamics if and only if conditions (37) and (38) hold. Using (38), we can rewrite (37) as follows:

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

44
Furthermore, any \((\beta, \gamma, \rho)\), the equilibrium of the incomplete-information economy gives an invertible mapping from \(\sigma \in (0, \infty)\) to \(\vartheta \in (0, \rho)\), whereas condition (38) 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 \((\omega_b, \omega_f)\) satisfies condition (39) along with \(\omega_b \in (0, \infty)\). Finally, the level of the informational friction that achieves this replication is obtained by inverting condition (38) to obtain \(\vartheta\), and thereby also \(\sigma\), as an implicit function of \(\omega_b\).

**Proof of Proposition 6**

Suppose that the individual’s policy function is

\[ a_{it} = g(L)z_{it} + h(L)x_{it}. \]

In the limit as \(V(\zeta_t) \to \infty\), agents effectively interpret \(z_{it}\) as a perfect signal of \(\zeta_t\) and as a completely uninformative about \(\xi_t\). As a result, the forecasts of both \(\xi_t\) and \(a_{t+1}\) only depend on \(x_{it}\). Furthermore, the forecast of the individual’s own future action can be decomposed as follows:

\[ E_{it}[a_{it+1}] = E_{it}[g(L)z_{it+1}] + E_{it}[h(L)x_{it+1}]. \]

The first component only depends on \(z_{it}\) and is given by

\[ E_{it}[g(L)z_{it+1}] = g(L)z_{it} - (1 - \rho)g(0)z_{it}, \]

which yields a simple solution to \(g(L)\):

\[ g(L) = \frac{\varphi}{1 - \rho \beta}. \]

On the other hand, using similar logic as in the proof of Proposition 2, we get:

\[ E_{it}[a_{it}] = \frac{\lambda \tau}{\rho \tau} \left( g(L) + h(L) \right) \frac{\frac{L}{1-\rho \lambda}}{L - \lambda} - (g(L) + h(L)) \frac{\lambda (1 - \rho \lambda)}{(1 - \lambda \lambda)(1 - \lambda \Lambda)} x_{it} \]

\[ E_{it}[a_{it+1}] = \frac{\lambda \tau}{\rho \tau} \left( \frac{h(L) + g(L)}{(1 - \lambda \Lambda)(1 - \lambda \lambda)} - \frac{(h(L) + g(L)) (1 - \rho \lambda)}{(1 - \rho \lambda)(L - \lambda)(1 - \lambda \lambda)} \right) x_{it} \]

\[ E_{it}[h(L)x_{it+1}] = \frac{\lambda \tau}{\rho \tau} \left( \frac{h(L)}{(1 - \lambda \Lambda)(1 - \lambda \lambda)} - \frac{h(L)(1 - \rho \lambda)}{(1 - \rho \lambda)(L - \lambda)(1 - \lambda \lambda)} \right) x_{it} \]

\[ + \frac{\lambda}{\rho} \left( \frac{h(L)(L - \rho)}{L(L - \lambda)} - \frac{h(L)(\lambda - \rho)}{\lambda(L - \lambda)} - \frac{\rho h(0)}{L} \right) \frac{1 - \rho \lambda}{1 - \lambda \Lambda} x_{it} \]

45
The fixed point for $h(L)$ is

$$h(z) = \Delta(z) + (\beta + \gamma)^2 \frac{\lambda}{\rho} \frac{\tau_n}{\tau_\eta} \left( \frac{h(z)}{(1 - \lambda z)(z - \lambda)} - \frac{h(\lambda)(1 - \rho z)}{(1 - \rho \lambda)(z - \lambda)(1 - \lambda z)} \right) + \beta \frac{\lambda}{\rho} \left( \frac{h(z)(z - \rho)}{z(z - \lambda)} - \frac{h(\lambda)(\lambda - \rho)}{\lambda(z - \lambda)} - \frac{\rho h(0)}{\lambda z} \right) \frac{1 - \rho z}{1 - \lambda z}$$

This leads to

$$C(z)h(z) = d(z)$$

where

$$C(z) \equiv z(1 - \lambda z)(z - \lambda) - \frac{\lambda}{\rho} \left( \beta(z - \rho)(1 - \rho z) + (\beta + \gamma) \frac{\tau_n}{\tau_\eta} z \right)$$

$$d(z) \equiv \Delta(z) z(z - \lambda)(1 - \lambda z) - \frac{1}{\rho} \left( \frac{\tau_n}{\tau_\eta} \frac{\lambda(\beta + \gamma)}{1 - \rho \lambda} + \beta(\lambda - \rho) \right) z(1 - \rho z) h(\lambda) - \beta(z - \lambda)(1 - \rho z) h(0)$$

$$\Delta(z) \equiv \frac{\varphi}{1 - \rho \beta} \frac{1}{1 - \lambda z} \left( 1 - \frac{\lambda}{\rho} \right) \gamma \rho$$

By eliminating the inside roots of $C(z)$, the solution to $h(z)$ follows.

**Proof of Proposition 10**

First, let us prove $g_k < \hat{g}_k$. Recall that \{\(g_k\)\} is given by

$$g_k = \rho^k \sum_{h=0}^{\infty} (\rho \gamma)^h \prod_{\tau=k}^{k+h} \lambda_{\tau}$$

Clearly,

$$0 < g_k < \rho^k \sum_{h=0}^{\infty} (\rho \gamma)^h \lambda_k = \frac{1}{1 - \gamma \rho} \rho^k \lambda_k = \frac{1}{1 - \gamma \rho} f_{1,k} = \hat{g}_k,$$

which proves the first property.

Next, let us prove that $\frac{\hat{g}_{k+1}}{g_k} > \frac{\hat{g}_{k+1}}{g_k} > \rho$. Since \{\(\lambda_k\)\} is strictly increasing,

$$\frac{\hat{g}_{k+1}}{g_k} = \frac{\lambda_{k+1}\rho^{k+1}}{\lambda_k \rho^k} = \frac{\lambda_{k+1}}{\lambda_k} \rho > \rho.$$
and
\[
\frac{g_{k+1}}{g_k} = \frac{\sum_{h=0}^{\infty} (\rho^k)^h \lambda_{k+1} \lambda_{k+2} \cdots \lambda_{k+h+1}}{\sum_{h=0}^{\infty} (\rho^k)^h \lambda_k \lambda_{k+1} \cdots \lambda_{k+h}} \rho > \rho,
\]

It then also follows that
\[
\frac{g_{k+1}}{g_k} = \frac{\lambda_{k+1} (\rho^{k+1} + \gamma g_{k+2})}{\lambda_k (\rho^k + \gamma g_k)} \frac{(\rho^{k+1} + \gamma g_{k+2})}{(\rho^k + \gamma g_k)} = \frac{\rho^k + \rho^{-1} \gamma g_{k+2}}{\rho^k + \gamma g_k} > \frac{\rho^k + \gamma g_{k+1}}{\rho^k + \gamma g_k} = 1,
\]

which proves that

\[
\frac{g_{k+1}}{g_k} > \frac{\hat{g}_{k+1}}{\hat{g}_k}.
\]

(Had \( \lambda_k \) been weakly increasing, all the inequalities would have been weak.)

Finally, note that, since \( \lambda_k \) is increasing and bounded from above by 1, it has to converge, to a number \( \lambda_\infty \leq 1 \). It follows that

\[
\lim_{k \to \infty} \frac{\hat{g}_{k+1}}{\hat{g}_k} = \frac{\lambda_\infty}{\lambda_\infty} \rho = \rho
\]

and

\[
\lim_{k \to \infty} \frac{g_{k+1}}{g_k} = \frac{\sum_{h=0}^{\infty} (\rho^h \lambda_\infty)^h}{\sum_{h=0}^{\infty} (\rho^h \lambda_\infty)^h} \rho = \rho.
\]

**Appendix B: Additional Variant with Idiosyncratic Shocks**

In the case studied in Subsection 6.2, the presence of higher-order uncertainty distorts the response of the aggregate outcome to aggregate shocks, while the assumption that each agent knew perfectly her own fundamental guarantees that the informational friction does not affect at all the response of individual outcomes to idiosyncratic shocks. This offered a sharp illustration of how our approach helps resolve the disconnect between micro and macro estimates of habit, adjustment costs, etc. We now expand on the robustness of this point to situations in which the agents lack perfect knowledge their own fundamentals.

We continue to assume that \( \xi_{it} \) is given by the sum of an aggregate and an idiosyncratic component, but prevent agent \( i \) from observing either \( \xi_{it} \) or its components. In particular, we specify the information structure as follows. First, we let each agent observe the same noisy signal \( x_{it} \) about the aggregate shock \( \xi_t \) as in our baseline model. Second, we let each agent observe the following noisy signal about the idiosyncratic shock \( \zeta_{it} \) :

\[
z_{it} = \zeta_{it} + \nu_{it},
\]
where $v_{it}$ is independent of $\zeta_{it}$, of $\xi_t$, 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 $1 - \lambda$ be the Kalman gain in the forecasts of the aggregate fundamental, so that

$$E_{it}[\xi_t] = \rho \lambda E_{it-1}[\xi_{t-1}] + (1 - \lambda) x_{it}$$

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

$$E_{it}[\zeta_t] = \rho \hat{\lambda} E_{it-1}[\zeta_{i,t-1}] + (1 - \hat{\lambda}) z_{it}$$

It is straightforward to extend the results of Section 5 to the current information structure. It can thus be shown that the equilibrium action is given by the following:

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

where $\vartheta$ is determined in the same manner as in our baseline model and where $u_{it}$ is a residual that is orthogonal to both $\zeta_{it}$ and $\xi_{it}$ and that captures the combined effect of all the idiosyncratic noises in the information of agent $i$.

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

$$a^*_{it} = \frac{\varphi}{1 - \rho \beta} \zeta_{it} + \frac{\varphi}{1 - \rho \delta} \xi_{it}.$$ 

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

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

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

From the proof of Proposition 3, it is straightforward to verify that $\vartheta > \lambda$, with $\vartheta \rightarrow \lambda$ as $\gamma \rightarrow 0$. Furthermore, following Mackowiak and Wiederholt (2009), it is natural to assume that $\hat{\lambda}$ is lower than $\lambda$, 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 letting $\gamma \rightarrow 0$, or abstracting from role higher-order uncertainty).

But once such interactions are taken into account, we have that $\vartheta$ can be higher than $\hat{\lambda}$ even if the latter exceeds $\lambda$. This reflects the role of higher-order uncertainty. Because GE effects are active...
only with aggregate shocks, because GE effects cause the equilibrium outcomes to track higher-order beliefs, and because higher-order beliefs are more inertial than first-order beliefs, the response to aggregate shocks can be more attenuated and more sluggish than the response to idiosyncratic shocks even if the agents know more about aggregate shocks than about idiosyncratic shocks.

Appendix C: 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:

\[
\text{Cost}_t = \Phi \left( \frac{I_t}{K_{t-1}} \right)
\] (40)

where \( I_t \) denotes the rate of investment, \( K_{t-1} \) denotes the capital stock inherited from the previous period, and \( \Phi \) 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).  \(^{22}\)

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:

\[
\text{Cost}_t = \Psi \left( \frac{I_t}{I_{t-1}} \right)
\] (41)

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 (40); 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

\(^{22}\) 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.
adjustment cost takes the form seen in condition (41). 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

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

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

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

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

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

subject to

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

Here, the fundamental shock, $\xi_t$, 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 $Z_t = \frac{I_t}{K_t}$ 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

$$Q_t q_t + \psi Z z_t = \chi \mathbb{E}_t \left[ A \xi_{t+1} + Q q_{t+1} + \psi Z (1 + Z) z_{t+1} \right].$$ (42)

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

$$\max_{X_{it}} E_{it} \left[ Q_{it} X_{it} - c X_{it} \right],$$

subject to

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

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

$$z_{it} = E_{it} [z_t + \sigma q_t].$$

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

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

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

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

**Frictionless Benchmark.** If all intermediate firms observe $\xi_t$ 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:

$$z_t = \varphi \xi_t + \delta E_t [z_{t+1}]$$

where

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

**Incomplete Information.** Suppose now that firms receive a noisy signal about the fundamental $\xi_t$ 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 (42), the aggregate investment dynamics follow

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

The investment dynamics can be understood as the solution to the dynamic beauty contest studied in
Section 3 by letting
\[
\varphi = \frac{\rho \chi A}{\psi Z}, \quad \beta = \chi, \quad \text{and} \quad \gamma = \chi Z.
\]
The following is then immediate.

**Proposition 11.** When information is incomplete, there exist \( \omega_f < 1 \) and \( \omega_b > 0 \) such that the equilibrium process for investment solves the following equation:
\[
z_t = \varphi \xi_t + \omega_f \mathbb{E}_t[z_{t+1}] + \omega_b z_{t-1}
\]
Finally 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:
\[
\max_{\{K_t, I_t\}} \sum_{t=0}^{\infty} \chi^t \mathbb{E}_0 \left[ \exp(\xi_t) AK_t - Q_t I_t - \Psi \left( \frac{I_t}{I_{t-1}} \right) I_t \right]
\]