Myopia and Anchoring*

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

August 23, 2018

Abstract

We consider a stationary setting featuring forward-looking behavior, higher-order uncertainty, and learning. 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 featuring habit persistence, adjustment costs, or momentum. This builds a bridge to both the DSGE literature and an emerging literature on bounded rationality. 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 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. We discuss additional applications to consumption, investment, and asset prices.

*We thank Alexander Kohlhas for an insightful conference discussion. We also acknowledge useful comments from Jaroslav Borovicka, Simon Gilchrist, Jennifer La'O, Stephen Morris, Mikkel Plagborg-Moller and seminar participants at Columbia, NYU, CUHK, the 2018 Duke Macro Jamboree, and the 2018 Cambridge/INET conference.

†MIT and NBER; angelet@mit.edu.

‡Yale University; zhen.huo@yale.edu.

1 Introduction

Modern macroeconomic models emphasize forward-looking expectations. For instance, the modern analogue of the IS curve relates current spending to expectations of future spending; the New Keynesian Philips Curve (NKPC) ties current inflation to expectations of future inflation and future output gaps; and the basic, rational-expectations, asset-pricing model gives the equilibrium price of an asset as the present discounted value of future dividends.

Such models face (at least) two kinds of challenges. First, they often predict implausibly strong responses to news about the distant feature, a problem emphasized in the recent literature on the forward-guidance puzzle.[1] Second, they have hard time capturing the persistence of the aggregate time series and in particular the property that actual outcomes appear to respond sluggishly to the underlying innovations, as if there is a strong backward-looking component in their laws of motion.

To address the first problem, an emerging literature sacrifices on rational expectations.[2] And to address the second problem, the DSGE literature sacrifices on the micro-foundations, augmenting the key equations of the theory with ad hoc backward-looking elements that find little support in the microeconomic data and in the kind of models used to understand such data.[3]

In this paper, we propose a simple alternative that addresses both problems, maintains rational expectations, and eases the aforementioned disconnect between micro and macro. This is achieved by accommodating incomplete information (or rational inattention), higher-order uncertainty, and learning. The key insight is that these features are akin to the combination of two distortions:

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

Furthermore, the predicted distortions are larger the stronger the general-equilibrium (GE) feedback in the economy is, which explains why these distortions may loom at the macro level even if they seem small in micro data. Last but not least, the proposed resolution appears to be quantitatively potent and is corroborated by evidence on expectations (Coibion and Gorodnichenko, 2012, 2015).

Framework. Our starting point is a game with a continuum of players and best responses that depend positively on the expectations of the future actions of others. Denote with ξ_t the exogenous, payoff-relevant fundamental and with a_t the equilibrium outcome (the average action). When information is complete, our setting reduces to a representative-agent model, in which a_t obeys the

[1] Del Negro, Giannoni, and Patterson (2015), McKay, Nakamura, and Steinsson (2016), and Angeletos and Lian (2016a).
[2] Gabaix (2016), Farhi and Werning (2017), and Woodford (2018).
[3] For instance, to capture the dynamics of inflation, Gali and Gertler (1999), Christiano, Eichenbaum, and Evans (2005), Smets and Wouters (2007), and many others replace the standard NKPC with the Hybrid NKPC, a variant that finds no support in the microeconomic data on prices and in the related menu-cost literature. Similarly, to match the dynamics of aggregate consumption, the macro literature has assumed a degree of habit that is much higher than that estimated in microeconomic data (Havranek, Rusnak, and Sokolova, 2017); and to match the dynamics of investment, the literature has assumed a form of adjustment cost that is incompatible with that found in the related micro literature.

following law of motion:

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

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

Condition (1) directly nests each of the two key equations of the New Keynesian model: the NKPC and the Dynamic IS curve. To capture the former, interpret ξ_t as the output gap, or the real marginal cost, and a_t as inflation. To capture the latter, interpret ξ_t as the real interest rate and a_t as aggregate spending. Alternatively, condition (1) can be read as an asset-pricing equation, with ξ_t standing for the current expectation of the next-period dividend and a_t for the price.

We depart from the representative-agent benchmark by allowing information to be incomplete. The informational friction 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). One way or another, the key is that we accommodate, not only imperfect observability of the underlying aggregate fundamental, but also higher-order uncertainty (uncertainty about the beliefs and actions of others). In other words, we allow the agents to face realistic doubts about what *others* know and how others will respond to any news about the current state and the future prospects of the economy.

Main Result. 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:

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

for some scalars $\omega_f < 1$ and $\omega_b > 0$. The first modification ($\omega_f < 1$) represents myopia towards the future. The second modification ($\omega_b > 0$) represents anchoring of the current outcome to the past outcome. Both distortions are shown to increase, not only with the level of noise, but also with the strength of the underlying GE effects.

The exact form of this result relies on strong assumptions about the process for ξ_t and the information structure. Nevertheless, the result encapsulates two general insights, whose robustness we document in Section 10 under a more flexible specification of our model.

The first insight abstracts from the dynamics of learning and focuses on the question of how the informational friction impacts the response of the economy to the underlying innovations. Two effects are at work here. First, the imperfect observation of these innovations arrests the response of the beliefs of the fundamental. Second, the higher-order uncertainty arrests the response of the expectations of the future outcomes at an even greater degree than that characterizing the beliefs of the fundamental. Both effects contribute towards making the economy respond to news about the future as if the news is discounted more heavily than in the frictionless benchmark, explaining the form of myopia accommodated in our paper. And since the second effect is more relevant when

the general-equilibrium (GE) feedback in the economy is stronger, the documented form of myopia increases with the strength of the GE feedback.

The second insight shifts attention to learning. Learning induces persistence and momentum, not only in the beliefs of the underlying fundamentals, but also in the expectations of the future outcomes. Furthermore, the persistence is stronger in the latter kind of beliefs than in the former, reflecting the interaction of learning with higher-order uncertainty. This explains why the current outcome appears to be anchored to the past outcome at a rate that, like our form of myopia, increases with the strength of the GE feedback.

Earlier, more special versions of these insights appear in Angeletos and Lian (2016a), for the first insight, and Woodford (2003), for the second.[4] Our contribution is, not only to blend the two insights and to study their interaction, but also and, perhaps most importantly, 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), offers a sharp illustration of how informational frictions help address the empirical challenges mentioned earlier, and paves the way to our main empirical application, that on inflation.

GE and Micro vs Macro. As already explained, our theory predicts that ω_f falls and ω_b increases with the strength of the GE feedback. In the context of the Dynamic IS curve, the relevant GE feedback is the Keynesian income-spending multiplier; in the context of the NKPC, it is the dynamic strategic complementarity in the price-setting decisions of the firms; and in the context of asset pricing, it has to do with how much current traders speculate against future traders.

The GE feedback matters because it regulates the importance of higher-order uncertainty. But this feedback is active only when agents respond to aggregate shocks. It follows that the documented forms of myopia and anchoring may loom large at the macro level even if they appear to be small in micro data. This provides a simple, unified explanation to why the macroeconomic estimates of the habit persistence in consumption are higher than their microeconomic counterparts (Havranek, Rusnak, and Sokolova, 2017), why the persistence of inflation is higher in the aggregate time series than in disaggregated data (Altissimo et al., 2010), and why asset-price momentum is more pronounced at the aggregate level (Jung and Shiller, 2005).

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 higher one often assumed in the DSGE literature. We next show that our theory helps rationalize existing estimates of

[4] It is worth noting that, whereas Woodford (2003) documents the second insight with numerical simulations, we are able to offer a sharp formalization of it. Mankiw and Reis (2002) and Sims (2003) also touch on the second insight, although they do not explain the roles of GE effects and higher-order beliefs.

the Hybrid NKPC, such as those found in Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2005). Importantly, this is achieved with an informational friction that also matches the evidence on inflation expectations provided by Coibion and Gorodnichenko (2012, 2015).

With this application, we not only illustrate the applied value of our observational-equivalence result, but also connect the theory to evidence on expectations. Note in particular that Coibion and Gorodnichenko (2012, 2015) treated inflation as an exogenous variable. In so doing, this work provided empirical moments that help gauge the level of the informational friction but did not quantify its impact on the *actual* inflation dynamics. Our paper contributes towards filling this gap.[5]

Applications to Consumption, Investment, and Asset Prices. 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 ease the disconnect between the micro and macro estimates of such habit. Our theory also implies 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 credit frictions.

Turning to investment, we show how a model in which it is costly to adjust the capital stock may end up looking like a model in which it is costly to adjust the rate of investment. This helps reconcile the DSGE literature with the literature that studies investment dynamics at the firm or plant level.

Finally, applying our results to an asset-pricing context helps, not only produce momentum along the lines of Kasa, Walker, and Whiteman (2014), but also rationalize irresponsiveness to information about fundamentals at long horizons. This in turn hints to a possible fragility of the predictions of the literature that emphasizes long-term risks to the accommodation of higher-order uncertainty.

Layout. Section 2 expands on the related literature. Section 3 introduces our baseline 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. Sections 8 and 9 consider the applications to consumption, investment, and asset prices. Section 10 expands on the robustness of our insights. Section 11 concludes.

[5] Related to our approach, a number of studies (Fuhrer and Moore, 1995; Roberts, 1997; Nunes, 2010; Adam and Padula, 2011; Fuhrer, 2012) use direct measures of expectations in survey data as proxies for the representative agent's inflation expectations when estimating the *standard* NKPC. As discussed in Mavroeidis, Plagborg-Møller, and Stock (2014), this approach is inconsistent with the micro-foundations of the NKPC, which is based on the assumption of symmetric information. Our approach avoids this problem.

2 Related Literature

Our paper builds heavily on the literature on informational frictions. Our main marginal contributions are the sharp observational-equivalence result; the bridge to the DSGE literature; and the empirical exercise in the context of inflation. An additional contribution is the analysis in Section 10; this develops a flexible yet tractable specification of the information structure, which overcomes the usual difficulties in characterizing the dynamics of higher-order beliefs.

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 contain the insight that incomplete information can rationalize myopia at the aggregate level. Conversely, Angeletos and Lian (2016a) offer the latter insight, but abstract from learning, thus also shutting down the kind of momentum we accommodate here. In addition, they consider a non-stationary setting, which rules out recurring shocks and is therefore inappropriate for the applied purposes of our paper.

Nimark (2008) is the first paper to combine incomplete information with Calvo-like sticky prices, thus anticipating the application found in Section 7 of this paper. In particular, that paper shares both our emphasis on forward-looking beliefs and the insight that incomplete information can rationalize a backward-looking component in the inflation dynamics. It does not, however, contain either our analytical results or our empirical explorations.

Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), Kasa, Walker, and Whiteman (2014), and Nimark (2017) study a similar asset-pricing application as the one found in Section 8 of our paper. We complement these papers by offering a sharp representation of how incomplete information can rationalize momentum and as well as myopia.

Mackowiak and Wiederholt (2009) argue that rational inattention can reconcile price rigidity at the macro level with price flexibility at the micro level and, relatedly, Carroll et al. (2018) argue that informational frictions can reconcile the micro and macro estimates of consumption habit. Although this message sounds similar to one of ours, the mechanism is distinct. In those papers, the difference between micro and macro is attributed to the idea that agents pay less attention to, or otherwise face greater first-order uncertainty about, aggregate shocks relative to idiosyncratic shocks. In our paper, instead, the difference is attributed to GE effects and higher-order uncertainty; it is therefore present even if the agents face the same level of first-order uncertainty about the two kinds of shocks.

By replacing rational expectations with alternative solution concepts, Gabaix (2016) and Farhi and Werning (2017) accommodate a similar form of myopia as the one found in the present paper and in the earlier work of Angeletos and Lian (2016a). They do not, however, produce our second feature, the anchoring of current outcomes to past outcomes and the momentum in beliefs. In terms of condition (2), they let $\omega_f < 1$ but restrict $\omega_b = 0$. A recent paper by Woodford (2018), on the other hand, accomplishes both objectives by combining finite planning horizons—in the form of a

truncation in the value-function iteration of the agents' decision problems—with a form of adaptive leaning. We view these approaches as complementary to ours, but prefer ours because it preserves rational expectations and it offers a simple yet powerful explanation of salient features of the data.

Related is also the literature on adaptive learning (Sargent, 1993; Evans and Honkapohja, 2012; Marcet and Nicolini, 2003). This literature, too, departs from rational expectations in order to accommodate, among other effects, a form of anchoring similar to ours; see, in particular, Carvalho et al. (2017) for an application in the context of inflation. Not only does our approach reconcile such anchoring with rational expectations, but it also disentangles the expectations of aggregate outcomes and the associated GE effects (which relate to higher-order beliefs) from the expectations of fundamentals and the associated PE effects (which relate to first-order beliefs).

3 The Abstract Framework

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

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

where ξ_t is the exogenous payoff-relevant fundamental, $\mathbb{E}_{it}[\cdot]$ 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 ξ_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 β and γ. In particular, β identifies the extent to which the optimal action of an agent depends on her expectations of her *own* future action, whereas γ 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:

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

where $\overline{\mathbb{E}}_t[.]$ 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, β controls the rate at which the future is discounted, and γ 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 $\overline{\mathbb{E}}_t[\cdot]$ 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:

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

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 ξ_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 ξ_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 ξ_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 10), we make the following two assumptions. First, we let the fundamental ξ_t follow an AR(1) process:

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

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

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

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 character-

ization 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 10, 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.

Rational Expectations and Common Knowledge. 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 ξ_t, but also that she is perfectly confident that every other agent knows ξ_t, that everyone knows that everyone knows, and so on, ad infinitum. Along with the simplifying assumption that there is no other information about the future fundamental, common knowledge of the current value of ξ_t induces also a common belief about the entire future path of it: everybody shares the same expectations about the future fundamentals, and this fact itself is common knowledge. Finally, the combination of such common knowledge with the rational-expectations solution concept implies that the agents can reach a perfect consensus about the future path not only of ξ_t but also of a_t, the endogenous outcome.

In a nutshell, setting $\sigma = 0$ prevents the agents from having any doubts about the awareness and the responsiveness of others, thus facilitating a high coordination. Conversely, letting $\sigma > 0$ accommodates such doubts and inhibits coordination. That's the essence of higher-order uncertainty.[6]

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

[6] 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. For complementary discussions of the modeling role of higher-order uncertainty, see Angeletos and Lian (2016b) and Morris and Shin (2003).

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 ξ_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, τ. 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 $\frac{\varphi}{1-\rho\delta}$.

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

$$a_t = \frac{\varphi}{1-\rho\delta}\xi_t = \Phi^*(L)\eta_t, \qquad \textit{with} \qquad \Phi^*(L) \equiv \frac{\varphi}{1-\rho\delta}\frac{1}{1-\rho L}, \tag{8}$$

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, ξ_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

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

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

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

Iterating this condition once gives

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

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:

$$a_t = \varphi \sum_{h=0}^{\infty} \gamma^h \overline{\mathbb{F}}_t^{h+1} [\xi_{t+h}] \tag{11}$$

where, for any random variable X, $\overline{\mathbb{F}}_t^h [X]$ is defined recursively by

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

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 $\frac{\varphi}{1-\rho\beta} = 1$ (this is completely innocuous) and rewrite condition (10) as follows:

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

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:

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

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

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

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 $\overline{\mathbb{E}}_t[\overline{\mathbb{E}}_{t+k}[\xi_{t+k}]]$ 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:

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

For any t and any $k \geq 2$, there are $k-1$ types of second-order beliefs about ξ_{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 explains the computational challenges that one typically faces when trying to analyze business-cycle models with incomplete information.

Note, however, that a complete characterization of the dynamics of higher-order beliefs is not strictly needed. In a rational-expectations equilibrium, the agents only need to forecast the endogenous aggregate outcome. The expectations (first-order beliefs) of the endogenous outcome may be expressed as functions of the higher-order beliefs of the fundamentals, but higher-order beliefs do not matter per. As in Lucas (1972), one can thus bypass the hierarchy of beliefs and instead study directly the fixed point between the actual outcome and the first-order beliefs of it.

This is a basic but important point. To understand the deeper logic and the robustness of certain results, the analyst may find it useful to dissect the rational-expectations equilibrium through the lenses of higher-order beliefs. For applied purposes, however, it may be more straightforward to solve directly for the rational-expectations fixed point.

This is where the specification of the process of fundamental and of the information structure we have assumed here and the methods developed in Huo and Takayama (2015) come handy. These methods imply that, insofar as the underlying fundamental and the available signals follow ARMA processes of arbitrary but finite order, the fixed point we are interested in is also an ARMA process of higher but finite orders. Under the particular assumptions we have made here, this turns out to be an AR(2) process, whose exact form is characterized below.

Proposition 2. *The equilibrium exists, is unique and is such that the aggregate outcome follow the*

following law of motion:

$$a_t = \Phi(L;\rho,\vartheta)\eta_t, \qquad \textit{with} \qquad \Phi(L;\rho,\vartheta) \equiv \left(1-\frac{\vartheta}{\rho}\right)\left(\frac{1}{1-\vartheta L}\right)\Phi^*(L), \tag{12}$$

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

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

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

$$a_t = \frac{\varphi}{1-\delta\rho}\overline{\mathbb{E}}_t[\xi_t], \tag{13}$$

which is the same as the complete-information outcome, modulo the replacement of ξ_t, the actual fundamental, with $\overline{\mathbb{E}}_t[\xi_t]$, the average first-order forecast of it.[7] Furthermore, using the Kalman filter, one can show that the average first-order forecast follows an AR(2) process given by

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

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 $\gamma = 0$.

What happens when $\gamma > 0$, that is, when we switch on the GE effect? Proposition 2 then applies with $\vartheta > \lambda$.[8] 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 property reflects the presence of higher-order uncertainty and its interaction with learning. 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.

In Section 10, we make this logic clear, and also highlight its robustness, by working out the dynamic structure of the entire belief hierarchy under a more flexible information structure than the one assumed here. For the time being, let us concentrate on the simpler information structure assumed

[7] To see this, note that, when $\gamma = 0$, condition (3) reduces to $a_{it} = \varphi\mathbb{E}_{it}[\xi_t] + \delta\mathbb{E}_{it}[a_{it+1}]$, with $\delta = \beta$. Iterating this forward gives $a_{it} = \varphi\mathbb{E}_{it}[\sum_k \delta^k \xi_{t+k}] = \frac{\varphi}{1-\delta\rho}\mathbb{E}_{it}[\xi_t]$. Aggregating gives the result.

[8] The fact that $\vartheta > \lambda$ is not obvious from looking at Proposition 2, but follows from the property that ϑ is increasing in γ, which is proved in Proposition 3.

here and let us note that, unlike the first-order beliefs, the relevant second- and higher-order beliefs do not follow AR(2) processes. There is therefore a certain "magic" behind the result that the equilibrium outcome follows an AR(2) even though the higher-order beliefs do not.

The closed-form characterization obtained in Proposition 2 allows us to shed light on how the information friction interacts with the GE effect in shaping the equilibrium dynamics. In particular, we can establish the following properties for the coefficient ϑ.[9]

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

(ii) ϑ is increasing in γ.

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

To interpret this result, note first that ϑ plays a dual role in the impulse response of a_t: a higher ϑ 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 ϑL. Accordingly, part (i) of the result 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 γ. This follows from two key points made earlier: that a higher γ 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 that, insofar the GE effect is non-zero, ϑ is strictly increasing in β as well.[10]

In the next two sections, we use the above comparative statics to shed light on how the informational friction and its interaction with GE effects drive the forms of myopia and anchoring we capture in our paper. We conclude this section with the following remark. By design of our baseline model, Proposition 2 confounds 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 result continues to hold, subject to the restriction that the value of ϑ 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 10 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.

[9] Note how the sharpness of the comparative statics provided in Proposition 3 contrasts the reliance of much of the related literature on numerical simulations.

[10] To understand the last finding, recall that β 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 β *indirectly* increases the dependance of the current outcome to forward-looking higher-order beliefs.

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:

$$a_t = \varphi \xi_t + \delta \omega_f \mathbb{E}_t \left[a_{t+1} \right] + \omega_b a_{t-1} \tag{15}$$

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 ω_f represents a higher discounting of the future, or less forward-looking behavior; a higher ω_b represents a greater anchoring of the current outcome to the past outcome, or more backward-looking behavior.

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

Proposition 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 (ω_f, ω_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 σ) and/or a stronger GE feedback (higher γ) maps to both greater myopia (lower ω_f) and greater anchoring (higher ω_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 σ) or the strength of the underlying GE effects (parameterized by γ).

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 by adding, respectively, habit persistent in consumption, adjustment costs to changing 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 permit the DSGE literature to match the sluggishness in the dynamic responses of consumption, investment, and inflation to identified shocks.[11] 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, or why they may exhibit momentum. 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 possibility of reconciling bounded foresight with rational expectations, but also a more promising account of the macroeconomic data.[12]

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 an incomplete-information economy only when ω_f and ω_b happen to lie together on a particular line in the (ω_f, ω_b) space.

Proposition 5. *There exists a function Ω such that the equilibrium dynamics of a hybrid economy can be replicated by that of an incomplete-information economy for some $\sigma > 0$ if and only if $\omega_b > 0$ and the following restriction is satisfied:*

$$\omega_f = \Omega\left(\omega_b; \delta, \rho\right) \tag{16}$$

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

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

[11] These add-ons also help the New Keynesian model generate positive comovement in the macroeconomic quantities.

[12] It might be possible to reconcile Gabaix (2016) and Farhi and Werning (2017) with $\omega_b > 0$ by modifying the exogenous default points in these works in a manner than helps mimic the effects of Bayesian learning. The approach taken in a recent paper by Woodford (2018) has that flavor.

from ω_f and ω_b. Suppose further that one knows δ and ρ 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 ω_f and ω_b can be used to identify σ.

Additional testable predictions, or overidentifying restrictions, can be obtained by looking at the forecasts of future outcomes. Let $\epsilon_t^k \equiv a_{t+k} - \overline{\mathbb{E}}_t [a_{t+k}]$ be the realized average k-period ahead forecast error. In the hybrid model, ϵ_t^k is serially uncorrelated, because the economy is populated by a rational representative agent. In our model, instead, ϵ_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 σ, this provides us with an additional restriction that can be used to identify σ 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:

$$a_{it} = \mathbb{E}_{it}[\varphi \xi_{it} + \beta a_{it+1} + \gamma a_{t+1}] \tag{17}$$

where

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

and where ζ_{it} is a purely idiosyncratic shock. We let the latter follow a similar AR(1) process as the aggregate shock: $\zeta_{it} = \rho \zeta_{it-1} + \epsilon_{it}$, where ϵ_{it} is i.i.d. across both i and t.[13]

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

$$a_{it} \;=\; \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[\xi_{it+k}\right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[a_{t+k+1}\right] \;=\; \mathrm{PE}_{it} + \mathrm{GE}_{it}$$

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

where

$$\text{PE}_{it} \equiv \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[\xi_{it+k} \right] \qquad \text{and} \qquad \text{GE}_{it} \equiv \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[a_{t+k+1} \right]$$

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

$$a_t = \overline{\text{PE}}_t + \overline{\text{GE}}_t,$$

where $\overline{\text{PE}}_t$ and $\overline{\text{GE}}_t$ 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 γ 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 a 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.

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 the typical agent from facing uncertainty about the beliefs of others unless she also faces uncertainty about her own (and common) fundamental. That is, our baseline specification confounds the first-order uncertainty of an agent about her own fundamental with her higher-order uncertainty. 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 ξ_{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 $\frac{\mathbb{V}(\zeta_{it})}{\mathbb{V}(\xi_t)} \to \infty$ (large idiosyncratic shocks).

In this variant, we can effectively separate the forecasts of one's own fundamental from the fore-

casts of the fundamentals of others: relative to the signal x_{it}, the own fundamental ξ_{it} contains negligible information about the current and the future values of ξ_t; and symmetrically, relative to ξ_{it}, x_{it} contains negligible information about the current and future values of ζ_{it}. It follows that

$$\mathsf{PE}_{it} \equiv \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[\xi_{it+k}\right] = \frac{1}{1-\rho\beta}\xi_{it},$$

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. *The individual and the aggregate outcomes are given by, respectively,*

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

where u_{it} is idiosyncratic noise, orthogonal to both ξ_t and ζ_{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 $\frac{1}{1-\rho\beta}\xi_t$; and a GE effect, which is denoted by $\overline{\mathsf{GE}}_t$. 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 aforecited paper complement each other towards generating more pronounced distortions at the macro level than at the micro level.[14]

7 Inflation and Aggregate Supply

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 Philips 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. Firms set prices optimally, but can adjust them only infrequently. Each period, a firm has the option to reset its price with probability $1-\theta$; otherwise, 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

[14] We verify all these intuitions in Appendix B with a variant that lets both kinds of shocks be observed with noise.

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:

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

subject to the demand function

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

where $Q_{t|t+k}$ is the stochastic discount factor between t and $t+k$, Y_{t+k} and P_{t+k} are, respectively, aggregate income and the aggregate price level in period $t+k$, P_{it} is the firm's price, as set in period t, $Y_{i,t+k|t}$ is the firm's quantity in period $t+k$, conditional on the firm not having changed its price since period t, and Ψ_{t+k} is the real marginal cost in period $t+k$.

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

$$p_{it}^* = (1 - \delta\theta) \sum_{k=0}^{\infty} (\delta\theta)^k \mathbb{E}_{it} [\psi_{t+k} + p_{t+k}], \tag{18}$$

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:

$$p_t = (1 - \theta) \int p_{it}^* + \theta p_{t-1} \tag{19}$$

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.[15]

[15] 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.

This permits us to restate condition (18) as

$$p_{it}^{*} - p_{t-1} = (1-\delta\theta)\sum_{k=0}^{\infty}(\delta\theta)^{k}\mathbb{E}_{it}[\psi_{t+k} + \pi_{t+k}], \tag{20}$$

By (19), on the other hand, we have $\pi_t \equiv p_t - p_{t-1} = (1-\theta)\int (p_{it}^{*} - p_{t-1})$. Combining and rearranging, we arrive at the following expression for inflation:

$$\pi_t = \frac{(1-\delta\theta)(1-\theta)}{\theta}\sum_{k=0}^{\infty}(\delta\theta)^{k}\overline{\mathbb{E}}_t\left[\psi_{t+k}\right] + \delta(1-\theta)\sum_{k=0}^{\infty}(\delta\theta)^{k}\overline{\mathbb{E}}_t\left[\pi_{t+k+1}\right]. \tag{21}$$

When information is complete, we can replace $\overline{\mathbb{E}}_t[\cdot]$ 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:

$$\pi_t = \kappa\psi_t + \delta\mathbb{E}_t[\pi_{t+1}], \tag{22}$$

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

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 (ψ_t and π_t to, respectively, ξ_t and a_t; and by letting

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

Note that the GE feedback, as measured by γ, is larger when prices are more flexible. Higher-price flexibility therefore reinforces the role of higher-order uncertainty, a point we revisit shortly.

To apply our analytical results, we finally assume that ψ_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:*

$$\pi_t = \kappa\psi_t + \omega_f\delta\mathbb{E}_t[\pi_{t+1}] + \omega_b\pi_{t-1} \tag{23}$$

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

Part (i) establishes that, when information is incomplete, it is *as if* inflation is governed by a variant of the NKPC that introduces myopia, in the form of $\omega_f < 1$, along with a backward-looking compo-

nent, 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 κ 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 relatively modest degree of price stickiness.

7.2 Testing the 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 (ω_f, ω_b) in the way described in Proposition 5, whereas unrestricted estimations of the Hybrid NKPC allow these parameters to be free. And second, our theory ties the pair (ω_f, ω_b) to the 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 (ω_f, ω_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 (ω_f, ω_b) that are compatible with our theory for some level of noise. To construct this locus, we only need to specify δ, θ, and ρ. We set $\delta = 0.99$, $\theta = 0.6$, and $\rho = 0.95$. The value of θ corresponds to a modest degree of price stickiness, broadly in line with textbook calibrations of the New Keynesian model and with the micro data. The value of ρ, 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 (ω_f, ω_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

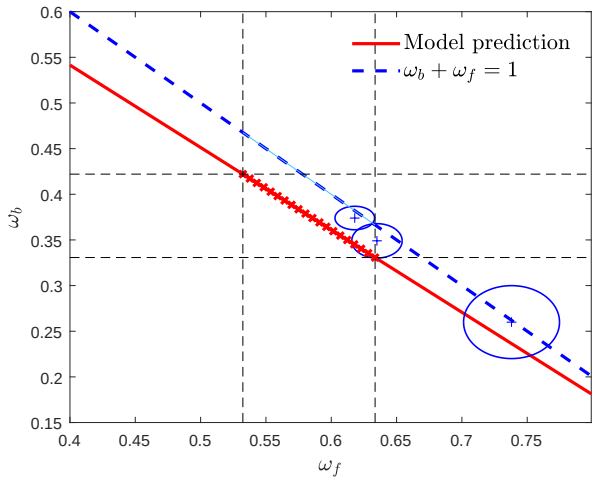


Figure 1: Testing the Theory

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 ω_f and ω_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:

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

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 than that in the aforecited paper, we can characterize it as a function of σ and of (δ, θ, ρ). 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 σ in our model. For any σ in this interval, we can then compute the pair (ω_f, ω_b) predicted by our theory.

We can thus map the evidence reported in Coibion and Gorodnichenko (2015) to a segment of the (ω_f, ω_b) locus we obtained earlier on. This segment is identified by the red crosses in Figure 1 and gives the pairs of (ω_f, ω_b) that are consistent with the confidence interval for K provided in Coibion and Gorodnichenko (2015). It is then evident from the figure that our model can pass jointly both the test of matching that evidence and the test of matching the existing estimates of the Hybrid NKPC.[16]

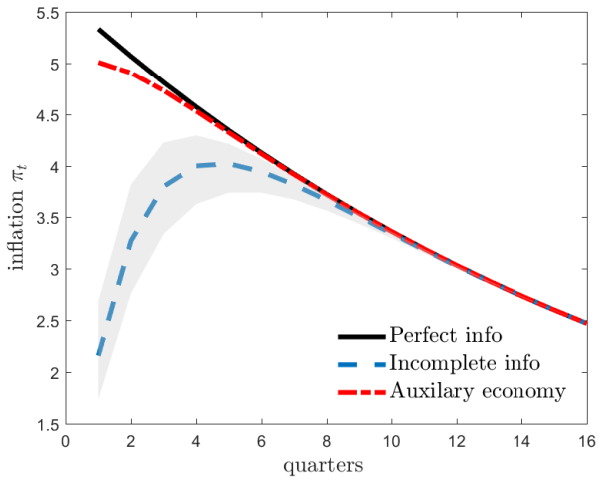


Figure 2: Impulse Response Function of Inflation

The quantitative implications of our theory for the dynamics of inflation are further illustrated in Figure 2. 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%

[16] 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 ω_f and the largest ω_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.

lower than its complete-information counterpart, and the peak of the inflation response is attained 5 quarters after impact rather than on impact. 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 appear to be unable to produce such sluggishness.[17]

Recall from equation (21), the inflation dynamics π_t in the incomplete information model can be decomposed into two components: beliefs about the present value of real marginal costs, $\varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t[\psi_{t+k}]$, and the present value of future inflation, $\gamma \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t[\pi_{t+k+1}]$. Note that the same decomposition can also be applied when agents have perfect information

$$\pi_t^* = \varphi \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\psi_{t+k} | \psi_t\right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\pi_{t+k+1}^* | \psi_t\right], \tag{25}$$

A natural question is when information is incomplete, which component contributes more to the anchoring of inflation. We proceed by defining the following auxiliary inflation $\widetilde{\pi}_t$

$$\widetilde{\pi}_t = \varphi \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t \left[\psi_{t+k}\right] + \gamma \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t \left[\pi_{t+k+1}^* | \psi_t\right]. \tag{26}$$

The difference between π_t^* and $\widetilde{\pi}_t$ measures the importance of beliefs about real marginal costs, and the difference between $\widetilde{\pi}_t$ and π_t measures the importance of beliefs about inflation. The red dashed line in Figure 2 corresponds to $\widetilde{\pi}_t$. Clearly, the inflation dynamics in the incomplete information differ from its complete-information counterpart is mainly because of the anchoring of beliefs about future inflation. These beliefs are tied with higher-order beliefs about real marginal costs, which display more inertia and more dampening effects compared with the first-order beliefs.

8 Consumption and Aggregate Demand

We now shift the focus from the supply block of the New Keynesian model to its demand block. In particular, we show how incomplete information and higher-order uncertainty 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.

[17] See, for example, Golosov and Lucas Jr (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, it contributes towards filling that gap.

Consider an economy populated by a large number of infinitely-lived households, indexed by $i \in [0,1]$. For any household i, let her preferences be

$$\sum_{t=0}^{\infty} \theta^t U(C_{it}, N_{it}),$$

where C_{it} denotes consumption, N_{it} 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 $\theta \in (0,1)$ is the subjective discount factor. We assume that the consumers trade a one-period riskless bond and let R_t denote one plus the real interest rate between periods t and $t+1$. We treat R_t as an exogenous random variable[18] and 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. We finally use lower-case variables to represent log-deviations from the steady state (e.g., $r_t \equiv \log R_t - \log \chi$).

Following similar steps as in Angeletos and Lian (2016a), we can express aggregate consumption in period t as follows:

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

where $\overline{E}_t$ 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;[19] and finally solving for c_t, we obtain the following equilibrium restriction on aggregate consumption:

$$c_t = -\sum_{k=0}^{\infty} \theta^k \overline{\mathbb{E}}_t[r_{t+k}] + (1-\theta) \sum_{k=1}^{\infty} \theta^{k-1} \overline{\mathbb{E}}_t[c_{t+k}] \tag{27}$$

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.

When information is complete, the average expectation, $\overline{\mathbb{E}}_t$, can be replaced by $\mathbb{E}_t$, the expectation conditional on the common information set. Using the Law of Iterated Expectations, it is then straightforward to show that condition (27) can be restated in recursive form as follows:

$$c_t = -r_t + \mathbb{E}_t \left[c_{t+1} \right]. \tag{28}$$

[18] 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.

[19] 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.

Note that this condition is invariant to the discount factor θ, which means that the response of aggregate demand to news about the future real rate is, perhaps surprisingly, the same regardless of how impatient or short-sighted the households are. This is because, in the frictionless, representative-agent benchmark, discounting has perfectly offsetting PE and GE effects.[20]

When instead information is incomplete, the Law of Iterated Expectations does not hold at the aggregate level and, as a result, condition (27) cannot be reduced to condition (28). Instead, it is best to read condition (27) as a dynamic beauty contest among the consumers. This game is nested in our framework by interpreting (ξ_t, a_t) as (r_t, c_t) and by letting

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

The property that the sum $\delta = \beta + \gamma$ is invariant to θ verifies the aforementioned claim that a lower θ has offsetting PE and GE effects in the frictionless benchmark. On the other hand, the property that γ decreases with ϑ reveals that a lower θ unambiguously increases the relevant GE effect, which is the Keynesian income-spending multiplier. The following result is then an immediate implication of our analysis, provided, of course, that the information structure is specified as in our abstract analysis.

Proposition 8. *When information is incomplete, there exist scalars $\omega_f < 1$ and $\omega_b > 0$ such that the equilibrium process for consumption solves the following equation:*

$$c_t = -r_t + \omega_f \mathbb{E}_t[c_{t+1}] + \omega_b c_{t-1} \tag{29}$$

Furthermore, a lower θ, which means a stronger income-spending multiplier, maps to a lower ω_f and a higher ω_b.

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).[21] Furthermore, the parameter θ, which was irrelevant under complete information, now emerges a key determinant of the as-if myopia and habit.

One does not have to interpret θ literally. For instance, we can readily extend the analysis to a perpetual-youth, overlapping-generations model along the lines Del Negro, Giannoni, and Patterson (2015) and, under appropriate assumptions, replicate the results reported above with θ replaced by $\chi\theta$, where χ is the survival probability. The latter can be thought of as a measure of how long the planning horizons of the individuals are. Alternatively, as noted in Farhi and Werning (2017), $1 - \chi$

[20] The property that discounting has offsetting PE and GE effects in the frictionless, representative-agent benchmark has been emphasized in McKay, Nakamura, and Steinsson (2016).

[21] There is a subtle difference. The form of habit assumed in those papers imposes $\omega_f + \omega_b = 1$. In our setting, instead, $\omega_f + \omega_b < \rho < 1$.

can be thought of as a proxy for the probability of binding liquidity constraints. For our purposes, the key observation is that such parameters may be irrelevant under complete information because they tend to have offsetting PE and GE effects, but have important interactions with incomplete information because they determine the strength of the relevant GE effect, which in turn regulates the importance of higher-order uncertainty.

The reason that Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) replaced condition (28) with condition (29) is quite simple. In the data, aggregate consumption responds little to monetary and other shocks on impact, but builds up force over time. Such a pattern is inconsistent with condition (28). Replacing the latter with condition (29) fixes the problem and improves the empirical performance of the model. The existing micro-foundation of condition (29), however, runs at a problem. If this condition is the product of habit persistence in preferences, one would expect estimates of the habit parameter ω_b 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.

9 Investment and Asset Prices

Although in the previous application we 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.

As yet another possible application, consider a log-linearized version of the standard asset-pricing

condition in an infinite horizon, representative-agent model:

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

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

By assuming a representative agent, the above condition conceals the importance of higher-order beliefs. A number of works have sought to unearth that role by considering variants with heterogeneously informed, short-term traders, in the tradition of Singleton (1987).[22] We can capture such a variant in our setting by modifying the equilibrium pricing condition as follows:

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

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

Let us now assume that the dividend is given by $d_{t+1} = \xi_t + u_{t+1}$, where ξ_t follows an AR(1) process and u_{t+1} is i.i.d. over time, and that the information of the typical trader can be represented by a series of private signals as in condition (7).[23] Applying our results, and using the fact that $\xi_t = \mathbb{E}_t[d_{t+1}]$, we then have that the component of the equilibrium asset price that is driven by ξ_t obeys the following law of motion, for some $\omega_f < 1$ and $\omega_b > 0$:

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

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

Kasa, Walker, and Whiteman (2014) have already emphasized how incomplete information and higher-order uncertainty can help explain momentum; related points are contained in Allen, Morris, and Shin (2006) and Bacchetta and van Wincoop (2006). Our result offers a sharp illustration of this insight and blends it with the insight regarding myopia. In the present context, the latter insight

[22] See, inter alia, Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), Kasa, Walker, and Whiteman (2014), and Nimark (2017).

[23] Here, we are abstracting from the complications of the endogenous revelation of information and we think of the signals in (7) as convenient proxies for all the information of the typical trader. One can also interpret this as a setting in which the dividend is observable (and hence so is the price, which is measurable in the dividend) and the assumed signals are the representation of a form of rational inattention. Last but not least, we have verified that the solution with endogenous information can be approximated very well by the solution obtained with exogenous information.

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

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

Allowing learning from prices adds more realism, but typically rules out an analytic characterization of the equilibrium.[24] Suppose, in particular, that the traders in our setting can perfectly observe the current price as well as the last-period dividend. In this case, the equilibrium pricing dynamics does not admit a finite state-space representation. To illustrate, set $\delta = 0.98, \rho = 0.95, \sigma_u = 2$, and $\sigma_\epsilon = \sigma_\nu = 5$, and approximate the equilibrium dynamics with an MA(100) process. The solid blue in Figure 3 gives the resulting IRF of the equilibrium price to an innovation in ξ_t. The dashed red line is obtained by taking our hybrid economy, which assumes away the learning from either the price or the past dividend, and recalibrating the level of the idiosyncratic noise so that the implied IRF is close as possible to the one obtained in the economy in which such learning is allowed. As evident in the figure, the hybrid economy does a very good job in replicating the dynamics of the latter economy.

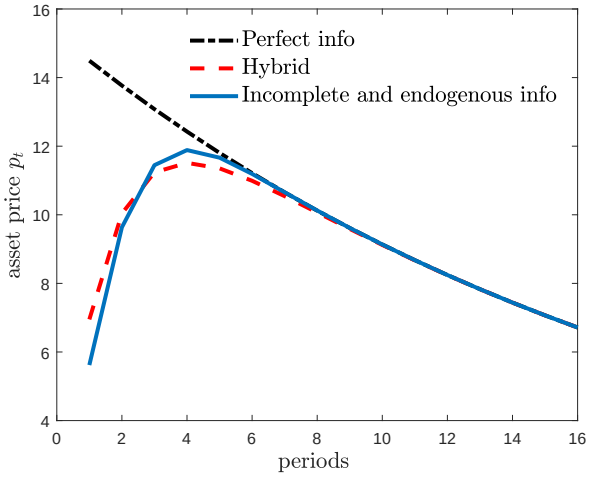


Figure 3: Impulse Response Function of Asset Price

We have verified that this similarity extends to a wide range of values for the parameters of the

[24] See Nimark (2017) and Huo and Takayama (2015) for a more detailed discussion.

assumed setting. This similarity may, of course, be broken by assuming a more complex stochastic process for the fundamental and a more convoluted learning dynamics. However, the analysis of the next section together with the example presented here illustrate why our analysis can be thought of as a convenient proxy of settings with endogenous information aggregation.

10 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, even if we account for the role of the informational friction on the beliefs of the fundamental, the absence of common knowledge arrests the response of the beliefs of the future beliefs and the future actions of others to any news about the future fundamentals, thus also arresting the equilibrium expectations of the future outcome (e.g., future inflation, future income, or future asset price). This explains why agents behave as if they discount the future more heavily than what it is rational. The second insight is that learning introduces extra persistence, or momentum, in the equilibrium beliefs and the equilibrium outcome relative to that in the underlying fundamental.

In this section, we elaborate on the robustness of these insights and on the interaction of two key elements of our theory, namely higher-order uncertainty and learning. To this goal, we modify the fundamental process and the information structure as follows. First, we assume that the fundamental ξ_t follows infinite MA process

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

The AR(1) process we have assumed earlier is included as a special case where $\rho_k = \rho^k$. Second, 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 $\epsilon_{i,t,t-k} \sim \mathcal{N}(0, (\tau_k)^{-2})$ is i.i.d. across i and t, uncorrelated across k, and orthogonal to the past, current, and future innovations in the fundamental. That is, whereas our baseline specification has the agents observe a signal about ξ_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 η_{t-k} are independent of the signals received about the shock $\eta_{t-k'}$. Second, it allows for more flexible learning dynamics in the sense that the precision τ_k does not have to be flat in k: the

quality of the incremental information received in any given period about a past shock may either 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 σ controls both objects at once.

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

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

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

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

We conclude that the IRF of the (average) first-order belief to any given innovation is given by the sequence $\mathbf{F}_1 = \{f_{1,k}\}_{k=0}^{\infty}$, where

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

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 λ_0 is (i.e., the less precise the initial information is), the larger the initial initial gap between the two IRFs (i.e., a larger the initial forecast error). And the slower λ_k increases with k (i.e., the slower the learning over time), the longer it takes for that gap (and the average forecast) to disappear.

These properties are intuitive and are shared by the specification studied in the rest of the paper. 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

σ. 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 (31) to period $t+1$ and taking the period-t average expectation, we get

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

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

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

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

$$f_{h,k} \equiv \frac{\partial\mathbb{E}\left[\overline{\mathbb{F}}_t^h\left[\xi_{t+h}\right]\middle|\,\eta_{t-k}\right]}{\partial\eta_{t-k}} = \lambda_k\lambda_{k+1}...\lambda_{k+h-1}\rho_{k+h-1} \tag{32}$$

Note next that $\rho_{k+h-1} = \frac{\partial\mathbb{E}[\xi_{t+h}|\eta_{t-k}]}{\partial\eta_{t-k}}$. It follows that the ratio $\frac{f_{h,k}}{\rho_{k+h-1}}$ 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 ration is identically 1 for all k and h. When, instead, information is incomplete, we have that

$$\frac{\frac{\partial}{\partial\eta_{t-k}}\mathbb{E}\left[\overline{\mathbb{F}}_t^h\left[\xi_{t+h}\right]\middle|\,\eta_{t-k}\right]}{\frac{\partial}{\partial\eta_{t-k}}\mathbb{E}\left[\xi_{t+h}\middle|\,\eta_{t-k}\right]} = \frac{f_{h,k}}{\rho_{k+h-1}} = \lambda_k\lambda_{k+1}...\lambda_{k+h-1}.$$

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 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 $a_t = \overline{\mathbb{E}}_t[\xi_t] + \gamma \overline{\mathbb{E}}_t[a_{t+1}]$, which in turn implies that $a_t = \sum_{h=1}^{\infty} \gamma^{h-1} \overline{\mathbb{F}}_t^h [\xi_{t+h-1}]$. From the preceding characterization of the higher-order beliefs, we have $\overline{\mathbb{F}}_t^h [\xi_{t+h-1}] = \sum_{k=0}^{\infty} f_{h,k} \eta_{t-k}$, with $f_{h,k}$ defined as in (32). It follows that

$$a_t = \sum_{k=0}^{\infty} g_k \eta_{t-k}$$

where, for all k,

$$g_k \equiv \frac{\partial \mathbb{E}\left[a_t \middle| \eta_{t-k}\right]}{\partial \eta_{t-k}} = \sum_{h=1}^{\infty} \gamma^{h-1} f_{h,k} = \left\{ \sum_{h=1}^{\infty} \gamma^{h-1} \lambda_k \lambda_{k+1} ... \lambda_{k+h-1} \rho_{k+h-1} \right\}. \tag{33}$$

This makes clear how the IRF of the equilibrium outcome is connected to the IRFs of the first- and higher-order beliefs. Importantly, the higher γ is, the more the dynamics of the equilibrium outcome tracks the dynamics higher-order beliefs relative to the dynamics of lower-order beliefs.

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

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

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

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

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

of others move *even* less than the forecast of the fundamental.

This is the crux of the forward-looking component of our observational-equivalence result (that is, the one regarding myopia). Note in particular that the extra discounting of the future remains present even if when if control for the impact of the informational friction on first-order beliefs. Indeed, replacing ξ_t' with ξ_t in the above shuts down the effect of first-order uncertainty. And yet, the extra discounting survives, reflecting the role of higher-order uncertainty. This complements the related points we make in Section 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 ξ_t, this belief happens to coincide with the average first-order belief in the original economy, and these facts are common knowledge. The equilibrium outcome in this economy is proportional to the subjective belief of ξ_t and is given by

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

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

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

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

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

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

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

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

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

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

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

The first term in the right-hand side of this recursion corresponds to the average expectation of the future fundamental. The second term corresponds the average expectation of the future outcome (the actions of others). The role of first-order uncertainty is captured by the fact that $f_{1,k}$ is lower than ρ_k. The role of higher-order uncertainty is captured by the presence of λ_k in the second term: it is as if the discount factor γ has been replaced by a discount factor equal to $\lambda_k \gamma$, which is strictly less than γ. This represents a generalization of the form of myopia seen in condition (34). There, learning was shut down, so that that λ_k and the extra discounting of the future were invariant in the horizon k. Here, the additional discounting varies with the horizon because of the anticipation of future learning (namely, the knowledge that λ_k will increase with k).

Consider next property (ii), namely the property that

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

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

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

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

The inequality captures the effect of learning on first-order beliefs. Had information being perfect, we would have had $\frac{\widehat{g}_{k+1}}{\widehat{g}_k} = \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}$; now, we instead have $\frac{\widehat{g}_{k+1}}{\widehat{g}_k} > \frac{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h+1}}{\sum_{h=0}^{\infty} \gamma^h \rho_{k+h}}$. This means that, in the variant economy, the impact of the shock on the equilibrium outcome can build force over time because, and only because, learning allows for a gradual build up in first-order beliefs.[25]

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

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

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

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,[26] 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.

Should we state the following corollary in terms of $\widehat{g}_k$, or discuss the special case where ξ_t follows an AR(1) process?

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 λ_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 ω_b, which captures the endogenous build up in the equilibrium dynamics, is positive because of learning and it is higher the higher the importance of higher-order uncertainty.

11 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

[26] Recall, this is by property (i) of Proposition 10.

frictions: myopia vis-a-vis the future; and anchoring of current outcomes to past outcomes.

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 as 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 NKPC, 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) is 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.

Two other interesting venues for future research regard asset pricing. One is to explore whether our approach can help explain the fact that asset-pricing pathologies appear to more pronounced at the aggregate data than in disaggregated data. Another is to explore the implications of higher-order uncertainty for the pricing of long-run risks.

References

Abel, Andrew B. and Olivier J. Blanchard. 1983. "An Intertemporal Model of Saving and Investment." *Econometrica* 51 (3):675–692. 11

Adam, Klaus and Mario Padula. 2011. "Inflation dynamics and subjective expectations in the United States." *Economic Inquiry* 49 (1):13–25. 5

Allen, Franklin, Stephen Morris, and Hyun Song Shin. 2006. "Beauty Contests and Iterated Expectations in Asset Markets." *Review of financial Studies* 19 (3):719–752. 2, 4, 9, 22

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

Alvarez, Fernando and Francesco Lippi. 2014. "Price Setting with Menu Cost for Multiproduct Firms." *Econometrica* 82 (1):89–135. 17

Angeletos, George-Marios and Chen Lian. 2016a. "Forward Guidance without Common Knowledge." *American Economic Review* forthcoming. 1, 1, 2, 3, 5, 8

———. 2016b. "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics* 2:1065–1240. 6, 5

Angeletos, George-Marios and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4):1103–1142. 3

Bacchetta, Philippe and Eric van Wincoop. 2006. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *The American Economic Review* 96 (3):552. 2, 4, 9, 22

Bachmann, Rüdiger, Ricardo J. Caballero, and Eduardo M. R. A. Engel. 2013. "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model." *American Economic Journal: Macroeconomics* 5 (4):29–67. 11

Caballero, Ricardo J. and Eduardo M. R. A. Engel. 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach." *Econometrica* 67 (4):783–826. 11

Carroll, Christopher D., Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White. 2018. "Sticky Expectations and Consumption Dynamics." Working Paper 24377, National Bureau of Economic Research. 2

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

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1):1–45. 3, 1, 5, 7.1, 8, 11

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

———. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–78. 1, 1, 7, 7.2, 7.2, 7.2

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." *FRB of New York mimeo* . 1, 8

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

Farhi, Emmanuel and Iván Werning. 2017. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *NBER Working Paper No. 23281* . 2, 2, 5, 12, 8

Fuhrer, Jeff. 2012. "The Role of Expectations in Inflation Dynamics." *International Journal of Central Banking* . 5

Fuhrer, Jeff and George Moore. 1995. "Inflation persistence." *The Quarterly Journal of Economics* 110 (1):127–159. 5

Gabaix, Xavier. 2016. "A Behavioral New Keynesian Model." *NBER Working Paper No. 22954* . 2, 2, 5, 12

Gali, Jordi and Mark Gertler. 1999. "Inflation dynamics: A structural econometric analysis." *Journal of Monetary Economics* 44 (2):195–222. 3, 1, 7, 7.1, 7.2

Gali, Jordi, Mark Gertler, and J David Lopez-Salido. 2005. "Robustness of the estimates of the hybrid New Keynesian Phillips curve." *Journal of Monetary Economics* 52 (6):1107–1118. 1, 7, 7.2, 7.2, 16

Golosov, Mikhail and Robert E Lucas Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2):171–199. 17

Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit Formation in Consumption: A Meta-Analysis." *European Economic Review* 95:142–167. 3, 1, 8

Hayashi, Fumio. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation." *Econometrica* 50 (1):213–224. 11

Huo, Zhen and Naoki Takayama. 2015. "Rational Expectations Models with Higher Order Beliefs." *Yale mimeo* . 4, 24

Jung, Jeeman and Robert J. Shiller. 2005. "Samuelson's Dictum and the Stock Market." *Economic Inquiry* 43 (2):221–228. 1, 9

Kasa, Kenneth, Todd B. Walker, and Charles H. Whiteman. 2014. "Heterogeneous Beliefs and Tests of Present Value Models." *The Review of Economic Studies* 81 (3):1137–1163. 1, 2, 9, 22

Khan, Aubhik and Julia K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76 (2):395–436. 11

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

Mackowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803. 2, 6.3, 11

Mankiw, N. Gregory and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4):1295–1328. 4

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

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock. 2014. "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve." *Journal of Economic Literature* 52 (1):124–88. 5

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The Power of Forward Guidance Revisited." *American Economic Review* 106 (10):3133–3158. 1, 20

Midrigan, Virgiliu. 2011. "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations." *Econometrica* 79 (4):1139–1180. 17

Morris, Stephen and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-fulfilling Currency Attacks." *American Economic Review* :587–597. 5

———. 2002. "Social Value of Public Information." *American Economic Review* 92 (5):1521–1534. 2, 3, 4, 5

———. 2003. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)*. Cambridge University Press. 6, 5

Nakamura, Emi and Jón Steinsson. 2013. "Price Rigidity: Microeconomic Evidence and Macroeconomic Implications." *Annual Review of Economics* 5 (1):133–163. 17

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

———. 2017. "Dynamic Higher Order Expectations." *Cornell Univeristy mimeo* . 2, 4, 22, 24

Nunes, Ricardo. 2010. "Inflation dynamics: the role of expectations." *Journal of Money, Credit and Banking* 42 (6):1161–1172. 5

Roberts, John M. 1997. "Is inflation sticky?" *Journal of Monetary Economics* 39 (2):173–196. 5

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

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

Singleton, Kenneth J. 1987. "Asset Prices in a Time-series Model with Disparately Informed, Competitive Traders." In *New Approaches to Monetary Economics*, edited by William A. Barnett and Kenneth J. Singleton. Cambridge University Press, 249–272. 9

Smets, Frank and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3):586–606. 3, 1, 5, 7.1, 8, 11

Tirole, Jean. 2015. "Cognitive Games and Cognitive Traps." *Toulouse School of Economicss mimeo* . 3

Vives, Xavier and Liyan Yang. 2017. "Costly Interpretation of Asset Prices." mimeo, IESE/University of Toronto. 15

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* . 1, 4, 2, 3, 4, 5

Appendix A: Proofs

Proof of Proposition 1

Follows directly from the analysis in the main text.

Proof of Proposition 2

Suppose that the agent's equilibrium policy function is given by

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

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

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

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

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

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

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

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

which leads to

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

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

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

Next, we characterize the beliefs of ξ_t, $a_{i,t+1}$, and a_{t+1}, that is, the beliefs that show up in the

best-response condition of the agent. The forecast of a random variable

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

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

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

Consider the forecast of the fundamental. Note that

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

from which it follows that

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

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

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

and the forecasts are

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

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

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

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

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

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

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

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

which can be transformed as

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

where

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

Note that $C(z)$ is a cubic equation and therefore contains with three roots. We will verify later that there are two inside roots and one outside root. To make sure that $h(z)$ is an analytic function, we choose $h(0)$ and $h(\lambda)$ 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 $h(L)$

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

where ϑ^{-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

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

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

properties hold:

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

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

Proof of Proposition 3

Note that

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

By the continuity of $C(z)$, it must 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 ϑ, 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 ϑ.

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

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

taking its derivative with respect to τ_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 ϑ is a decreasing function of τ_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 ϑ increases with both β and γ.

When $\gamma = 0$, the three roots of $C(z)$ become β, λ and λ^{-1}. For this case, the outside root is independent of β and γ.

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

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

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

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

To match the hybrid model, we need

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

Combining (38) and (39), and solving for the coefficients of ω_f and ω_b, we infer that the two economies generate the same dynamics if and only if the following two conditions hold:

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

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

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

It is immediate to check that $\omega_f < 1$ and $\omega_b > 0$ if $\vartheta \in (0, \rho)$, which in turn is necessarily true for any $\sigma > 0$; and that $\omega_f = 1$ and $\omega_b = 0$ if $\vartheta = \rho$, which in turn is the case if and only if $\sigma = 0$. This completes the proof of part (i). Part (ii) follows from Proposition 3 together with the fact that a higher ϑ maps to a lower ω_f and a higher ω_b.

Proof of Proposition 5

As already noted, the hybrid and the incomplete-information economies generate the same dynamics if and only if conditions (40) and (41) hold. Using (41), we can rewrite (40) as follows:

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

Furthermore, any (β, γ, ρ), the equilibrium of the incomplete-information economy gives an invertible mapping from $\sigma \in (0, \infty)$ to $\vartheta \in (0, \rho)$, whereas condition (41) gives an invertible mapping from $\vartheta \in (0, \rho)$ to $\omega_b \in (0, \infty)$. It follows that there exists a $\sigma \in (0, \infty)$ such that the equilibrium dynamics of the incomplete-information economy replicates that of the hybrid economy if and only if the pair (ω_b, ω_f) satisfies condition (42) along with $\omega_b \in (0, \infty)$. Finally, the level of the informational friction that achieves this replication is obtained by inverting condition (41) to obtain ϑ, and thereby also σ, as an implicit function of ω_b.

Proof of Proposition 6

Suppose that the individual's policy function is

$$a_{it} = g(L)z_{it} + h(L)x_{it}.$$

In the limit as $\frac{\mathbb{V}(\zeta_{it})}{\mathbb{V}(\xi_t)} \to \infty$, agents effectively interpret z_{it} as a perfect signal of ζ_{it} and as a completely uninformative about ξ_t. As a result, the forecasts of both ξ_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:

$$\mathbb{E}_{it}\left[a_{it+1}\right] = \mathbb{E}_{it}[g(L)z_{it+1}] + \mathbb{E}_{it}[h(L)x_{it+1}].$$

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

$$\mathbb{E}_{it}[g(L)z_{it+1}] = g(L)z_{it} - (1-\rho)g(0)z_{it},$$

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:

$$\begin{aligned}
\mathbb{E}_{it}\left[a_t\right] &= \frac{\lambda}{\rho}\frac{\tau_u}{\tau_\eta}\frac{(g(L)+h(L))\frac{L}{1-\rho L} - (g(\lambda)+h(\lambda))\frac{\lambda}{1-\rho\lambda}}{L-\lambda}\frac{(1-\rho L)}{(1-\lambda L)}x_{it} \\
\mathbb{E}_{it}\left[a_{t+1}\right] &= \frac{\lambda}{\rho}\frac{\tau_u}{\tau_\eta}\left(\frac{h(L)+g(L)}{(1-\lambda L)(L-\lambda)} - \frac{(h(\lambda)+g(\lambda))(1-\rho L)}{(1-\rho\lambda)(L-\lambda)(1-\lambda L)}\right)x_{it} \\
\mathbb{E}_{it}[h(L)x_{it+1}] &= \frac{\lambda}{\rho}\frac{\tau_u}{\tau_\eta}\left(\frac{h(L)}{(1-\lambda L)(L-\lambda)} - \frac{h(\lambda)(1-\rho L)}{(1-\rho\lambda)(L-\lambda)(1-\lambda L)}\right)x_{it} \\
&\quad + \frac{\lambda}{\rho}\left(\frac{h(L)(L-\rho)}{L(L-\lambda)} - \frac{h(\lambda)(\lambda-\rho)}{\lambda(L-\lambda)} - \frac{\rho}{\lambda}\frac{h(0)}{L}\right)\frac{1-\rho L}{1-\lambda L}x_{it}
\end{aligned}$$

The fixed point for $h(L)$ is

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

This leads to

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

where

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

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

Proof of Proposition 10

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

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

Clearly,

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

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

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

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

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

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

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

It remains to show that the term on the right-hand side is greater than 1. To proceed, we start with the following observation. If $\theta_1 \geq \theta_2 > 0$, and $\frac{y_2}{y_1+y_2} \geq \frac{x_2}{x_1+x_2}$, then

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

Note that

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

and

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

Based on the observation, we will show that

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

by induction. We first establish the following

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

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

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

Now suppose that

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

We want to show

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

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

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

and

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

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

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

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

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

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

Appendix 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 ξ_{it} is given by the sum of an aggregate and an idiosyncratic component, but prevent agent i from observing either ξ_{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 ξ_t as in our baseline model. Second, we let each agent observe the following noisy signal about the idiosyncratic shock ζ_{it} :

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

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

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

$$\mathbb{E}_{it}[\xi_t] = \rho\lambda \mathbb{E}_{it-1}[\xi_{t-1}] + (1 - \lambda)\, x_{it}$$

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

$$\mathbb{E}_{it}[\zeta_{it}] = \rho\widehat{\lambda}\mathbb{E}_{it-1}[\zeta_{i,t-1}] + (1-\widehat{\lambda})z_{it}$$

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:

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

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

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

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

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

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

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

From the proof of Proposition 3, it is straightforward to verify that $\vartheta > \lambda$, with $\vartheta \to \lambda$ as $\gamma \to 0$. Furthermore, following Mackowiak and Wiederholt (2009), it is natural to assume that $\widehat{\lambda}$ is lower than λ, because the typical agent is likely to allocate more attention to idiosyncratic shocks than to aggregate shocks. This guarantees a lower distortion at the micro level than at the macro level even if we abstract from GE interactions (which amounts to letting $\gamma \to 0$, or abstracting from role higher-order uncertainty).

But once such interactions are taken into account, we have that ϑ can be higher than $\widehat{\lambda}$ even if the latter exceeds λ. 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:

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

where I_t denotes the rate of investment, K_{t-1} denotes the capital stock inherited from the previous period, and Φ is a convex function. This specification gives the level of investment as a decreasing function of Tobin's Q. It also generates aggregate investment responses that are broadly in line with those predicted by more realistic, heterogeneous-agent models that account for the dynamics of investment at the firm or plant level (Caballero and Engel, 1999; Bachmann, Caballero, and Engel, 2013; Khan and Thomas, 2008).[27]

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

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

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

In the sequel, we set up a model of aggregate investment with two key features: first, the adjustment cost takes the form seen in condition (43); and second, the investments of different firms are strategic complements because of an aggregate demand externality. We then augment this model with incomplete information and show that it becomes observationally equivalent to a model in which the adjustment cost takes the form seen in condition (44). This illustrates how incomplete information can merge the gap between the different strands of the literature and help reconcile the dominant DSGE practice with the relevant microeconomic evidence on investment.

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

[27] 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.

aggregate investment is given by

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

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

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

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

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

subject to

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

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

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

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

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

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

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

subject to

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

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

optimal choice of X_{it} is

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

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

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

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

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

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

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

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

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

where

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

Incomplete Information. Suppose now that firms receive a noisy signal about the fundamental ξ_t as in Section 3. Here, we make the same simplifying assumption as in the NKPC application. We assume that firms observe current z_t, but preclude them from extracting information from it. Together with the pricing equation (45), the aggregate investment dynamics follow

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

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

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

The following is then immediate.

Proposition 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:*

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

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:

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