Dampening General Equilibrium: Incomplete Information and Bounded Rationality *

George-Marios Angeletos† **Chen Lian**‡

February 13, 2022

Abstract

We review how realistic frictions in information and/or rationality arrest general equilibrium (GE) feedbacks. In one specification, we maintain rational expectations but remove common knowledge of aggregate shocks. In another, we replace rational expectations with Level-k Thinking or a smooth variant thereof. Two other approaches, heterogeneous priors and cognitive discounting, capture the same essence while offering a gain in tractability. Relative to the full-information rational-expectation (FIRE) benchmark, all these modifications amount to attenuation of GE effects, especially in the short run. This in turn translates to either under- or over-reaction in aggregate outcomes, depending on whether GE feedbacks are positive or negative in the first place. We review a few applications, with emphasis on monetary and fiscal policy. We finally discuss how the available evidence on expectations, along with other considerations, can help guide the choice among the various alternatives, as well as between them and FIRE.

Keywords: General Equilibrium, Incomplete Information, Higher-Order Beliefs, Level-k Thinking, Reflective Equilibrium, Coordination, Beauty Contests.

JEL Codes: D50, D83, D84, E52, E62, E7.

---

*This chapter was prepared for the *Handbook of Economic Expectations* and subsumes an older paper of ours, entitled "Dampening General Equilibrium: from Micro to Macro" (Angeletos and Lian, 2017). We have benefited from the comments of various colleagues, especially those of the editors, Ruediger Bachmann, Wilbert van der Klaauw, and Giorgio Topa. Angeletos acknowledges the support of the National Science Foundation under Grant Number SES-1757198.

†MIT and NBER; angelet@mit.edu.

‡UC Berkeley and NBER; chen_lian@berkeley.edu.

1 Introduction

General equilibrium (GE) feedbacks—and the presumed ability of economic agents to understand them—are a central piece of modern macroeconomics. In some contexts (e.g., a Keynesian economy at the zero lower bound), they reinforce partial equilibrium (PE) effects, acting as "macroeconomic multipliers" that amplify exogenous shocks or raise policy effectiveness. In other contexts (e.g., competition over limited resources), they offset PE effects, helping stabilize aggregate outcomes or curtailing policy effectiveness. Either way, GE mechanisms limit the usefulness of PE intuitions and empirical works that focus on identifying PE effects.

In this paper, we review and synthesize recent research that studies how frictions in information and rationality arrest perceived and actual GE feedbacks. Our starting point is the standard modeling practice in macroeconomics, which combines a strong solution concept, Rational Expectations Equilibrium (REE), and a strong informational assumption, "full" or "complete" information Following the literature, we refer to this combination as FIRE, a shortcut for Full Information Rational Expectations. But we also emphasize that, for our purposes, this combination has a very specific meaning: the "FI" half translates to common knowledge of aggregate shocks and the "RE" half translates to common knowledge of rationality. This hints at a deep connection between all the approaches considered in this paper, a connection that we make clear as we proceed.

Our main take-home lesson can be summarized as follows: relaxing either the "FI" or the "RE" half of FIRE, at least in the four ways considered in this chapter, amounts to anchoring the expectations of the responses of others to aggregate shocks or policy shifts and, thereby, dampening the GE effect of such shocks. Turning this lesson around, we conclude that the standard practice "overestimates" the potency of GE feedbacks—and crucially, this "bias" is largest when these feedbacks are stronger to start with. This serves both as a warning for the state of the art and as a direction for future research: to obtain reliable predictions for policy counterfactuals, it is important to make sure that the theory fits not only the usual type of microeconomic evidence (e.g., higher MPCs) but also survey evidence on the adjustment of beliefs.

But what are the four departures from FIRE that support this lesson? In what ways are they similar and in what ways are they different? What does the evidence tell us about them? And how should we choose among them, or between them and FIRE? The rest of this Introduction previews answers offered to these questions.

Four theories and their common ground. The first one, studied in Section 3, maintains rational expectations but accommodates incomplete information, along the lines of Morris and Shin (2002, 2006), Woodford (2003), Nimark (2017) and Angeletos and Lian (2018). The defining property of this approach is not the noise in the observation of the underlying shocks per se but rather the absence of common knowledge about them. Basically, the key is that agents worry that others may not share the same information with them. This anchors expectations of the behavior of others to the common prior, causing all agents to behave *as if* the GE feedback was lower to start with.[1]

[1] Rational Inattention a la Sims (2003) and Sticky Information a la Mankiw and Reis (2002) can be viewed as subsets of the first approach, subject to the following qualification. These works have often emphasized the inertia in first-order beliefs, as opposed to that in higher-order beliefs. But GE attenuation is driven exclusively by the latter friction, which is what the literature cited in the main text focuses on. See also the review in Angeletos and Lian (2016).

The second approach, studied in Section 4, goes in the opposite direction, maintaining full information but replacing rational expectations with Level-k Thinking (Farhi and Werning, 2019; Iovino and Sergeyev, 2021) or two close variants of it, Reflective Equilibrium (Garcıa-Schmidt and Woodford, 2019) and Cognitive Hierarchy (Camerer, Ho, and Chong, 2004). This amounts to assuming that the typical agent believes that other agents are less sophisticated than themselves and, as a result, less able to comprehend strategic interactions and GE feedbacks. And since the best response to such a belief is to act *as if* the actual GE feedback were smaller, this approach, too, translates to GE attenuation relative to FIRE.

Our description of the above two approaches suggests a deep connection between them. Indeed, both approaches can be viewed as complementary relaxations of strong common-knowledge assumptions: the first relaxes common knowledge of the shock, the second relaxes common knowledge of rationality. Furthermore, both approaches can be understood as anchoring higher-order beliefs to a certain default point: the common prior about the underlying shock in the first approach and the level-0 belief in the second approach. One subtle difference is that the common prior in the first approach is disciplined by the objective truth, while the level-0 belief in the second approach is a free parameter, which the analyst must choose judiciously. This difference and a minor "bug" aside,[2] the two approaches are close cousins to each other.

The third approach, Heterogeneous Priors, employed by Angeletos and La'O (2009) and Angeletos and Sastry (2021) and reviewed in Section 5, builds a bridge between the above two approaches. It decouples the first approach's relevant friction—the lack of common knowledge—from the pure noise in the observation of the underlying fundamentals. It allows for a sharp translation between the depth of knowledge and depth of rationality. And it allows a gain in tractability while preserving the common essence of the other two approaches.

The last quality is shared by Gabaix (2020)'s Cognitive Discounting, the fourth and final approach reviewed in this chapter. This concept introduces a form of under-extrapolation, which ultimately produces a similar kind of GE attenuation as the other three approaches.

Applications. To simplify the exposition, most of the analysis in this chapter employs a static framework. But we show how the logic extends to dynamic settings, in which behavior is forward looking and the GE feedback runs across multiple periods. And we use this translation in Section 6 to review a number of topical applications, including why bounded rationality arrests the power of forward guidance at the zero lower bound (Angeletos and Lian, 2018; Farhi and Werning, 2019), or how it influences the relative merits of government spending and consumption taxes as means of regulating aggregate demand (Vimercati, Eichenbaum, and Guerreiro, 2021). We also discuss how to recast under this chapter's umbrella earlier applications that study business cycles and unemployment fluctuations (Angeletos and La'O, 2010; Venkateswaran, 2014), a literature that studies "competition neglect" (Camerer and Lovallo, 1999; Greenwood and Hanson, 2015), and even some recent decision-theoretic work on "narrow bracketing" (Lian, 2021).

[2] As explained in Section 4, this "bug" refers to the following issue: the plain-vanilla Level-k Thinking makes an ambiguous prediction when the GE feedback is negative, but this ambiguity goes away with an appropriate "smooth" variant of it, namely Reflective Equilibrium. We first pointed out this issue in Angeletos and Lian (2017), where we also drew an analogy to the relation between Cobweb and Tatonnement in Walrasian analysis.

Connecting the theory to evidence on expectations. Having emphasized the common ground and applicability of the various approaches under consideration, Section 7 clarifies the following three points, which relate to the available evidence on expectations.

First, all the approaches are tightly connected to under-reaction of average expectations to news. Such under-reaction is evident in surveys of expectations. This offers not only support for the common ground of these theories but also a way to quantify the degree of GE attenuation.

Second, although GE attenuation equals under-reaction of *expectations* relative to actual outcomes, it does not necessarily translate to under-reaction of *actual* outcomes relative to FIRE. To put it differently, the same friction can either dampen or amplify aggregate fluctuations relative to FIRE. Which of the two cases obtains depends on whether the GE feedback is positive or negative to start with. When it is positive (i.e., it reinforces the PE effect), GE attenuation translates to dampened volatility and sluggish adjustment. And when it is negative (i.e., when it offsets the PE effect), GE attenuation translates to amplification and overshooting.

Finally, while all the approaches under consideration share similar testable predictions about the joint dynamics of aggregate outcomes and average expectations, they make distinct testable predictions about *individual* expectations. In particular, consider the following two empirical questions: first, what is the correlation between *average* forecast errors in one period and innovations of average forecasts in previous periods; and second, what is the corresponding correlation for *individual* forecasts. Our first approach (incomplete information) allows the answer to the first question to be positive but restricts the answer to the second question to be zero, because the forecast errors of rational agents cannot be predictable on the basis of their *own* past information. By contrast, all other approaches (Level-k Thinking, Cognitive Discounting, etc) allow the answer to the second question to be positive but also restrict that the answers to the two questions be exactly the same: if belief under-reaction is due to a systematic bias, it oughts to manifest equally in the predictability of individual and average forecast errors.

The available survey evidence suggests that, at least for professional forecasters, the reality is closer to the first scenario than the second. See Coibion and Gorodnichenko (2012, 2015) for average forecasts, Bordalo et al. (2020) and Broer and Kohlhas (2021) for individual forecasts, and Angeletos, Huo, and Sastry (2021) for a synthesis. This, and a few additional considerations that we explain in due course, seem to favor incomplete information over the alternatives. That said, laboratory experiments offer ample support for Level-k Thinking and its variants,[3] including in settings that resemble macroeconomic interactions ("beauty contests").

Combining these observations with our main lesson (that all the approaches produce the same GE attenuation), we conclude that the first-order task is to move away from FIRE in the *common* direction of these approaches as opposed to running a horserace between them. Notwithstanding this point, we offer some additional guidance on the relative costs and benefits of each approach, on which ones are most tractable, and on which ones are most suitable for stationary environments (regular business cycles) versus non-stationary environments (unprecedented experiences).

[3] Nagel (1995); Stahl and Wilson (1994, 1995); Crawford, Costa-Gomes, and Iriberri (2013); Kneeland (2015); Mauersberger and Nagel (2018); Broer and Kohlhas (2021).

2 Framework

In this section, we introduce an abstract setting, which stylizes PE and GE effects. We next show how it nests a simplified but micro-founded New Keynesian economy. We finally define and characterize the FIRE benchmark, which is the point of departure for the rest of the paper.

2.1 PE and GE in a Nutshell

There is a continuum of agents $i \in [0,1]$, each making a decision $c_i \in \mathbb{R}$. Anticipating our upcoming micro-foundation, we interpret c_i as individual spending. The optimal spending depends on an individual-specific fundamental θ_i (e.g., an agent-specific preference shock or an agent-specific interest rate) and her expectation of the aggregate spending:

$$c_i = \theta_i + \alpha E_i[c], \tag{1}$$

where α is a fixed scalar, $c = \int_{i\in[0,1]} c_i di$ is aggregate spending (an endogenous random variable), and $E_i[\cdot]$ is agent i's expectation (which, for now, does not have to be rational).

An agent's fundamental has both an aggregate and an idiosyncratic component: $\theta_i = \theta + \epsilon_i$, where θ is the aggregate component and ϵ_i is the idiosyncratic one. Implicit in (1) is the assumption that every agent knows perfectly her own fundamental (θ_i), regardless of what she knows about the aggregate shock (θ) and how she forms beliefs about the actions of others and aggregate activity (c). This lets us focus on how the latter kind of beliefs are formed and in particular on how they vary with the depth of the agents' knowledge and rationality.

The aggregate counterpart of (1) is given by

$$c = \underbrace{\theta}_{\text{PE}} + \underbrace{\alpha \bar{E}[c]}_{\text{GE}}. \tag{2}$$

where $\theta = \int \theta_i di$ and $\bar{E}[c] = \int E_i[c] di$ are the cross-sectional averages of θ_i and $E_i[c]$, respectively. This allows us to decompose the change in c triggered by any change in θ into a partial equilibrium (PE) and a general equilibrium (GE) effect. More specifically, the first term in (2) represents the average PE effect, for it captures the direct impact of θ on c, holding constant the average expectations of aggregate spending; and similarly the second term in (2) captures the average GE effect, or the indirect effect of θ on c via the feedback from aggregate spending to individual spending.

Note that the GE effect depends on the scalar α, which we treat as a fixed and commonly known parameter, and on the average expectation of the response of others, which is our focal point. In game-theoretic terms, α identifies the degree and form of strategic interaction; condition (2) is indeed mathematically the same as the best response in the kind of beauty-contest games studied in Morris and Shin (2002) and Angeletos and Pavan (2007). But for our purposes, α translates to the strength and direction of the GE feedback. When $\alpha > 0$, the GE effect works in the same direction as the PE effect, helping amplify it. When instead $\alpha < 0$, the GE effect works in the opposite direction of, and attenuates, the PE effect. In both cases, a higher $|\alpha|$ means a large absolute strength of the GE effect, but the direction of the GE effect, relative to the PE effect, switches sign with α.

Throughout, we restrict $\alpha \in (-1,1)$ so that (2) is a contraction mapping. This in turn guarantees the following two key properties: first, that there will be a unique REE no matter the information structure; and second, that the level-k thinking solution converges to the unique REE as the depth of reasoning goes to infinity.[4]

2.2 Micro-foundation: A Simplified New-Keynesian Model

We now show how (1) may stylize a New Keynesian economy. To start with, suppose that there is a continuum of infinitely-lived consumers, as in Angeletos and Lian (2018), Garcıa-Schmidt and Woodford (2019), and Farhi and Werning (2019). After the familiar log-linearization, the optimal consumption of any consumer i in any period t is given by the following relation:

$$c_{i,t} = (1-\beta) a_{i,t} - \beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k E_{i,t} \left[i_{t+k} - \pi_{t+k+1} \right] \right\} + \left(1-\beta\right) \left\{ \sum_{k=0}^{+\infty} \beta^k E_{i,t} \left[y_t \right] \right\} + \sigma\beta\varrho_{i,t}, \tag{3}$$

where $\beta \in (0,1)$ is the subjective discount factor, $a_{i,t}$ is the consumer's asset position at the start of period t, y_t is aggregate income, i_t and π_{t+1} are the nominal interest rate and inflation between t and $t+1$, and $\varrho_{i,t}$ is a discount rate shock. The corresponding aggregate shock is denoted by ϱ_t and serves the usual role: it is a proxy for an aggregate demand shock.[5]

Equation (3) is basically the permanent income hypothesis, adjusted to accommodate shocks to preferences and interest rates. By aggregating it across i and imposing market clearing (more precisely, by assuming that agents themselves understand that $y_t = c_t$), we arrive at the following:

$$c_t = -\beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[i_{t+k} - \pi_{t+k+1} \right] \right\} + \left(1-\beta\right) \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[c_{t+k} \right] \right\} + \sigma\beta\varrho_t. \tag{4}$$

Under FIRE, we further have $E_{i,t}[\cdot] = \mathbb{E}_t[\cdot]$. The above can then be reduced to a representative consumer's Euler equation (plus the appropriate transversality condition). Away from FIRE, the above clarifies how aggregate spending in one period depends on expectations of monetary policy, inflation, and aggregate spending in all future periods.

Consider next the supply side of the economy. For simplicity, we assume that firms remain fully rational and fully informed, even though the consumers might not be so. In particular, we assume that firm behavior gets summarized in a standard hybrid NKPC:

$$\pi_t = \kappa c_t + \psi_{-1}\pi_{t-1} + \psi_{+1}\mathbb{E}_t \left[\pi_{t+1} \right], \tag{5}$$

[4] When instead $\alpha \notin (-1,1)$, a "pandora box" opens: with rational expectations, there can be multiple equilibria; and with Level-k Thinking, the sensitivity of the level-k outcome to a free parameter (the arbitrarily chosen level-0 outcome) explodes to infinity as $k \to \infty$. The study of these possibilities is outside the scope of this chapter.

[5] Implicit in (3) is the simplifying assumption that all agents face the same interest rate and receive the same income.

where $\psi_{-1}, \psi_{+1} \geq 0, \psi_{-1} + \psi_{+1} \leq 1$, $\pi_{-1} = 0$. Note that this embeds the full-information rational expectation operator, and that this contains both a forward-looking and a backward-looking component, as in Clarida, Gali, and Gertler (1999) and much of the applied New Keynesian literature.

Finally, consider monetary policy. To be able to reduce this infinite horizon economy to our simple static framework, we assume that monetary policy replicates flexible-price outcomes for all $t \geq 1$ (but not for $t = 0$). This translates to $c_t = i_t - \pi_{t+1} = 0$ for all $t \geq 1$. For $t = 0$, on the other hand, we impose the following Taylor rule:

$$i_0 = \phi_c c_0 + \phi_\pi \pi_0. \tag{6}$$

Note that this nests an interest-rate peg, or the ZLB, with $\phi_c = \phi_\pi = 0$. And more generally it lets ϕ_c and ϕ_π parameterize how "passive" or "accommodative" monetary policy is (with higher values for ϕ_c and ϕ_π mapping to less accommodation).

Combining the above assumptions about monetary policy with equations (4) and (5), and solving out for the inflation dynamics, we obtain the following equilibrium restriction at $t = 0$:

$$c_0 = \left(1 - \beta - \beta\sigma\left(\phi_c + \frac{\kappa}{1 - \psi_{+1}\chi}\left(\phi_\pi - \chi\right)\right)\right)\bar{E}_0\left[c_0\right] + \sigma\beta\varrho_0, \tag{7}$$

where $\chi \equiv \frac{1 - \sqrt{1 - 4\psi_{+1}\psi_{-1}}}{2\psi_{+1}} \in (0, 1)$. This is readily nested in (2) with

$$\theta \equiv \sigma\beta\rho \qquad \text{and} \qquad \alpha \equiv 1 - \beta - \beta\sigma\left(\phi_c + \frac{\kappa}{1 - \psi_{+1}\chi}\left(\phi_\pi - \chi\right)\right).$$

That is, the aggregate fundamental equals the exogenous aggregate demand shock rescaled, and the degree of strategic interaction is pinned down by parameters that regulate aggregate demand (β, σ), aggregate supply $(\kappa, \psi_{-1}, \psi_{+1})$, and monetary policy (ϕ_c, ϕ_π).

Let us first consider the case of an interest rate peg, or fully passive monetary policy, in the sense of $\phi_c = \phi_\pi = 0$. In this case, the formula for α reduces to

$$\alpha \;=\; \underbrace{1 - \beta}_{\text{Keynesian cross}} \;+\; \underbrace{\kappa\frac{\beta\sigma\chi}{1 - \psi_{+1}\chi}}_{\text{inflation-spending spiral}} \;>\; 0.$$

That is, the GE effect is positive (equivalently, the consumers' decisions are strategic complements) and comes from two parts. First, the "Keynesian cross:" a consumer's optimal consumption increases with other consumers' consumption, which in turn determines her own income. And second, the "inflation-spending spiral:" a higher aggregate spending today leads to an increase in inflation, which other things equal depresses the real interest rate and stimulates aggregate spending. These points explain why α increases both with $1 - \beta$, which herein measures the MPC or the slope of the Keynesian cross, and with κ, the slope of the Phillips curve.

As soon as monetary policy is active, in the sense that $\phi_c > 0$ and/or $\phi_\pi > 0$, a third GE effect comes into the picture. Higher aggregate spending now causes the monetary authority to raise the nominal interest rate, either directly (when $\phi_c > 0$), or indirectly via inflationary pressures (when $\phi_\pi > 0$). Other things equal, the increase in the nominal interest rate discourages private spending. It follows that this GE effect works in the opposite direction than the aforementioned two ones: it is the source of strategic substitutability among the consumers. Furthermore, when the Taylor rule is sufficiently steep (in the sense of a sufficiently high value for ϕ_c and/or ϕ_π), this negative GE effect may dominate, so that the overall GE feedback becomes negative ($\alpha < 0$). To put it differently, a sufficiently active, or hawkish, monetary policy turns consumer decisions from strategic complements to strategic substitutes.

2.3 Full Information Rational Expectations (FIRE)

Our benchmark is defined by imposing Rational Expectations Equilibrium (REE) along with full information, and indeed common knowledge of, the aggregate fundamental θ.

In any REE, equilibrium strategies, or the mappings from shocks to outcomes, are commonly known. Here, this translates to c being a commonly known function of θ. Along with the assumption that θ itself is commonly known, this guarantees that c is commonly known, too. That is

$$E_i\left[c\right] = \mathbb{E}\left[c\right] = c, \tag{8}$$

where $\mathbb{E}\left[\cdot\right]$ is the full-information rational expectations operator Using this in (2), we have:

Proposition 1. *There is a unique equilibrium under FIRE and it is such that*

$$c = \underbrace{\theta}_{PE} + \underbrace{\frac{\alpha}{1-\alpha}\theta}_{GE} = \underbrace{\frac{1}{1-\alpha}}_{GE\ multiplier}\theta. \tag{9}$$

In short, common knowledge of the underlying shock together with REE implies common knowledge of the responses of others to the shock. It is *as if* agents get together in the same room, perfectly coordinate their responses to the shock, and effectively act as a single and the same agent. It is in this sense that the GE adjustment is "perfect" under FIRE.

2.4 Beyond FIRE

The benchmark studied above combines rational expectations with full information. This is basically the same as imposing common knowledge of rationality and of information: everybody is both fully rational and fully informed, everybody knows this fact, everybody knows that everybody knows this fact, and so on. We say "basically" because of the following subtlety: in general, the outcomes that are consistent with common knowledge of rationality (a.k.a. rationalizable outcomes) are a *superset* of the REE outcomes. But as long as there is a unique rationalizable outcome, which is the case in our setting for $\alpha \in (-1, +1)$, the two sets coincide. For our purposes, the "RE" half of FIRE is therefore synonymous to common knowledge of rationality—and similarly, the "FI" half is synonymous to common knowledge of θ and of the entire payoff structure.

These observations explain the two routes we take in the next two sections. In Section 3, we depart from FIRE by maintaining rational expectations, or common knowledge of rationality, but relaxing full information, or common knowledge of θ. In Section 4, we do the converse.

3 Incomplete Information

As anticipated, in this section we relax the "FI" half of FIRE. More specifically, we maintain common knowledge of rationality but remove common knowledge of the underlying shock.

3.1 Removing Common Knowledge by Adding Idiosyncratic Noise

In our FIRE benchmark, everybody knew θ perfectly and this fact was common knowledge. We now replace this assumption with the following alternative, which captures the essence of a large literature on incomplete information and higher-order uncertainty (e.g., Morris and Shin, 2002; Woodford, 2003; Angeletos and Lian, 2016, 2018):

Assumption 1. *(i) Nature draws θ from $\mathcal{N}\left(0,\sigma_\theta^2\right)$.*

(ii) Agents have imperfect and heterogeneous information about θ. In particular, an agent's information set is given by (θ_i, s_i), where s_i is a sufficient statistic of the agent's information about θ (and about others' information thereof) and is given by

$$s_i = \theta + \epsilon_i,$$

where $\epsilon_i \sim \mathcal{N}\left(0,\sigma^2\right)$ is orthogonal to θ and i.i.d. across i.

(iii) The above facts are common knowledge.

The first part specifies the agents' common prior about θ. The second part specifies each agent's private information. The last part makes sure that the subjective priors about all exogenous shocks coincide with the objective truth, a requirement embedded in the REE concept.

A few remarks are worth making about this assumption. First, the Gaussian specification is useful but not essential: it keeps the analysis tractable but does not drive the main insights. Second, the signal s_i is meant to summarize all the information that agent i has about the underlying aggregate shock, and thereby also about the information of others and the resulting aggregate activity. Third, the idiosyncratic noise in this signal could reflect not only the absence of relevant information but also cognitive limitations in processing such information, as in the literature on rational inattention (Sims, 2003; Mackowiak and Wiederholt, 2009). Finally, the fact that the noise is idiosyncratic as opposed to aggregate amounts to introducing higher-order uncertainty: every agent is uncertain not only about θ itself but also about what others know about θ, and hence also about what others know about others, and so on. As we will explain in due course, it is the second kind of uncertainty that translates to GE attenuation.

Our task now is to study how the informational friction changes equilibrium behavior relative to FIRE. To do this, let us first fix the language: by "equilibrium" in this section we mean a linear REE, or equivalently strategy of the form $c_i = a\theta_i + bs_i$ that solves (1) under rational expectations, for some coefficients a and b. Note that this imposes not only rational expectations but also linearity, which simplifies the subsequent analysis. But the linearity restriction is without any loss: a non-linear REE does not exist in our setting, because $\alpha \in (-1,1)$ guarantees that there is a unique REE and Assumption 1 guarantees that this equilibrium is linear.

3.2 Main Lesson: GE Attenuation

From Assumption 1, we have that, for every i,

$$E_i[\theta] = \lambda s_i \qquad \text{with} \qquad \lambda \equiv \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma^2} \in (0,1]. \tag{10}$$

Note that λ is a monotone transformation of σ_θ^2/σ^2, the signal-to-noise ratio in an agent's information. And because varying the level of noise σ in $[0,+\infty)$ is equivalent to varying the coefficient λ in $[0,1)$, we can think of λ as an exogenous parameter.

Aggregating the above, we get

$$\bar{E}[\theta] = \lambda\theta. \tag{11}$$

This allows one to interpret λ as a measure of how "attentive" or "aware" agents are on average, namely how much the average expectation of θ varies with innovations in θ. But whereas this interpretation emphasizes the degree of knowledge about the underlying aggregate shock, it is best to interpret λ as a measure of how well agents can predict one another's behavior. That is, as a measure of the extent of common knowledge—or equivalently as a measure of how effectively they can coordinate with one another.

This point will become clearer in Section 3.3, where we expand on the role of higher-order beliefs, and in Section 5.1, where we decouple the relevant lack of common knowledge from noisy information about θ. But the following lemma, which follows from the above equation along with the fact that c must be proportional to θ in equilibrium, helps capture the essence.

Lemma 1. *In any equilibrium, the average rational expectation of c satisfies*

$$\bar{E}[c] = \lambda c. \tag{12}$$

This helps explain how a lower λ impedes coordination: under FIRE ($\lambda = 1$), behavior is perfectly coordinated in the sense that the typical agent expects others to respond as much as herself; and away from this benchmark, behavior is imperfectly coordinated in the sense that the typical agent worries that others will under-react.

How does this translate in dynamic settings? Let $\{\partial c_{t+\tau}/\partial\theta_t\}_{\tau=0}^{\infty}$ denote the impulse response function (IRF) of an aggregate outcome to an exogenous shock (e.g., that of aggregate output or inflation to a monetary shock) and let $\{\partial \bar{E}_{t+\tau}[c_{t+\tau}]/\partial\theta_t\}_{\tau=0}^{\infty}$ denote the corresponding IRF for the agents' average expectation of the same outcome. FIRE imposes $\partial \bar{E}_{t+\tau}[c_{t+\tau}]/\partial\theta_t = \partial c_{t+\tau}/\partial\theta_t$ for all τ: subjective and objective IRFs coincide. By introducing incomplete information, we have effectively imposed $|\partial \bar{E}_{t+\tau}[c_{t+\tau}]/\partial\theta_\tau| < |\partial c_{t+\tau}/\partial\theta_\tau|$ for all τ, or at least for τ small enough ("short run"). This dynamic translation is useful when connecting the theory to expectations evidence in Section 7.

Let us now return to our static setting and consider what Lemma 1 implies for actual behavior. By aggregating best responses, we obtain the following joint restriction on the actual outcome and the average expectation thereof:

$$c = \theta + \alpha \bar{E}[c].$$

Next, replacing $\bar{E}[c]$ from Lemma 1, we get

$$c = \theta + \lambda\alpha c. \tag{13}$$

This is the same fixed-point equation as that characterizing the FIRE outcome for an economy in which α is replaced by $\lambda\alpha$. In other words, it is *as if* the strategic interaction, or the GE feedback, has been attenuated towards 0 in proportion to the informational friction. We therefore reach the follow result, which contains the main lesson of this section:

Proposition 2. *There is a unique equilibrium such that*

$$c = \underbrace{\theta}_{PE} + \underbrace{\frac{\alpha\lambda}{1-\alpha\lambda}\theta}_{GE} = \underbrace{\frac{1}{1-\alpha\lambda}}_{GE\ multiplier}\theta, \tag{14}$$

By the same token, the following properties are true:

- *The aggregate outcome is the same as in a "twin" FIRE economy in which the GE parameter α in equatio (2) has been replaced by $\lambda\alpha$.*
- *No matter whether $\alpha < 0$ or $\alpha > 0$, the absolute size of the GE effect is reduced, and the more so the smaller the degree of common knowledge, as measured by λ.*
- *When the GE feedback is positive ($\alpha > 0$), the friction translates to* under-reaction *of the actual outcomes to the underlying innovations, relative to FIRE; and when the GE feedback is negative ($\alpha < 0$), it translates to* over-reaction *relative to FIRE.*

By varying λ between 0 and 1, we can thus span all the values between the PE effect θ and the FIRE outcome $\frac{1}{1-\alpha}\theta$. For λ close to zero (meaning a sufficiently large departure from common knowledge), the total impact of the aggregate shock θ is arbitrarily close to its PE effect. But as λ increases (meaning a higher degree of common knowledge), the GE effect becomes more and more important. Importantly, all these properties hold true no matter whether the GE effect amplifies the PE effect ($\alpha > 0$) or dampens the PE effect ($\alpha < 0$). But how this translates in terms of the responsiveness of the aggregate outcome to the underlying innovations naturally depends on the direction of the GE feedback: the friction translates to *under*-reaction (relative to FIRE) when the feedback is positive and to *over*-reaction when the feedback is negative.

The last point underscores how the friction under consideration is distinct from pure inattention, sparsity, or adjustment costs. These forces amount to attenuation of the PE effect and translate to under-reaction regardless of α (indeed even if $\alpha = 0$). The friction under consideration, instead, amounts to attenuation of the GE effect and its translation to observables depends crucially on the sign and magnitude of α.

We will illustrate how this basic insight sheds new light on topical policy questions in Section 6. But let us offer here a first hint of what it may mean in applied terms. The HANK literature has emphasized that liquidity constraints and other frictions contribute to high MPCs and thereby to larger GE multipliers for monetary and fiscal policy. This was stylized in the example of Section 2, in which α was indeed related to the MPC. In this context, the above result suggests that the Keynesian multiplier may be muted in the short run, when it is likely that there is little common knowledge about the underlying innovations and one another's behavior (λ is small to start with), but may kick in over time as agents become more aware of the underlying innovations and one another's behavior (λ gets larger over time due to learning, as discussed in Section 5).

3.3 From Rational Expectations to Higher-Order Beliefs (HOBs)

We now "dig deeper" to show how the insight developed above relates to a certain property of higher-order beliefs (the beliefs of the beliefs of others). For this purpose, we will borrow heavily from game theory, and we will echo Morris and Shin (1998, 2002, 2006) and Woodford (2003). But this does not mean that the above insight, or the accommodation of incomplete information more generally, requires that the economic agents themselves be "game theorists," i.e., engage in strategic or higher-order reasoning. Instead, the typical consumer or firm can be a "statistician" in the tradition of Muth and Lucas: they only need to have a statistical model of how c and θ comove. The REE concept—with or without full information—imposes that the statistical model in people's minds is consistent with actual behavior, but is deliberately agnostic about how exactly agents learn this model in the first place.

We emphasize this point because it separates the approach taken here from that taken in the next section: by design, Level-k Thinking and its variants *require* that agents engage in deductive or higher-order reasoning. But for now, we put this point aside and proceed to translate our GE attenuation result from a property of the REE fixed point to the following property of higher-order beliefs: under incomplete information, higher-order beliefs (the beliefs of the beliefs of others) tend to react less to innovations than first-order beliefs (the beliefs of the fundamentals themselves).

To illustrate this property and its likely robustness, we proceed as follows. Go back to the best-response condition (1) but refrain, at least for the time being, from any assumption about what an agent knows about the underlying aggregate shock and the information, rationality or behavior of others. Aggregating this condition across i, we obtain

$$c = \theta + \alpha \bar{E}[c].$$

Note that this condition encapsulates individual rationality (optimality) for arbitrary expectations of c. Next, suppose that every agent believes that *other* agents are individually rational. Then, every agent can infer that the above condition holds, which implies that the average expectations of c must satisfy

$$\bar{E}[c] = \bar{E}[\theta] + \alpha\bar{E}[\bar{E}[\theta]] = \bar{E}[\theta] + \alpha\bar{E}^2[c],$$

where $\bar{E}^2[\cdot] \equiv \bar{E}[\bar{E}[\cdot]]$ denotes second-order beliefs. Iterating ad infinitum, we can express that the average expectations of c in terms of the higher-order beliefs of fundamentals:

$$\bar{E}[c] = \left(\sum_{h=1}^{\infty} \alpha^{h-1}\bar{E}^h\left[\bar{\theta}\right]\right), \tag{15}$$

where $\bar{E}^h[\theta] \equiv \bar{E}[\bar{E}^{h-1}[\theta]]$ for all $h \geq 1$, and with the convention $\bar{E}^0[\theta] \equiv \theta$.

The following remarks help clarify what we have done:

1. In the above construction, "iterating ad infinitum" amounts to imposing common knowledge of rationality: the first iteration requires that agents know that others are rational, the second iteration requires that agents know that others are rational, and so on.

2. Had we stopped this process at some finite order K, instead of letting $K \to \infty$, we would have obtained

$$\bar{E}[c] = \left(\sum_{h=1}^{K} \alpha^{h-1}\bar{E}^h\left[\bar{\theta}\right]\right) + \alpha^K\bar{E}^{K+1}[c].$$

   This restricts the subjective expectations of the aggregate outcome only up to a point, letting $\bar{E}^{K+1}[c]$ indeterminate (a "free variable"). As it will become clear in the next section, Level-K thinking amounts to (i) stopping at some finite K and (ii) making a specific assumption about the value of $\bar{E}^{K+1}[c]$. Here instead we dispense of this "free variable" by letting $K \to \infty$, or by imposing "infinite rationality" in the precise sense of common knowledge of rationality.

3. In general, common knowledge of rationality is weaker than REE. But as mentioned earlier, the two notions are equal in our context, thanks to the restriction $\alpha \in (-1, 1)$. It follows that condition (15) can be read as follows: it describes the rational expectations of c for *arbitrary* information structures, including full information.

With the last point in mind, let us now explain how condition (15) specializes in the knife-edge of full information—and what changes away from this benchmark.

With full information, beliefs of all orders collapse to the underlying fundamental:

$$\bar{E}^h[\theta] = \bar{E}^{h-1}[\theta] = ... = \bar{E}^1[\theta] = \theta. \tag{16}$$

This in turn implies that $\bar{E}[c] = c$, which can be read interchangeably as follows: (i) the expected response of c to any innovation in θ coincides with the actual response; (ii) agents face no uncertainty about one another's responses; and (iii) agents perfectly coordinate their responses.

When information is noisy, these properties remain true as long as all the noise is aggregate, that is, as long as all the information is "public" (in the precise sense of common knowledge). But they cease to hold as soon as once information is private, or the noise is idiosyncratic. And it is precisely the latter scenario that represents a "real" departure from FIRE.

Consider in particular the information structure assumed in the previous analysis. With this structure, we have that

$$\bar{E}[\theta] = (1-\lambda)\mu_\theta + \lambda\theta,$$

where μ_θ corresponds to the prior mean (itself normalized to 0 for convenience), θ corresponds to the shock, and $\lambda \in (0,1)$ parameterizes the precision of the available private information, or the degree of common knowledge. The above implies that

$$\bar{E}^2[\theta] = (1-\lambda)\mu_\theta + \lambda\bar{E}[\theta] = (1-\lambda^2)\mu_\theta + \lambda^2\theta$$

and, by induction,

$$\bar{E}^h[\theta] = (1-\lambda^h)\mu_\theta + \lambda^h\theta$$

for all $h \geq 1$. We therefore see that, in response to any innovation in θ, beliefs of higher-order move less than beliefs of lower-order—to put it differently, higher-order beliefs are more anchored to the prior and less responsive to news.

This property is a robust implication of incomplete information and has been extensively discussed in the literature (Morris and Shin, 1998, 2002, 2003; Abreu and Brunnermeier, 2003; Woodford, 2003; Angeletos and Lian, 2016). The point here was to show how this property translates to GE attenuation, as well as to pave the way for some additional insights in the sequel. But keep in mind our earlier point: although this game-theoretic translation offers important insight, in the present approach agents themselves could still be "statisticians."

4 Bounded Rationality

In this section, we recoup the "FI" half of FIRE but relax its "RE" half. In particular, we start by adapting the plain-vanilla version of Level-k Thinking from Farhi and Werning (2019) to our setting. For this version, GE attenuation obtains necessarily when $\alpha > 0$ but not when $\alpha < 0$. We then explain how this bug is resolved by two "smooth" variants, Reflective Equilibrium (Garcıa-Schmidt and Woodford, 2019) and Cognitive Hierarchy (Camerer, Ho, and Chong, 2004).

4.1 Level-k Thinking

Level-k Thinking, also known as Limited-Depth Thinking, was developed in the experimental literature; see Nagel (1995) and Stahl and Wilson (1994, 1995) for early contributions and Crawford, Costa-Gomes, and Iriberri (2013) for a survey. It was later imported into macroeconomics by Garcıa-Schmidt and Woodford (2019), albeit with a twist that we explain in the next subsection, and by Farhi and Werning (2019), which we follow here.

The concept is defined in a recursive manner. Let $k \in \{0, 1, \cdots\}$. For $k = 0$, behavior is exogenously specified; and for any $k \geq 1$, behavior is given by the best response to the belief that all other agents are level-$(k-1)$ thinkers. A higher k represents "deeper thinking" in the sense of incorporating more rounds of strategic interaction.

Putting this into math, let $\hat{c}^k$ and c^k denote, respectively, the conjectured (subjective) and the actual (objective) value of the aggregate outcome, c, when all agents are level-k thinkers. For each $k \geq 1$, the actual outcome c^k is the best response to the conjecture that the outcome will be $\hat{c}^k$; that is, $c^k = \theta + \alpha \hat{c}^k$. The conjecture in turn is given by the actual outcome for level-$(k-1)$ thinkers; that is, $\hat{c}^k = c^{k-1}$. Starting for an arbitrary $\hat{c}^1 = c^0$, we then have that the level-1 outcome is given by $c^1 \equiv \theta + \alpha c^0$ and, by induction, for any $k \geq 1$, the level-k outcome is given by

$$c^k = \theta + \alpha c^{k-1} = \sum_{j=0}^{k-1} \alpha^j \theta + \alpha^k c^0. \tag{17}$$

Because $\alpha \in (-1, 1)$, we immediately have that $c^k \to \frac{1}{1-\alpha}\theta = c^{\text{FIRE}}$ as $k \to \infty$ regardless of c^0. That is, as the depth of thinking grows without bound ("agents become infinitely rational"), the aggregate outcome converges to its FIRE counterpart regardless of the analyst's specification of c^0. But for any $k < \infty$, c^k remains anchored to c^0. This begs the question of what's the "right" parameterization of c^0.

This choice matters greatly for our purposes. So far, the value of c^0, and indeed the entire cross-sectional distribution of c_i^0, is a "free variable." In the experimental literature, it is common to let c_i^0 be i.i.d. across i, drawn from a uniform distribution with full support over the set of feasible actions. The alternative proposed by Farhi and Werning (2019) is to set $c_i^0 = 0$ for all i, or equivalently

$$c^1 = \theta.$$

This translates as follows: "level-1 thinkers understand PE but not GE" (i.e., they behave as if there were no GE feedback). It follows that level-2 thinkers play $c^2 = \theta + \alpha\theta$, level-3 thinkers play $c^3 = \theta + (\alpha + \alpha^2)\theta$, and so on.

These iterations resemble adaptive expectations, except that the adjustment in beliefs happens instantaneously, not over real time, and on the basis of *hypothetical* outcomes (i.e., conjectures about the behavior of others), not on the basis of the observation of *actual* past outcomes. Furthermore, under the assumption that $c^1 = \theta$, these iterations can be interpreted as rounds of GE feedbacks: level-1 agents understand only PE; level-2 agents incorporate a single round of the GE feedback; level-3 agents incorporate two rounds; and so on.

Under this prism, FIRE translates to "infinite rounds of GE feedback", and bounded rationality translates to "finite rounds of GE feedback." This in turn suggests that Level-k Thinking should generate a similar form of GE attenuation as incomplete information, with lower sophistication (lower k) here translating to less information (lower λ) there. As it turns out, this logic is valid when the GE feedback is positive ($\alpha > 0$), but has to be qualified when the feedback is negative ($\alpha < 0$).

Proposition 3. *Suppose that the economy consists of a continuum of level-k agents, for some $k \geq 1$. Suppose* *further that level-0 agents are anchored to the prior, i.e.,* $c^0 = 0$*. Then, the aggregate outcome is given by*

$$c = \underbrace{\theta}_{PE} + \underbrace{\frac{\alpha - \alpha^k}{1-\alpha}}_{GE}\theta = \underbrace{\frac{1-\alpha^k}{1-\alpha}}_{GE\ multiplier}\theta.$$

Furthermore, letting GE^k *denote the GE effect in this economy and* GE^{FIRE} *its FIRE counterpart,*[6] *the following is true:*

- *When* $\alpha > 0$, $|GE^k|$ *is strictly increasing in* k *and bounded from above by* $|GE^{FIRE}|$.
- *When instead* $\alpha < 0$, *the above statement holds only for* k *odd. For* k *even, the opposite is true:* $|GE^k|$ *is strictly decreasing in* k *and bounded from below by* $|GE^{FIRE}|$.

In other words, GE attenuation obtains in two cases: when the GE feedback is positive and the depth of thinking k is any finite number; or when the feedback is negative and k is odd. In both of these cases, level-k thinkers under-estimate the response of others, not only relative to the FIRE counterpart but also relative to their *own* behavior; they therefore behave *as if* the GE feedback was smaller, i.e., as in a FIRE economy in which α is compressed towards zero. But the opposite is true when the GE feedback is negative and k is even. In this case, level-k thinkers *over*-estimate the responses of others and, as a result, behave as if the GE feedback is larger.

We illustrate what's going on in Figure 1. The left panel features strategic complementarity ($\alpha > 0$), and the right one features strategic substitutability ($\alpha < 0$). In either panel, the solid blue lines represent the agent's decision rule (1) before and after the shock, namely for $\theta = 0$ and $\theta = 1$, respectively. $c^0 = 0$ identifies the level-0 outcome, which itself coincides with the pre-shock REE outcome. c^∞ identifies the post-shock REE outcome. The level-1 response, or equivalently the PE effect of the shock, is captured by the vertical shift from point X to point Y. The frictionless GE effect ("infinite rounds") is captured by the shift from Y to Z. Note that the GE effect amplifies the PE effect when $\alpha > 0$ and offsets it when $\alpha < 0$. Finally, the dashed arrows represent the rounds of level-k thinking: level-1 is captured by the shift from X to Y (because level-1 coincides with PE); level-2 is captured by the shift from Y to Y'; and so on. It is then evident that Level-k Thinking captures incomplete GE adjustment when $\alpha > 0$, but opens the door to "GE overshooting" when $\alpha < 0$ and k is even.

Like others (Garcıa-Schmidt and Woodford, 2019; Angeletos and Sastry, 2021), we view this overshooting as a bug, not a feature of Level-k Thinking. In Sections 4.3 below, we thus explain how to fix this bug by a few appropriate amendments of this concept. With these fixes in place, and provided that one continues to assume that the level-1 outcome coincides with the PE effect, bounded rationality produces essentially the same kind of GE attenuation as that produced by incomplete information.[7] The deeper connection between the two approaches can indeed be understood by revisiting the discussion of higher-order beliefs from Section 3.3.

[6] That is, $GE^k \equiv \frac{\alpha-\alpha^k}{1-\alpha}\theta$ and $GE^{\text{FIRE}} \equiv \frac{\alpha}{1-\alpha}\theta$.

[7] As a matter of fact, the above "bug" could have emerged even in our first approach, had we allowed for a more convoluted information structure.

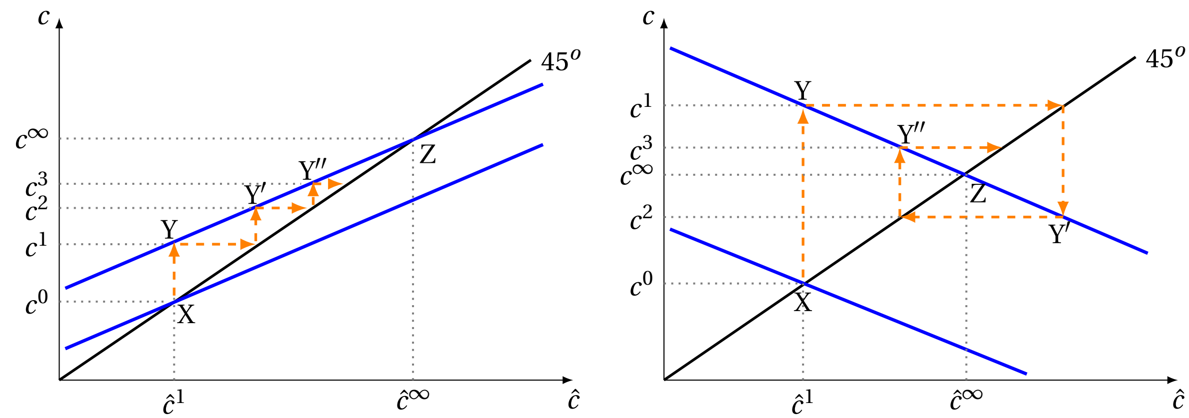

Figure 1: Level-k Thinking

4.2 Parenthesis: Back to Higher-order Beliefs

Recall that under rational expectations, and with or without full information, the aggregate outcome can be expressed as the following infinite sum:

$$c = \sum_{h=0}^{\infty} \alpha^h \bar{E}^h[\theta],$$

with $\bar{E}^0[\theta] \equiv \theta$, $\bar{E}^1[\theta] \equiv \bar{E}[\theta]$, $\bar{E}^2[\theta] \equiv \bar{E}[\bar{E}[\theta]]$ and so on. Under FIRE, we have $\bar{E}^h[\theta] = \theta$ for all h and the above specializes to

$$c = c^{\text{FIRE}} = \sum_{h=0}^{\infty} \alpha^h \theta.$$

When we drop the "FI" half of FIRE along the lines of Section 3, we get $\bar{E}^h[\theta] = \lambda^h \theta$ for all $h \geq 0$ and therefore

$$c = c^{\text{inco}} = \sum_{h=0}^{\infty} \alpha^h \lambda^h \theta,$$

where $\lambda \in (0,1)$ parameterizes the degree of common knowledge. Relative to FIRE, this amounts to discounting higher-order beliefs at an exponential rate. Finally, when we replace the "RE" half of FIRE with Level-k Thinking, we get

$$c = c^k = \sum_{h=0}^{k-1} \alpha^h \theta.$$

Relative to FIRE, this amounts to keeping the beliefs of order $h \leq k-1$ intact and truncating the beliefs of order $h \geq k$.

Under this prism, both incomplete information and level-k thinking translate to "arresting the response of higher-order beliefs to the underlying shock." In the first case, this arresting takes the form of exponential discounting over h. In the second case, it takes the form of a truncation, or a step function over h. The second approach is less "smooth" than the first one, but the essence is the same, especially if the analyst feels comfortable to exchange the notion of less sophistication (lower k) for the notion of less information (lower λ).

4.3 Reflective Equilibrium and Cognitive Hierarchy

The last paragraph hints that a "smooth" variant of Level-k Thinking may avoid the type of overshooting identified in the right panel of Figure 1. Two such variants are readily available in the form of Reflective Equilibrium and Cognitive Hierarchy. The former was introduced by Garcıa-Schmidt and Woodford (2019) in the context of the New Keynesian model; the latter was developed by Camerer, Ho, and Chong (2004) and has been widely used in the behavioral and experimental literature (Crawford, Costa-Gomes, and Iriberri, 2013).

Let us start with Reflective Equilibrium. The basic idea is to preserve the essence of Level-k Thinking but recast the discrete iterations described in Section 4.1 with a continuous analogue. In particular, the depth of thinking k is now treated as a continuous variable in $(0,\infty)$ and the cognitive process is modeled as follows. Let $\hat{c}(k)$ and $c(k)$ denote, respectively, the conjectured and actual values of the aggregate outcome, as functions of the depth of thinking.[8] The actual outcome is given by the average best response to the conjecture:

$$c(k) = \theta + \alpha\hat{c}(k). \tag{18}$$

And the conjecture is given by as the solution to the following ODE:

$$\frac{d\hat{c}(h)}{dh} = c(h) - \hat{c}(h) \quad \forall h \in [0,k] \tag{19}$$

with the initial condition $\hat{c}(0) = 0$.

Equation (19) requires that, as we move from step h to step $h + dh$, the conjecture is adjusted upwards if the conjecture at step h fell short of the corresponding outcome, and downwards otherwise. This captures the same essence as Level-k Thinking; but whereas that concept specifies these steps as discrete and requires that the conjecture in each step is given by the outcome in the previous step, Reflective Equilibrium allows the steps to be infinitesimally small and the adjustment to be continuous. As result, the bug is gone: the conjecture never overshoots the FIRE outcome.

One can visualize this with the help of Figure 1 from Section 4.1. There (with plain-vanilla Level-k Thinking), as we varied k in $\{0,1,2,...\}$, the conjectures $\hat{c}^k$ and the outcomes c^k were jumping around in discrete steps. Here (with Reflective Equilibrium), as we vary k continuously in $(0,\infty)$, the pair $(\hat{c}^k, c^k)$ varies continuously along the segment between points Y and Z, guaranteeing that the outcome never overshoots relative to FIRE.[9] Translating this in terms of GE attenuation, we have the following result:

Proposition 4 (Reflective Equilibrium)**.** *Suppose that agents are boundedly rational in the sense of Reflective Equilibrium, as specified above. There exists a function* $\delta : [0,+\infty) \times (-1,1) \to [0,1)$ *such that, when the depth of*

[8] These objects are also functions of θ and α, but these dependencies are suppressed to ease the notation.

[9] This brings to mind two older concepts developed to describe the off-equilibrium adjustment in Walrasian economies, namely Cobweb dynamics and Tatonnement. The first featured discrete adjustments, like Level-k Thinking. The second featured continuous adjustment, like Reflective Equilibrium. For a formalization of this parallel, see Angeletos and Lian (2017). Here, it suffices to say that Reflective Equilibrium avoids the overshooting property of Level-K Thinking for basically the same reason that Tatonnement avoided a similar overshooting property of Cobweb dynamics.

thinking is k, the aggregate outcome is given by

$$c = \underbrace{\theta}_{PE} + \underbrace{\frac{\delta(k,\alpha)\alpha}{1-\delta(k,\alpha)\alpha}\theta}_{GE} = \underbrace{\frac{1}{1-\delta(k,\alpha)\alpha}}_{GE\ multiplier}\theta.$$

Furthermore, $\delta(k,\alpha)$ is strictly increasing in k, starting from 0 at $k=0$ and converging to 1 as $k\to\infty$, regardless of the sign of α.

Reflective equilibrium therefore not only fixes the bug of the plain vanilla Level-k Thinking but also provides a tight connection to incomplete information: the two approaches generate the same kind of GE attenuation, with $\delta(k,\alpha)$ taking the place of λ. More succinctly, a lower depth of thinking under the present approach translates to a lower degree of common knowledge under our first approach (and vice versa).

Similar points apply to Cognitive Hierarchy. Motivated by the heterogeneity of observed choices and reported beliefs in experiments, this concept modifies the plain-vanilla Level-k Thinking in two ways. First, actual heterogeneity is introduced in the depth of thinking. And second, agents are allowed to understand the presence of such heterogeneity, albeit up to a point: an agent still thinks that she is at least as sophisticated as anybody else. With a judicious parameterization of the objective heterogeneity and of the subjective beliefs thereof, the distance between the average conjecture $\hat{c}$ and the corresponding best-response outcome c is a continuous function of a single parameter τ, which can be interpreted as the average sophistication in the population. The details are spelled out in Appendix A but the bottom line is this: relative to Reflective Equilibrium, this concept allows for heterogeneity but produces basically the same predictions for aggregate beliefs and aggregate outcomes.

5 Additional Variants and Dynamic Extensions

In this section, we review the third approach, by Angeletos and La'O (2009) and Angeletos and Sastry (2021), which captures the common essence of the above two approaches while also being easier to use. We then discuss how the insights extend to dynamic settings, in particular the roles played by learning and forward looking behavior. We finally explain why the fourth approach, by Gabaix (2020), can be viewed as a close cousin of the other three approaches.

5.1 A Bridge: Heterogeneous Priors, or Shallow Reasoning

We now discuss the approach taken by Angeletos and La'O (2009) and Angeletos and Sastry (2021). This approach mimics *both* incomplete information and Level-k Thinking (and its variants), while also being highly tractable and easily adaptable to richer settings.

The basic idea in Angeletos and La'O (2009) is to use heterogeneous priors to decouple the degree of common knowledge, or the extent of the coordination friction, from the precision of the available information. Think of this as revisiting the analysis in Section 3 and letting λ be a separate parameter from σ, the level of noise. Angeletos and Sastry (2021) further simplify that approach, highlight its tight connection to Level-k Thinking, and conclude that it captures interchangeably bounded depth of knowledge and bounded depth of rationality. Accordingly, they refer to this approach as "shallow reasoning." But one could also call it "anchored beliefs" or "lack of common knowledge of others' attentiveness."

Here, we follow the version of Angeletos and Sastry (2021), which amounts to the following joint assumption on the agents' actual information about θ and on their subjective beliefs about the information of others:

Assumption 2. *All agents know θ and all agents are individually rational in the sense that they play according to the best response condition (1). However, this fact is not common knowledge. Instead, every agent believes (i) that the above fact is true for only a fraction λ of other agents, and that these agents share the same beliefs with herself; and (ii) that the remaining fraction $1-\lambda$ of agents are uninformed about θ and that they therefore play $c_i = 0$.*

This approach resembles incomplete information in that it preserves common knowledge of rationality but drops common knowledge of the shock. But the latter property is no longer accomplished by adding non-trivial noise in the agents' own observation of θ. Instead, it is accomplished by introducing heterogeneous priors about one another's information. In a nutshell, all agents may *themselves* be perfectly informed about θ and nevertheless worry that other agents are *not*.

At the same time, this approach resembles Level-k Thinking and its two variants in the following regard: there, agents thought that they were smarter than the others; here, they think they are better informed than the others. Furthermore, it is straightforward to translate such beliefs about the informativeness of others to beliefs about the rationality of others: the proposed concept is indeed equivalent, in terms of both beliefs and behavior, to a variant of Cognitive Hierarchy in which the actual distribution of types has mass 1 at $k=\infty$, while the corresponding subjective beliefs have mass $1-\lambda$ at $k=\infty$ and λ at $k=0$.

In a nutshell, one can interpret λ interchangeably as the depth of knowledge and the depth of reasoning. Crucially, though, the present approach avoids the iterative nature of Level-k Thinking and instead preserves the fixed-point nature of REE: agents are statisticians, albeit with a mis-specified model about the information of others. To put it differently, the present approach boils down to Bayesian Nash Equilibrium with heterogeneous priors.

Let us see what this means in practice. Let c be the actual outcome, or equivalently the action of the typical agent. Since this agent believes that only a fraction λ of the population are like herself (i.e., have the same information and play the same action), and that the remaining fraction is inactive, this agent's expectation of the aggregate outcome is given by

$$E_i[c] = (1-\lambda)\,0 + \lambda c = \lambda c.$$

The best response to this belief is

$$c_i = \theta + \alpha\lambda c. \tag{20}$$

Aggregating and solving for c, we get the following result:

Proposition 5. *Suppose that agents are shallow reasoners in the sense described in Assumption 2. The aggregate outcome is given by*

$$c = \underbrace{\theta}_{PE} + \underbrace{\frac{\alpha\lambda}{1-\alpha\lambda}\theta}_{GE} = \underbrace{\frac{1}{1-\alpha\lambda}}_{GE\ multiplier}\theta.$$

Note that this is the exact same characterization as that provided in Proposition 2 for our incomplete-information model. The difference, as anticipated in the beginning of this section, is that λ is now disconnected from how much agents know about the underlying shock and instead is more directly tied to the degree of common knowledge. At the same time, this approach preserves the bounded-rationality flavor of Level-k Thinking, for the reasons already explained. We conclude that the present approach bridges the other two approaches—and as evident from above analysis, it is more tractable than both of them.

5.2 Dynamics I: Learning

So far the analysis has been confined to a static framework in order to deliver the key insights in a sharp and transparent manner. In applications, however, it is often central to allow for two different kinds of dynamics. The first (discussed in this subsection) has to do with learning over time, or the dynamics of beliefs. The second (discussed in the next subsection) has to do with intertemporal payoff interdependences, which make the decisions forward-looking actions.

To isolate the role of learning, we continue to assume that the GE feedback is static but let decisions be repeated over time, on the basis of time-varying fundamentals and beliefs. More specifically, we index time by $t \in \{0, 1, \cdots\}$ and specify the optimal behavior at each t as follows:

$$c_{i,t} = \theta_{i,t} + \alpha E_{i,t}[c_t], \tag{21}$$

where $c_{i,t}$ is the action (e.g., spending) of agent i in period t, $\theta_{i,t}$ is her payoff-relevant fundamental, c_t and θ_t are the corresponding aggregates, and $E_{i,t}$ is the agent's expectation in period t.

This is basically the same kind of dynamic setting as that considered in Woodford (2003) and Angeletos and La'O (2010). In the former, $c_{i,t}$ maps to the price set by a firm, $\theta_{i,t}$ to its nominal marginal cost, and $E_{i,t}[c_t]$ to its expectation of the aggregate price level; and in the latter, $c_{i,t}$ translates to the production of a firm, $\theta_{i,t}$ to its productivity, and $E_{i,t}[c_t]$ to its expectations of aggregate output (or aggregate demand). Putting aside the possible interpretations, the key observation here is that (21) rules out any feedback from the outcomes at $\tau \neq t$ to optimal behavior at t, which retains the static nature of the GE feedback and isolates the adjustment in beliefs as the only interesting dynamic element.

To illustrate, consider an aggregate shock to θ_0. With time passing, the agent receives more information about θ_0 (e.g., her own fundamental $\theta_{i,t}$) and gradually learns about θ_0. Because of such learning, the effective degree of information friction about θ_0 may decrease (equivalently, the effective degree of λ may increase) with the passage of time. For example, in the environment where the GE effect works in the same direction as the PE effect, such learning will let the GE multiplier increase with the age of the shock.

The basic idea that gradual learning can help generate sluggish responses is, of course, not new; a version of it can be found, inter alia, in Sims (2003) and Mankiw and Reis (2002). But the subtler point first recognized by Woodford (2003), further pushed by Angeletos and Huo (2021), and echoed above is that the sluggishness in the aggregate outcomes implied by any given informational friction is larger the larger the macroeconomic complementarity or the GE feedback. This in turn helps explain why the form of sluggishness documented here is closely connected to the notion of GE attenuation, as well as why it may be more pronounced at the macro level than at the micro level (Angeletos and Huo, 2021), or more succinctly why macro hump shapes can be consistent micro jumps (Auclert, Rognlie, and Straub, 2020).

Another important question is how exactly learning, or the adjustment in beliefs, takes place. Learning under full rationality, or Bayesian learning, is an important benchmark but not the only option. We refer the reader to Baley and Veldkamp (2021) for a more detailed treatment of Bayesian learning with incomplete information and to Carroll and Wang (2021) for an alternative that emphasizes the spreading of beliefs from one agent to another.

5.3 Dynamics II: Forward-Looking Behavior

In (2), the GE feedback is static; but modern macroeconomics emphasize forward looking behavior and intertemporal GE interactions. To illustrate, let us revisit the New Keynesian economy of Section 2.2. There, we assumed that the monetary policy replicated flexible-price outcomes for all $t \geq 1$. That assumption let us fix $E_{i,0}[c_t] = c_t = 0$ for all $t \geq 1$ and all i, muting the roles forward-looking behavior and dynamic GE interactions, and allowing us to reduce the entire economy to a static game at $t = 0$. But once we relax that assumption, the economy translates to a dynamic economy featuring not only a different decision at each t but also a GE feedback across t.

To see this more clearly, suppose that the discount rate shock follows an AR(1) process with persistence ρ and that prices are perfectly rigid, so that $\pi_t = 0$ for all t. The optimal consumption of an agent can then be expressed as follows:

$$c_{i,t} = -\beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k E_{i,t}\left[i_{t+k}\right] \right\} + \left(1-\beta\right) \left\{ \sum_{k=0}^{+\infty} \beta^k E_{i,t}\left[c_{t+k}\right] \right\} + \frac{\beta\sigma}{1-\rho}\varrho_{i,t}. \tag{22}$$

Next, assume the following Taylor rule for monetary policy:

$$i_t = \phi c_t, \tag{23}$$

where $\phi > 0$ parameterizes how aggressively monetary policy responds to fluctuations in aggregate demand. Replacing the above into (22) and aggregating, we reach the following joint restriction between aggregate demand and expectations thereof:

$$c_t = \theta_t + \alpha \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t [c_{t+k}] \right\}, \tag{24}$$

where $\theta_t \equiv \frac{\beta\sigma}{1-\rho}\varrho_t$ and $\alpha \equiv 1-\beta-\beta\phi$ can be interpreted as, respectively, the aggregate demand shock and the degree of dynamic macroeconomy complementarity. We then see clearly the following properties. First, the mathematical structure and interpretation of (24) is similar to the static analogue used in the rest of our chapter, except that now the outcome (aggregate spending) in any given period depends on the average expectations of the outcome not only in the same period but also in all future periods. And second, this dependence, or the overall dynamic GE feedback, is pinned down by the difference between the slope of the Keynesian cross $(1-\beta)$ and the slope of the Taylor rule (ϕ).

The intuition for the last property is basically the same as that in Section 2.2: a larger MPC contributes towards a positive GE feedback, or strategic complementarity $(\alpha > 0)$; a less accommodative monetary policy contributes towards a negative GE feedback, or strategic substitutability $(\alpha < 0)$. What changes is that these feedbacks, or the strategic interactions, run across time.

Although this change is relatively minor at the conceptual level, it can raise non-trivial challenges at the computational level. This is particularly true for the incomplete-information approach; see Nimark (2008, 2017), Huo and Takayama (2021), and Angeletos and Huo (2021) for both a detailed discussion of the challenge and some ways to tame them. Level-k Thinking is significantly more manageable, as illustrated in Farhi and Werning (2019), because it shuts down any adjustment in beliefs. But if one is willing to abstract from learning, one can further maximize tractability by employing the approach in Angeletos and Sastry (2021). In the present context, for example, this approach translates to replacing the average subjective expectations $\bar{E}_t [c_{t+k}]$ in equation (24) with $\lambda\mathbb{E}_t [c_{t+k}]$, where $\mathbb{E}_t[\cdot]$ is the FIRE expectation operator and $\lambda \in (0,1)$ measures the depth of knowledge or rationality. That is, equation (24) becomes

$$c_t = \theta_t + \lambda\alpha \left\{ \sum_{k=0}^{+\infty} \beta^k \mathbb{E}_t [c_{t+k}] \right\}. \tag{25}$$

Two properties are then evident: the insight about GE attenuation goes through; and the above approach is just as tractable as the FIRE benchmark.[10]

5.4 Cognitive Discounting

We now turn to the approach taken by Gabaix (2020). This departs from FIRE by assuming that the *perceived* law of motion of the aggregate state of the economy is less persistent than the *actual* law of motion. Suppose in

[10] One can further show that the exact same equation obtains with rational expectations and incomplete information as long as there is no learning, or at least as long as agents do not anticipate that there will be learning in the future. This point is proved in Angeletos and Lian (2018) and makes clear that the key simplification of Angeletos and Sastry (2021) is to abstract from learning, thus also bypassing the technical challenge of keeping track of the dynamics in the hierarchy of beliefs (Townsend, 1983).

particular that a model's true law motion is given by

$$x_t = Ax_{t-1} + B\epsilon_t$$

where A and B are matrices, x_t is the state variable, and ϵ_t is the innovation. Gabaix (2020) assumes that the "behavioral agent" incorrectly perceives the law of motion to be $x_t = m \cdot (Ax_{t-1} + B\epsilon_t)$, for some exogenous scalar $m \in (0,1)$ that is interpreted as a "cognitive" discount factor.

To illustrate how this translates to GE attenuation, consider the New Keynesian example of the previous section. As long as the exogenous shock follows an AR(1) process, the FIRE outcome also follows an AR(1): it satisfies $c_t = \rho c_{t-1} + b^{\text{FIRE}}\epsilon_t$, where ϵ_t stands for the innovation in θ_t, ρ is the actual persistence of θ_t, and $b^{\text{FIRE}} > 0$ is a scalar pinned down by σ, β, ρ and ϕ. One can guess and verify that a similar property is true in the presence of cognitive discounting $(m < 1)$:

$$c_t = \rho c_{t-1} + b\epsilon_t,$$

albeit for some $b \neq b^{\text{FIRE}}$. With this guess, the *perceived* law of motion becomes

$$c_t = \hat{\rho} c_{t-1} + \hat{b}\epsilon_t,$$

with $\hat{\rho} \equiv m\rho$ and $\hat{b} \equiv mb$. Together with the fact that future innovations are unpredictable, this implies that the subjective expectations satisfy $\bar{E}_t[c_{t+k}] = \hat{\rho}^k c_t = m^k \rho^k c_t$, while the objective, rational, expectations satisfy $\mathbb{E}_t[c_{t+k}] = \rho^k c_t$. It follows that

$$\bar{E}_t[c_{t+k}] = m^k \mathbb{E}_t[c_{t+k}], \tag{26}$$

which makes clear the following point: the present variant shares with all the other approaches in this chapter the property that the average expectations of the actions of others, or of the aggregate outcome, respond less to the underlying innovations than the actual outcome.

One subtle difference is that this kind of under-reaction is predicted to increase with the horizon of forecasts, while this is not the case for other approaches in this chapter. With Angeletos and Sastry (2021), for example, we have $\bar{E}_t[c_{t+k}] = \lambda \mathbb{E}_t[c_{t+k}]$, where $\lambda \in (0,1)$ is invariant to the forecast horizon k. Furthermore, when agents are incompletely informed but rational and there is learning over time, we effectively have that $\bar{E}_t[c_{t+k}] = \mu_k \mathbb{E}_t[c_{t+k}]$, where $\mu_k \in (0,1)$ is an *increasing* function of k, opposite to (26). The intuition is that, when agents expect both themselves and others to learn, they also expect the friction in coordination to ease (formally, to reach higher levels of common knowledge). Notwithstanding this subtle difference, cognitive discounting is a close cousin of the other approaches in this chapter. And it shares with Angeletos and Sastry (2021) both a high degree of tractability and the treatment of the agents as "statisticians" rather than as "game theorists."

6 Applications

In this section, we illustrate how the abstract insights translate to specific applications.

6.1 Forward Guidance at the Zero Lower Bound

Consider a New Keynesian economy in a liquidity-trap recession, so that the central bank's hands are tied by the Zero Lower Bound (ZLB) on interest rates. The central bank may still attempt to stimulate aggregate demand by committing to keep interest rates low after the economy exits the trap and the ZLB ceases to bind. The basic New Keynesian model predicts that this kind of forward guidance can be extremely effective. But this does not appear to have been the case during the Great Recession, so this prediction became known as the "forward guidance puzzle" (Del Negro, Giannoni, and Patterson, 2015; McKay, Nakamura, and Steinsson, 2016).

To understand this puzzle and our proposed resolution thereof, let us revisit the environment behind equation (24), pick an arbitrary $T \geq 2$ and make the following assumptions about monetary policy: $i_t = 0$ for $t \leq T-1$ (this captures the role of the ZLB); i_T is chosen and announced at $t = 0$ (this captures the role of forward guidance); and finally for all $t \geq T+1$, i_t is such that the flexible-price outcome, or $c_t = 0$, is henceforth obtained (this captures "return to normalcy"). The question of interest boils down to the following: how does c_0 varies with i_T, and in particular how does this response depend on T, the horizon of forward guidance?

To ease the exposition, let us address this question while abstracting from the role of inflation: prices are perfectly rigid, so that $\pi_t = 0$ for all t. Together with the above assumptions about monetary policy, this implies that equation (4) reduces to

$$c_t = -\beta\sigma\left\{\beta^{T-t}\bar{E}_t\left[i_T\right]\right\} + \left(1-\beta\right)\left\{\sum_{k=0}^{T-t}\beta^k \bar{E}_t\left[c_{t+k}\right]\right\}, \tag{27}$$

for all $t \leq T$ (and $c_t = 0$ thereafter). With FIRE, we further have $\bar{E}_t\left[i_T\right] = i_T$ and $\bar{E}_t\left[c_{t+k}\right] = c_{t+k}$, so that the response of c_t to the news about i_T can be calculated as follows:

$$\frac{\partial c_t}{\partial i_T} = \underbrace{-\beta\sigma\beta^{T-t}}_{\text{PE}} + \underbrace{\left(1-\beta\right)\left\{\sum_{k=0}^{T-t}\beta^k \frac{\partial c_{t+k}}{\partial i_T}\right\}}_{\text{GE}}. \tag{28}$$

From this expression, we can see that, although the PE effect of i_T on c_t *decreases* with T, reflecting discounting at the individual level, the GE effect actually *increases* with T, because a large horizon translates to more periods over which the Keynesian feedback between income and spending can operate (or equivalently to a larger response in expected permanent income). Furthermore, these two conflicting effects happen to offset each other perfectly in the representative-agent version of the New Keynesian model, so that the following is true here:

Proposition 6. *Under FIRE,* $\frac{\partial c_0}{\partial i_T} = -\sigma$ *for all T.*

More succinctly, forward guidance about interest rates far in the future is predicted to be as effective as conventional monetary policy, which seems counterfactual. And if we relax the simplifying assumption of perfectly rigid prices, the puzzle gets even worse: under FIRE, the response of c_0 to any news about i_T increases with the horizon T. This is because allowing for non-rigid prices amounts to introducing an additional GE feedback between aggregate spending and inflation, which only adds to the income-spending multiplier.

The bottom line is that the puzzle is driven by GE feedbacks. But then it is an immediate corollary of this chapter's main theme that this puzzle is resolved, or at least eased, if we accommodate some bounds on the agents' depth of knowledge or rationality. For example, if we adopt the approach of Angeletos and Sastry (2021), we have that $\bar{E}_t[i_T] = i_T$ but $\bar{E}_t[c_{t+k}] = \lambda c_{t+k}$, so that (28) changes to the following:

$$\frac{\partial c_t}{\partial i_T} = \underbrace{-\beta\sigma\beta^{T-t}}_{\text{PE}} + \underbrace{\lambda\left(1-\beta\right)\left\{\sum_{k=0}^{T-t}\beta^k \frac{\partial c_{t+k}}{\partial i_T}\right\}}_{\text{GE}}. \tag{29}$$

That is, the GE effect is effectively discounted by a factor equal to λ, helping the PE logic to "reign supreme." Indeed, solving the above fixed point problem, one can easily see that the power of forward guidance is attenuated and that the degree of attenuation increases not only with the friction in knowledge/rationality but also with the horizon T.

Proposition 7. *Suppose that there is a bound on agents' depth of knowledge or rationality, in the form of Angeletos and Sastry (2021). Then,*

$$\frac{\partial c_0}{\partial i_T} = \phi\left(\lambda, T\right) \left.\frac{\partial c_0}{\partial i_T}\right|_{FIRE}, \tag{30}$$

where $\phi(\lambda, T)$ is strictly bounded between 0 *and* 1, *strictly increasing in λ for any $T \geq 2$, and strictly decreasing in T for any $\lambda \in [0, 1)$.*

Given what has been shown in this chapter, it should be no surprise that this result readily translates to rational expectations once information is incomplete (Angeletos and Lian, 2018), to Level-k Thinking (Farhi and Werning, 2019), and to Cognitive Discounting Gabaix (2020). As one moves across these approaches, the math changes a bit, but the essence is the same.

The result as stated above directly extends to the full version of the New Keynesian model, which allows prices to adjust slowly. For a textbook calibration of that model and a modest value of the novel friction ($\lambda = .75$), Angeletos and Lian (2018) find that the attenuation effect can be quite sizable: at a horizon of 5 years ($T = 20$), the power of forward guidance could be as low as *one tenth* of its FIRE counterpart.[11] This number is of course sensitive not only to λ but also to all other parameters that regulate the potency of GE feedbacks, such as the slope of the Keynesian cross and the slope of the Phillips curve. By adopting a textbook calibration, Angeletos and Lian (2018) overestimates the latter but also underestimates the former. Accordingly, Farhi and

[11] $\lambda = .75$ can be read as follows: every agent expects the typical other agent to be inattentive and not respond with probability 25%. Calibrating λ to the type evidence considered in Coibion and Gorodnichenko (2012, 2015) could have justified an even smaller value.

Werning (2019) have emphasized that the documented attenuation effect is likely to be more pronounced in HANK economies that properly account for liquidity constraints and high MPCs than in a RANK economy. This circles back to our earlier point about the synergies between the type of frictions considered in this chapter and those traditionally considered in the literatures on incomplete markets and heterogeneity. For additional examples of this point, see Angeletos and Huo (2021) and Auclert, Rognlie, and Straub (2020).

6.2 Fiscal Policy

We now turn to fiscal policy. In particular, we consider the question studied in Vimercati, Eichenbaum, and Guerreiro (2021): which of the two fiscal tools, an increase in government spending or a reduction in consumption taxes, is more effective in stimulating aggregate demand when the economy is at the ZLB. Both types of fiscal stimuli are assumed to be financed by an increase in future lump sum taxes, so neither of them involves a labor distortion. The key issue, instead, is how expectations respond to the one versus the other.

Because consumption taxes enter individual consumption in the exact same way as real interest rates, our preceding insights about forward guidance translate as follows: bounded rationality, or lack of common knowledge, reduces the stimulative power of future tax cuts but not so much that of current tax cuts, because the latter have a more pronounced, direct, PE effect. And since current taxes, unlike current monetary policy, are not constrained by the ZLB, this policy tool remains basically as effective as in FIRE—which is one key lesson in Vimercati, Eichenbaum, and Guerreiro (2021).

But what about the other tool, government spending? Suppose, for simplicity, that the initial level of public debt is zero, so that the average asset position of the households is also zero. By aggregating the individual consumption functions, we now get that, unsurprisingly, aggregate demand depends on real interest rates and *disposable* permanent income:

$$c_t = -\beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[i_{t+k} - \pi_{t+k+1} \right] \right\} + (1-\beta) \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[y_{t+k} - T_{t+k} \right] \right\},$$

where T_t is the lump sum tax in period t. By the intertemporal budget constraint of the government, on the other hand, the present discounted value of taxes must equal the present discounted value of government spending. To the extent that consumers understand this basic fact, the above equation is replaced by

$$c_t = -\beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[i_{t+k} - \pi_{t+k+1} \right] \right\} + (1-\beta) \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[y_{t+k} - G_{t+k} \right] \right\},$$

where G_t is the government spending at t. Finally, by market clearing, $y_t - G_t = c_t$. And provided that consumers understand this property, too, the above reduces to

$$c_t = -\beta\sigma \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[i_{t+k} - \pi_{t+k+1} \right] \right\} + (1-\beta) \left\{ \sum_{k=0}^{+\infty} \beta^k \bar{E}_t \left[c_{t+k} \right] \right\}, \tag{31}$$

which is *exactly* the same as equation (4) before.

This immediately implies that the following property extends from FIRE to all the alternatives studied in this chapter:

Proposition 8. *Holding constant the average expectations of real interest rates, aggregate private spending is invariant to the path of government spending, regardless of the depth of knowledge and rationality.*

By the same token, if prices are rigid and the monetary authority does not change interest rates in response to fiscal policy, the fiscal multiplier is one, again regardless of the depth of knowledge and rationality. This is indeed a corollary of Proposition 8 from Angeletos and Lian (2018).[12]

This lesson is different from that reported in Vimercati, Eichenbaum, and Guerreiro (2021): that paper argues that, holding real rates constant, government spending crowds out private spending once agents are boundedly rational in the sense of Level-k Thinking. The apparent contradiction is due to a subtle difference in the specification of level-1 beliefs. Let us explain.

In Section 4.1, and following Farhi and Werning (2019), we assumed that level-1 agents expect others not to respond to the underlying shock ($\hat{c}^1 = c^0 = 0$) and therefore act as if the shock had only a PE effect ($c^1 = \theta$). In the present context, this translates to the belief that the path of aggregate private spending is invariant to the path of government spending, or $\hat{c}_t^1 = 0$ for all t. By contrast, in Vimercati, Eichenbaum, and Guerreiro (2021) the level-1 belief is that the path of aggregate *output* is invariant to the path of government spending, which translates to $\hat{c}_t^1 = -G_t$. That is, in response to an increase in government spending, level-1 consumers now expect other consumers to cut down their spending, which in turn motivates them to cut down their own spending despite the fact that government spending has no direct PE effect. As a result, the fiscal multiplier is lower than its FIRE counterpart for level-1 thinkers; and by induction, it remains so for any finite level of reasoning.

Which specification of level-1 beliefs is more plausible? The following remark offers some theoretical guidance: only the first specification is consistent with rational expectations and lack of common knowledge of the path of government spending. But the question is ultimately an empirical one: one can imagine conducting surveys that elicit the relevant subjective beliefs, whether in the messy real world or in a controlled experiment.

6.3 Other Applications

We now discuss how the insights can find application outside the New Keynesian context—or even outside macroeconomics.

To start with, consider the incomplete-information RBC model of Angeletos and La'O (2010). This boils down to a game among the firms, with the best response of firm i given by

$$y_{i,t} = (1-\alpha)\chi a_{i,t} + \alpha E_{i,t}\left[y_t\right] \tag{32}$$

[12] That proposition, and similarly an extension in Vimercati, Eichenbaum, and Guerreiro (2021), allows for prices to be non-rigid. This does not upset Proposition 8 above, but changes the prediction about fiscal multipliers, because government spending can now feed into inflation. For careful treatments of this case, see Section 6 of Angeletos and Lian (2018) and Section 3 of Vimercati, Eichenbaum, and Guerreiro (2021).

where $y_{i,t}$ is its output, $a_{i,t}$ is its TFP, $y_t = \int y_{i,t} di$ is aggregate output, and $\chi > 0$ and $\alpha < 1$ are fixed scalars. In this context, the empirically relevant case is $\alpha > 0$;[13] and this chapter's lesson about GE attenuation translates to under-reaction of aggregate output to aggregate TFP shocks. For some parameterizations, this attenuation can be sufficiently large that aggregate employment changes little, or even decreases, in response to a positive TFP shock, helping reconcile the RBC model with the evidence in Gali (1999).

Another flexible-price application is found in Venkateswaran (2014). Under our lens, that paper can be summarized as follows: incomplete information attenuates the GE feedback inside the Diamond-Mortessen-Pissarides model; and because that feedback is negative, this translates to larger unemployment fluctuations than under FIRE, helping ease the so-called Shimer puzzle.

Next, and moving from macroeconomics to IO, consider the the works of Camerer and Lovallo (1999) and Greenwood and Hanson (2015) on "competition neglect." The latter concept refers to the idea that firms may fail to recognize that a favorable demand or supply shock will lead additional competitors to enter the industry, and as a result may cause the firms to over-expand in the short run. Under the prism of our analysis, the concept of competition neglect is the same as the concept of GE attenuation. And in this particular context, GE attenuation translates to over-reaction, or the kind of over-investment documented in Greenwood and Hanson (2015), simply because the GE feedback works in the opposite direction to the PE effect.

Finally, let us turn to single-agent decision problems, the focus of Lian (2021). That paper shows how the theoretical tools and insights developed for *inter*-personal coordination frictions (as in Section 3) can be then used to study *intra*-personal decisions and to provide a new way to think about narrow bracketing and mental accounting. This is formalized by assigning different decisions to different selves, who in turn share the same objective (the decision maker's utility) but make their respective decisions on the basis of different, non-nested, information.

7 Discussion: Similarities, Differences, and Empirical Backdrop

The key similarities of the various approaches considered in this chapter can be summarized as follows: they all can be understood as relaxations of one or another common-knowledge assumption; relative to FIRE, they all amount to discounting higher-order beliefs, or equivalently to arresting the expected response of others to aggregate shocks; and they all ultimately translate to GE attenuation. We next expand on some of their key differences, of both conceptual and empirical nature. And we try to offer further guidance on when and how to use these approaches.

[13]In principle, α could be of either sign because it encapsulates two conflicting GE feedbacks: when other firms produce more, aggregate income goes up, which feeds to higher demand and higher returns for each firm; but they also push up the cost of labor or other inputs, which contributes in the opposite direction.

7.1 Key Differences

Let us focus on our main two approaches, incomplete information and Level-K Thinking (or its variants). One key difference has already been emphasized: while Level-k Thinking requires that agents be "game theorists" (to reason recursively), incomplete information allows them to be "statisticians," in the tradition of Lucas (1972) and Muth (1961).

This cuts both ways. In the macroeconomic context, it does not seem a priori compelling—at least to us—to model the consumers or the firms as strategic players that actively engages in the kind of iterative reasoning articulated in Section 4. Instead, it seems more natural to model their thinking about the economy as a statistical model, which is what the incomplete information approach allows. On the other hand, the RE half of this approach, namely the assumption that the subjective statistical model in people's mind is the same as the objective truth, is hard to defend in the context of unusual circumstances, such as the Great Recession and the ZLB on monetary policy.

Another difference regards the scope for learning. Like other departures from rational expectations, Level-k Thinking begs the question of what agents do in the face of systematic discrepancies between their expectations and the realized outcomes. In the experimental literature, the estimated depth of thinking often increases with the rounds of play. Translating this into our context may suggest that the GE attenuation diminishes with the passage of time. But there is a problem: for this idea to be operational in a stationary environment, it would have to be that agents are at once "shallow thinkers" vis-a-vis recent shocks and "deep thinkers" vis-a-vis old shocks. Clearly, this makes no sense in the context of Level-k Thinking, at least insofar one takes this concept literally: an agent's reasoning capacity and her understanding of the economy may vary with her education and experience, but may not fluctuate with the age of recurring aggregate shocks.

Incomplete information (and its sibling, Rational Inattention) naturally avoids these problems. First, by maintaining rational expectations, this approach makes sure that agents are never presented with the conundrum of systematic discrepancies between their expectations and the realized outcome. And second, by allowing agents to accumulate more information about any given shock with the passage of time since the shock occurs, it naturally predicts that the degree of common knowledge is specific to each shock and increases with the age of the shock. More succinctly, incomplete information predicts that agents behave *as if* they are once shallow thinkers vis-a-vis recent shocks and deep thinkers vis-a-vis old shocks.

Last but not least, the two approaches share the same predictions about the time-series properties of *average* forecasts but have distinct predictions about the corresponding properties of *individual* forecasts. The same applies to cognitive discounting. We elaborate on this point in the next section and use it to highlight the combination of the survey evidence reported in Coibion and Gorodnichenko (2012, 2015), Bordalo et al. (2020), Angeletos, Huo, and Sastry (2021) and elsewhere seems to favor incomplete information over all the alternatives.

All considered, we would offer the following guidelines: Incomplete Information, its siblings, Rational Inattention and Sticky Information, and the bounded-rationality alternatives of Angeletos and Sastry (2021) and

Gabaix (2020) seem best suited for the study of "normal" business cycles and "systematic" policy rules. Furthermore, if one is willing to abstract from learning, the latter two approaches capture the essence while buying a lot of tractability. But we agree with Farhi and Werning (2019) that Level-k Thinking or some variants thereof may make sense in the context of unprecedented experiences, as well as that it offers another, plausible, and indeed complementary, way to model bounded rationality.

At the end of the day, it really depends on what the analyst wishes to achieve. If the main goal is to capture and quantify the degree of GE attenuation, the various approaches reviewed in this chapter are basically interchangeable. The key discipline from the data will then come in the form of measuring how much average expectations comove with realized outcomes, conditionally on a given policy shift or some other aggregate shocks. In our static, one-shock framework, this means measuring the ratio between the empirical counterparts of $\partial \bar{E}[c]/\partial\theta$ and $\partial c/\partial\theta$; and in a more realistic, dynamic, multi-shock context, it translates to the ratio between the empirical counterparts of $\{\partial \bar{E}_t[c_{t+\tau}]/\partial\epsilon_t\}_{\tau=0}^{\infty}$ and $\{\partial \mathbb{E}_t[c_{t+\tau}]/\partial\epsilon_t\}\}_{\tau=0}^{\infty}$.[14] Whether this ratio is matched in the theory by choosing a low degree of common knowledge, λ, a low depth of thinking, k, or a low cognitive discount factor, m, looks like a "detail" with regard to the aforementioned goal. But if the goal shifts to understanding and quantifying the adjustment of expectations over time, a theory that accommodates some form of learning seems necessary, and then incomplete information seems to be in a better position than the alternatives.

7.2 Empirical Backdrop: Under-reaction in Average vs Individual Forecasts

We now expand on the point made above regarding the available evidence on average and individual forecasts.

Under all the approaches considered in this chapter, GE attenuation is a direct consequence of the fact that agents underestimate one another's responses. A common testable prediction of all these approaches is therefore the following: relative to actual outcomes such as inflation and GDP, the *average* expectations of these outcomes should under-react in response to identified shocks, or to past revisions of expectations.

This prediction and some close variants of it find empirical support in surveys of expectations (Coibion and Gorodnichenko, 2012, 2015; Angeletos, Huo, and Sastry, 2021). The available evidence on average forecasts therefore clearly favors these approaches over FIRE. But such evidence does not help distinguish between the various approaches. To accomplish this, we must look into *individual* forecasts.

If the first approach (incomplete information but rational expectations) is right, there can be under-reaction in average forecasts but not in individual forecasts: an agents' forecast errors should be unpredictable by her own past forecast revisions, no matter what her information is. And if instead any of the other approaches is right, there should be the same level of under-reaction at the aggregate and the individual level.

What does the evidence say about this? While there is, as already noted, evidence of under-reaction in average forecasts, that's not the case for individual forecasts: if anything, individual forecasts appear to over-react a bit (Bordalo et al., 2020). This seems to favor incomplete information over the alternative: among the two general ways of relaxing FIRE and getting GE attenuation, that of dropping the "FI" half but maintaining

[14] In the latter case, one must read "ratio" as a measure of distance between the two IRFs. See, for example, Angeletos, Huo, and Sastry (2021).

the “RE” half is more consistent with the available evidence on expectations.

That said, the relevant available evidence on individual forecasts regards exclusively analysts and professional economists, leaving open the question of what goes on with regular consumers and firms. Furthermore, Level-k Thinking has found ample support in laboratory experiments, and so does the broader idea of “anchored” beliefs. All in all, we therefore view that the evidence is supportive of relaxing both the “FI” and the “RE” halves of the standard practice, and we iterate our recommendation from the previous section. Researchers may feel comfortable choosing their modeling strategy among the presented alternatives on the basis of the following two criteria: whether the main theme is GE attenuation or the dynamic adjustment in expectations; and which approach is most flexible and tractable.

8 Conclusion

GE effects, such as competitive pressures or Keynesian multipliers, are at the core of both theoretical and quantitative macroeconomics: they limit the usefulness of PE intuitions, and they often account for a significant fraction of the overall quantitative effects of aggregate shocks or policy interventions.

In this chapter, we have operationalized the notion that the standard modeling practice, which assumes FIRE, tends to overstate the potency of such GE effects. We did so by showing that plausible relaxations of either the “FI” or the “RE” half of this benchmark amount to anchoring the beliefs about the behavior of others and arresting GE feedback loops.

We illustrated this through specific examples. But the broader message can be summarized as follows. Perhaps there is more to “simplistic” PE intuitions than the macroeconomic theorist is trained to see. Perhaps the prevailing structural interpretations of the data are misleading because we have “overestimated” GE effects. And perhaps the policy maker should put more emphasis on policies that work through salient PE effects, as opposed to multi-layer GE effects, if she wishes to steer the economy in the short run.

Our framework was deliberately simple and abstract. This helped deliver the key insights in a transparent and flexible manner. It allowed us to shed light not only on the common ground of the various approaches but also on some key differences among them. And it let us offer the following guideline for how future researchers may choose among the different approaches: all approaches are equally good for the purpose of capturing the apparent under-reaction in average expectations and the resulting GE attenuation; the approaches of Angeletos and Sastry (2021) and Gabaix (2020) accomplish this goal with maximal tractability; incomplete information gains ground in stationary settings with learning; and Level-k Thinking becomes more appealing in non-stationary or unfamiliar environments.

Finally, we would like to emphasize that the type of frictions studied here also helps accommodate “animal spirits” within a unique equilibrium framework. Angeletos and La’O (2013), Benhabib, Wang, and Wen (2015) and Huo and Takayama (2021) accomplished this by maintaining rational expectations but introducing incomplete information, making room for extrinsic fluctuations in higher-order beliefs. Furthermore, the resulting business fluctuations have the following quality: they resemble those of aggregate demand shocks, and capture salient features of the data, without a strict reliance on nominal rigidity and accommodative monetary policy.[15] But one can engineer basically the same fluctuations by replacing incomplete information with Level-k Thinking and by allowing the level-0 belief to be random, or by making an analogous modification of the other two approaches discussed in this chapter. Therefore, the key question is once again an empirical one: above, we asked how much expectations move with fundamentals; here, we must ask how much expectations move for extrinsic reasons.

All in all, the approaches reviewed in this chapter, and related work on rational inattention, hold promise to offer a better account of the joint stochastic behavior of observed expectations and observed outcomes, and a more reliable framework for macroeconomic policy evaluation. Recent quantitative explorations, such as those found in Angeletos and Huo (2021), Auclert, Rognlie, and Straub (2020), Chahrour and Ulbricht (2021), Mackowiak and Wiederholt (2015), Melosi (2017), and Qiu (2019), are encouraging; but the verdict is still open. See also Milani (2021).

[15] On this matter, see also Angeletos, Collard, and Dellas (2018), Angeletos, Collard, and Dellas (2020), Chahrour, Nimark, and Pitschner (2021), Ilut and Saijo (2020), and Chahrour and Gaballo (2021).

Appendix A: Cognitive Hierarchy

As mentioned in the main text, Cognitive Hierarchy (Camerer, Ho, and Chong, 2004) modifies the plain-vanilla version of Level-k Thinking described in Section 4.1 in two regards. First, it introduces heterogeneity in sophistication: an agent's own level k is given by an i.i.d. draw from some distribution with full support over $\{0,1,2,\cdots\}$. And second, agents understand that there is such heterogeneity, albeit up to a point: a level-k agent believes that any other agent is a level-h thinker, where h is drawn in an i.i.d. manner from a distribution with full support over $h \in \{0,1,\cdots,k-1\}$.

Note that this amounts to assuming that, no matter how low her k might be, an agent always believes that she is at the top of the pecking order: every agent believes that she is the smartest person in the world. But this belief is incorrect for any finite k: in actuality, there is always someone smarter. Also note that, so far, this allows for a countable infinity of free parameters (in addition to c_0): the analyst must exogenously specify, not only the objective distribution of k in the population (denoted below by f), but also the subjective counterpart for each level k (denoted below by g_k).

To escape this embarrassment of riches, Camerer, Ho, and Chong (2004) make the following additional assumptions. First, the actual, or objective, distribution of types is assumed to be a Poisson distribution over $k \in \{0,1,...\}$. In our continuum-player context, this assumption translates as follows: the actual mass of level-k agents in the population is $f(k) \equiv e^{-\tau}\tau^k/k$, where $\tau \in \mathbb{R}_+$ is the Poisson parameter. And second, the subjective belief of a level-k thinker about the distribution of types in the population is assumed to be a truncated version of the objective one: for any $k \geq 1$, a level-k agent believes that the mass of level-h agents is given by

$$g_k(h) = \frac{f(h)}{\sum_{l=0}^{k-1} f(l)} = \frac{e^{-\tau}\tau^h/h}{\sum_{l=0}^{k-1} e^{-\tau}\tau^l/l}$$

for $h \leq k-1$ and by $g_k(h) = 0$ for $h \geq k$.

With these assumptions in place, there are only two free parameters left: c^0 and τ. This is the same number of free parameters as in the plain-vanilla Level-k Thinking and Reflective Equilibrium alike, with τ here taking the place of k in the previous analysis. Indeed, note that a higher value of τ shifts both the actual distribution of types and all the subjective counterparts towards higher types, which verifies that τ has basically the same meaning as k before: it parameterizes the average degree of sophistication in the economy.

We close the concept by assuming $c^0 = 0$. As before, this translates to "level-0 thinkers are unresponsive," or "level-1 understand PE but not GE;" and it leaves τ as the only free parameter.

The outcome can then be solved recursively, following a similar procedure as in Section 4.1, although with more tedious algebra because of all the heterogeneity. Let c^k denote the behavior of a level-k agent and $\hat{c}^k$ her subjective expectation of aggregate activity. Level-0 and level-1 thinkers behave as in Section 4.1: $c^0 = 0$ and $c^1 = \theta$. But this is no longer the case for higher-level thinkers. In particular, because level-2 thinkers believe that the population consists of a mass $g_2(0)$ of level-0 thinkers and a mass $g_2(1)$ of level-0 thinkers, their conjecture about c is given by

$$\hat{c}^2 = g_2(0)c^0 + g_2(1)c^1 = g_2(1)\theta,$$

and their behavior in turn is given by $c^2 = \theta + \alpha\hat{c}^2$. Iterating, we get that, for any $k \geq 2$, a level-k thinker expects the aggregate outcome to be

$$\hat{c}^k = \sum_{h=0}^{k-1} g_k(h)\, c^h$$

and best responds by playing $c^k = \theta + \alpha\hat{c}^k$. The *actual* aggregate outcome is finally given by the *objective* average behavior in the population:

$$c \equiv \sum_{k=0}^{+\infty} f(k)\, c^k.$$

At the end of the day, the above procedure maps any triplet of values for θ, τ and α to c, that is, for the average expectation and the actual outcome. Because c^k are proportional to θ at each step of this procedure, it is evident that c must be proportional to θ. The next result characterizes this mapping in terms of the realized outcome; a similar result applies for the average expectation.

Proposition 9 (Cognitive Hierarchy)**.** *Suppose that agents are boundedly rational in the sense of cognitive hierarchy, as specified above. There exists a function $\kappa : \mathbb{R}^+ \times (-1,1) \to \mathbb{R}$ such that the aggregate outcome is given by*

$$c = \underbrace{\theta}_{PE} + \underbrace{\kappa(\tau,\alpha)\frac{\alpha}{1-\alpha}\theta}_{GE}.$$

In the above, τ parameterizes the average level of sophistication, $\frac{\alpha}{1-\alpha}\theta$ corresponds to the GE feedback under FIRE, and $\kappa(\tau,\alpha)$ measures the fraction of this effect that survives under cognitive hierarchy.

Unfortunately, a closed-form characterization of κ is not possible, due to the complexity of the iterative procedure behind it. However, because κ is a function of only two parameters, it is easy to do an extensive numerical exploration of it. On the basis of it, the following property appears to be true: for any $\alpha \in (-1,+1)$, $\kappa(\tau,\alpha)$ is bounded between 0 and 1, is strictly increasing in τ, converges monotonically to 1 as $\tau \to \infty$. In other words, the bug of the plain-vanilla level-k thinking is gone, and the GE feedback is necessarily arrested regardless of its sign.[16]

Appendix B: Proofs

Proof of Proposition 1. This follows from equations (1), (2), and (8).

Proof of Lemma 1. This follows from the argument in the main text.

[16] Having said that, it is useful to remind that we have restricted $\alpha \in (-1,1)$, which was necessary and sufficient for the corresponding REE to be the unique rationalizable outcome (equivalently, for the iteration of best responses to define a contraction mapping). If instead we let $\alpha < -1$, $\kappa(\tau,\alpha)$ ceases to be monotone in τ, that is, GE overshooting can still happen. By contrast, this possibility is absent from the incomplete-information model of Section 3, even when $\alpha < -1$.

Proof of Proposition 2. This follows from equation (13).

Proof of Proposition 3. This follows from equation (17).

Proof of Proposition 4. From equations (18) and (19), we know

$$\frac{d\hat{c}(h)}{dh} = \theta - (1-\alpha)\,\hat{c}(h) \quad \forall h \in [0,k],$$

which has solution

$$\hat{c}(h) = \frac{\theta}{1-\alpha}\left(1 - e^{-(1-\alpha)h}\right).$$

From equation (18), we know

$$\begin{aligned} c(k) &= \theta + \alpha \frac{\theta}{1-\alpha}\left(1 - e^{-(1-\alpha)k}\right) \\ &= \theta + \frac{\delta(k,\alpha)\alpha}{1-\delta(k,\alpha)\alpha}\theta, \end{aligned}$$

where

$$\delta(k,\alpha) = \frac{\frac{1}{1-\alpha}\left(1 - e^{-(1-\alpha)k}\right)}{1 + \frac{\alpha}{1-\alpha}\left(1 - e^{-(1-\alpha)k}\right)}.$$

This proves Proposition 4.

Proof of Proposition 5. This follows from equation (20).

Proof of Proposition 6. Observe that $\frac{\partial c_T}{\partial i_T} = -\sigma$. Proposition 6 then follows from backward induction based on equation (28).

Proof of Proposition 7. From equation (29), we know that

$$\frac{\partial c_T}{\partial i_T} = -\frac{\beta\sigma}{1-\lambda\left(1-\beta\right)}$$

and, for all $t \leq T-1$,

$$\frac{\partial c_t}{\partial i_T} = \beta \frac{\partial c_{t+1}}{\partial i_T} + \lambda\left(1-\beta\right)\frac{\partial c_t}{\partial i_T} = \frac{\beta}{1-\lambda\left(1-\beta\right)}\frac{\partial c_{t+1}}{\partial i_T}.$$

As a result,

$$\frac{\partial c_0}{\partial i_T} = -\left(\frac{\beta}{1-\lambda\left(1-\beta\right)}\right)^T \sigma.$$

It follows that the result holds with

$$\phi(\lambda, T) = \left(\frac{\beta}{1-\lambda\left(1-\beta\right)}\right)^T.$$

Proof of Proposition 8. This follows directly from equation (31).

Proof of Proposition 9. This follows follows directly from the derivation in the text.

References

Abreu, Dilip and Markus K Brunnermeier. (2003). "Bubbles and Crashes." *Econometrica* 71 (1):173–204.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. (2018). "Quantifying confidence." *Econometrica* 86 (5):1689–1726.

———. (2020). "Business-cycle anatomy." *American Economic Review* 110 (10):3030–70.

Angeletos, George-Marios and Zhen Huo. (2021). "Myopia and Anchoring." *American Economic Review* 111 (4):1166–1200.

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry. (2021). "Imperfect macroeconomic expectations: Evidence and theory." *NBER Macroeconomics Annual* 35 (1):1–86.

Angeletos, George-Marios and Jennifer La'O. (2009). "Incomplete Information, Higher-order Beliefs and Price Inertia." *Journal of Monetary Economics* 56:S19–S37.

———. (2010). "Noisy Business Cycles." In *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, 319–378.

———. (2013). "Sentiments." *Econometrica* 81 (2):739–779.

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

———. (2017). "Dampening General Equilibrium: From Micro to Macro." *NBER Working Paper No. 23379* .

———. (2018). "Forward Guidance without Common Knowledge." *American Economic Review* 108 (9):2477–2512.

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

Angeletos, George-Marios and Karthik Sastry. (2021). "Managing Expectations: Instruments vs. Targets." *Quarterly Journal of Economics* .

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. (2020). "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model." *NBER Working Paper 26647* .

Baley, Isaac and Laura Veldkamp. (2021). "Bayesian learning." *Handbook of Economic Expectations* .

Benhabib, Jess, Pengfei Wang, and Yi Wen. (2015). "Sentiments and Aggregate Demand Fluctuations." *Econometrica* 83 (2):549–585.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. (2020). "Overreaction in Macroeconomic Expectations." *American Economic Review* 110 (9):2748–82.

Broer, Tobias and Alexandre Kohlhas. (2021). "Forecaster (mis-) behavior." .

Camerer, Colin and Dan Lovallo. (1999). "Overconfidence and Excess Entry: An experimental Approach." *The American Economic Review* 89 (1):306–318.

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong. (2004). "A cognitive hierarchy model of games." *The Quarterly Journal of Economics* 119 (3):861–898.

Carroll, Christopher and Tao Wang. (2021). "Epidemiological Expectations in Economics." *Handbook of Economic Expectations* .

Chahrour, Ryan and Gaetano Gaballo. (2021). "Learning from House Prices: Amplification and Business Fluctuations." *The Review of Economic Studies* .

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner. (2021). "Sectoral Media Focus and Aggregate Fluctuations." *American Economic Review* 111 (12):3872–3922.

Chahrour, Ryan and Robert Ulbricht. (2021). "Robust predictions for DSGE models with incomplete information." *American Economic Journal: Macroeconomics* .

Clarida, Richard, Jordi Gali, and Mark Gertler. (1999). "The science of monetary policy: a new Keynesian perspective." *Journal of Economic Literature* 37 (4):1661–1707.

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

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

Crawford, Vincent P, Miguel A Costa-Gomes, and Nagore Iriberri. (2013). "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature* 51 (1):5–62.

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

Farhi, Emmanuel and Iván Werning. (2019). "Monetary Policy, Bounded Rationality, and Incomplete Markets." *American Economic Review* 109 (11):3887–3928.

Gabaix, Xavier. (2020). "A behavioral New Keynesian model." *American Economic Review* 110 (8):2271–2327.

Gali, Jordi. (1999). "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review* 89 (1):249–271.

Garcıa-Schmidt, Mariana and Michael Woodford. (2019). "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review* 109 (1):86–120.

Greenwood, Robin and Samuel G Hanson. (2015). "Waves in Ship Prices and Investment." *The Quarterly Journal of Economics* 130 (1):55–109.

Huo, Zhen and Naoki Takayama. (2021). "Rational Expectations Models with Higher Order Beliefs." *Yale mimeo* .

Ilut, Cosmin and Hikaru Saijo. (2020). "Learning, Confidence, and Business Cycles." *Journal of Monetary Economics* .

Iovino, Luigi and Dmitriy Sergeyev. (2021). "Quantitative Easing without Rational Expectations." *Boconni mimeo* .

Kneeland, Terri. (2015). "Identifying Higher-Order Rationality." *Econometrica* 83 (5):2065–2079.

Lian, Chen. (2021). "A theory of narrow thinking." *The Review of Economic Studies* .

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

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

———. (2015). "Business cycle dynamics under rational inattention." *The Review of Economic Studies* 82 (4):1502–1532.

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

Mauersberger, Felix and Rosemarie Nagel. (2018). "Levels of reasoning in Keynesian Beauty Contests: a generative framework." *Handbook of computational economics* 4:541–634.

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

Melosi, Leonardo. (2017). "Signalling effects of monetary policy." *The Review of Economic Studies* 84 (2):853–884.

Milani, Fabio. (2021). "Expectational Data in DSGE Models." *Handbook of Economic Expectations* .

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

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

———. (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.

———. (2006). “Inertia of Forward-looking Expectations.” *The American Economic Review* :152–157.

Muth, John F. (1961). “Rational expectations and the theory of price movements.” *Econometrica: Journal of the Econometric Society* :315–335.

Nagel, Rosemarie. (1995). “Unraveling in Guessing Games: An Experimental Study.” *The American Economic Review* 85 (5):1313–1326.

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

———. (2017). “Dynamic Higher Order Expectations.” *Cornell Univeristy mimeo* .

Qiu, Zhesheng. (2019). “Level-k dsge and monetary policy.” *City University of Hong Kong Mimeo* .

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

Stahl, Dale O and Paul W Wilson. (1994). “Experimental Evidence on Players’ Models of Other Players.” *Journal of Economic Behavior & Organization* 25 (3).

———. (1995). “On Players’ Models of Other Players: Theory and Experimental Evidence.” *Games and Economic Behavior* 10 (1):218–254.

Townsend, Robert M. (1983). “Forecasting the Forecasts of Others.” *The Journal of Political Economy* :546–588.

Venkateswaran, Venky. (2014). “Heterogeneous Information and Labor Market Fluctuations.” *NYU mimeo* .

Vimercati, Riccardo Bianchi, Martin S Eichenbaum, and Joao Guerreiro. (2021). “Fiscal Policy at the Zero Lower Bound without Rational Expectations.” *NBER Working Paper 29134* .

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* .