From RANK to HANK, without FIRE*

George-Marios Angeletos Joao Guerreiro Dalton Rongxuan Zhang

December 2, 2025

Abstract

We offer guidance and tools for augmenting macroeconomic models with a realistic friction in expectations and for quantifying its implications. We start by distilling the common essence of a few popular alternatives to FIRE (Full Information Rational Expectations), including noisy or sticky information, rational or behavioral inattention, level-k thinking and cognitive discounting. We then develop a new specification, which captures the same essence while maximizing tractability. We show how this works in both RANK (the Representative-Agent New Keynesian model) and HANK (Heterogeneous-Agent New Keynesian models). We further clarify the separate partial- and general-equilibrium effects of informational frictions or bounded rationality; emphasize the interaction of such frictions with the Intertemporal Keynesian Cross and other forms of strategic complementarity; review concrete lessons for business cycles and macroeconomic policy; and discuss how to discipline the theory with appropriate survey evidence.

*Prepared for the 13th World Congress of the Econometric Society. Angeletos: Northwestern University and NBER; angeletos@northwestern.edu. Guerreiro: UCLA and Federal Reserve Bank of Minneapolis; jGuerreiro@econ.ucla.edu. Zhang: Northwestern University; r.zhang@u.northwestern.edu. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

Most analyses of business cycles and macroeconomic policy assume away not only any deviation from rational expectations but also any friction in information and attention. Even the emerging literature on Heterogeneous Agents New Keynesian (HANK) models almost invariably abstracts from frictions in expectations. This abstraction, known as Full Information Rational Expectations (FIRE), is far from innocuous. A growing theoretical literature shows how sensitive our models' predictions can be to incomplete information or bounded rationality, while a parallel empirical literature provides ample evidence in support of such deviations from FIRE.

In this paper, we review, synthesize and further advance some—but certainly not all—of the key developments in this area, with the following three goals in mind:

1. To distill the *common essence* of a seemingly diverse set of alternative relaxations of FIRE, including noisy or sticky information, rational or behavioral inattention, level-k thinking, and cognitive discounting.

2. To show how this common essence can be accommodated in a *tractable* manner not only in the baseline New Keynesian model but also in HANK models, which offer a better account of household behavior and of general-equilibrium (GE) interactions.

3. To illustrate the concrete applied lessons that these research endeavors deliver.

4. To offer guidance and tools for future research, with an eye towards quantification.

To these goals, we employ two closely related frameworks. Both frameworks allow for a flexible departure from FIRE, but differ on how "seriously" they account for incomplete markets and heterogeneity.

Framework and general insights

The textbook, representative-agent New Keynesian model (RANK) assumes not only FIRE but also complete markets and infinite horizons. Our baseline model, employed in Sections 2-4, relaxes the latter assumptions in a minimalistic way: by allowing for overlapping generations (OLG) of finitely-lived agents, as in Blanchard (1985). This is a familiar proxy for liquidity constraints: finite horizons accommodate a high marginal propensity to consume (MPC) out of cash-in-hand. This in turn facilitates a sharp analysis of how any departure from FIRE *interact*s with the slope of the Keynesian cross, or the strategic complementarity in private spending—echoing Angeletos and Lian (2018), Farhi and Werning (2019) and Angeletos and Huo (2021). In Section 5, we then proceed to show how the methods and insights developed in our baseline setting can be extended to richer HANK models—building on Auclert, Rognlie, and Straub (2020) and Bardóczy and Guerreiro (2023).

Throughout, we proceed in two steps. First, we show how to characterize equilibrium *beyond* FIRE, and indeed for essentially arbitrary subjective expectations, via an appropriate generalization of the Intertemporal Keynesian Cross (IKC). Next, we specialize to specific departures from FIRE, nesting and synthesizing the related literature, and spell out a few concrete lessons.

The first step—our generalization of the IKC away from FIRE—is particularly straightforward in our baseline setting (Sections 2-4), but remains highly tractable in richer HANK models as well (Section 6). The key is to separate the agents' subjective expectations from their objective counterparts and to account for the following property: without FIRE, agents are making predictable errors, and predictable revisions, when forecasting the economy's response to shocks—where by "predictable" we mean in the analyst's, or the econometrician's, eyes. Moving on the second step, we specialize to the following models of informational friction and/or bounded rationality:

- Noisy but rational expectations, along the lines of Mankiw and Reis (2002), Woodford (2003), Sims (2003, 2010), Maćkowiak and Wiederholt (2009, 2015), Angeletos and La'O (2010) and Angeletos and Huo (2021).
- Sparsity (Gabaix, 2014); level-k Thinking (Garcıa-Schmidt and Woodford, 2019; Farhi and Werning, 2019; Iovino and Sergeyev, 2023); and cognitive discounting (Gabaix, 2020).
- A new parametric form of inattention, which maximizes tractability while also disentangling two key forces: (i) how much agents know about the underlying shocks per se and (ii) how they reason about general equilibrium effects and hence about the economy's response to shocks.

Although these approaches differ from one another in notable and testable ways, which we point out at the end of this Introduction, they all share the following core:

- They account for a salient feature of the evidence on expectations, which instead no FIRE model can account for: in response to aggregate shocks (e.g., a monetary expansion or a cost-push shock), the average *forecasts* of aggregate outcomes (e.g., of income and inflation) respond less, and with delay, relative to the *actual* outcomes themselves.
- They cause consumers and firms to behave *as if* they underestimate GE feedbacks and *as if* they are less forward looking relative not only to the FIRE counterfactual but also relative to how they actually respond to idiosyncratic shocks.
- They predict that the distortion of the aggregate dynamics relative to FIRE increases, not only when agents are less informed or less rational (appropriately defined), but also when the underlying GE feedbacks are stronger (e.g., when the Keynesian cross is steeper).

Applied lessons

How do the high-level insights reviewed above inform our understanding of business cycles, or the analysis of monetary and fiscal policy? We review seven concrete lessons (in Section 5):

1. Monetary policy affects real economic activity with long and variable lags.

2. The inflation response to monetary or other demand shocks lags that of real output.

3. Adverse supply shocks are more inflationary, and propagate faster, when they are more salient.

4. It is *as if* there are large adjustment costs at the macro level but not at the micro level—aggregate outcomes may exhibit inertia and humps, even if consumers and firms respond swiftly to idiosyncratic shocks.

5. Forward guidance about future monetary policy, even when fully credible, may fail to stimulate aggregate demand.

6. Government spending multipliers are smaller, while transfer multipliers are larger.

7. The indeterminacy problem of the New Keynesian model is lessened, and two controversial ideas—the Fiscal Theory of the Price Level and neo-Fisherian effects—find less or no space.

Testable differences and empirical discipline

As already mentioned, all the non-FIRE theories considered here share the following testable prediction at the macro level: *average* expectations underreact, or adjust with delay, to *aggregate* news. But different theories make different predictions about whether a similar pattern is present at the micro level, namely in the reaction of *individual* forecasts to *individual* news. Under rational expectations, an agent's forecast errors may be predicted by information available to *other* agents, or to the econometrician, but not by her *own* information. Theories that emphasize informational frictions but preserve rational expectations thus rule out underreaction at the individual level, even though they allow for underreaction at the aggregate level. By contrast, theories that emphasize misspecified beliefs—such as level-k thinking, shallow reasoning, and cognitive discounting—predict the same degree of underreaction at both the individual and the aggregate level.

This discussion identifies a sharp testable difference between the theories that relax the FI half of FIRE and those that relax the RE half. Turning to the available evidence, there is ample support for underreaction in average forecasts but much less for underreaction in individual forecasts. In fact, Bordalo et al. (2020) and Kohlhas and Broer (2023) argue that individual forecasts often overreact to own information, which suggests that the belief misspecification is of the opposite kind than that assumed by cognitive discounting or level-k thinking. All in all, the available survey evidence therefore favors some version of dispersed noisy information and slow learning, along the lines of Woodford (2003) and Mankiw and Reis (2002), perhaps augmented with overconfidence and/or over-extrapolation (see Angeletos, Huo, and Sastry, 2021; Bordalo et al., 2020; Liao, 2025, for further discussion).

That said, insofar as one is focused on macroeconomic implications, the different theories remain close substitutes to one another. One may therefore feel free to pick among them, or devise new substitutes thereof, with the following three main criteria in mind: (1) tractability; (2) a meaningful distinction between direct/PE and indirect/GE effects; and (3) a good fit to the IRF of average forecast errors to aggregate shocks. The first two criteria are self-explanatory; the third one zeroes in on the empirical moments that discipline the core theoretical mechanism. Our preferred specification, which blends "sticky expectations" (Mankiw and Reis, 2002; Auclert, Rognlie, and Straub, 2020; Carroll et al., 2020) with "shallow reasoning" (Angeletos and Sastry, 2021), meets all these criteria.

Related Literature

One of our paper's main goals—the synthesis of a seemingly diverse literature—is shared by Angeletos, Huo, and Sastry (2021). Relative to that paper, we shift emphasis from the testable differences of the various theories to their common core and push forward a new specification of inattention, which can be seamlessly transferred from RANK to HANK. This in turn adds to a small but growing literature that pushes the research frontier on HANK without FIRE, such as Auclert, Rognlie, and Straub (2020), Pfäuti and Seyrich (2022), Guerreiro (2023), Bardóczy and Guerreiro (2023), Cai (2024), Ilut, Luetticke, and Schneider (2025), and Eichenbaum, Guerreiro, and Obradovic (2025). On the other hand, this paper is not a substitute for more comprehensive reviews of different subsets of this literature, such as Sims (2010) and Maćkowiak, Matejka, and Wiederholt (2023) on rational inattention, Angeletos and Lian (2016) on coordination and higher-order uncertainty, Gabaix (2019) on behavioral inattention, or Coibion and Gorodnichenko (2026) on surveys and information treatments. Last but not least, our paper's emphasis on how frictions in information or rationality can arrest general-equilibrium feedbacks echoes Angeletos and Lian (2017, 2018) and Farhi and Werning (2019).

2 Environment

Our baseline model extends the textbook New Keynesian model in two ways. First, we allow for a flexible departure from FIRE on both the demand and the supply side. Second, we use an OLG structure to proxy for HANK: as in Farhi and Werning (2019) and Angeletos, Lian, and Wolf (2024a), mortality risk parameterizes the average marginal propensities to consume in a manner that not only resembles state-of-the-art HANK models but also connects with the relevant microeconomic evidence. Altogether, the baseline model will facilitate a sharp analysis of the interaction of the two kinds of friction, while also setting the stage for Section 6, where we will accommodate a richer HANK structure.

Throughout, we index time by $t \in \{0, 1, ...\}$, work with the log-linearized equations around the flexible-price, steady state, and let lower-case variables denote log-deviations of the corresponding upper-case variables from this steady state. We also use the following, standard, notation: c_t and y_t are, respectively, aggregate spending and income at t; r_t is the expected real rate of interest between t and $t+1$; p_t is the price level at t and π_t is the inflation rate between t and $t+1$; $E_{i,t}[\cdot]$ is the subjective expectation of household or firm i at t, $\overline{E}_t[\cdot] \equiv \int E_{i,t}[\cdot]$ is the average expectation in the cross section of the population, and $\mathbb{E}_t^*[\cdot]$ is the FIRE counterpart. Finally, to simplify the exposition, we assume that the only exogenous aggregate shocks are two kinds of innovations in monetary policy: unanticipated changes in current interest rates; and news about future monetary policy. Additional demand or supply shocks can easily be accommodated.

2.1 Households and Aggregate demand

The micro-foundations of the demand side of our economy follow closely Angeletos and Huo (2021) and Angeletos, Lian, and Wolf (2024a). Here, we briefly discuss the key model ingredients and then derive the applicable version of the IKC, which is the most central equation in this paper.

The economy is populated by overlapping generations of perpetual youth households. In each period, there is a unit continuum of households, indexed by $i \in [0, 1]$. A household survives from one period to the next with probability $\omega \in (0, 1]$; whenever it dies, it is replaced by a new household, with the same index i. Following Blanchard (1985), we let households save and borrow via actuarially fair annuities. We furthermore let labor supply be intermediated by labor unions and firm profits be distributed to households in proportion to their labor income, so that every household works the same amount of time and receives the same income, except for an idiosyncratic shock.[1] Finally, and more crucially for our purposes, we let households have arbitrary information and arbitrary beliefs, except for the following three simplifications: each household knows her current wealth; subjective expectations satisfy the law of iterated expectations; and finally "consumer sentiment" about income and interest rates coincides with the average forecast of the corresponding aggregates. The last assumption means that $\int E_{i,t}[y_{i,t+k}]di = \overline{E}_t y_{t+k}$ and $\int E_{i,t}\left[r_{i,t+k}\right]di = \overline{E}_t r_{t+k}$, where $\overline{E}_t x_{t+k} \equiv \int E_{i,t} x_{t+k} di$ henceforth denotes the period-t average forecast of the value of aggregate variable x in period $t+k$.

[1] In particular, household i receives income $Y_{i.t} = Y_t \exp\{z^y_{i,t}\}$ and faces a rate of return $\frac{1}{\omega} R_t \exp\{z^r_{it}\}$ conditional on survival, where Y_t is aggregate income, R_t is the real, one-period, interest rate, and (z^r_{it}, z^y_{it}) are idiosyncratic shocks. Such shocks preclude households from learning the aggregate state from their own income and rates of returns.

Under these assumptions, household optimality together with market clearing yields the following fixed-point relation on aggregate spending and income:

$$y_t = c_t = \left(1-\beta\omega\right)\sum_{k=0}^{\infty}(\beta\omega)^k \overline{E}_t y_{t+k} - \sigma\beta\omega\sum_{k=0}^{\infty}(\beta\omega)^k \overline{E}_t r_{t+k}. \tag{1}$$

where β is the discount factor (also the reciprocal of the steady-state gross real interest rate), ω is the survival probability (a proxy for liquidity frictions), and σ is the elasticity of intertemporal substitution. A detailed derivation of equation (1) is provided in Appendix A, but the economics behind it is quite simple. The first part of this equation ($y_t = c_t$) is just market clearing, while the second part ($c_t = ...$) is basically the aggregate consumption function: it gives c_t as a function of the average forecasts of aggregate income and of real interest rates in all future periods.[2] Combining the two parts yields a fixed-point relation between aggregate spending and income. The solution to this fixed-point relation returns the equilibrium path of y_t (equivalently, of c_t) as a function of the paths of the real interest rate and the demand shock—similarly to what the Euler or DIS equation does in RANK. In fact, all we have done here is to unpack this equation in terms of the aggregate consumption function plus market clearing, and to extend it in two crucial ways: finite horizons; and arbitrary expectations.

Put differently, equation (1) is a version of the Intertemporal Keynesian Cross (IKC). Compared to the one developed in Auclert, Rognlie, and Straub (2024), the key novelty here is to relax FIRE.[3] We can thus pose the following question: how does the departure from FIRE interacts with liquidity constraints, or with the "slope" of the IKC, as parameterized here by ω? In Sections 4-5, we will offer a sharp answer to this question by leveraging the simplicity of equation (1); in Section 6, we will then extend the analysis to richer HANK models.

With this in mind, let us zero in on how ω enters equation (1). When $\omega = 1$, we nest the infinite-horizon Permanent Income Hypothesis. In this benchmark, the marginal propensity to consume (MPC) out of current income, or cash-in-hand, is $1-\beta$ (≈5% per year) and future income is discounted at the same rate as the risk-free rate. Relative to this benchmark, $\omega < 1$ accommodates not only a higher MPC out of current income (say, in the order of 30%) but also heavier discounting of future income. Taking our model literally, these properties are the product of mortality risk. But the same properties are produced by liquidity constraints in realistic HANK models. This explains the sense in which OLG mimics HANK—and why a smaller ω can be reinterpreted as tighter liquidity constraints.[4]

[2]What is seemingly missing from this expression is the dependence of c_t on aggregate private wealth; but the latter is zero here, because there is neither capital nor government debt.

[3]There are two additional, less important, differences: expectations of future outcomes enter in a simple exponential-discounting from, thanks to the perpetual-youth specification; and there are no backward-looking terms (y_t does not depend on y_{t-1}, y_{t-2}, etc), because there is no MPC heterogeneity (so the wealth distribution is not a relevant state variable) and aggregate wealth is zero (so this is not a state variable either). These simplifications are relaxed in Section 6.

[4]See Farhi and Werning (2019) and Angeletos, Lian, and Wolf (2024a) for more thorough discussions of this point.

2.2 Firms and Aggregate Supply

The micro-foundations of our supply block follows the textbook New Keynesian model, except that we once again relax FIRE. There is a unit-mass continuum of monopolistically competitive firms, who set prices subject to the standard Calvo friction, hire labor on a spot market, and produce according to a technology that is linear in labor. Like consumers, firms may have incomplete information and/or misspecified beliefs. For the time being, we make only two simplifying assumptions: first, whenever a firm adjusts price, she knows perfectly the past price level p_{t-1}, although she may not necessarily extract information from it; and second, "producer sentiment", as measured by $\int E_{i,t}\left[mc_{i,t+k}\right]dj$, coincides with $\overline{E}_t[mc_{t+k}]$, the average expectation of the aggregate real marginal cost.

Under these assumptions, inflation obeys the following generalized Phillips curve:

$$\pi_t = \kappa\left\{\sum_{k=0}^{\infty}(\beta\theta)^k\overline{E}_t[y_{t+k}]\right\} + \tfrac{1-\theta}{\theta}\sum_{k=1}^{\infty}(\beta\theta)^k\overline{E}_t\pi_{t+k}, \tag{2}$$

where θ is the familiar Calvo parameter (the probability of that a firm is unable to adjust prices in any given period) and κ is the slope of the Phillips curve with respect to real output (a function of θ and other supply-side parameters, such as the Frisch elasticity of labor supply). The derivation of equation (2) is relegated to Appendix A, but the economics behind it is straightforward. When a firm gets to adjust its price to a new level, the optimal price depends on its expectations about nominal marginal costs—or equivalently about real marginal costs and inflation—during the lifespan of this new price. Real marginal costs are in turn pinned down by real output, plus an exogenous supply shock. It follows that inflation in any given period is pinned down by the firms' average expectations of real output and inflation, in the precise form shown in equation (2).

There is a clear parallel between equations (1) and (2): the former unpacks the textbook Euler equation and accommodates arbitrary subjective expectations among the consumers, the latter does the analogous exercise for the textbook NKPC. In Section 6, we will extend our analysis to more general demand structures, represented by more flexible IKC. A similar point applies on the supply side: one can readily translate our analysis in Section 6 from the IKC context to the generalized Phillips curves considered in Auclert, Rognlie, and Straub (2023), nesting menu costs and firm heterogeneity.

2.3 Monetary Policy

Without serious loss of generality, we assume that the central bank varies the nominal interest rate so as to implement the following state-dependent rule for the real interest rate:

$$r_t = \phi y_t - m_t \tag{3}$$

where $\phi \in \mathbb{R}$ parameterizes how accommodative or hawkish the central bank is and where m_t represents an exogenous monetary policy shock.[5] We have signed this shock so that $m_t > 0$ means a monetary expansion, i.e., an interest-rate cut. Finally, we will later consider two "flavors" of this shock: an unanticipated change in current rates; and news, or forward guidance, about future rates.

2.4 The famous three equations, with and without FIRE

To sum up, our baseline model boils down to three equations: the IKC (1), the Phillips curve (2), and the monetary policy rule (3). These are the New Keynesian model's famous three equations, extended to allow for *both* a departure from FIRE ($\overline{E}_t \neq \mathbb{E}_t$) and finite lives ($\omega < 1$).

Suppose now that we replace $\overline{E}_t$, the average subjective forecast, with its FIRE counterpart, $\mathbb{E}_t$. We can then apply the law of iterated expectations (which holds for $\mathbb{E}_t$ but not necessarily for $\overline{E}_t$) to restate equations (1) and (2) in a recursive form. That is, we can replace (1) with $y_t = -\sigma r_t + \mathbb{E}_t[y_{t+1}]$, which is the familiar Euler equation plus market clearing, and (2) with $\pi_t = \kappa y_t + \mathbb{E}_t[\pi_{t+1}]$, which is the standard NKPC. The FIRE restriction of our model therefore coincides with RANK, despite the accommodation of $\omega < 1$. The following points are then evident:

1. Once we impose FIRE, ω drops out of the AD relation. As we will explain later, this is because ω has exactly offsetting partial-equilibrium and general-equilibrium effects under FIRE. A similar point applies to θ : holding κ constant, θ appears in our generalized NKPC but drops out of its FIRE counterpart. This basic observation anticipates our later insights about how these two parameters—which regulate the strategic complementarity in, respectively, the spending decisions of the consumers and the pricing decisions of the firms—interact with incomplete information and/or bounded rationality.

2. Replacing (3) into (2) yields the following fixed-point relation on output:

$$y_t = (1-\beta\omega)(1-\sigma\phi)\sum_{k=0}^{\infty}(\beta\omega)^k \overline{E}_t[y_{t+k}] + \sigma\beta\omega\sum_{k=0}^{\infty}(\beta\omega)^k \overline{E}_t[m_{t+k}]. \tag{4}$$

   Clearly, there is no essential difference between $\phi = 0$ and $\phi \neq 0$: varying ϕ amounts to tilting the slope of the Keynesian cross. Most of our analysis will thus concentrate on $\phi = 0$, but our methods and insights readily extend to $\phi \neq 0$.

3. While we have allowed monetary shocks in equation (3), we have abstracted from discount-rate and cost-push shocks in equations (1) and (2). This is an innocuous simplification. More sub-

[5] See Appendix C for how to accommodate, not only a conventional Taylor rule, but also a much more flexible specification, where the central bank could be conditioning its instrument on the entire history of output and inflation.

stantial is the abstraction from endogenous fluctuations in government debt. Relaxing this assumption adds an endogenous state variable (Angeletos, Lian, and Wolf, 2024a; Aguiar, Amador, and Arellano, 2024), but could be handled with the methods of Section 6.

4. Some of the empirical literature on the Phillips curve has relaxed FIRE in the following ad hoc way: taking the standard NKPC, $\pi_t = \kappa y_t + \beta \mathbb{E}_t \pi_{t+1}$, and replacing $\mathbb{E}_t \pi_{t+1}$ with $\overline{E}_t \pi_{t+1}$, the consensus forecast as measured in surveys. This approach is at odds with the micro-foundations of the NKPC. Similarly, one can not take the Euler equation $y_t = -\sigma r_t + \mathbb{E}_t[y_{t+k}]$ and just replace $\mathbb{E}_t[y_{t+k}]$ with $\overline{E}_t[y_{t+k}]$. In both cases, one has to unpack the representative-agent equations, reveal the multiple-agent interaction behind them, and properly account for subjective beliefs.

3 FIRE

In this section, we review the FIRE version of our model, which, as just explained, yields the same aggregate outcomes as RANK. We first characterize these outcomes for two cases: unanticipated monetary shocks; and news shocks. We then decompose them to partial-equilibrium (PE) and general-equilibrium (GE) components, setting the stage for the subsequent analysis away from FIRE.

3.1 Monetary shocks and forward guidance under FIRE

FIRE amounts to imposing (i) households have common knowledge about the exogenous shocks and (ii) form expectations according to Rational Expectations Equilibrium. We will later relax the first assumption ("full information"), the second assumption ("infinite rationality"), or both. For now, we ask how aggregate spending and output respond to innovations in monetary policy under FIRE.

To simplify the exposition, we henceforth let $\phi = 0$; as noted earlier, this is without serious loss of generality and merely fixes ideas by representing monetary policy as an exogenously specified path for the real rate. Next, we let m_t follow an AR(1) process with two kinds of innovations:

$$m_t = \rho m_{t-1} + \varepsilon_t + \eta_t, \tag{5}$$

where $\rho \in (0,1)$ and (ε_t, η_t) are random variables, independent of one another and i.i.d. over time. Finally, we let ε_t and η_t become commonly known at dates t and $t-\ell$, respectively; in words, ε_t is an unanticipated innovation in current rates, whereas η_t is news revealed ℓ periods earlier ("forward guidance"). This yields $\mathbb{E}_t[m_{t+h}] = \rho^h m_t + \sum_{k=1}^{\min\{h,\ell\}} \rho^{h-k} \eta_{t+k}$, while iterating the Euler equation yields $c_t = -\sigma \sum_{h=0}^{\infty} \mathbb{E}_t[r_{t+h}] = \sigma \sum_{h=0}^{\infty} \mathbb{E}_t[m_{t+h}]$.[6] Combining, we arrive at the following result.

[6]Here, we implicitly select equilibrium by imposing $\lim_{h\to\infty} \mathbb{E}_t[c_{t+h}] = 0$. See Angeletos et al. (2025) for a discussion of

Proposition 1. *Under FIRE, aggregate spending and output are given in equilibrium by*

$$c_t = y_t = \frac{\sigma}{1-\rho}\left(m_t + \sum_{h=1}^{\ell} \eta_{t+h}\right) \tag{6}$$

Two properties are notable. First, the economy's response to a 1% real rate cut is the same whether this cut happens now or far in the future. This is a version of the Forward Guidance Puzzle (Del Negro, Giannoni, and Patterson, 2015). Second, the economy's response is invariant to ω : under FIRE, our proxy for HANK behaves like RANK. This is an example of the irrelevance result in Werning (2015). Behind both of these properties lies a delicate balancing of partial equilibrium (PE) and general equilibrium (GE) forces, whose separate roles will shine once we move away from FIRE. In anticipation of this, we next show how to decompose the FIRE outcomes into PE and GE components.

3.2 PE-GE decomposition

Let the interest rate faced by households i be $r_{i,t} = r_t - m_{i,t} = -m_t - m_{i,t}$, where $m_{i,t}$ is i.i.d. across i and independent of m_t. Next, let $m_{i,t} = \rho m_{i,t} + \varepsilon_{i,t} + n_{i,t-\ell}$, where $\varepsilon_{i,t}$ and $\eta_{i,t-\ell}$ are the idiosyncratic counterparts of ε_t and $\eta_{t-\ell}$. In equilibrium, i's spending can then be expressed as follows:

$$c_{i,t} = c_t + (1-\beta\omega)\, a_{i,t} + \frac{\sigma\beta\omega}{1-\beta\omega\rho} m_{i,t} + \frac{\sigma\beta\omega}{1-\beta\omega\rho} \sum_{h=1}^{\ell} (\beta\omega)^h \eta_{i,t+h} \tag{7}$$

where $a_{i,t}$ is i's wealth and c_t is aggregate spending (as in Proposition 1). For any $h \geq 0$, $\frac{\sigma\beta\omega}{1-\beta\omega\rho}(\beta\omega)^h$ measures how much i's spending increases with an *idiosyncratic* 1% rate cut h periods ahead. For any h, this coefficient decreases with $\beta\omega$: if a household is less patient or has a shorter horizon, she responds less to future interest rates. Furthermore, this coefficient diminishes with h and vanishes as $h \to \infty$: news about idiosyncratic fundamentals in the far future are heavily discounted.

Consider now i's response to an *aggregate* interest rate cut at horizon $h \geq 0$. This stimulates i's spending via two channels: by lowering her own interest rates; and by stimulating others' spending, which in turn raises her own income. The first channel is the same as above: $\frac{\sigma\beta\omega}{1-\beta\omega\rho}(\beta\omega)^h$ measures, not only i's response to an idiosyncratic shock, but also the partial-equilibrium (PE) effect of aggregate shock. The second channel captures the GE feedback via aggregate spending and income: when everyone learns at t that interest rates will fall at $t+h$, aggregate spending increases at all dates between t and $t+h$. This raises i's permanent income by an amount that increases with h. A higher h thus translates to a larger GE feedback, which counteracts the smaller PE effect.

the indeterminacy problem of the New Keynesian and of the "right" equilibrium selection.

Fact 1. *Under FIRE, the equilibrium response of aggregate spending and income to monetary policy shock can be decomposed into PE and GE effects as follows:*

$$c_t = y_t = \underbrace{\left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\right)^{-1}}_{\text{GE multiplier}} \underbrace{\frac{\sigma\beta\omega}{1-\beta\omega\rho}}_{\text{PE effect}} m_t + \sum_{h=1}^{\ell} \underbrace{\left[\left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\right)(\beta\omega)^h\right]^{-1}}_{\text{GE multiplier}} \underbrace{\frac{\sigma\beta\omega}{1-\beta\omega\rho}(\beta\omega)^h}_{\text{PE effect}} \eta_{t+h}$$

Let us zero in on the role of ω. When consumers have shorter horizons or hit borrowing constraints more frequently, they not only care less about the long run but also respond more strongly to income fluctuations in the short run. This explains why the GE feedback between spending and income is stronger when exactly the PE effect of monetary policy is weaker. That the two forces move in opposite directions is intuitive and extends to quantitative HANK models (e.g., Kaplan, Moll, and Violante, 2018); that the two forces perfectly offset each other is a knife-edge implication of the combination of FIRE and the particular demand structure assumed here. Summing up:

Fact 2. *Under FIRE, a lower ω (short horizons or tighter liquidity) implies both a smaller PE effect and a large GE effect, with the two perfectly offsetting each other in our setting.*

Figure 1 illustrates this fact for two scenarios: a unanticipated monetary expansion in period 0 (left panel); and a news about a monetary expansion 12 quarters later (right panel). In each case, we draw the Impulse Response Function (IRF) of aggregate output, along with the underlying PE component, for $\omega = 1$ and $\omega = .7$. The former nests RANK, the latter proxies HANK with an aggregate MPC of about 30%. In RANK, the PE component accounts for almost the entire response to the unanticipated monetary expansion, and for the lion's share of the response of the news shock. Moving to HANK reduces the PE effect but also increases the GE multiplier, leaving the total effect the same. The PE-GE split is therefore irrelevant so far, but will play a central role in the sequel, where we relax FIRE.

4 Beyond FIRE

In this section, we review how incomplete information and/or bounded rationality modify the predictions of the New Keynesian model, making them not only more palatable but also consistent with evidence on expectations. The relevant empirical backdrop here is the type of belief inertia documented in, inter alia, Coibion and Gorodnichenko (2012), Coibion and Gorodnichenko (2015), and Angeletos, Huo, and Sastry (2021): in response to identified policy or other aggregate shocks, average (or "consensus") forecasts react more sluggishly than actual outcomes, or than the econometrician's forecasts thereof. This fact rejects FIRE in favor of the alternative theories considered in this section.

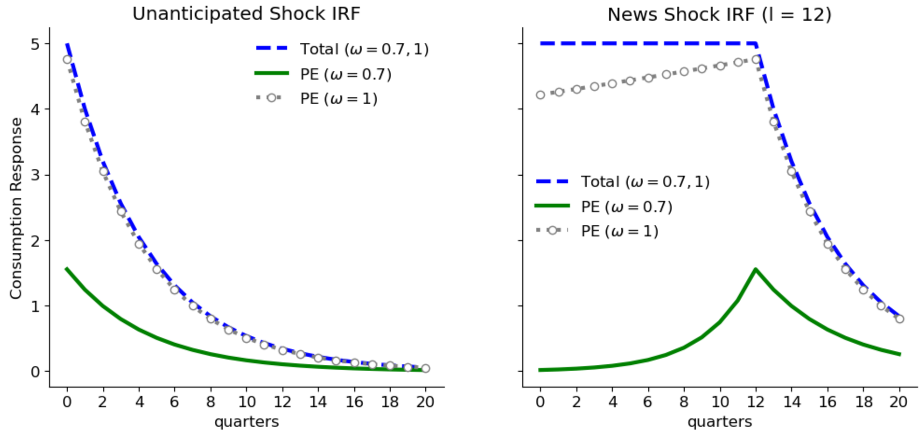

Figure 1: PE and GE Responses to two Monetary Policy Shocks, Under FIRE

We start by introducing incomplete (i.e., noisy and heterogeneous) information as in Woodford (2003) and Angeletos and Huo (2021). We go on to unpack the different forces at work, to formalize the sense in which that this approach and certain forms of bounded rationality are close cousins, and to develop our preferred specification, which maximizes tractability while preserving the essence.

4.1 Incomplete information

In this section, we relax the FI half of FIRE while maintaining the RE half. In particular, we momentarily shut down news shocks, focusing on conventional monetary shocks; and we assume that each consumer's information in period t is given by $\mathscr{I}_{i,t} = \{x_{i,s}\}_{s=-\infty}^{t}$, where $x_{i,t} = m_t + u_{i,t}$ is the signal (new information) received in period t and where $u_{i,t} \sim \mathcal{N}(0, \sigma_u^2)$ is noise, i.i.d. across both i and t. This structure introduces, not only first-order uncertainty (i.e., uncertainty about, or inattention to, the policy shock itself), but also higher-order uncertainty (i.e., uncertainty about how *other* agents, or the economy as whole, may respond to the shock). The solution of the equilibrium involves a non-trivial fixed point between expectations and outcomes, but the next result, which is proved in Angeletos and Huo (2021), offers a useful representation.

Proposition 2. *Under incomplete information (as specified above), equilibrium spending responds to monetary shocks in the same way as a RANK economy with the following modified Euler equation:*

$$c_t = -\sigma r_t + \alpha_f \mathbb{E}_t[c_{t+1}] + \alpha_b c_{t-1}, \tag{8}$$

for some $\alpha_f < 1$ and $\alpha_b > 0$ that are functions of σ_u and ω. It is thus as if *consumers are both less forward-looking* ($\alpha_f < 1$) *and more backward-looking* ($\alpha_b > 0$)*. Furthermore, the distortions increase, not only with the level of noise, but also with the deviation from PIH (they fall with ω).*

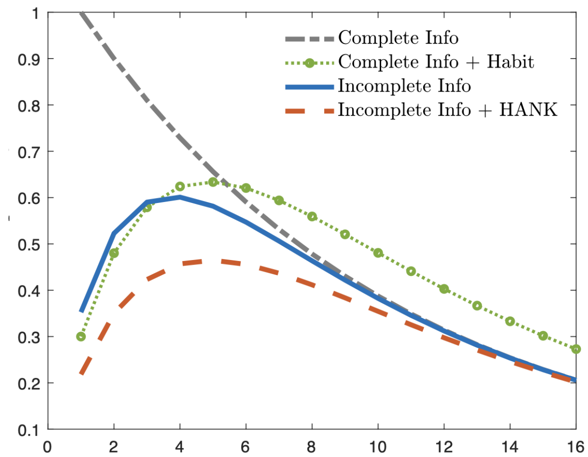

Figure 2: Consumption and Output Responses to an Interest Rate Cut, without FIRE

We will elaborate on the mechanisms behind this result and on its connection to other forms of bounded rationality in Sections 4.2–4.4. For now, we visualize the result in Figure 2, which is borrowed from Angeletos and Huo (2021). This figure compares the IRF of aggregate spending to an unanticipated monetary expansion across four scenarios. “Complete Info” corresponds to the textbook RANK; this delivers a steep, front-loaded response. “Complete Info + Habit” maintains FIRE but introduces consumption habit (i.e., a cost to adjusting consumption), and calibrates this habit to a similar level as that estimated in a large DSGE literature; this captures, by design, the kind of hump-shaped aggregate dynamics observed in the data. “Incomplete Info” replaces habit with noisy information (along the lines described above) and calibrates the level of noise to survey evidence; evidently, this produces a *quantitatively* similar hump shape as the previous case. Finally, “Incomplete Information + HANK” interacts $\sigma_u > 0$ with $\omega < 1$; this reinforces the informational friction, reduces the effect of the monetary shock at all horizons, and produces a more pronounced hump shape.

Incomplete information thus resembles adjustment costs, but there are a few important differences. First, incomplete information provides a *unified* theory for the more ad hoc, equation-by-equation, “patches” used in the DSGE: the same primitive friction can now replicate the patterns generated, not only by habit in consumption, but also by adjustment costs to investment, by the hybrid version of the NKPC, and even by momentum in asset prices (see Angeletos and Huo, 2021, for a detailed exposition of these additional applications). Second, although the economy as whole may appear to be subject to large adjustment costs in response to the monetary or other aggregate shocks, an individual may still respond swiftly to idiosyncratic shocks. This helps close the discomforting gap between the macroeconomic and the microeconomic estimates of habit and other adjustment costs (Angeletos and Huo, 2021); equivalently, it reconciles sluggish responses (“humps”) at the macro level with quick, front-loaded, response (“jumps”) at the micro level (Auclert, Rognlie, and Straub, 2020).

Finally, the as-if adjustment cost is not a fixed parameter: the effective friction varies, not only with the level of noise, but also with the MPC (or the slope of the Keynesian cross) and with policy parameters (such as ϕ once we relax $\phi = 0$). This is because such "deeper" parameters regulate the GE feedback, or the strategic complementarity, which in turn interacts with the informational friction.

To elucidate the various forces at work, the remainder of Section 4 proceeds as follows. Section 4.2 introduces a specification that shuts down learning and, instead, zeroes in on PE and GE channels of inattention (equivalently, on first- and higher-order uncertainty). Section 4.3 shows how this specification nests, as special cases, various forms incomplete information and bounded rationality. Section 4.4 adds back learning and explains how this, too, operates via both PE and GE channels. We conclude the discussion here by emphasizing that the lessons of 2 are not just a theoretical possibility: when disciplined by survey evidence on expectations, the theory can indeed produce quantitatively large inertia in both aggregate consumption and inflation (Angeletos and Huo, 2021; Cai, 2024).

4.2 A flexible specification—maximizing tractability, separating PE and GE forces

In this subsection, we abstract from learning, which amounts to setting $\alpha_b = 0$ in Proposition 2, and zero in on the PE and GE forces behind $\alpha_f < 1$. To this goal, we introduce the following information structure, which—as it will be made clear—has both conceptual and computational advantages.[7]

Assumption 1. *Following any innovation in monetary policy, a fraction $\lambda^p \in [0,1]$ of the population learns it immediately and perfectly, while the rest never learn it. Furthermore, everyone believes that the fraction of agents that learn the innovation is $\lambda^g \in [0,1]$, where λ^g may or may not coincide with λ^p.*

By letting $\lambda^p < 1$, we accommodate first-order uncertainty, that is, uncertainty about the policy shock itself. By letting $\lambda^g \neq \lambda^p$, we separately parametrize higher-order uncertainty, that is, uncertainty about how much *others* will adjust their beliefs and behavior to the shock. In simpler terms, λ^p measures consumers' attention to monetary policy, while λ^g measures how much they adjust their expectations of aggregate spending—and hence also of their own income—in response to monetary shocks. As shown next, the first friction attenuates the direct or PE effect of monetary policy, while the second friction arrests the GE multiplier.

Proposition 3. *Under Assumption 1, equilibrium spending and income satisfy*

$$c_t = y_t = \underbrace{\left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1}}_{\textit{GE multiplier}} \cdot \underbrace{\frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda^p}_{\textit{PE effect}} m_t. \tag{9}$$

[7] A close cousin of Assumption 1 was introduced in Angeletos and Sastry (2021), albeit in a more abstract context.

To understand this result, let us take a brief detour: consider a simpler, static example, in which an individual's optimal spending is $c_i = E_i\left[\alpha m + \gamma c\right]$, where m is the policy shock, $c \equiv \int c_j dj$ is aggregate spending, and $\alpha > 0$ and $\gamma \in (0,1)$ are exogenous parameters. Think of this interchangeably as a static Keynesian cross and as a "beauty contest" (Morris and Shin, 2002). Aggregating across i gives

$$c = \alpha \overline{E}[m] + \gamma \overline{E}[c]. \tag{10}$$

That is, the actual c depends, not only on the forecasts of the policy shock, but also on the forecasts of c, and thereby on the forecasts of the forecasts of others. By the first half of Assumption 1,

$$\overline{E}[m] = \lambda^p\, m, \tag{11}$$

which shows that λ^p governs—and can be measured by—how much the average forecasts of m covary with m. Consider next the forecasts of c. To characterize them, let us guess and verify that, in equilibrium, $c_i = bE_i[m]$ for all i and some coefficient b. Then, $c \equiv \int c_i di = b\overline{E}[m]$ and therefore $\overline{E}[c] = \overline{E}[b\overline{E}[m]] = b\overline{E}[\overline{E}[m]]$, which reveals the role of the forecasts of the forecasts of others. By the second half of Assumption 1, $\overline{E}[\overline{E}[m]] = \lambda^g \overline{E}[m]$; intuitively, because agents believe that only a fraction λ^g of other agents are attentive to the shock, higher-order beliefs (the forecasts of the forecasts of others) respond less than first-order beliefs (the forecasts of m). By the same token,

$$\overline{E}[c] = b \cdot \lambda^g \overline{E}[m] = \lambda^g \cdot b\overline{E}[m] = \lambda^g c, \tag{12}$$

which shows that λ^g governs—and can be measured by—how much the forecasts of aggregate spending covary with actual spending. Finally, substituting (11) and (12) into (10) and solving for b, we get

$$c = \underbrace{\left(1 - \gamma\lambda^g\right)^{-1}}_{\text{GE multiplier}} \cdot \underbrace{\alpha\lambda^p}_{\text{PE effect}}\; m. \tag{13}$$

Letting $\lambda^p < 1$ is thus equivalent to reducing α, the PE elasticity of c to m, whereas letting $\lambda^g < 1$ is equivalent to reducing γ, the strategic complementarity or the GE feedback. In words, own inattention arrests the PE effect, whereas lack of confidence about others' attention arrests the GE feedback.

Proposition 3 verifies that the above logic extends from the static example to our dynamic economy; in fact, equation (13) directly translates to equation (9), with $\alpha = \frac{\sigma\beta\omega}{1-\beta\omega\rho}$ and $\gamma = \frac{1-\beta\omega}{1-\beta\omega\rho}$. Intuitively, because behavior is forward-looking, both the PE and the GE effects increase with ρ. But the meaning and the effect of λ^p and λ^g remain exactly the same as in the static example. What is more, as can be seen in the Appendix, the solution of the dynamic economy is just as tractable as the solution of the static example. This underscores an important advantage of Assumption 1: the equilibrium dynamics can be solved via essentially the same guess-and-verify method (a.k.a. method of undetermined coefficients) as standard FIRE models. Crucially, this advantage remains when we add learning (in Section 4.4) or move to richer HANK models (in Section 6).

We conclude by circling back to Proposition 2 and verifying the claim made in the beginning of this section: so far, we have effectively restricted $\alpha_b = 0$ and have elucidated the forces behind $\alpha_f < 1$.

Corollary 1. *Under Assumption 1, the modified Euler equation (8) stated in Proposition 2 applies with $\alpha_b = 0$ and $\alpha_f = A(\lambda^p, \lambda^g, \omega)$ for some increasing function A, with $A(\cdot) = 1$ if and only if $\lambda^p = \lambda^g = 1$. That is, $\alpha_f < 1$ is the combined footprint of the PE and GE attenuation described above. Furthermore, α_f is a steeper function of λ^p and λ^g when ω is lower: the same informational friction has more bite when consumers have shorter horizons and higher MPCs.*

4.3 An equivalence result for a few popular theories

We now demonstrate how Assumption 1 replicates various forms of incomplete information and/or bounded rationality. To this end, we focus on the case of a single innovation occurring at $t = 0$, continue to abstract from learning over time, and consider the following seven specifications:

1. **Noisy information (Woodford, 2003) or rational inattention (Sims, 1993), minus learning.** Each household observes a noisy private signal $m_i = m_0 + u_i$, where $u_i \sim \mathcal{N}(0, \sigma_u^2)$. This is basically the same specification as that introduced in Section 4.1, except that we are abstracting from dynamic learning. Rational inattention corresponds to letting households choose σ_u optimally, subject to some cost. FIRE, on the other hand, corresponds to $\sigma_u \to 0$.

2. **Overconfidence (Kohlhas and Broer, 2023).** The actual signal structure remains as above, but households have misspecified beliefs about it. While the true precision of everyone's signal is σ_u^{-2}, each household believes that her own precision is σ_{own}^{-2} and that the precision of others is $\sigma_{\text{others}}^{-2}$, where σ_{own}^{-2} exceeds both σ_u^{-2} and $\sigma_{\text{others}}^{-2}$.

3. **Level-k thinking (Garcıa-Schmidt and Woodford, 2019; Farhi and Werning, 2019; Iovino and Sergeyev, 2023).** Instead of an informational friction, we now introduce a bound on rationality: households form forecasts about aggregate outcomes by starting from steady state and iterating best responses (or the IKC) k times. Formally, level-k outcomes solve the following recursion:

$$c_t^k = (1-\beta\omega) \sum_{h=0}^{\infty} (\beta\omega)^h c_{t+h}^{k-1} - \sigma\beta\omega \sum_{h=0}^{\infty} (\beta\omega)^h r_{t+h}, \; \forall t, k$$

with $r_t = -\rho^t m_0$ and $c_t^0 = 0\ \forall\tau$. To understand this concept, note that level-1 households observe the shock perfectly but believe that others are level-0 thinkers, so their forecasts of aggregate spending are $E^1_{i,t}[c_t] = 0$; level-1 households thus respond *as if* the shock was purely idiosyncratic and the GE feedback was absent. Level-2 households best-respond to the belief that all other households are level-1, which amounts to incorporating one round of GE feedback. Level-3 households best-respond to the belief that others are level-2, and so on. It follows that level-k outcomes embed $k-1$ rounds of GE feedback, and that FIRE corresponds to $k \to \infty$.

4. **A representative inattentive agent.** There is a representative household, who understands the economy's structure and behaves as in RANK, except that her beliefs about the policy shock are distorted to $E_t m_t = \mu \mathbb{E}_t m_t$, for some $\mu \in (0,1)$.

5. **Sparsity** (Gabaix, 2014; Guerreiro, 2023). Households obey the FIRE consumption function, subject to the following distortion: they (mis)perceive their income and their interest rate to be, respectively, $E_{i,t} y_{i,t+h} = \mu_y \mathbb{E}_t y_{i,t+h}$ and $E_{i,t} r_{i,t+h} = \mu_r \mathbb{E}_t r_{i,t+h}$, where $(\mu_y, \mu_r) \in (0,1)^2$ are fixed scalars interpreted as the product of behavioral (non-rational) inattention.[8]

6. **Cognitive discounting** (Gabaix, 2020). Households cognitively discount future state variables, in effect underestimating the persistence in those variables. Here, this means that, while the true persistence of m_t is ρ, households misperceive it to be $\hat{\rho} = \mu\rho$, for some $\mu \in (0,1)$.

7. **Confusion of aggregate and idiosyncratic shocks.** Households face idiosyncratic shocks: $r_{i,t} = r_t + \xi^r_{i,t}$ and $y_{i,t} = y_t + \xi^y_{i,t}$, where $\xi^r_{i,t}$ and $\xi^y_{i,t}$ are idiosyncratic, independent AR(1) processes, with persistences ρ_r and ρ_y, respectively. Furthermore, households observe $r_{i,t}$ and $y_{i,t}$ perfectly, and respond to aggregate shocks as if they were idiosyncratic.[9]

The question of interest is how aggregate spending responds to the underlying policy shock, in each of the above scenarios. We answer this question below, leveraging the flexibility of Assumption 1.

Proposition 4 (Observational Equivalence)**.** *The equilibrium response of aggregate spending to policy shock in any of the several forms of inattention or bounded rationality described above coincides with that implied by Assumption 1 and characterized in Proposition 3, for some (λ^p, λ^g). In particular:*

1. *Noisy Information maps to $\lambda^p = \lambda^g = \lambda \equiv \frac{\sigma_u^{-2}}{\sigma_u^{-2}+\sigma_\epsilon^{-2}}$.*

2. *Overconfidence maps to $\lambda^p = \frac{\sigma_{own}^{-2}}{\sigma_{own}^{-2}+\sigma_\epsilon^{-2}}$ and $\lambda^g = \frac{\sigma_{others}^{-2}}{\sigma_{others}^{-}+\sigma_\epsilon^{-2}} < \lambda^p$.*

[8] That is, (μ_y, μ_r) may be the solution of a meta-problem weighting the costs and benefits of attention, similarly to the case of rational inattention mentioned earlier. For our purposes, such endogeneity is neither necessary nor detrimental.

[9] The last assumption can be reconciled with rational expectations by letting the variances of $\xi^r_{i,t}$ and $\xi^y_{i,t}$ be arbitrarily large relative to that of the aggregate shocks, and by assuming away any other information about the latter.

3. *Level-k Thinking maps to* $\lambda^p = 1 > \lambda^g = \Lambda(k)$, *for an increasing function* Λ, *with* $\lim_{k\to\infty}\Lambda(k) = 1$.
4. *A representative inattentive agent maps to* $\lambda^p = \mu < 1$ *and* $\lambda^g = 1$.
5. *Sparsity maps to* $\lambda^p = \mu_r < 1$ *and* $\lambda^g = \mu_y < 1$.
6. *Cognitive discounting maps to* $\lambda^p = \lambda^g = \Lambda(\mu) < 1$, *for an increasing function* Λ, *with* $\Lambda(1) = 1$.
7. *Confusion of aggregate and idiosyncratic shocks maps to* $\lambda^p = \Lambda_p(\rho_r/\rho)$ *and* $\lambda^g = \Lambda_g(\rho_y/\rho)$, *for some increasing functions* Λ_p *and* Λ_g, *with* $\Lambda_p(1) = \Lambda_g(1) = 1$.

Corollary 2. *Relative to FIRE, a representative inattentive agent exhibits a smaller PE effect but the same GE multiplier. Conversely, level-k thinking arrests the GE feedback without changing the PE effect. All other theories affect both channels, either equally or differentially. Finally, translated in terms of Proposition 2, all these theories accommodate* $\alpha_f < 1$ *but restrict* $\alpha_b = 0$.

Proposition 4 shows how the different theories map to different configurations of λ^p and λ^g under Assumption 1. Corollary 2 summarizes which theories attenuate PE vs GE effects and connects to Proposition 2. The textbook version of noisy information imposes $\lambda^p = \lambda^g = \lambda$, where λ is the signal-to-noise ratio. Adding overconfidence allows $\lambda^g < \lambda^p$ and clarifies that the aggregate dynamics depend on the *perceived* precision. Level-k Thinking arrests the GE feedback while preserving the PE effect, so it maps to $\lambda^p = 1 > \lambda^g$. Conversely, a scenario featuring a representative inattentive agent exhibits the full GE feedback, despite an attenuated PE effect. These two scenarios are thus extreme opposite of each other and help isolate the two distortions, while all other theories blends them.

Beyond the specific theories exactly nests above, our formulation can proxy for a wide range of other approaches. For instance, Woodford (2018) models agents as engaging in finite planning: they solve their intertemporal decision problems only up to a finite horizon, thus failing to incorporate not only the full dynamic consequences of their own choices but also the full equilibrium dynamics. Under our prism, this amounts to $\lambda^p < 1$ and $\lambda^g < 1$. Likewise, Mei and Wu (2024) consider agents endowed with a finite depth of equilibrium reasoning: they internalize only a limited number of causal feedback rounds in the general-equilibrium mapping. This corresponds to $\lambda^p = 1$ and $\lambda^g < 1$, similarly to level-k thinking. Summing up, Assumption 1 offers not only a unifying lens but also a more direct pathway to both the theory's essence and its empirical footprint (more on this in Section 7).

4.4 Learning

The second half of Corollary 2 reminds that Assumption 1 and all of its cousins reviewed in the previous section allow $\alpha_f < 1$ (extra discounting of the future) but rule out $\alpha_b > 0$ (anchoring to past outcomes). We now explain how learning is the key to $\alpha_b > 0$ and hence also to the hump shapes seen in Figure 2. We first review two popular forms of learning, and then discuss how to add a form of learning to Assumption 1, preserving its tractability and versatility.

Noisy learning as in Woodford (2003) and Angeletos and Huo (2021). As in Subsection 4.2, household i receives a new signal each period, given by $x_{i,t} = m_t + u_{i,t}$, where $u_{i,t} \sim \mathcal{N}(0, \sigma_i^2)$ is the noise. Letting $g \in (0,1)$ denote the Kalman gain and L the lag operator, we can express the dynamic evolution of the average forecasts of m_t as follows:

$$\overline{E}_t[m_t] \; = \; g\, m_t + (1-g) \underbrace{\rho \overline{E}_{t-1}[m_{t-1}]}_{\overline{E}_t[m_{t-1}]} = \frac{g}{1-\rho\left(1-g\right)L} m_t \tag{14}$$

Since m_t is an AR(1), $\overline{E}_t[m_t]$ is an AR(2), which means that the average expectations exhibit hump-shaped dynamics. The intuition is straightforward: as agents accumulate more information, their forecasts gets closer to the true m_t; but as time passes, the true m_t itself converges back to zero; it follows that, following a positive innovation, the forecasts of m_t initially increase but eventually fall.

While this discussion offers crucial intuition, it tells only one half of the story: aggregate spending depends, not only on the forecasts of m_t, but also on the forecasts of aggregate spending itself. We now explain why this is the key to understanding why the value of α_b in Proposition 2 increases as we lower ω—or, equivalently, why the hump in aggregate spending gets more pronounced as we increase the MPC. First, recall the optimal consumption at t depends on expectations of income at $t+h$, for all $h \geq 0$. Next, note that forecasting future income is the same as forecasting aggregate spending, which in turn amounts to forecasting the forecasts of others (Townsend, 1983). Furthermore, higher-order beliefs adjust more sluggishly than first-order beliefs, because it takes more time to reach *consensus* than to reach individual knowledge of any given shock (Rubinstein, 1989; Morris and Shin, 1997, 2006). Finally, because agents care more about the behavior of others, and thus also about the forecasts of others, when the strategic complementarity in their spending decisions is larger (equivalently, when the Keynesian cross is steeper), a smaller ω amplifies the importance of higher-order beliefs. Combining these observations, we conclude that a smaller ω gives rise to a more pronounced hump in aggregate spending and income—equivalently, a smaller ω maps to a larger α_b in Proposition 2.

Sticky information: same essence, some differences. Another popular form of learning is sticky information as in Mankiw and Reis (2002). In this framework, households update their information with probability λ and stay uninformed with probability $1-\lambda$. This gives rise to a simple recursion on average expectations: $\overline{E}_t[m_t] = \lambda m_t + (1-\lambda)\overline{E}_{t-1}[m_t]$. Clearly, this is equivalent to (14), with the Kalman gain g replaced by λ. That is, sticky information and noisy information produce identical dynamics for average first-order beliefs. As clarified in Angeletos and Lian (2016, Section 8.5) and Angeletos and Huo (2021, Online Appendix H), this exact coincidence does *not* extend to higher-order beliefs and aggregate outcomes, but the difference is only quantitative, not qualitative. In particular, the two models share the prediction that the hump shape in aggregate spending is more pronounced, not only when the informational friction is larger, but also when the strategic complementarity is stronger, or when the Keynesian cross is steeper. For all practical purposes, the two models are therefore one and the same insofar as one focuses on average forecasts and aggregate outcomes. Finally, Coibion and Gorodnichenko (2012) point out that the two models make different predictions about the cross section of beliefs, and go on to argue that noisy information performs better in this regard. However, this difference is not relevant for the more narrow purposes of our paper.

Adding learning to Assumption 1. We previously showed that Assumption 1 captures, succinctly and tractably, the PE and GE effects of noisy information and of related forms of bounded rationality, while abstracting from learning. We now extend this assumption to accommodate learning. We do so in a way that not only preserves tractability but also separately parameterizes the speed of adjustment in the agents' knowledge about the underlying policy shock and in their perceptions about the economy's adjustment to it.

Assumption 2. *Following any innovation, an agent either knows the innovation perfectly or doesn't have any information about it. Furthermore, the actual fraction of agents who know the innovation h periods after is λ_h^p, whereas the perceived fraction is λ_t^g, where*

$$\lambda_{h+1}^p = (1-G^p)\lambda_h^p + G^p \qquad \text{and} \qquad \lambda_{h+1}^g = (1-G^g)\lambda_h^g + G^g \qquad \forall h \geq 0$$

and where $(\lambda_0^p, \lambda_0^g, G^p, G^g) \in (0,1)^4$. Finally, agents understand these dynamics.

The λ's have the same meaning as in Assumption 1. But now the associated belief distortions are allowed to diminish over time. G^p parameterizes how quickly the distortion diminishes in the forecasts of the policy shock (which here coincide with the forecasts of the real interest rate), whereas G^g parameterizes the speed of adjustment in the forecasts of aggregate spending and income.

How does this relate to Bayesian learning with noisy signals (as in Woodford, 2003)? The essence is the same, but there is no "infinite regress" (Townsend, 1983) and the solution is straightforward.

Proposition 5. *Under Assumption 2, the IRF of aggregate output to a monetary-policy innovation is obtained by solving the following modified IKC*

$$y_t = (1-\beta\omega)(1-\sigma\phi)\sum_{h=0}^{\infty}(\beta\omega)^h \tilde{\lambda}_{t,h}^g y_{t+h} + \sigma\beta\omega\sum_{h=0}^{\infty}(\beta\omega)^h \lambda_t^p m_{t+h}, \tag{15}$$

where $\tilde{\lambda}_{t,h}^g \equiv \lambda_{t+h}^g \frac{\lambda_t^p}{\lambda_{t+h}^p}$ and $m_t = \rho^t \varepsilon_0$ for all t, h, and where ε_0 is the innovation.

The proof, which can be found in the Appendix, uses a similar guess-and-verify argument as that used in Proposition 3. The key technical lesson here is that solving the model under Assumption 2 is just as easy as solving it under perfect foresight. Section 6 will extend this lesson to richer HANK models. This underscores the tractability afforded by Assumption 2.

Turning now to the economic meaning of Proposition 5, note that our earlier scenario, with inattention but no learning, is nested with $\tilde{\lambda}_t^g = \lambda^g$ and $\lambda_t^p = \lambda^g$ for all t. Relative to this case, the accommodation of learning has two implications. First, the effective discounting of future incomes and future rates now decays with the passage of time since the occurrence of the innovation. By the same token, the equilibrium spending gets closer and closer to its FIRE counterpart over time. Second, at any given point of time, the GE multiplier is less attenuated. Even if only a few agents are informed today, they believe that more agents will become informed and react to the shock in the future; they thus expect their future income to move further in the direction of the FIRE outcome, even if it doesn't move much now; they therefore also find it optimal to raise their spending relative to the case without learning (but not relative to FIRE). We will illustrate these predictions in our upcoming application to forward guidance (Section 5.2). Finally, note that sticky information a la Mankiw and Reis (2002) is nested with $\lambda_0^p = \lambda_0^g = \lambda$ and $G^p = G^g = \lambda$, for some $\lambda \in (0,1)$. Relative to this case, our generalization disentangles the actual and the perceived speeds of learning.

5 Applications

Having elucidated the common mechanisms behind a diverse set of theories, we now show how these mechanisms help understand salient business cycle facts, how they change policy predictions, and how they fix certain "pathologies" of the standard paradigm.

5.1 Monetary policy affects real activity with a lag, inflation peaks even later

In response to identified monetary shocks, real activity peaks multiple quarters later, and inflation peaks with even greater delay (e.g., Christiano, Eichenbaum, and Evans, 2005; Aruoba and Drechsel, 2024). Standard DSGE models replicate these facts with ad hoc features, such as consumption habit and price indexation. Frictions in expectations offer a compelling alternative. We already demonstrated this point for aggregate spending (recall Figure 2); so, we now focus on inflation.

To start with, let the friction be present *only* in the supply side: firms receive noisy private signals $x_{i,t} = m_t + u_{i,t}$, where $u_{i,t} \sim \mathcal{N}(0,\sigma^2)$ is the noise, whereas households are perfectly informed. Equilibrium output is then $y_t = \frac{\sigma}{1-\rho} m_t$, while equilibrium inflation solves equation (2). The latter has the same mathematical structure as the IKC (1), with π_t in place of y_t and the real marginal cost κy_t in place of r_t. Translating Proposition 2 to the present context, we infer that inflation satisfies

$$\pi_t = \kappa y_t + \alpha^f \beta \mathbb{E}_t[\pi_{t+1}] + \alpha^b \pi_{t-1},$$

where $\alpha^f < 1$ and $\alpha^b > 0$ are again the products of the informational friction. Thus, it is *as if* inflation obeys a Hybrid NKPC—and inflation exhibits a hump even if output does not.

Allowing the consumers to have incomplete information re-introduces the hump in aggregate output seen earlier in Figure 2. The hump in output translates to a hump in the real marginal costs, which in turn reinforces the hump in inflation. All in all, both output and inflation have a hump, but inflation peaks after output, just as in the data. Angeletos and Huo (2021) indeed argue that the theory performs well not only qualitatively but also quantitatively: when the level of noise is calibrated to survey evidence on forecast errors, the inflation inertia is just as large as that observed in the data.

We conclude with two observations. First, although we have illustrated the above points assuming noisy information a la Woodford (2003) and leveraging the results of Angeletos and Huo (2021), the same points could have been delivered more straightforwardly with Assumption 2. Second, just as the bite of the informational friction on output depends on households' horizons and the slope of the Keynesian cross, its bite on inflation depends on firms' horizons and the strategic complementarity in pricing. Through this channel, a lower θ (higher price flexibility) translates, perhaps paradoxically, to *more* inertia in inflation, for essentially the same reason that a lower ω (higher MPC) translates to more inertia in aggregate spending. See also Angeletos and Huo (2021), Afrouzi and Yang (2021), Werning (2022) and Cai (2024), which emphasize how relaxing FIRE interacts with market concentration, the specification of nominal rigidity, and the policy response to inflationary pressures.

5.2 Forward Guidance

As shown in Section 3, FIRE gives rise to the so-called forward guidance puzzle: news of a cut in real interest rates far in the future stimulates the economy by the same amount as a commensurate cut today. This was true despite the fact that the PE effect of forward guidance was small for realistic ω, and was explained by a large GE multiplier.[10] Moving away from FIRE, we then argued that incomplete information and its cousins arrest the GE feedback, and the more so the stronger this feedback is. It stands to logic that the forward guidance puzzle is lessened away from FIRE. Angeletos and Lian (2018) and Farhi and Werning (2019) demonstrated this idea by introducing, respectively, incomplete

[10]In Section 3, we illustrate this puzzle by considering news at $t = 0$ of a monetary expansion at $t = T$, holding real rates constant at $t \leq T - 1$. The standard puzzle is actually worse, due to the zero lower bound (ZLB) on nominal rates and the resulting feedback loop between spending, inflation, and real rates—the "deflationary spiral" in reverse. See Del Negro, Giannoni, and Patterson (2015), McKay, Nakamura, and Steinsson (2016) and Angeletos and Lian (2018).

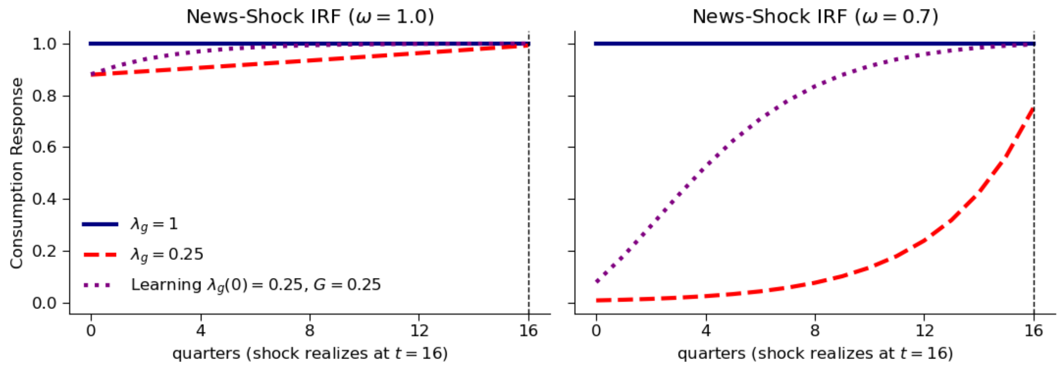

Figure 3: Unanticipated Monetary Policy Shocks Under Bounded Rationality

information and level-k thinking. Here, we first replicate the common insight of these papers with the help of Assumption 1, and then go on to study the role of learning with the help of Assumption 2.

Figure 3 revisits the policy experiment of the right panel of Figure 1: news at $t = 0$ of a persistent interest rate cut starting at $t = 16$. In particular, we now plot the output response under three informational scenarios and two values for ω. The left panel of Figure 3 sets $\omega = 1$ (RANK), the right panel set $\omega = .7$ (HANK). In each panel, the solid blue line imposes FIRE, the dashed red line imposes Assumption 1 with $\lambda^p = 1$ and $\lambda^g = .25$, and finally the dotted purple line imposes Assumption 2 with $\lambda_0^p = 1$, $\lambda_0^g = .25$, and $G^g = .25$. In the last two scenarios, households are immediately and fully aware of the news (which preserves the full PE effect), but they are also unsure about the attentiveness or the rationality of others and they therefore underestimate the response of aggregate spending to the news (which arrests the GE feedback). In the one scenario (dashed red line), this underestimation remains constant over time; in the other (dotted purple line), it decays over time, thanks to learning.

The following patterns are evident from the figure. First, letting $\lambda^g < 1$ lessens the forward-guidance puzzle: the output response is smaller than that under FIRE, and the more so the further back we go from $t = 12$. As already explained, this is because $\lambda^g < 1$ arrests the GE feedback, which itself is larger when the horizon of forward guidance is longer. Second, the resolution of the puzzle is very pronounced in HANK ($\omega = .7$), while it is rather modest in RANK ($\omega = 1$). This is because HANK allows for a larger GE multiplier and therefore also for a larger bite of the informational friction. Finally, learning speeds up the adjustment via both a mechanical and a forward-looking channel: as time passes, consumers learn more and adjust their spending in the direction of FIRE; but even before they can learn, consumers raise their spending in anticipation of the faster boom in the future.

A quantitative assessment of the role of learning in the context of forward guidance remains an open question for future work. Other related endeavors include Iovino and Sergeyev (2023) on quantitative easing and Angeletos and Sastry (2021) on optimal policy communication.

5.3 Fiscal multipliers

In this section, we turn attention to fiscal policy. We first illustrate how inattention can arrest the GE effects of government spending—echoing Angeletos and Lian (2018), Woodford and Xie (2019) and Bianchi-Vimercati, Eichenbaum, and Guerreiro (2024). We then show how inattention can reinforce the failure of Ricardian equivalence—echoing Gabaix (2020), Woodford and Xie (2022) and Eichenbaum, Guerreiro, and Obradovic (2025). Combining these insights suggests that inattention likely reduces government spending multipliers but increases transfer multipliers.

The Government-Spending Multiplier. We now extend the model to incorporate government spending and taxation, with the aim of elucidating how inattention alters the propagation of fiscal shocks. Aggregate output satisfies $y_t = c_t + g_t$, where government spending follows the process $g_t = \rho_g g_{t-1} + \varepsilon_t$, where ε_t denotes the spending shock and ρ_g the persistence of spending. To streamline the analysis, we assume a balanced budget each period, so that taxes satisfy $t_t = g_t$, and we posit a monetary policy rule of the form $m_t = \phi y_t$, which links the real interest rate to contemporaneous output.

Combining the aggregate consumption function, the the monetary policy rule, and market clearing, we obtain the following IKC:

$$y_t = c_t + g_t = \left(1 - \beta\omega\left(1 - \sigma\phi\right)\right) \sum_{k=0}^{\infty} \left(\beta\omega\right)^k \overline{E}_t[y_{t+k}] - \left(1 - \beta\omega\right) \sum_{k=0}^{\infty} \left(\beta\omega\right)^k \overline{E}_t[t_{t+k}] + g_t. \tag{16}$$

Note that a hawkish or active monetary policy (in the sense of $\phi > 0$) attenuates the Keynesian multiplier, whereas the opposite is true for an accommodative or passive monetary policy (in the sense of $\phi < 0$). By Assumption 1, $\overline{E}_t[t_{t+k}] = \overline{E}_t\left[g_{t+k}\right] = \lambda^p \mathbb{E}_t\left[g_{t+k}\right]$ and $\overline{E}_t[y_{t+k}] = \lambda^g \mathbb{E}_t\left[y_{t+k}\right]$. Replacing in (16) and solving for y_t, we conclude that that the government-spending multiplier is

$$\frac{dy_t}{dg_t} = \frac{1 - \varrho_g \lambda^p}{1 - \left(\varrho_g + \varphi\right)\lambda^g}, \tag{17}$$

where $\varrho_g \equiv \frac{1-\beta\omega}{1-\beta\omega\rho_g}$ increases with the persistence of the fiscal shock and $\varphi \equiv \frac{\beta\omega}{1-\beta\omega\rho_g}\sigma\phi$ decreases with the degree of monetary policy accommodation (i.e., it increases with ϕ).

Consider first the case in which monetary policy maintains a constant real interest rate ($\varphi = 0$). Under FIRE ($\lambda^p = \lambda^g = 1$), we obtain a unit multiplier, just as in Woodford (2011) and Bilbiie (2011). Consider next the case with lack of common knowledge (Angeletos and Lian, 2018) or level-k thinking (Bianchi-Vimercati, Eichenbaum, and Guerreiro, 2024), which by Proposition 4 can be mapped to $\lambda^p = 1$ and $\lambda^g < 1$. In this case, the fiscal multiplier becomes less than one. The mechanism is intuitive. Agents internalize that fiscal expansions necessitate higher taxes, inducing a negative wealth effect that depresses spending. Simultaneously, they anticipate the positive general-equilibrium feedback from higher output, which works in the opposite direction. Under FIRE, these two forces exactly offset, yielding a multiplier of one and leaving consumption unchanged. But when agents fail to fully internalize the aggregate consequences of fiscal policy—that is, when their perceptions of the general-equilibrium feedback are muted relative to the perceived fiscal burden—the wealth effect dominates. As a result, consumption contracts and the multiplier falls below unity.

The opposite scenario becomes possible if λ^p is sufficiently small relative to λ^g. Intuitively, if agents underestimate the tax burden of government spending enough relative to its stimulating effect on output, they may end up increasing their consumption. However, this scenario is less likely when $\phi < 0$, such as in the ZLB context.[11] The reason is that the expansionary force of government spending now operates through two general-equilibrium channels—higher income and lower real rates—and inattention dampens both of them. Accordingly, Angeletos and Lian (2018) argue that, in the ZLB context, frictions in information and rationality reduce the effectiveness of government spending in essentially the same way as they reduce the effectiveness of forward guidance for monetary policy.

The Transfer Multiplier. We now extend the model to examine the general-equilibrium effects of government transfers ("stimulus checks"). The central insight is that departures from FIRE introduce a distinct mechanism through which Ricardian Equivalence fails, thereby allowing transfers to influence economic activity in a meaningful way.

For simplicity, we assume that fiscal policy entails a one-time aggregate transfer. In the spirit of rational confusion models (see Proposition 4), we assume that the transfer received by individual i is the sum of the aggregate transfer, ϵ_0, plus an idiosyncratic transfer τ_i, which has a zero cross-sectional mean. People observe their total transfer $\epsilon_0 + \tau_i$, however only a fraction λ^p is attentive to the aggregate transfer policy ϵ_0, while the remaining $1-\lambda^p$ attribute their entire transfer to the idiosyncratic component. Agents believe a fraction λ^g is attentive, and no agent observes y_0. We assume that monetary policy maintains a constant real interest rate. Furthermore, for simplicity, we suppose that the transfer is financed with taxes at time 1, which implies that $y_t = 0$ for $t \geq 1$. These assumptions imply two distortions in expectations. First, agents underweight the magnitude of future tax burden, i.e., time 1 taxes are $t_1 = \beta^{-1}\epsilon_0$, but since only λ^p individuals are informed of the policy, it follows that the average expectation is given by $\overline{E}_0[t_1] = \lambda^p t_1$. Second, as before, individuals underestimate the general-equilibrium effect of the policy on income: $\overline{E}_0\left[y_0\right] = \lambda^g y_0$.[12]

[11] When nominal rates are constrained by the ZLB, the inflationary pressure of higher government spending translate to lower real rates. This explains why $\phi < 0$ proxies for the ZLB.

[12] To simplify the exposition, we let the fiscal shock be transitory, so the Keynesian multiplier operates merely within period 0. To make sure that the informational friction has a bite despite this simplification, we then preclude people from observing current income. The essence remains the same in more general setting where agents observe current income but shocks are persistent and the Keynesian multiplier operates over multiple periods, as in our earlier analysis; see in particular Eichenbaum, Guerreiro, and Obradovic (2025) for the application to the transfer multiplier.

In this economy, the transfer multiplier is given by:

$$\frac{dy_0}{d\epsilon_0} = \frac{\left(1-\beta\omega\right)}{1-\left(1-\beta\omega\right)\lambda^g}\left(1-\lambda^p\omega\right). \tag{18}$$

When $\lambda^g = \lambda^p = 1$, this multiplier reduces to the FIRE benchmark (e.g., Angeletos, Lian, and Wolf, 2024a,b). In this benchmark FIRE, the multiplier is strictly positive if and only if $\omega < 1$, reflecting the failure of Ricardian Equivalence due to finite horizons. Relative to this benchmark, the behavioral distortions discussed above have two distinct consequences. First, when agents are inattentive to taxes $\lambda^p < 1$, they discount the future tax burden of the current transfer and therefore consume a larger proportion of the transfer relative to FIRE. This effect increases the overall transfer multiplier. Second, when agents are inattentive to the GE consequences of the policy $\lambda^g < 1$, they respond less to the increases in equilibrium income associated with the transfer, leading to a dampening of the transfer multiplier. Eichenbaum, Guerreiro, and Obradovic (2025) develop a quantitative HANK model calibrated to microdata on planned propensity to spend out of aggregate transfers and show that the first force tends to dominate, implying that the transfer multiplier is larger with inattention.

5.4 Other lessons

We conclude this section be reviewing briefly a few additional lessons.

1. **Supply shocks are more inflationary when they are more salient.** Liu and Zhang (2024) show that when information about inflation becomes more salient, firms rationally confuse positive demand shock for negative supply shocks, and therefore raise their prices more in response to demand shocks. What is more, this mechanism is amplified by the strategic complementarity in the firms' pricing choices. All in all, this helps generate a steeper Phillips curve. In Liu (2025), a similar result obtains via a different mechanism—the interaction of consumer confusion and search frictions. When consumers are better informed about aggregate supply shocks, they are less likely to misinterpret aggregate cost-push shocks for idiosyncratic price increases and hence also less likely to search for a competitor, so firms in pass through costs more fully.

2. **The indeterminacy problem of the New Keynesian model is lessened or resolved.** Gabaix (2020) shows that cognitive discounting enlarges the region of parameters where the Taylor principle holds. This is basically an implication of adding discounting ($\alpha_f < 1$) in the Euler equation. Angeletos and Lian (2023), on the other hand, show that finite social memory of past payoff-irrelevant contingencies, or small appropriate noise as in the global-games literature (Morris and Shin, 2001, 2003), can help select a unique equilibrium regardless of whether

the Taylor principle holds or not. This is because all sunspot or "backward-looking" solutions of the New Keynesian model hinge on a fragile infinite chain: consumers may respond to sunspots today only if they expect other consumers to keep responding to them in the infinite future.

3. **Two controversial ideas—the FTPL and neo-Fisherian effects—find less space.** The equilibrium selected by finite memory or small noise in Angeletos and Lian (2023) is RANK's conventional solution (also known as its forward-looking or MSV solution), even when monetary policy is "passive" (i.e., the Taylor principle is violated). This solution preserves Ricardian equivalence and rules out deficit-driven fluctuations in output and inflation, thus also ruling out a regime with "active" fiscal policy. Put differently, there is no space left for the Fiscal Theory of the Price Level (FTPL) within RANK.[13] Garcıa-Schmidt and Woodford (2019), on the other hand, shows that "reflective equilibrium"—a close cousin of level-k thinking—can rule out a neo-Fisherian scenario, whereby interest rate hikes trigger inflationary booms.

6 HANK without FIRE

In this section, we adapt our analysis to a general class of HANK models, leveraging recent advances in the sequence-space approach (Auclert, Rognlie, and Straub, 2020). Although the essence remains largely the same, there is an important gain: by accommodating realistic behavior at the microeconomic (or PE) level, we can now model more accurately the relevant macroeconomic (or GE) feedbacks and thereby also the effects of informational frictions and/or bounded rationality. Our focus here is, not to carefully evaluate this gain, but rather to illustrate how exactly one can go after it.

We start in Section 6.1 with a brief review of the perfect-foresight version of the Intertemporal Keynesian Cross (IKC) developed in Auclert, Rognlie, and Straub (2024); solving that IKC is equivalent to characterizing the economy's linearized impulse response functions under FIRE (Boppart, Krusell, and Mitman, 2018). In Section 6.2, we next show how to adapt the IKC methodology away from FIRE, allowing for essentially arbitrary subjective expectations; this echoes our derivation of equation (1) in Section 2.1, except that now the underlying microfoundations are richer and, by consequence, the applicable versions of the aggregate consumption function and the IKC are more complicated (but still tractable and insightful). We conclude in Section 6.3 by illustrating how the IKC can be solved under our preferred specification of inattention; we thus revisit the response of the economy to a monetary policy shock, under the same departure from FIRE as in our baseline analysis, but now with a more realistic HANK microfoundation.

[13] Moving to HANK naturally allows fiscal deficits to drive aggregate demand and thereby output and inflation. This in turn can generate FTPL-like predictions, but now the underlying theoretical mechanism is different, more robust, and empirically grounded; see Angeletos, Lian, and Wolf (2024b), Angeletos et al. (2025) and Rachel and Ravn (2025).

6.1 The IKC with FIRE (equivalently, with perfect foresight)

Consider the perfect-foresight response of aggregate consumption to perturbations in aggregate income and interest rates. As shown in Auclert, Rognlie, and Straub (2024), this response is characterized by a sequence of functionals $\{\mathscr{C}_t^{\text{PF}}\}_{t=0}^{\infty}$, each of which describes aggregate consumption at any given date as a function of aggregate income and interest rates in *all* dates:

$$C_t = \mathscr{C}_t^{\text{PF}}\left(\{x_s\}_{s=0}^{\infty}\right) \qquad \text{where} \qquad \{x_s\}_{s=0}^{\infty} = \{Y_s, r_s\}_{s=0}^{\infty}. \tag{19}$$

Although the functional $\mathscr{C}_t^{\text{PF}}$ is derived from the underlying microfoundations, for our purposes it can be treated as a primitive. Compute the following partial derivatives, evaluated in steady state:

$$M_{t,s} = \left.\frac{\partial \mathscr{C}_t^{\text{PF}}\left(\{x_s\}_{s=0}^{\infty}\right)}{\partial Y_s}\right|_{\{x_s\}=\{x^{ss}\}} \qquad \text{and} \qquad M_{t,s}^r = \left.\frac{\partial \mathscr{C}_t^{\text{PF}}\left(\{x_s\}_{s=0}^{\infty}\right)}{\partial r_s}\right|_{\{x_s\}=\{x^{ss}\}}.$$

We then have the following linear approximation: $d\boldsymbol{C} = \boldsymbol{M}\cdot d\boldsymbol{Y} + \boldsymbol{M}^r\cdot d\boldsymbol{r}$, where $d\boldsymbol{X} = [dX_0, dX_1, \ldots]$ denotes deviations from steady state, $\boldsymbol{M} \equiv \left[M_{t,s}\right]$ is the sequence-space Jacobian with respect to income (or the matrix of intertemporal MPCs), and $\boldsymbol{M}^r \equiv \left[M_{t,s}^r\right]$ is the Jacobian with respect to interest rates. Imposing market clearing, we conclude that the dynamic response of aggregate output to any changes in real rates must solve the following fixed-point relation:

$$d\boldsymbol{Y} = \boldsymbol{M}\cdot d\boldsymbol{Y} + \boldsymbol{M}^r\cdot d\boldsymbol{r}. \tag{20}$$

This relation, known as the Intertemporal Keynesian Cross (IKC), is the analogue of the Dynamics IS equation and is the key to solving HANK models with the sequence-space methods. If we specify monetary policy as an exogenous impulse $d\boldsymbol{r}$, the equilibrium response of aggregate spending and output to monetary policy shock is obtained by solving (20) for $d\boldsymbol{Y}$ as a function of $d\boldsymbol{r}$. More generally, one can combine equation (20) with additional equilibrium restrictions, such as a Taylor rule for monetary and a generalized Phillips curve, to obtain the impulse response of output, inflation and interest rates to various exogenous shocks.

6.2 The IKC without FIRE

We now adapt the IKC methodology to more flexible models of expectations. This relaxes the coincidence between subjective and objective beliefs, thus also allowing for the following mechanism: unlike FIRE, the agents' forecast errors, or forecast revisions, can vary systematically (i.e., predictably) with shocks to policy or other fundamentals. This mechanism will manifest below in an appropriate modification of the IKC. At the same time, we continue to restrict attention to first-order dynamics; in the present context, this means that we allow for flexible dynamics in average forecasts but abstract from dynamics in higher moments (e.g., variation in subjective uncertainty).

The aggregate consumption function without FIRE. We start with the following assumption, which states that aggregate consumption at any given date can be expressed as a function of the realized income and interest rates thus far, the contemporaneous expectations regarding future income and interest rates, and the entire history of such expectations in the past:

Assumption 3 (General Demand Structure)**.** *There exists a functional $\mathscr{C}$ such that, for all $t \geq 0$, aggregate consumption at date t is given by*

$$C_t = \mathscr{C}\left(\left\{x_s, \left\{\overline{E}_s x_{s+h}\right\}_{h=1}^{\infty}\right\}_{s=-\infty}^{t}\right), \tag{21}$$

where $x_s \equiv (y_s, r_s)$ and where $\overline{E}_s[x_{s+h}] \equiv \int E_{i,t}[x_{s+h}]di$.

While the dependence of aggregate consumption on current fundamentals and current expectations is immediate, the inclusion of past fundamentals and past expectations warrants further discussion. This backward-looking dependence implicitly captures the influence of lagged economic conditions on the endogenous state variables—such as asset holdings—that influence current consumption decisions. That is, the dependence of c_t on the history $\left\{x_s, \left\{\overline{E}_s x_{s+h}\right\}_{h=1}^{\infty}\right\}_{s=-\infty}^{t-1}$ captures how past realizations and expectations shape the state of the economy at date t.[14]

To illustrate this point, we briefly return to our baseline OLG setting and show how to derive the relevant consumption functional in that setting. In so doing, we also clarify why the backward-looking terms could be sidestepped in this setting but not more generally.

Revisiting our OLG setting. Consider our baseline model, where aggregate consumption is given by

$$c_t = (1-\beta\omega)\left(\overline{a}_t + \sum_{h=0}^{\infty}(\beta\omega)^h \overline{E}_t\left[y_{t+h}\right]\right) - \sigma\beta\omega\sum_{h=0}^{\infty}(\beta\omega)^h \overline{E}_t\left[r_{t+h}\right], \tag{22}$$

while aggregate assets satisfy

$$\overline{a}_t = \beta^{-1}\left(y_{t-1} + \overline{a}_{t-1} - c_{t-1}\right). \tag{23}$$

In our earlier analysis, we used market clearing to replace $\overline{a}_t = 0$ in (22) and obtain an IKC that lacked backward-looking terms. Here, we instead iterate on (23) and express $\overline{a}_t$ as a function of the economy's history; replacing into (22) then yields a variant IKC, one containing backward-looking terms.

[14] Implicit in equation (22) is the assumption that agents observe their concurrent income and interest rates, stated more explicitly in Assumption 4 below.

In particular, take c_{t-1} from equation (22) evaluated at date $t-1$ and replace it in (23) to obtain

$$\overline{a}_t = \omega\overline{a}_{t-1} + \omega y_{t-1} - (1-\beta\omega)\sum_{h=1}^{\infty}(\beta\omega)^h \overline{E}_{t-1}\left[y_{t-1+h}\right] + \sigma\beta\omega\sum_{h=0}^{\infty}(\beta\omega)^h \overline{E}_{t-1}\left[r_{t-1+h}\right].$$

Next, iterate backwards and use $\lim_{T\to\infty}\omega^T\overline{a}_{t-T} = 0$ to express date-t aggregate assets as

$$\overline{a}_t = \sum_{s=-\infty}^{-1}\omega^{-s}y_{t+s} - \sum_{s=-\infty}^{-1}\omega^{-s}\left\{\left(\frac{1-\beta\omega}{\omega}\right)\sum_{h=1}^{\infty}(\beta\omega)^h\overline{E}_{t+s}\left[y_{t+s+h}\right] + \sigma\beta\sum_{h=0}^{\infty}(\beta\omega)^h\overline{E}_{t+s}\left[r_{t+s+h}\right]\right\},$$

Finally, replacing in the consumption function above, we obtain

$$\begin{aligned} c_t = (1-\beta\omega)&\left\{\sum_{s=-\infty}^{-1}\omega^{-s}y_{t+s} - \sum_{s=-\infty}^{-1}\omega^{-s}\left\{\left(\frac{1-\beta\omega}{\omega}\right)\sum_{h=1}^{\infty}(\beta\omega)^h\overline{E}_{t+s}\left[y_{t+s+h}\right] + \sigma\beta\sum_{h=0}^{\infty}(\beta\omega)^h\overline{E}_{t+s}\left[r_{t+s+h}\right]\right\}\right\} \\ &+ (1-\beta\omega)\sum_{h=0}^{\infty}(\beta\omega)^h\overline{E}_t\left[y_{t+h}\right] - \sigma\beta\omega\sum_{h=0}^{\infty}(\beta\omega)^h\overline{E}_t\left[r_{t+h}\right], \end{aligned}$$

which is directly nested in (21). In Appendix C, we also show how a full HANK model with uninsurable idiosyncratic income risk and borrowing constraints fits into the same aggregate consumption specification. The basic idea is that past realizations and past expectations generally enter current aggregate demand via their effects on relevant state variables.

Deriving the generalized IKC. We now return to the general consumption function in equation (21) and demonstrate how to obtain a tractable way of accommodating more realistic expectations formation. As a preliminary step, it is instructive to observe that the perfect foresight consumption function employed in the standard IKC formulation—equation (19)—can be viewed as a special case of the general belief-dependent consumption function. Specifically, under the assumption that agents hold steady-state expectations prior to the realization of an "MIT shock" at date $t=0$ and possess perfect foresight thereafter, the consumption function $\mathscr{C}_t^{\text{PF}}$ satisfies:

$$\mathscr{C}_t^{\text{PF}}\left(\{x_s\}_{s=0}^{\infty}\right) = \mathscr{C}\left(\left\{\{x^{ss}\}_{h=0}^{\infty}\right\}_{s=-\infty}^{-t-1}, \left\{\{x_{t+s+h}\}_{h=0}^{\infty}\right\}_{s=-t}^{0}\right).$$

The perfect-foresight IKC coefficients can be computed from the general consumption function using this fact. Given the generalized consumption function, the perfect-foresight IKC coefficients can be computed assuming that all expectations from time zero on move by the same amount, i.e.,

$$M_{t,s} = \sum_{\tau=0}^{\min\{t,s\}}\left[\frac{\partial\mathscr{C}(\cdot)}{\partial\overline{E}_\tau\left[Y_s\right]}\right]_{\left\{\overline{E}_\tau[x_{\tau+h}]\right\}=\{x^{ss}\}} \quad \text{and} \quad M^r_{t,s} = \sum_{\tau=0}^{\min\{t,s\}}\left[\frac{\partial\mathscr{C}(\cdot)}{\partial\overline{E}_\tau\left[r_s\right]}\right]_{\left\{\overline{E}_\tau[x_{\tau+h}]\right\}=\{x^{ss}\}}, \quad \forall t,s\in\{0,1,\ldots\},$$

where we henceforth use the convention $\overline{E}_t[x_s] \equiv x_s$ whenever $s \leq t$. (This convention reflects our assumption that agents observe their own assets, income and returns, although they don't extract information from them about the aggregate shocks.)

The principal complexity in extending the IKC beyond FIRE arises from the fact that, under arbitrary subjective expectations, aggregate consumption depends not only on the realized trajectory of aggregate variables, but also on the sequence of subjective beliefs that agents hold at each point in time regarding the infinite future. Auclert, Rognlie, and Straub (2020) discuss how the additional coefficients that determine the response of aggregate consumption to changes in beliefs can be derived analytically from the original sequence-space Jacobians computed under perfect foresight. This insight facilitates tractable extensions of the IKC framework to settings with realistic expectations.

Our analysis here follows Bardóczy and Guerreiro (2023). We begin by assuming that expectations prior to date 0 are fixed at their steady-state values: $\overline{E}_t x_{t+h} = x^{ss}$ for all $t < 0$. At time 0, an MIT shock occurs, generating deterministic deviations in the paths of aggregate output and interest rates. However, agents do not possess perfect foresight following the shock. Instead, they may form incorrect beliefs and subsequently revise these beliefs, in a manner that is predictable to the analyst (or the econometrician). The generalized IKC presented below accounts for how this kind of systematic errors and revisions in the agents' subjective beliefs influence the actual equilibrium dynamics.

Denote the average forecast revision as $\overline{FR}_t[x_{t+h}] \equiv \overline{E}_t[x_{t+h}] - \overline{E}_{t-1}[x_{t+h}]$.[15] The rational expectation of this object conditional on the MIT shock is identically zero when agents have FIRE, but may systematically vary with the shock when agents have incomplete information and/or mis-specified beliefs. This is the kind of predictability mentioned above. We then have the following characterization of the fixed point between actual outcomes, subjective expectations, and predictable revisions.

Proposition 6 (Generalized IKC under Realistic Expectations)**.** *The generalized IKC is given by:*

$$d\boldsymbol{C} = \boldsymbol{M} \cdot \overline{E}_0[d\boldsymbol{Y}] + \sum_{t \geq 1} \mathscr{R}_t \cdot \overline{FR}_t[d\boldsymbol{Y}] + \boldsymbol{M}^r \cdot \overline{E}_0[d\boldsymbol{r}] + \sum_{t \geq 1} \mathscr{R}_t^r \cdot \overline{FR}_t[d\boldsymbol{r}], \tag{24}$$

where the forecast-revision Jacobians are given by

$$\mathscr{R}_t = \begin{bmatrix} \mathbf{0}_{t\times t} & \mathbf{0}_{t\times\infty} \\ \mathbf{0}_{\infty\times t} & \boldsymbol{M} \end{bmatrix} \quad \text{and} \quad \mathscr{R}_t^r = \begin{bmatrix} \mathbf{0}_{t\times t} & \mathbf{0}_{t\times\infty} \\ \mathbf{0}_{\infty\times t} & \boldsymbol{M}^r \end{bmatrix},$$

with $\mathbf{0}_{r\times c}$ denoting a zero matrix with dimensions $r \times c$.

To further elucidate the structure of the Jacobians in Proposition 6, we now offer an intuitive decomposition based on the timing of belief updates. For simplicity, consider an exogenous perturba-

[15] Note that the forecast error $x_t - \overline{E}_{t-1}[x_t]$ is a special case of the forecast revision $\overline{FR}_t[x_t]$.

tion to future income dY_s; an analogous analysis applies for interest rate shocks.

Suppose that, at date 0, people hold incorrect expectations about income at some date s (i.e. $\overline{E}_0[dY_s] \neq dY_s$) and that these expectations remain unchanged until a forecast revision at some date $\tau \leq s$ (i.e., $\overline{FR}_\tau[dY_s] \neq 0$). The change in aggregate consumption at any $t \leq s$ can be decomposed to

$$dC_t = dC_t^{\text{anticipated}} + dC_t^{\text{unanticipated}},$$

where $dC_t^{\text{anticipated}}$ captures the response to the initial expectation $\overline{E}_0[dY_s]$ and $dC_t^{\text{unanticipated}}$ reflects the adjustment triggered by the forecast revision at τ. This decomposition clarifies the operational role of the Jacobians $\boldsymbol{M}$ and $\mathscr{R}_\tau$ in capturing, respectively, the effects of anticipated and unanticipated components of the shock. We now examine each component of the decomposition in detail.

Consider first the effect of the initial expectation $\overline{E}_0[dY_s]$, under the assumption that all subsequent revisions are zero. This corresponds to a persistent shift in beliefs such that $\overline{E}_t[dY_s] = \overline{E}_0[dY_s]$ for all $t \geq 0$. Under this scenario, agents experience a once-and-for-all update to their expectations at 0 and maintain perfect foresight thereafter. Consequently, the consumption response at t coincides with the response implied by the perfect-foresight IKC:

$$dC_t^{\text{anticipated}} = M_{t,s} \cdot \overline{E}_0[dY_s], \quad \text{for all } t \geq 0.$$

Next, consider the consumption response to the forecast revision at date τ about time s income $\overline{FR}_\tau[dY_s]$. Since agents do not anticipate making forecast errors, consumption behavior is unaffected prior to date τ. At time τ, agents are surprised and revise their beliefs accordingly. Importantly, because the forecast revision is predictable by the analyst but not by the agents, the consumption response mimics the response to a new MIT shock, as if $FR_\tau[dY_s]$ were a new innovation in fundamentals. It follows that

$$dC_t^{\text{unanticipated}} = \begin{cases} 0 & \text{if } t < \tau \\ M_{t-\tau,s-\tau} \cdot \overline{FR}_\tau[dY_s] & \text{if } t \geq \tau. \end{cases}$$

This shows that the forecast-revision Jacobians $\mathscr{R}_\tau$ in Proposition 6 can be constructed directly from the original Jacobians $\boldsymbol{M}$, appropriately shifted to reflect the timing of the unanticipated realizations.

Stacking these responses across all dates s and horizons h yields the generalized IKC (24). This exercise thus demonstrates the contribution of both anticipated and unanticipated changes in expectations to aggregate dynamics, and highlights how the original Jacobians $\boldsymbol{M}$ and $\boldsymbol{M}^r$ can be re-purposed to account for arbitrary belief dynamics in a transparent and tractable manner.

Using the Generalized IKC to compute IRFs. The generalized IKC (24) can be used to construct impulse response functions (IRFs) under different models of expectation formation, much like the original IKC could be used to construct IRFs under FIRE.

To this end, we continue to consider the case of a one-time innovation hitting the economy at $t = 0$, and seek to construct the IRF of consumption dC_t and aggregate output $dY_t = dC_t$. Constructing this IRF amounts to taking the *objective* expectation—i.e., the analyst's expectation—of the solution of the IKC, conditional on the assumed innovation. That is, we compute

$$\mathbb{E}_0[dC_t] \equiv \mathbb{E}[dC_t|\epsilon_0],$$

where ϵ_0 denotes the exogenous innovation—e.g., an unexpected change in the period-0 interest rate, or news about interest rate in future periods—and $\mathbb{E}[\cdot]$ denotes the FIRE operator. Note then that, while each agent in the economy does not expect to make future revisions, the analyst may well expect such revisions: we have assumed $E_{i,t}[x_{s'} - E_{i,s}[x_{s'}]] = 0$ for all $i, t, s \geq t, s' \geq s$, but have allowed $\mathbb{E}_0[x_{s'} - E_{i,s}[x_{s'}]] \neq 0$ and therefore also $\mathbb{E}_0\left[d\boldsymbol{x} - \overline{E}[d\boldsymbol{x}]\right] \neq 0$ for all $s \geq 0, s' \geq s$.[16]

This is the precise point where the deviation from FIRE has a bite. Indeed, suppose momentarily that the shock is such that $\mathbb{E}_0\left[\overline{FR}_t\left[x_s\right]\right] = 0$, which means that average forecast errors are unpredictable in the eyes of the analyst. In this knife-edge case, condition (24) gives

$$\mathbb{E}_0[d\mathbf{C}] \equiv \boldsymbol{M}\mathbb{E}_0[d\boldsymbol{Y}] + \boldsymbol{M}^r\mathbb{E}_0[d\boldsymbol{r}],$$

which means that the analyst expects the economy to respond to the period-0 innovation *exactly* as in FIRE (or perfect foresight). That is, the IRF to a given shock can differ from their FIRE counterparts only insofar as average forecast errors are predictable in the analyst's eyes (i.e., under the objective expectations and conditional on the given shock).

Apart from illustrating the meaning and use of the generalized IKC (24), this discussion underscores the following points, which directly extend from our baseline analysis to the present context:

- Empirically, what is key for aggregate dynamics—at least in the class of linearized economies considered here—is not whether the econometrician can predict *individual* forecast errors conditional on an individual's own information or own revisions, but rather whether *average* forecast errors are predictable.

- Theoretically, the model is closed by specifying the underlying shock and what the corresponding IRF of the average forecast revisions.

[16]Note that this implies that $E_{i,t}\left[FR_{i,s}\left[x_{s'}\right]\right] = 0$, i.e., forecast revisions are unpredictable by the agents themselves, although they could be predictable by the analyst $\mathbb{E}_0\left[FR_{i,s}\left[x_{s'}\right]\right] \neq 0$.

Needless to say, these points go hand-to-hand. The empirical macroeconomist may concentrate on the IRF of *average* forecast errors/revisions to identified shocks—similarly to Coibion and Gorodnichenko (2012) and Angeletos, Huo, and Sastry (2021), and unlike Bordalo et al. (2020). The theoretical macroeconomist may in turn use such evidence to guide the choice of the model of inattention or bounded rationality that needs to be "pasted" upon the preferred version of the IKC. We next illustrate how this can be done in practice, blending the IKC of this section with the insights of Section 4.

6.3 Response to a Monetary Shock

As in Section (4), we abstract from any feedback from real economic activity to real rates ($\phi = 0$) and let $r_t = -m_t$, where m_t follows an AR(1). We next abstract from any innovation in period $t \geq 1$, and concentrate on the IRF to ϵ_0, the time-0 innovation, starting from steady state ($m_{-1} = 0$).

Under FIRE, we have $\mathbb{E}_0[m_t] = m_t = \rho^t \epsilon_0$. Here we instead let agents be inattentive to the shock, as well as skeptical about others' attentiveness and/or responsiveness to it.

Assumption 4. *All agents observe the concurrent values of their own assets, income and rates of return, but do not extract information about the underlying aggregate innovations. Furthermore, whenever an innovation occurs, a randomly-selected fraction λ^p of agents—the "attentive" ones—learns the innovation immediately and perfectly, and adjust their expectations about the future in accordance to FIRE, while the rest—the "inattentive" ones—never adjust expectations. Finally, everyone believes that other agents are attentive with probability λ^g.*

This is basically the same as Assumption 1 in Section 4.2, except for a minor twist. There, we had allowed agents to be uncertain about their concurrent income and interest rates. Here, we allow agents to observe and condition their consumption on these objects, thus guaranteeing that borrowing constraints are never violated. At the same time, we preserve the deviation from FIRE by explicitly assuming that agents do not extract information about the *aggregate* shocks from their own conditions. As noted earlier, this could be rationalized by letting individual income and individual interest rates to be contaminated with large, transitory, idiosyncratic shocks—this translates to large noise in the information contained in individual outcomes about aggregate outcomes. For simplicity, we furthermore assume that the allocation of attention is uncorrelated with MPC or other characteristics in the cross section—or else the the IKC must account for this correlation, as in Guerreiro (2023). Finally, one does not have to take Assumption 4 too literally: what matters ultimately for aggregate dynamics is the responsiveness of average beliefs, which is herein parameterized by λ^p and λ^g.

Now let us guess and verify the existence of an equilibrium (i.e., a fixed point to the IKC). In particular, let $dx_t = \chi_t \cdot \epsilon_0$, for some $\{\chi_t\}$ that identifies the IRFs of output and interest rates to ϵ_0. Clearly,

$\chi_t^r = \rho^t$, that is, the real-rate IRF is just the exogenously specified AR(1) process; what we need to solve for is $\{\chi_t^Y\}$, the output IRF. By Assumption 4, the attentive agents forecast $E_{i,t}[m_s] = m_s = \rho^s \epsilon_0$ for all t and all $s > t$, whereas for the inattentive ones forecast $E_{i,t}[m_s] = 0$. By the same token,

$$\overline{E}_t[dr_s] = -\overline{E}_t[m_s] = \lambda^p m_s = \lambda^p \rho^s \epsilon_0 \ \forall t, s > t.$$

This captures the inattentiveness to the shock itself, or the unresponsiveness of first-order beliefs.

In order to derive the generalized IKC under these assumptions, it will be useful to define the myopic sequence space Jacobians to income and interest rates,

$$\mathscr{E} = \begin{bmatrix} M_{0,0} & 0 & 0 & \dots \\ M_{1,0} & M_{0,0} & 0 & \dots \\ M_{1,0} & M_{1,0} & M_{0,0} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \quad \text{and} \quad \mathscr{E}^r = \begin{bmatrix} M_{0,0}^r & 0 & 0 & \dots \\ M_{1,0}^r & M_{0,0}^r & 0 & \dots \\ M_{1,0}^r & M_{1,0}^r & M_{0,0}^r & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix},$$

respectively. These Jacobians represent the responses of the inattentive to income and interest rates.

Proposition 7. *Under assumption 4, the generalized IKC is given by*

$$d\boldsymbol{Y} = d\boldsymbol{C} = \tilde{\boldsymbol{M}} \cdot d\boldsymbol{Y} + \tilde{\boldsymbol{M}}^r \cdot d\boldsymbol{r}, \tag{25}$$

where the generalized sequence-space Jacobians to income and interest rates are given by

$$\tilde{\boldsymbol{M}} = \lambda^g \boldsymbol{M} + \left(1-\lambda^g\right)\mathscr{E} \qquad \text{and} \qquad \tilde{\boldsymbol{M}}^r = \lambda^p \boldsymbol{M}^r + \left(1-\lambda^p\right)\mathscr{E}^r + \left(\lambda^p - \lambda^g\right)(\boldsymbol{M} - \mathscr{E})\mathscr{M}^0 \mathscr{E}^r,$$

respectively, with $\mathscr{M}^0$ such that $\mathscr{M}^0(\mathbf{I} - \mathscr{E}) = \mathbf{I}$ and $\mathbf{I}$ being the identity operator. Furthermore, as long as an equilibrium exists, the impulse responses of interest rates and output to ϵ_0 are given by

$$d\boldsymbol{r} = \varrho \cdot \epsilon_0 \qquad \text{and} \qquad d\boldsymbol{Y} = \chi^Y \cdot \epsilon_0,$$

respectively, where $\varrho = \{\rho^s\}_{s=0}^{\infty}$ and χ solves the linear system: $\left(I - \tilde{\boldsymbol{M}}\right)\chi = \tilde{\boldsymbol{M}}^r \cdot \varrho$.

The results in Proposition 7 are intuitive. First, note that the effects of income on consumption are mediated by the matrix $\tilde{\boldsymbol{M}}$ which is a convex combination of the FIRE IKC $\boldsymbol{M}$ and the myopic IKC $\mathscr{E}$, with weights given by λ^g, which reflects the fact that people expect λ^g individuals to be attentive and responsive to these general-equilibrium changes. The remaining fraction $1-\lambda^g$ are thought to be inattentive and therefore will only respond to changes in equilibrium income when they observe them, i.e., they are expected to behave myopically. The first two terms of $\tilde{\boldsymbol{M}}^r$ show that a similar logic applies to the partial-equilibrium response to interest rates, except that now mediated by the true fraction of attentive individuals, λ^p. When $\lambda^p = \lambda^g$, the response to interest rates is given solely by these two terms. However, when $\lambda^p \neq \lambda^g$, there is an additional term, which reflects the fact that attentive individuals mis-forecast the change in equilibrium output as a consequence of the myopic responses of inattentive people. Finally, we can think of χ^Y, the impulse response function of output, as the solution of inverting $\left(I - \tilde{\boldsymbol{M}}\right)^{-1} \tilde{\boldsymbol{M}}^r \cdot \varrho$.[17]

This formulation underscores two key insights that we have discussed before. The direct transmission of the monetary policy shock through household consumption responses is governed by $\tilde{\boldsymbol{M}}^r$, which interpolates between the rational expectations and myopic cases based on the degree of partial equilibrium belief distortion, λ^p. The amplification mechanism of aggregate demand operates through the Keynesian cross, whose slope is determined by $\tilde{\boldsymbol{M}}$. As λ^g declines, the forward-looking response of aggregate consumption weakens, and backward-looking components increasingly dominate the propagation of shocks. This decomposition clarifies how deviations from FIRE shape both the strength and temporal structure of the macroeconomic transmission mechanism.

Both of these features echo the analysis presented in Section 4. There are, however, two substantive generalizations. First, the framework accommodates a broader class of forward-looking effects: expectations of future outcomes can influence period-t consumption in a flexible, non-parametric manner, rather than being restricted to the exponential structure imposed by the OLG environment previously considered. Second, and more fundamentally, the framework allows for backward-looking effects. That is, aggregate consumption may respond not only to expectations about the future but also to the realized history of past outcomes. This extension is reflected in the modified backward-looking MPCs (namely, $\tilde{M}_{t,t-h}$ for $h \geq 0$), which, in contrast to their FIRE counterparts, now capture the behavioral response of households to past realized surprises.

Figure 4 presents the IRFs of output following a monetary policy shock under various models of belief formation. The dark blue line corresponds to the benchmark case of full-information rational expectations (FIRE), while the red dashed line illustrates the response under inattention with $\lambda^p = \lambda^g = 0.25$. The orange dotted line shows the IRF under a model of sticky expectations, assuming a stickiness parameter $\lambda = 0.25$. Finally, the purple dash-dotted line depicts the output response under an alternative specification of cognitive discounting, where $\overline{E}_t\left[x_{t+h}\right] = m^h x_{t+h}$ and we assume $m = 0.25$. We normalize the size of the shock such that the time 0 output response under FIRE is one.

Consistent with the findings of the previous section, we observe that inattention with learning generates a hump-shaped response of aggregate demand—a feature absent in the other displayed models. This behavior arises from the interaction between attenuated forward-looking behavior and

[17] Formally this is not precise since the matrix $I - \tilde{\boldsymbol{M}}$ is not invertible, but Auclert, Rognlie, and Straub (2024) show how this nuisance can be addressed, preserving the tractability and economic interpretation of the solution.

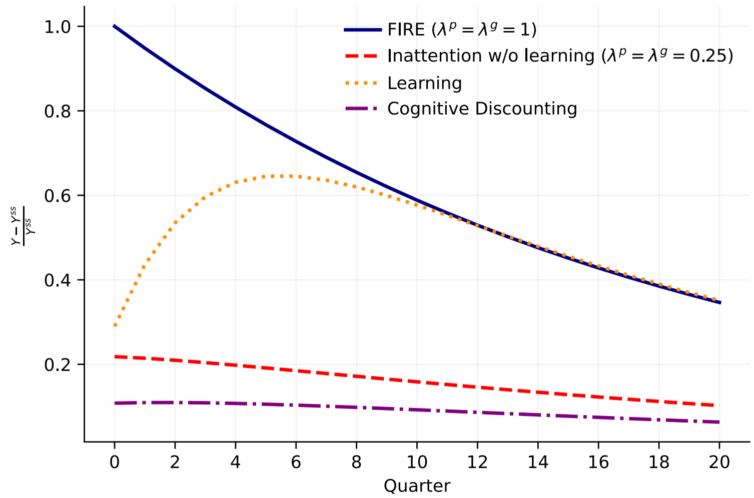

Figure 4: IRF of Output to a Monetary Policy Shock

gradual forecast revisions, which delays the propagation of monetary policy shocks. These results are consistent with Auclert, Rognlie, and Straub (2020), who show that sticky information generates a hump-shaped consumption IRF in a medium-scale HANK model, and underscore the versatility of the computational framework developed above. The generalized IKC (24) is agnostic about the expectation formation process, enabling the analyst to accommodate a wide range of belief distortions in a tractable and unified manner. The case considered here is merely an example of the much larger class of belief specifications that the framework can accommodate.

Beyond real interest rate pegs. We have posited a monetary authority that sets the real interest rate. The extension to a more general policy specification is immediate. One need only combine the IKC in Proposition 6 with the general-equilibrium relations that pin down the equilibrium real rate as a function of the monetary shock and output. As shown in Appendix C, the IKC then retains the same formal structure as in Proposition 6, with the proviso that the mappings $\boldsymbol{M}$ and $\boldsymbol{M}^r$ are redefined to incorporate the induced general-equilibrium feedback.

Forward guidance The generalized IKC framework also lends itself naturally to the analysis of alternative policy experiments, such as counterfactual interest rate paths or changes in monetary policy design. One of the key advantages of the sequence-space representation is its modularity: alternative policy scenarios can be implemented directly by modifying the paths of $d\boldsymbol{r}$ (or $d\boldsymbol{m}$), without the need to re-solve the full general-equilibrium model or introduce additional state variables associated with news shocks or more general policy regimes.

This flexibility stems from the invariance of the generalized IKC to the specification of the policy shock process, which allows the analyst to compute the IRFs under different such specifications through the same Jacobians. As a result, the framework facilitates efficient comparative statics and

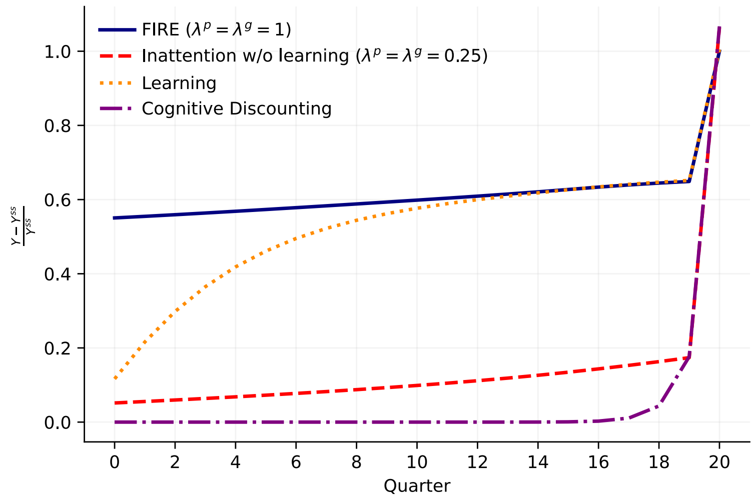

Figure 5: IRF of Output to Forward Guidance

policy analysis, even in settings with complex belief distortions or rich belief dynamics.

Figure 5 computes the IRF of output following an announcement of a future interest rate cut, to occur at time $t = 20$, under various models of belief formation. The dark blue line corresponds to the benchmark case of full-information rational expectations (FIRE), while the red dashed line illustrates the response with inattention with $\lambda^p = \lambda^g = 0.25$. The orange dotted line shows the IRF under a model with learning, which formally is a sticky expectations model with coefficient $\lambda = 0.25$. Finally, the purple dash-dotted line depicts the output response under an alternative specification of cognitive discounting, where $\overline{E}_t[x_{t+h}] = \lambda^h x_{t+h}$ and we assume $m = 0.25$. We normalize the size of the shock such that the time 20 output response under FIRE is one.

First, all alternative belief specifications display a muted initial response relative to the FIRE benchmark. This attenuation reflects the limited ability of agents to anticipate the future policy action under boundedly rational expectations. Second, the model with learning (orange dotted line), exhibits a sharply increasing response as the realization of the policy shock approaches. This dynamic pattern reflects agents' ability to progressively incorporate information over time, leading to a response profile that converges toward the FIRE outcome.

In contrast, the other two models—the basic inattention model ($\lambda = 0.25$, red dashed line) and the cognitive discounting model (purple dash-dotted line)—show persistent under-reaction to the forward guidance shock. Their IRFs remain well below the FIRE benchmark throughout the horizon, indicating that the behavioral frictions embedded in these belief structures substantially impair the anticipatory component of aggregate demand. In both cases, forward-looking expectations are too attenuated to generate significant pre-emptive responses, thereby weakening the effectiveness of forward guidance.

7 Theory Meets Evidence

This section circles back to the motivating evidence on expectations. We first ask whether the various theories reviewed in this paper have testable differences, and whether we should care about them. We then zero in on how to discipline one's preferred theory.

7.1 Testable differences—and should we care?

Throughout, we have emphasized a set of predictions that are shared by various kinds of informational friction and bounded rationality. But are there testable differences between these theories? And what's the "best" theory we should add to the mainstream business-cycle paradigm?

We offer a relatively sharp answer to the first question and a more nuanced one to the second. On the one hand, the available evidence on expectations seems to favor dispersed noisy information, along the lines of Woodford (2003), perhaps augmented with overconfidence and/or over-extrapolation (Angeletos, Huo, and Sastry, 2021; Bordalo et al., 2020; Kohlhas and Broer, 2023). This claim is based on the following three observations:

1. In response to aggregate shocks, the gap between actual outcomes and the average forecast thereof is large on impact but decays with the time since a shock has hit the economy (Coibion and Gorodnichenko, 2012; Angeletos, Huo, and Sastry, 2021). This pattern is inconsistent with the basic forms of level-k thinking, cognitive discounting, or sparsity reviewed earlier as well as with any kind *non-stationary* learning (e.g., agents growing more sophisticated or learning how to play equilibrium). Instead, it requires some form of *stationary* learning, along the lines of noisy or sticky information, or Assumption 2.[18]

2. While *average* forecasts *under-react* to aggregate shocks (Coibion and Gorodnichenko, 2012; Angeletos, Huo, and Sastry, 2021), individual forecasts do not follow a comparable pattern; if anything, *individual* forecasts appear to *over-react* to individual information as measured by past revisions (Bordalo et al., 2020; Bianchi, Ilut, and Saijo, 2024; Bianchi, Ludvigson, and Ma, 2024; Halperin and Mazlish, 2025). This pattern rejects level-k thinking and cognitive discounting, because these theories predict a commensurate under-reaction in average and individual forecasts. But it is consistent with dispersed noisy information, because this theory predicts *more* under-reaction in average forecasts. The same is true for sticky information and for our preferred specification, Assumption 2.

[18] Coibion and Gorodnichenko (2012) further argue that the evidence favors noisy information à la Woodford (2003) and Angeletos and Huo (2021) over sticky information à la Mankiw and Reis (2002): consistent with the former and in contrast to the latter, the cross-sectional dispersion of forecasts does *not* systematically increase in response to aggregate shocks. However, we see the two theories as close substitutes to each other, for the reasons discussed in Section 4.4.

On the other hand, the following three observations justify a more "forgiving," and more flexible, attitude towards all the theories under consideration:

3. There is now ample evidence that real-world firms and consumers are unaware of valuable public information, follow heuristics, and are sensitive to nudges, framing and narratives. The "noise" in expectations therefore has to be interpreted flexibly—as the product of not only dispersed private information (Lucas, 1972) or rational inattention (Sims, 2003) but also as a representation of bounded rationality.

4. Although some theories let attention be endogenous to policy, there is not enough evidence to discipline this endogeneity. Treating the level of noise, or the λ's and the G's in Assumption 2, as exogenous parameters is thus a useful benchmark (although by no means the "truth").

5. Although both of the facts about forecast errors mentioned above are important, only the first one is directly relevant for the class of linearized models used in this paper and much of the macroeconomics literature: within this class, the economy's IRFs to any aggregate shock is disciplined *exclusively* from the predictability of *average* forecasts (the focal point of Coibion and Gorodnichenko (2012) and Angeletos, Huo, and Sastry (2021)), and not by the predictability of individual forecasts (the focal point of Bordalo et al. (2020)).

These observations motivate the more nuanced answer we offer to the second question raised earlier. Insofar as one abstracts from the macroeconomic effects of time-varying uncertainty or disagreement, one may pick among the available theories, or devise new substitutes thereof, with the following three main criteria in mind: (1) tractability; (2) a meaningful distinction between direct/PE and indirect/GE effects; and (3) a good fit to the IRF of average forecast errors to aggregate shocks. Our preferred specification, Assumption 2, meets all these criteria.

7.2 Calibrating the friction

We now turn to a different question: how to discipline the key parameters in Assumption 2. The basic answer is this: by the covariation between actual outcomes and corresponding average forecasts (as measured in surveys), conditional on the shock of interest.

To fix ideas, return to the setting of Section 4, suppose that the only uncertainty is a policy shock realized at $t = 0$, and focus on the agents' expectations of r_t. In this case, we have that $\overline{E}_h[r_h] = \lambda_h^p r_h$ for all $h \geq 0$. This means that $\{\lambda_h^p\}$ can be identified from the "ratio" of two IRFs: the IRF of the average subjective forecasts of r_t to an identified monetary shock; and the IRF of the actual r_t to the same shock. A similar point applies to $\{\lambda_h^g\}$: this can be identified by the IRF of actual output and of the corresponding average forecasts to the same shock. In principle, one could therefore estimate a completely flexible time pattern for the two λ's. In practice, though, it seems more reasonable to impose a parsimonious structure, as that in Assumption 2.

Note next that, at least in principle, one can estimate a separate set of λ's for each identified shock—and one can thus shed light on whether there is more or less informational friction in response to, say, monetary policy or TFP shocks. The key is to exploit the covariation between actual outcomes and average forecasts, or average forecast errors, *conditional* on identified shocks. See Coibion and Gorodnichenko (2012) and Angeletos, Huo, and Sastry (2021) for concrete applications of this approach. A different approach that relies on the *unconditional* moments of average forecast errors (e.g., Coibion and Gorodnichenko, 2015; Bordalo et al., 2020) is harder to map to the theory, unless (i) the informational friction happens to be of the same size and persistence for all the aggregate shocks or (ii) the relevant data happen to be dominated by a single shock.

A third approach, rooted in the rational-inattention tradition (Sims, 2003; Maćkowiak, Matejka, and Wiederholt, 2023), seeks an appealing middle ground. It disciplines the cross-shock and cross-variable variation in the information parameters (λ's) by positing that these parameters are selected as the outcome of a meta-problem that trades off the costs and benefits of attention (see, e.g., Guerreiro, 2023). The guiding premise is that attention responds to incentives and is context dependent.

A small but expanding empirical literature corroborates such a middle ground. Chiang (2024) and Flynn and Sastry (2024) present evidence consistent with attention varying over the business cycle for professional forecasters and firms, respectively. Mitman et al. (2022) and Guerreiro (2023), drawing on the Survey of Consumer Expectations, show that attention covaries systematically with household characteristics—especially with exposure to shocks. Link et al. (2025) measure households' attention to macroeconomic shocks and document a systematic relation to exposure to aggregate conditions. Weber et al. (2025) use randomized experiments across time and countries to show that households are more attentive to inflation when it is high and less attentive when it is low. Finally, Coibion, Gorodnichenko, and Kumar (2018) and Afrouzi, Flynn, and Yang (2024) demonstrate that heterogeneity in firms' incentives helps account for heterogeneity in their attention to inflation. These findings suggest, not only that attention is heterogeneous, but also that it is endogenous to the economic environment, in particular to macroeconomic policy. While our methods can be extended to accommodate such endogeneity, the latter's quantitative importance remains an important open question.

8 Conclusion

In this paper, we reviewed and synthesized a class of appealing alternatives to FIRE, including noisy or sticky information, rational inattention, sparsity, level-k thinking and cognitive discounting. We distilled the common essence of these theories and developed a new specification (Assumption 2), which *tractably* and *flexibly* accommodates this essence in a large class of linear(ized) macroeconomic models, ranging from the textbook RANK to realistic HANK. We emphasized the interaction of the friction in information or rationality with other forces that regulate the GE feedback, or the strategic complementarity, such as the magnitude of the MPC and the slope of the Keynesian cross. We reviewed a variety of applied lessons for business cycles, monetary policy and fiscal policy. And we briefly discuss how to discipline the theory with survey evidence on expectations.

Throughout, we focused on how frictions in information or rationality impact the propagation of traditional demand and supply shocks, such as shocks to monetary and fiscal policy or shocks to firm costs, leaving outside the topic of how such frictions can also give rise to noise- or sentiment-driven fluctuations (Angeletos and La'O, 2010, 2013; Benhabib, Wang, and Wen, 2015; Bui, Huo, Levchenko, and Pandalai-Nayar, 2015; Huo and Takayama, 2015; Ilut and Saijo, 2021; Lorenzoni, 2009; Flynn and Sastry, 2025). Combining the two mechanisms, one can reach a certain sweet spot: the economy can simultaneously feature subdued responses to fundamentals such as shocks to TFP and monetary or fiscal policy, and excess volatility due to forces akin to animal spirits. Furthermore, as emphasized in Angeletos, Collard, and Dellas (2018, 2020), such sentiment-driven fluctuations can generative positive co-movement in hours, consumption and investment without either TFP shocks or nominal rigidity—put differently, they can help rationalize demand-driven business cycles along the flexible-price core of the New Keynesian framework. This in turn can not only improve this framework's overall empirical performance but also explain why the actual business cycles may be best described as the product of shocks to "aggregate demand" even if monetary policy is optimally set. Further exploring the *joint* quantitative relevance of the two key mechanisms described here—fundamental inertia and non-fundamental volatility—remains a fruitful direction for future research.

We conclude with an obvious warning: notwithstanding the advantages of our preferred specification (Assumption 2), this is not a panacea. First, the precise translation between Assumption 2 and its various cousins (noisy information, level-k thinking, etc) is not invariant to policy parameters. Second, this assumption abstracts from the likely endogeneity of the informational friction, which instead is a central theme of the literature on Rational Inattention (Sims, 2003; Maćkowiak, Matejka, and Wiederholt, 2023). Third, this assumption rules out confusion—rational or otherwise—of different aggregate shocks (Chahrour and Gaballo, 2018; Liu and Zhang, 2024). Fourth, this assumption also rules out "diagnostic expectations" and related forms of over-extrapolation (Bordalo et al., 2019, 2020; Angeletos, Huo, and Sastry, 2021; Bianchi, Ilut, and Saijo, 2024). Finally, our analysis abstracts from the role of policy communication (e.g., Amador and Weill, 2010; Chahrour, 2014; Bassetto, 2019; Angeletos and Sastry, 2021) and the related evidence on information treatments (e.g., Coibion, Gorodnichenko, and Ropele, 2020; Coibion, Gorodnichenko, and Weber, 2022; Coibion and Gorodnichenko, 2026). While these limitations may take center stage in some contexts, it seemed reasonable to abstract from them for the questions in this paper. Furthermore, while the available survey evidence helps calibrate the relevant information parameters along the lines discussed in the previous section, it offers relatively little guidance on counterfactuals (i.e., on how these parameters may or may not change with changes in the monetary policy rule or the structure of the economy). Finally, one can readily combine Assumption 2 with misspecified beliefs about the structure of the economy or the underlying shock process; for instance, over-extrapolation can be accommodated by letting the perceived persistence $\hat{\rho}$ differ from the actual persistence ρ, and then discipline the difference from the relevant fact in Angeletos, Huo, and Sastry (2021). All in all, we therefore believe that our paper's lens and tools provide researchers with an efficient technology for augmenting macroeconomic models with realistic adjustment frictions in expectations, for understanding their PE and GE effects, and for quantifying their bite on actual outcomes. But there is ample room for further research in understanding how the additional forces just mentioned interact with those at the heart of this paper.

References

Afrouzi, Hassan, Joel P Flynn, and Choongryul Yang. 2024. "What can measured beliefs tell us about monetary non-neutrality?" *NBER Working Paper* .

Afrouzi, Hassan and Choongryul Yang. 2021. "Dynamic rational inattention and the Phillips Curve." *CESifo Working Paper* .

Aguiar, Mark, Manuel Amador, and Cristina Arellano. 2024. "Pareto improving fiscal and monetary policies: Samuelson in the New Keynesian model." *NBER Working Paper* .

Amador, Manuel and Pierre-Olivier Weill. 2010. "Learning from prices: Public communication and welfare." *Journal of Political Economy* 118 (5):866–907.

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):3030–3070.

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. 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* .

———. 2018. "Forward guidance without common knowledge." *American Economic Review* 108 (9):2477–2512.

———. 2023. "Determinacy without the Taylor principle." *Journal of Political Economy* 131 (8):2125–2164.

Angeletos, George-Marios, Chen Lian, and Christian K Wolf. 2024a. "Can deficits finance themselves?" *Econometrica* 92 (5):1351–1390.

———. 2024b. "Deficits and inflation: HANK meets FTPL." *NBER Working Paper* .

Angeletos, George-Marios, Chen Lian, Christian K Wolf, and Dalton R Zhang. 2025. "On the (Im)Plausibility of fiscal dominance in RANK." *Working Paper* .

Angeletos, George-Marios and Karthik A Sastry. 2021. "Managing expectations: Instruments versus targets." *Quarterly Journal of Economics* 136 (4):2467–2532.

Aruoba, S Borağan and Thomas Drechsel. 2024. "The long and variable lags of monetary policy: Evidence from disaggregated price indices." *Journal of Monetary Economics* 148:103635.

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* .

———. 2023. "New pricing models, same old Phillips Curves?" *Quarterly Journal of Economics* 139 (1):121–186.

———. 2024. "The Intertemporal Keynesian Cross." *Journal of Political Economy* 132 (12):4068–4121.

Bardóczy, Bence and Joao Guerreiro. 2023. "Unemployment insurance in macroeconomic stabilization with imperfect expectations." *Working Paper* .

Bassetto, Marco. 2019. "Forward guidance: Communication, commitment, or both?" *Journal of Monetary Economics* 108:69–86.

Benhabib, Jess, Pengfei Wang, and Yi Wen. 2015. "Sentiments and aggregate demand fluctuations." *Econometrica* 83 (2):549–585.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo. 2024. "Diagnostic business cycles." *Review of Economic Studies* 91 (1):129–162.

Bianchi, Francesco, Sydney C Ludvigson, and Sai Ma. 2024. "What hundreds of economic news events say about belief overreaction in the stock market." *NBER Working Paper* .

Bianchi-Vimercati, Riccardo, Martin Eichenbaum, and Joao Guerreiro. 2024. "Fiscal stimulus with imperfect expectations: Spending vs. tax policy." *Journal of Economic Theory* 217:105814.

Bilbiie, Florin O. 2011. "Nonseparable preferences, Frisch labor supply, and the consumption multiplier of government spending: One solution to a fiscal policy puzzle." *Journal of Money, Credit and Banking* 43 (1):221–251.

Blanchard, Olivier J. 1985. "Debt, deficits, and finite horizons." *Journal of Political Economy* 93 (2):223–247.

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT shocks in heterogeneous-agent economies: The impulse response as a numerical derivative." *Journal of Economic Dynamics and Control* 89:68–92.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in macroeconomic expectations." *American Economic Review* 110 (9):2748–2782.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J. Terry. 2019. "Real credit cycles." *Working Paper* .

Bui, Ha, Zen Huo, Andrei Levchenko, and Nitya Pandalai-Nayar. 2015. "Noisy Global Chains." *miméo* .

Cai, Michael. 2024. "Explaining the macroeconomic inertia puzzle." *Working Paper* .

Carroll, Christopher D, Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N White. 2020. "Sticky expectations and consumption dynamics." *American Economic Journal: Macroeconomics* 12 (3):40–76.

Chahrour, Ryan. 2014. "Public communication and information acquisition." *American Economic Journal: Macroeconomics* 6 (3):73–101.

Chahrour, Ryan and Gaetano Gaballo. 2018. "Learning from prices: Amplification and business fluctuations." *Working Paper* .

Chiang, Yu-Ting. 2024. "Attention and a paradox of uncertainty." *Working Paper* .

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of Political Economy* 113 (1):1–45.

Coibion, Olivier and Yuriy Gorodnichenko. 2012. "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy* 120 (1):116–159.

———. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review* 105 (8):2644–2678.

———. 2026. *Expectations matter: The new causal macroeconomics of surveys and experiments.* Princeton University Press.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. "How do firms form their expectations? New survey evidence." *American Economic Review* 108 (9):2671–2713.

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele. 2020. "Inflation expectations and firm decisions: New causal evidence." *Quarterly Journal of Economics* 135 (1):165–219.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2022. "Monetary policy communications and their effects on household inflation expectations." *Journal of Political Economy* 130 (6):1537–1584.

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2015. "The forward guidance puzzle." *FRB of New York Working Paper* .

Eichenbaum, Martin, Joao Guerreiro, and Jana Obradovic. 2025. "Ricardian non-equivalence." *Working Paper* .

Farhi, Emmanuel and Iván Werning. 2019. "Monetary policy, bounded rationality, and incomplete markets." *American Economic Review* 109 (11):3887–3928.

Flynn, Joel P and Karthik Sastry. 2024. "Attention cycles." *NBER Working Paper* .

———. 2025. "Contagious business cycles." *Working Paper* .

Gabaix, Xavier. 2014. “A sparsity-based model of bounded rationality.” *Quarterly Journal of Economics* 129 (4):1661–1710.

———. 2016. “Behavioral macroeconomics via sparse dynamic programming.” *NBER Working Paper* .

———. 2019. “Behavioral inattention.” In *Handbook of Behavioral Economics*, vol. 2. North-Holland, 261–343.

———. 2020. “A behavioral New Keynesian model.” *American Economic Review* 110 (8):2271–2327.

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.

Guerreiro, Joao. 2023. “Belief disagreement and business cycles.” *Working Paper* .

Halperin, Basil and J Zachary Mazlish. 2025. “Overreaction and forecast horizon: Longer-term expectations overreact more, shorter-term expectations drive fluctuations.” *Working Paper* .

Huo, Zen and Naoki Takayama. 2015. “Higher Order Beliefs, Confidence, and Business Cycles.” miméo, Yale University.

Ilut, Cosmin and Hikaru Saijo. 2021. “Learning, confidence, and business cycles.” *Journal of Monetary Economics* 117:354–376.

Ilut, Cosmin L, Ralph Luetticke, and Martin Schneider. 2025. “HANK’s response to aggregate uncertainty in an estimated business cycle model.” *NBER Working Paper* .

Iovino, Luigi and Dmitriy Sergeyev. 2023. “Central bank balance sheet policies without rational expectations.” *Review of Economic Studies* 90 (6):3119–3152.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante. 2018. “Monetary policy according to HANK.” *American Economic Review* 108 (3):697–743.

Kohlhas, Alexandre and Tobias Broer. 2023. “Forecaster (mis-)behavior.” *Review of Economics and Statistics* .

Liao, Junyi. 2025. “Adaptive expectations and reactions to information.” *Working Paper* .

Link, Sebastian, Andreas Peichl, Oliver Pfäuti, Chris Roth, and Johannes Wohlfart. 2025. “Attention to the macroeconomy.” *Working Paper* .

Liu, Xiaojie. 2025. "Consumer information friction and monetary non-neutrality: A sequential search approach." *Working Paper* .

Liu, Xiaojie and Dalton Rongxuan Zhang. 2024. "Demand or supply? Expectations and inflation under confusion." *Working Paper* .

Lorenzoni, Guido. 2009. "A Theory of Demand Shocks." *American Economic Review* 99 (5):2050–84.

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

Maćkowiak, Bartosz, Filip Matejka, and Mirko Wiederholt. 2023. "Rational inattention: A review." *Journal of Economic Literature* 61 (1):226–273.

Maćkowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal sticky prices under rational inattention." *American Economic Review* 99 (3):769–803.

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

Mankiw, N. Gregory and Ricardo Reis. 2002. "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* :1295–1328.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The power of forward guidance revisited." *American Economic Review* 106 (10):3133–3158.

Mei, Pierfrancesco and Lingxuan Wu. 2024. "Thinking about the economy, deep or shallow?" *Working Paper* .

Mitman, Kurt, Tobias Broer, Alexandre Kohlhas, and Kathrin Schlafmann. 2022. "Expectation and wealth heterogeneity in the macroeconomy." *Working Paper* .

Morris, Stephen and Hyun Song Shin. 1997. "Approximate common knowledge and co-ordination: Recent lessons from game theory." *Journal of Logic, Language and Information* 6 (2):171–190.

———. 2001. "Rethinking multiple equilibria in macroeconomic modeling." In *NBER Macroeconomics Annual 2000, Volume 15*. MIT Press, 139–182.

———. 2002. "Social value of public information." *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.” *American Economic Review* :152–157.

Pfäuti, Oliver and Fabian Seyrich. 2022. “A behavioral heterogeneous agent New Keynesian model.” *DIW Berlin Discussion Paper* .

Rachel, Lukasz and Morten O Ravn. 2025. “Brothers in arms: Monetary-fiscal interactions without Ricardian equivalence.” *Working Paper* .

Rubinstein, Ariel. 1989. “The electronic mail game: Strategic behavior under “almost common knowledge”.” *American Economic Review* :385–391.

Sims, Christopher A. 1993. “Rational expectations modeling with seasonally adjusted data.” *Journal of Econometrics* 55 (1–2):9–19.

———. 2003. “Implications of rational inattention.” *Journal of Monetary Economics* 50 (3):665–690.

———. 2010. “Rational inattention and monetary economics.” *Handbook of Monetary Economics* 3:155–181.

Townsend, Robert M. 1983. “Forecasting the forecasts of others.” *Journal of Political Economy* 91 (4):546–588.

Weber, Michael, Bernardo Candia, Hassan Afrouzi, Tiziano Ropele, Rodrigo Lluberas, Serafin Frache, Brent Meyer, Saten Kumar, Yuriy Gorodnichenko, Dimitris Georgarakos et al. 2025. “Tell me something I don’t already know: Learning in low- and high-inflation settings.” *Econometrica* 93 (1):229–264.

Werning, Iván. 2015. “Incomplete markets and aggregate demand.” *NBER Working Paper* .

———. 2022. “Expectations and the rate of inflation.” *NBER Working Paper* .

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* .

———. 2011. “Simple analytics of the government expenditure multiplier.” *American Economic Journal: Macroeconomics* 3 (1):1–35.

———. 2018. “Monetary policy analysis when planning horizons are finite.” In *NBER Macroeconomics Annual 2018, Volume 33*. University of Chicago Press.

Woodford, Michael and Yinxi Xie. 2019. “Policy options at the zero lower bound when foresight is limited.” In *AEA Papers and Proceedings*, vol. 109. 433–437.

———. 2022. “Fiscal and monetary stabilization policy at the zero lower bound: Consequences of limited foresight.” *Journal of Monetary Economics* 125:18–35.

Online Appendices for
From RANK to HANK, without FIRE

George-Marios Angeletos   Joao Guerreiro   Dalton Rongxuan Zhang

A Aggregate demand and supply without FIRE

In this appendix, we show how to derive the building blocks of our baseline model—equations (1) and (2), our versions of the IKC and the Phillips curve, or of "aggregate demand" and "aggregate supply", respectively. Throughout, we will use uppercase variables to indicate levels; unless indicated otherwise, lowercase variables denote log-deviations from the economy's deterministic steady state.

Equation (1): aggregate demand without FIRE

We start with our model's household block, which is basically the same OLG setting as that studied in Section VII o Angeletos and Huo (2021).[19] The economy is populated by a unit continuum of households. A household survives from one period to the next with probability $\omega \in (0,1]$ and is replaced by a new one whenever it dies. Households have standard separable preferences regarding consumption and labor, and do not consider the utility of future households that replace them. The expected utility of any (alive) household i in period $t \in \{0,1,\dots\}$ is hence

$$\mathbb{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega)^k\left[u(C_{i,t+k})-v(L_{i,t+k})\right]\right],$$

where $C_{i,t+k}$ and $L_{i,t+k}$ denote household i's consumption and labor supply in period $t+k$ (conditional on survival), $u(C) \equiv \frac{C^{1-\frac{1}{\sigma}}-1}{1-\frac{1}{\sigma}}$, $v(L) = \iota\frac{L^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}$.

Households can save and borrow through an actuarially fair, risk-free, real annuity, backed by government bonds. Conditional on survival, households receive a real return R_t/ω. Households furthermore receive labor income and dividend income $W_tL_{i,t}$ and Q_t (both in real terms). We furthermore assume that all households receive identical shares of dividends, and abstract from heterogeneity in labor supply, with labor supply intermediated by labor unions that demand identical hours worked from all households $L_{i,t} = L_t$. Therefore, $Y_t \equiv W_tL_t + Q_t$ is the household's total real income. This

[19] Angeletos and Huo (2021, Section VII) accommodate multiple types of households, with different life expectancies (different ω's) and therefore different MPCs. Here, we assume a single type of households, all with the same ω below. See also Angeletos and Lian (2018) for the special case of $\omega = 1$ (i.e., infinite horizons, albeit with incomplete information).

setup implies that the steady state of our model is the same as its RANK counterpart. In particular, the steady-state rate of interest (in the steady state around which we log-linearize) is β^{-1}.

$$A_{i,t+1} = \underbrace{\frac{R_t}{\omega}}_{\text{annuity}} \left(A_{i,t} + Y_{i,t} - C_{i,t}\right),$$

where $A_{i,t}$ denotes household i's real wealth at the beginning of date t (exclusive of social fund payments). Log-linearizing the household budget yields, for $k \geq 0$,

$$a_{i,t+k+1} = \frac{1}{\beta\omega}\left(a_{i,t+k} + y_{i,t+k} - c_{i,t+k}\right),$$

Household i's optimal consumption in period $t+k$ for $k \geq 0$ satisfies the following Euler equation:

$$u'\left(C_{i,t}\right) = E_{i,t}\left[\beta\omega\frac{R_t}{\omega}u'\left(C_{i,t+1}\right)\right].$$

This is the same as the textbook counterpart, except that $E_{i,t}$ need not coincide with FIRE. After log-linearization, this becomes

$$c_{i,t} = -\sigma r_t + \mathbb{E}_t\left[c_{i,t+1}\right].$$

Log-linearizing i's intertemporal budget constraint, substituting the above, and solving for $c_{i,t}$, we arrive at the following expression for $i's$ optimal consumption function (log-linearized around the steady state):

$$c_{i,t} = \left(1-\beta\omega\right)\left(a_{i,t} + \mathbb{E}_t\left[\sum_{k=0}^{\infty}\left(\beta\omega\right)^k\left(y_{i,t+k}\right)\right]\right) - \beta\sigma\omega\mathbb{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega)^k r_{t+k}\right]$$

Finally, by adding shocks to intertemporal preferences (or life expectancy), we can readily generalize the above to

$$c_{i,t} = \left(1-\beta\omega\right)\left(a_{i,t} + \mathbb{E}_t\left[\sum_{k=0}^{\infty}\left(\beta\omega\right)^k\left(y_{i,t+k}\right)\right]\right) - \beta\sigma\omega\mathbb{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega)^k r_{t+k}\right] + \nu_{i,t}, \tag{26}$$

where $\nu_{i,t}$ is an exogenous shock.

Equation (26) is a natural extension of the Permanent Income Hypothesis (PIH) and of the consumption function assumed in RANK. In this benchmark, equation (26) together with market clearing is equivalent to the familiar Euler equation. Here, we have unpacked the *individual* consumption function that lies behind the *aggregate* Euler equation, and have extended it with arbitrary subjective expectations. To see the connection between equation (26) and the familiar Euler equation, note that

RANK is nested by letting $\omega = 1$, abstracting from idiosyncratic shocks, and imposing $E_{i,t} = \mathbb{E}^*_t$, where $\mathbb{E}^*_t$ stands for the FIRE operator. Equation (26) then becomes

$$c_t = (1-\beta)\left(a_t + \mathbb{E}^*_t\left[\sum_{k=0}^{\infty}\beta^k y_{t+k}\right]\right) - \sigma\beta\mathbb{E}^*_t\left[\sum_{k=0}^{\infty}\beta^k r_{t+k}\right] + v_t.$$

Next, combine this equation with market clearing ($y_t = c_t$ and $a_t = 0$) and separate the period-t terms on the right-hand side from all future terms to arrive at

$$c_t = (1-\beta)c_t - \sigma\beta r_t + \underbrace{\mathbb{E}^*_t\left[(1-\beta)\sum_{k=1}^{\infty}\beta^k c_{t+k}\right] - \sigma\beta\mathbb{E}^*_t\left[\sum_{k=1}^{\infty}\beta^k r_{t+k}\right]}_{\mathbb{E}^*_t[c_{t+1}-v_{t+1}]} + v_t$$

Finally, rearrange to get

$$c_t = -\sigma r_t + \mathbb{E}^*_t[c_{t+1}] + \tilde{v}_t,$$

where $\tilde{v}_t \equiv v_t - \mathbb{E}^*_t[v_{t+1}]$. This explains, not only how the consumption function (26) relates to the familiar Euler equation in RANK, but also how that this equation has a subtler meaning than the one found in textbook treatments: the *aggregate* Euler equation is, in effect, a recursive representation of the Intertemporal Keynesian Cross (IKC) under FIRE. Let us now construct the IKC *away* from FIRE. Temporally, we allow for the possibility that interest rates and incomes may differ across individuals. Aggregating (26) across i, we obtain the following expression for aggregate consumption:

$$c_t = (1-\beta\omega)\,a_t + (1-\beta\omega)\sum_{k=0}^{\infty}(\beta\omega)^k \underbrace{\int E_{i,t}\left[y_{i,t+k}\right]di}_{\text{avg. sentiment about } y} - \sigma\beta\omega\sum_{k=0}^{\infty}(\beta\omega)^k \underbrace{\int E_{i,t}\left[r_{i,t+k}\right]di}_{\text{avg. sentiment about } r} + v_t \qquad (27)$$

where $x_t \equiv \int x_{i,t}di$ for any $x \in \{c,a,y,r,v\}$. In this equation, $\int E_{i,t}[y_{i,t+k}]di$ and $\int E_{i,t}\left[r_{i,t+k}\right]di$ capture the role of "consumer sentiment" in shaping aggregate spending in the economy. The following subtlety is then important to note. The sentiment variables seen in equation (27) are the *average* expectations of *individual* income and real rates. These are distinct from the following variants, which measure the *average* expectations of the corresponding *aggregate* variables:

$$\overline{E}_t y_{t+k} \equiv \int E_{i,t}\left[y_{t+k}\right]di \qquad \text{and} \qquad \overline{E}_t r_{t+k} \equiv \int E_{i,t}\left[r_{t+k}\right]di. \qquad (28)$$

In equation (27), we ask each consumer how she feels about her *own* economic outlook; in the variants defined above, we instead ask each consumer how she feels about the aggregate outlook.

Although conceptually distinct, these two kinds of consumer sentiment are positively correlated in practice. In line with much of the related literature, we thus abstract from any difference between these measures: we let $\int E_{i,t}\left[y_{i,t+k}\right]di = \overline{E}_t y_{t+k}$ and $\int E_{i,t}\left[r_{i,t+k}\right]di = \overline{E}_t r_{t+k}$. Using this restriction in

(27) and imposing market clearing ($c_t = y_t$ and $a_t = 0$), we arrive equation (1), which is the applicable version of the DIS equation or the IKC away from FIRE.

Equation (2): aggregate supply without FIRE

Let us now turn attention to the supply side of our baseline model, which mirrors that in Section VI of Angeletos and Huo (2021). As usual, inflation is given by

$$\pi_t = (1-\theta)\left(p_t^* - p_{t-1}\right), \tag{29}$$

where $p_t^* \equiv \int p_{j,t}^* dj$ is the average reset price and θ is the Calvo parameter. For any firm j that gets to adjust its price in period t, the optimal rest price equals a markup over the expected nominal marginal cost during the lifespan of the new price. Expressed in log-deviations, the optimal reset price is

$$p_{j,t}^* = \left(1-\beta\theta\right) E_{j,t}\left[\sum_{k=0}^{+\infty}(\beta\theta)^k\left(mc_{j,t+k} + p_{t+k}\right)\right] \tag{30}$$

where $E_{j,t}$ is the firm's subjective expectation, $mc_{j,t}$ is firm's real marginal cost, p_t is the aggregate price level, and θ is the probability that a firm can not adjust its price in any given period. Subtracting p_{t-1} from both sides of (30), using the simplifying assumption p_{t-1} is known at t, aggregating across j, and combining with (29), we reach the following version of the NKPC:

$$\pi_t = \tilde{\kappa}\left\{\sum_{k=0}^{\infty}(\beta\theta)^k \int E_{i,t}\left[mc_{i,t+k}\right] dj\right\} + \frac{1-\theta}{\theta}\sum_{k=1}^{\infty}(\beta\theta)^k \overline{E}_t \pi_{t+k} \tag{31}$$

where $\tilde{\kappa} = \frac{(1-\theta)(1-\beta\theta)}{\theta}$ is the slope of the New Keynesian Phillips curve (NKPC) with respect to marginal costs and where $\int E_{i,t}\left[mc_{i,t+k}\right] dj$ captures "producer sentiment" about marginal costs. Similarly to the consumer counterpart, this is the *average* expectation about *individual* costs, which in principle could differ from the average expectations sentiment about *aggregate* costs. But we again assume that the two forms of sentiment coincide, so that we can replace $\int E_{i,t}\left[mc_{i,t+k}\right] dj$ with $\overline{E}_t[mc_{t+k}]$, the average expectation of the aggregate real marginal cost. Finally, we express the aggregate real marginal cost as $mc_t = \zeta y_t + \tilde{u}_t$, where $\zeta > 0$ is the elasticity of the real marginal cost with respect to output and $\tilde{u}_t$ is an exogenous TFP or cost-push shock. Marginal cost depends on aggregate output because, in the standard New Keynesian model, it is tied to the real wage, which in turn varies with output through households' labor supply decisions. Combining and defining $\kappa \equiv \tilde{\kappa}\zeta$ and $u_t \equiv \sum_{k=0}^{\infty}(\beta\theta)^k \overline{E}_t[u_{t+k}]$, we arrive at equation (2), the version of the NKPC stated in the main text.

B Proofs for Sections 3–5

In this Appendix we collect the proofs of all the formal results stated in Sections 3–5. The proofs for Section 6 are contained in Appendix C below.

Proof of Proposition 1

The proof follows from the main text.

Proof of Proposition 2

This follows from Proposition 3 and Section VII in Angeletos and Huo (2021).

Proof of Proposition 3

In this proof, we solve for the impulse response function to an unanticipated monetary policy innovation. See Section 6 for further discussion of the general stochastic economy. The proof proceeds in two steps. First, we show that the impulse response function can be expressed as a modified intertemporal Keynesian cross (IKC). Then, we use this equivalence to derive the impulse responses.

First Part We begin by noting that the optimal consumption rule of both informed and uninformed households satisfies

$$c_{i,t}^{\text{type}} = (1-\beta\omega)\left(a_{i,t}^{\text{type}} + E_{i,t}^{\text{type}}\left[\sum_{k=0}^{\infty}(\beta\omega)^k y_{t+k}\right]\right) - \sigma\beta\omega E_{i,t}^{\text{type}}\left[\sum_{k=0}^{\infty}(\beta\omega)^k r_{t+k}\right],$$

where type $\in \{I, U\}$ denotes informed and uninformed households. Aggregating across types and using market clearing, aggregate consumption satisfies

$$c_t = \lambda^p c_t^A + (1-\lambda^p)\, c_t^I = (1-\beta\omega)\,\overline{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega)^k y_{t+k}\right] - \sigma\beta\omega\overline{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega)^k r_{t+k}\right].$$

We proceed by a guess-and-verify approach. Guess that each household follows

$$c_{i,t} = \alpha_1 a_{i,t} + \alpha_2 E_{i,t}^{\text{type}}[m_t],$$

with (α_1, α_2) to be determined in equilibrium. With bonds in zero net supply, $c_t = \alpha_2 \overline{E}_t[m_t]$ and $E_{i,t}^{\text{type}}[c_t] = \alpha_2 E_{i,t}^{\text{type}}\overline{E}_t[m_t]$. Since each household believes that any other agent learns the realization of m_t with probability λ^g, $E_{i,t}^{\text{type}}\overline{E}_t[m_t] = \lambda^g E_{i,t}[m_t]$. At the same time, the actual fraction of informed households that observe m_t is λ^p, so $\overline{E}_t[m_t] = \lambda^p m_t$. This implies that expected future income and interest rates satisfy $\overline{E}_t[c_{t+k}] = \lambda^g \alpha_2 \overline{E}_t[m_{t+k}] = \lambda^g c_{t+k}$ and $\overline{E}_t[m_{t+k}] = \lambda^p m_{t+k}$. Substituting these expressions into the aggregate consumption function gives

$$c_t = (1-\beta\omega)\left[\sum_{k=0}^{\infty}(\beta\omega)^k \lambda^g c_{t+k}\right] + \sigma\beta\omega\left[\sum_{k=0}^{\infty}(\beta\omega)^k \lambda^p m_{t+k}\right],$$

The economy is therefore equivalently characterized by a modified IKC, featuring dampened expected paths for income and interest rates.

Second Part Now to solve for the impulse responses, we substitute the conjecture into the modified IKC. Using the law of motion for the monetary policy shock, $r_{t+k} = -m_{t+k} = -\rho^k m_t$, the aggregate consumption condition becomes

$$\alpha_2 \lambda^p m_t = (1-\beta\omega)\left[\sum_{k=0}^{\infty}(\beta\omega\rho)^k \alpha_2 \lambda^p \lambda^g m_t\right] + \sigma\beta\omega\left[\sum_{k=0}^{\infty}(\beta\omega\rho)^k \lambda^p m_t\right].$$

Solving gives $\alpha_2 = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}$. Finally, matching coefficients in the individual policy rule confirms $\alpha_1 = 1-\beta\omega$. Therefore, the aggregate consumption response satisfies

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda^p m_t.$$

Proof of Corollary 1

From the proof of Proposition 3, aggregate demand under Assumption 1 satisfies

$$c_t = (1-\beta\omega)\left[\sum_{k=0}^{\infty}(\beta\omega)^k \lambda^g c_{t+k}\right] + \sigma\beta\omega\left[\sum_{k=0}^{\infty}(\beta\omega)^k \lambda^p m_{t+k}\right].$$

With market clearing $y_t = c_t$, rearranging the above equation in a recursive form gives

$$y_t = \frac{\sigma\beta\omega\lambda^p}{\left(1-(1-\beta\omega)\lambda^g\right)} m_t + \frac{\beta\omega}{\left(1-(1-\beta\omega)\lambda^g\right)} E_t y_{t+1}.$$

By Proposition 3 and using $E_t m_{t+1} = \rho m_t$, we obtain

$$E_t y_{t+1} = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda^p \rho m_t.$$

Substituting this expression into the previous equation and using $r_t = -m_t$ yields

$$y_t = -\sigma r_t + \alpha_f \mathbb{E}_t y_{t+1},$$

where

$$\alpha_f = A(\lambda^p, \lambda^g) = \frac{\beta\omega\left(\lambda^p + \rho\right) - \left(1 - \left(1 - \beta\omega\right)\lambda^g\right)}{\beta\omega\,\lambda^p\,\rho}.$$

Setting $\left(\lambda^p, \lambda^g\right) = (1,1)$ gives $\alpha_f = A(1,1) = 1$ and hence $y_t = -\sigma r_t + E_t y_{t+1}$. Finally, note that

$$\frac{\partial \alpha_f}{\partial \lambda^p} = \frac{1 - \beta\omega\rho - (1-\beta\omega)\lambda^g}{\beta\omega\,\rho\,(\lambda^p)^2} > 0, \qquad \frac{\partial \alpha_f}{\partial \lambda^g} = \frac{1-\beta\omega}{\beta\omega\,\lambda^p\,\rho} > 0.$$

And

$$\frac{\partial^2 \alpha_f}{\partial \lambda^p\,\partial\omega} = \frac{\lambda^g - 1}{\beta\,\rho\,(\lambda^p)^2\,\omega^2} \leq 0, \qquad \frac{\partial^2 \alpha_f}{\partial \lambda^g\,\partial\omega} = -\frac{1}{\beta\,\rho\,\lambda^p\,\omega^2} < 0.$$

Hence the slope of A in either λ^p or λ^g is larger when ω is smaller. That is, the same information friction has a stronger effect when households have higher marginal propensities to consume.

Proof of Proposition 4

Here we spell out the details of each of the set of popular theories discussed in Section 4.3 and we show how they can all be mapped to a pair of coefficients (λ^p, λ^g), as in our preferred specification. That is, in all cases, the aggregate consumption response is given by

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda^p m_t. \tag{32}$$

What changes across the different theories is merely the specific value of the pair (λ^p, λ^g). Throughout, we employ the method of undetermined coefficients to solve for the equilibrium response.

(1) Noisy information without learning. Each household observes a private signal $m_i = m_0 + u_i$, where $u_i \sim \mathcal{N}(0, \sigma_u^2)$, and forms a posterior belief

$$E_{i,t}[m_0] = \frac{\sigma_\epsilon^{-2}}{\sigma_\epsilon^{-2} + \sigma_u^{-2}} m_i \equiv \lambda m_i.$$

Aggregating across households, the average belief about the monetary policy innovation is

$$\overline{E}_t[m_0] = \frac{\sigma_\epsilon^{-2}}{\sigma_\epsilon^{-2} + \sigma_u^{-2}} m_0 \equiv \lambda m_0.$$

We proceed by a guess-and-verify approach, conjecturing that aggregate consumption takes the form $c_t = \alpha \overline{E}_t[m_t]$ with α to be determined in equilibrium. Substituting this conjecture into the aggregate consumption function,

$$\alpha \overline{E}_t[m_t] = \left(1-\beta\omega\right)\overline{E}_t\left[\sum_{k=0}^{\infty}\left(\beta\omega\rho\right)^k \alpha \overline{E}_t[m_t]\right] + \sigma\beta\omega\overline{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega\rho)^k m_t\right]$$

To solve for α, we need to determine $\overline{E}_t\overline{E}_t[m_{t+k}]$. With dispersed information (and hence a lack of common knowledge), the average belief about others' average expectation becomes $\overline{E}_t\overline{E}_{t+k}[m_{t+k}] = \lambda\overline{E}_{t+k}[m_{t+k}] = \lambda\rho^k\overline{E}_t[m_t]$. Since there is no learning over time, $\overline{E}_{t+k}[m_{t+k}] = \rho^k\overline{E}_t[m_t]$. Substituting these relations simplifies the system to

$$a\alpha\overline{E}[m_t] = \frac{1-\beta\omega}{1-\beta\omega\rho}\overline{E}_t\alpha\lambda m_t + \frac{\sigma\beta\omega}{1-\beta\omega\rho}\overline{E}_t m_t.$$

Solving for α gives

$$\alpha = \left(1-\frac{1-\beta\omega}{1-\beta\omega\rho}\lambda\right)^{-1}\frac{\sigma\beta\omega}{1-\beta\omega\rho},$$

and hence

$$c_t = \alpha\overline{E}_t[m_t] = \left(1-\frac{1-\beta\omega}{1-\beta\omega\rho}\lambda\right)^{-1}\frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda m_t.$$

This corresponds to equation (P4) with $\lambda^p = \lambda^g = \lambda = \frac{\sigma_u^{-2}}{\sigma_u^{-2}+\sigma_\epsilon^{-2}}$. Therefore, noisy signals symmetrically attenuate both partial equilibrium and general equilibrium effects.

(2) Overconfidence. As before, household i observes a private signal $m_i = m_0 + u_i$, with $u_i \sim \mathcal{N}(0, \sigma_u^2)$. The true signal precision is σ_u^{-2}. However, under overconfidence, each household believes its own signal is more precise, assigning $\sigma_{\text{own}}^{-2} > \sigma_u^{-2}$. Bayesian updating then gives the perceived posterior

$$E_{i,t}[m_0] = \lambda^p m_i, \qquad \text{with} \qquad \lambda^p = \frac{\sigma_{\text{own}}^{-2}}{\sigma_{\text{own}}^{-2}+\sigma_\epsilon^{-2}}.$$

Aggregating across agents gives the average first-order expectations as

$$\overline{E}_t[m_0] = \lambda^p m_0.$$

To solve for equilibrium, we again need to determine the average second-order expectations $\overline{E}_t\overline{E}_t[m_0]$. Since, for any $j \neq i$, each household i believes that j receives a signal with precision $\sigma_{\text{others}}^{-2}$, individual second-order expectations satisfy

$$E_{i,t}\left[E_{j,t}[m_0]\right] = E_{i,t}\left[\lambda^g m_0\right], \qquad \text{with} \qquad \lambda^g = \frac{\sigma_{\text{others}}^{-2}}{\sigma_{\text{others}}^{-2}+\sigma_\epsilon^{-2}} < \lambda^p.$$

Furthermore, since i is uninformed about the idiosyncratic noise u_j in j's information, $E_{i,t}[m_j] =$

$E_{i,t}[m_0]$. Combining and aggregating across i, we conclude that average second-order expectations satisfy

$$\overline{E}_t\overline{E}_t[m_0] = \lambda^g \overline{E}_t[m_0].$$

Compared to the textbook case studied above (with noisy information but no overconfidence), the main change here is therefore to disentangle the responsiveness of second-order beliefs from that of first-order beliefs. Finally, applying the same method of undetermined coefficients as in the previous case yields

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\lambda^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\lambda^p m_t.$$

Hence, Proposition 3 holds with the pair (λ^p, λ^g) defined above and overconfidence amplifies the partial equilibrium effect relative to the general equilibrium effect, with $\lambda^p > \lambda^g$.

(3) Level-k thinking. Let c_t^k denote the consumption of a level-k thinker in period t. First, for $k = 0$, level-0 agents are entirely unresponsive and continue to consume at the steady-state level $c_t^0 = 0$ for all t. Next, for $k \geq 1$, we assume that level-k agents best respond to the belief that all other agents are level-$(k-1)$ thinkers. More precisely, they choose consumption under the correct belief that real rates will be $r_{t+k} = -m_{t+k} = -\rho^k m_t$ and the incorrect belief that their income will be $y_t^k = c_t^{k-1}$. Formally

$$c_t^k = (1-\beta\omega)\Big[\sum_{h\geq 0}(\beta\omega)^h c_{t+h}^{k-1}\Big] + \sigma\beta\omega\Big[\sum_{h\geq 0}(\beta\omega)^h m_{t+h}\Big].$$

For example, for $k = 1$ thinker, the perceived income path is flat ($c_t^0 = 0$). Substituting the flat income path into the level-1 thinker consumption function gives

$$c_t^1 = \frac{\sigma\beta\omega}{1-\beta\omega\rho} m_t.$$

Similarly, for $k = 2$ thinker, their perceived income at time $t+h$ is $y_{t+h}^2 = c_{t+h}^1 = \frac{\sigma\beta\omega}{1-\beta\omega\rho} m_{t+h}$. Substituting this into the level-1 thinker consumption function gives

$$c_t^2 = \left(1 + \frac{1-\beta\omega}{1-\beta\omega\rho}\right)\frac{\sigma\beta\omega}{1-\beta\omega\rho} m_t.$$

By induction,

$$c_t^k = \left(\sum_{s=0}^{k-1}\left(\frac{1-\beta\omega}{1-\beta\omega\rho}\right)^s\right)\frac{\sigma\beta\omega}{1-\beta\omega\rho} m_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\Lambda(k)\right)^{-1}\frac{\sigma\beta\omega}{1-\beta\omega\rho} m_t,$$

where

$$\lambda^g = \Lambda(k) = \begin{cases} 0, & k = 0, \\ \frac{1-\left(\frac{1-\beta\omega}{1-\beta\omega\rho}\right)^{k-1}}{1-\left(\frac{1-\beta\omega}{1-\beta\omega\rho}\right)^{k}}, & k > 0\,. \end{cases}$$

Thus $\lambda^p = 1$, $\lambda^g = \Lambda(k) \uparrow 1$ as $k \to \infty$. Hence, as the depth of reasoning increases, agents internalize a larger portion of the general equilibrium feedback, and the equilibrium converges to the FIRE benchmark as the depth of reasoning approaches infinity.

(4) Representative inattentive agent. There is a representative household who perfectly understands the structure of the economy but holds distorted beliefs about the policy shock. Specifically, let $\mathbb{E}_t$ denote the FIRE operator and E_t the household's subjective expectation. The household believes

$$E_t[m_t] = \mu \mathbb{E}_t m_t, \qquad \mu \in (0,1).$$

With a representative household, the law of iterated expectations implies that average beliefs equal individual beliefs under the same (subjective) expectation. Hence, under the linear conjecture $c_t = aE_t[m_t]$, $E_t c_{t+k} = aE_t E_{t+k}[m_{t+k}] = c_{t+k}$. That is, households hold correct expectations about future income, thanks to common knowledge. Substituting the conjecture into the aggregate consumption condition,

$$c_t = (1-\beta\omega)\sum_{k\geq 0}(\beta\omega)^k c_{t+k} + \sigma\beta\omega\sum_{k\geq 0}(\beta\omega)^k \mu m_{t+k},$$

and solving by the same undetermined coefficients procedure as before yields

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\mu m_t,$$

Therefore, the general equilibrium effect remains identical to that under FIRE, while the partial-equilibrium effect is distorted. The mapping to our two-parameter representation is $\lambda^p = \mu$, $\lambda^g = 1$. In conclusion, a representative inattentive agent dampens the partial equilibrium response while preserving the full general equilibrium effect.

(5) Sparsity. It is natural to suppose that households must "try hard" to acquire or process information about income and interest rates in order to form forecasts. Sparsity provides such a formulation. Following Gabaix (2014, 2016) and Guerreiro (2023), we assume that households must learn about distinct objects separately and choose attention to minimize a loss that combines (i) second-order utility losses from inattention relative to full information rational expectations benchmark, $C_{\{y,r\}}(\mu)$, and (ii) a cognitive cost $G_{\{y,r\}}(\mu)$. Accordingly, attention to output and to interest rates is chosen via

$$\mu^g = \text{argmin}_{\mu\in[0,1]} C_y(\mu) + G_y(\mu) \qquad \mu^p = \text{argmin}_{\mu\in[0,1]} C_r(\mu) + G_r(\mu)$$

Once (μ^g, μ^r) are optimally pinned down, equilibrium forecasts take the same form as before:

$$\overline{E_t} y_{t+h} = \mu^g y_{t+h}, \qquad \overline{E_t} r_{t+h} = \mu^p r_{t+h}.$$

Substituting the optimally chosen $\mu^{g,p}$ into the aggregate consumption condition yields

$$c_t = (1-\beta\omega)\left[\sum_{k=0}^{\infty} (\beta\omega)^k \mu^g y_{t+k}\right] - \sigma\beta\omega\left[\sum_{k=0}^{\infty} (\beta\omega)^k \mu^p r_{t+k}\right].$$

That is, aggregate consumption is analogous to that in a FIRE economy, but with dampening factors μ^g and μ^p applied to future output and interest rates, respectively. Using standard techniques of matching coefficients under the FIRE equilibrium then yields

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\mu^g\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\mu^p m_t.$$

Hence, this maps to equation (P4) with $\lambda^p = \mu^p$, $\lambda^g = \mu^g$. In other words, Endogenous attention attenuates the partial equilibrium and general equilibrium effect according to the optimal allocation of cognitive effort.

(6) Cognitive discounting. Another well-known deviation from rationality is cognitive discounting (Gabaix (2020)), under which households discount future variables toward their steady-state values. More concretely, they discount the law of motion of state variables, thereby underestimating their persistence. In our framework, this means households perceive

$$\hat{\rho} = \mu\rho, \qquad \mu \in (0,1),$$

so they believe policy shocks are less persistent than they truly are. Therefore, their prediction about future income is $E_t m_{t+k} = (\mu\rho)^k m_t$, where E_t denotes the household's subjective expectation. Now applying the same guess-and-verify approach with $c_t = \alpha E_t m_t$. Note that $E_t y_{t+k} = \alpha(\mu\rho)^k m_t = \mu^k y_{t+k}$, so the aggregate consumption function becomes,

$$c_t = (1-\beta\omega)\left[\sum_{k=0}^{\infty} (\beta\omega\mu)^k y_{t+k}\right] - \sigma\beta\omega\left[\sum_{k=0}^{\infty} (\beta\omega\mu)^k r_{t+k}\right].$$

Solving by the same undetermined coefficients method as before yields

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho\mu}\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho\mu} m_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\Lambda(\mu)\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\Lambda(\mu) m_t.$$

This is equivalent to (P4) with $\lambda^p = \lambda^g = \Lambda(\mu) = \frac{1-\beta\omega\rho}{1-\beta\omega\rho\mu} \in (0,1]$. In other words, cognitive discounting symmetrically attenuates both partial equilibrium and general equilibrium feedback.

(7) Confusion between idiosyncratic and aggregate shocks. In this economy, households now face both aggregate and idiosyncratic shocks:

$$r_{i,t} = r_t + \xi^r_{i,t}, \qquad y_{i,t} = y_t + \xi^y_{i,t},$$

where the idiosyncratic components follow independent AR(1) processes with persistences (ρ_r, ρ_y). Households observe their idiosyncratic shocks perfectly, but they confuse aggregate shocks as if they were idiosyncratic. In other words, they believe that future income and interest rates follow

$$y_{i,t+k} = \rho_y^k y_{i,t}, \qquad r_{i,t+k} = \rho_r^k r_{i,t},$$

rather than the true aggregate processes. Given these beliefs, each household's consumption evolves according to

$$c_{i,t} = (1-\beta\omega)\sum_{k\geq 0}(\beta\omega\rho_y)^k y_{i,t} + \sigma\beta\omega\sum_{k\geq 0}(\beta\omega\rho_r)^k m_{i,t}.$$

Aggregating across households gives

$$c_t = (1-\beta\omega)\sum_{k\geq 0}(\beta\omega\rho_y)^k y_t + \sigma\beta\omega\sum_{k\geq 0}(\beta\omega\rho_r)^k m_t.$$

Applying the standard matching of undetermined coefficients as in the FIRE benchmark yields

$$c_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho_y}\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho_r} m_t = \left(1 - \frac{1-\beta\omega}{1-\beta\omega\rho}\Lambda_g(\rho_y)\right)^{-1} \frac{\sigma\beta\omega}{1-\beta\omega\rho}\Lambda_p(\rho_r) m_t,$$

where

$$\Lambda_p(\rho, \rho_r) = \frac{1-\beta\omega\rho}{1-\beta\omega\rho_r}, \qquad \Lambda_g(\rho, \rho_y) = \frac{1-\beta\omega\rho}{1-\beta\omega\rho_y}.$$

Note that if $\rho_r > \rho$ or $\rho_y > \rho$, the partial equilibrium or general equilibrium responses exceed those under FIRE. In other words, confusing idiosyncratic and aggregate shocks can amplify macroeconomic propagation, akin to over-extrapolation Angeletos, Huo, and Sastry (2021), whenever persistence of idiosyncratic shocks is larger than that of the aggregate shocks. The same is possible if we replace cognitive discounting with over-extrapolation or diagnostic expectations.

Proof of Corollary 2

The result follows directly from Proposition 4. In the representative inattentive agent case, $\lambda^p < 1$ and $\lambda^g = 1$, attenuating only the partial equilibrium channel. Under level-k reasoning, $\lambda^p = 1$ and $\lambda^g < 1$, dampening the general equilibrium feedback while preserving the partial equilibrium effect. All remaining theories correspond to $(\lambda^p, \lambda^g) \in (0,1]^2$ and therefore imply $\alpha_f < 1$.

Proof of Proposition 5

The proof uses a simple guess-and-verify process, similar to that used in Proposition 3. Suppose that the equilibrium satisfies $y_t = \alpha_t \overline{E}_t m_t$, for some deterministic $\{\alpha_t\}_{t=0}^{\infty}$. We can then compute the forecasts of future income as follows:

$$\overline{E}_t \left[y_{t+h} \right] = \overline{E}_t \left[\alpha_{t+h} \overline{E}_{t+h} m_{t+h} \right] = \alpha_{t+h} \overline{E}_t \overline{E}_{t+h} m_{t+h} = \alpha_{t+h} \lambda_t^p \lambda_{t+h}^g m_{t+h}, \tag{33}$$

where the last step uses Assumption 2. Note that this step amounts to replacing the second-order forecast operator $\overline{E}_t \overline{E}_{t+h}$ with the product $\lambda_t^p \lambda_{t+h}^g$. Intuitively, the average forecast at t of the average forecast at $t+h$ translates to the fraction of informed agents at t (i.e., λ_t^p) times what these agents believe that the fraction of informed agents will be at $t+h$ (i.e., λ_{t+h}^g). Next, note that the (first-order) forecasts of the underlying shock satisfy

$$\overline{E}_t [m_{t+h}] = \lambda_t^p m_{t+h}. \tag{34}$$

We can thus rewrite (33) as follows:

$$\overline{E}_t \left[y_{t+h} \right] = \alpha_{t+h} \lambda_t^p \lambda_{t+h}^g \frac{\overline{E}_{t+h} m_{t+h}}{\lambda_{t+h}^p} = \lambda_{t+h}^g \frac{\lambda_t^p}{\lambda_{t+h}^p} \alpha_{t+h} \overline{E}_{t+h} m_{t+h} = \lambda_{t+h}^g \frac{\lambda_t^p}{\lambda_{t+h}^p} y_{t+h}, \tag{35}$$

where the last steps uses our guess that $y_{t+h} = \alpha_{t+h} \overline{E}_{t+h} m_{t+h}$. Replacing (35) and (34) into the original, FIRE version of the IKC (1), we then arrive at the modified IKC (15). The argument is completed by noting that solving (15) is equivalent to verifying our initial guess.

Proof of Spending Multiplier (17)

Leveraging Proposition 4, we know that various deviations from FIRE can be summarized as

$$\overline{E_t} y_{t+k} = \lambda^g y_{t+k}, \qquad \overline{E_t} y_{t+k} = \lambda^p y_{t+k}.$$

Substituting these expectations into equation (16), together with $y_t = c_t + g_t$ and the balanced budget $t_t = g_t$ yields

$$y_t - g_t = c_t = \left(1-\beta\omega\left(1-\sigma\phi\right)\right)\sum_{k=0}^{\infty}\left(\beta\omega\right)^k \lambda^g y_{t+k} - \left(1-\beta\omega\right)\sum_{k=0}^{\infty}\left(\beta\omega\right)^k \lambda^p g_{t+k}.$$

We proceed by a guess-and-verify approach, conjecturing that the spending multiplier is constant and satisfies $\frac{dy_t}{dg_t} = \psi$ for all t. Then,

$$\psi = 1 + \left(1-\beta\omega\left(1-\sigma\phi\right)\right)\sum_{k=0}^{\infty}\left(\beta\omega\right)^k \lambda^g \psi \rho_g^k - \left(1-\beta\omega\right)\sum_{k=0}^{\infty}\left(\beta\omega\right)^k \lambda^p \rho_g^k$$

This simplifies to

$$\psi = 1 + \frac{1-\beta\omega\left(1-\sigma\phi\right)}{1-\beta\omega\rho_g}\lambda^g\psi - \left(\frac{1-\beta\omega}{1-\beta\omega\rho_g}\right)\lambda^p$$

Defining $\varrho_g \equiv \frac{1-\beta\omega}{1-\beta\omega\rho_g}$ and $\varphi \equiv \frac{\beta\omega}{1-\beta\omega\rho_g}\sigma\phi$, the expression further simplifies to the desired representation

$$\psi = \frac{1-\varrho_g\lambda^p}{1-\left(\varrho_g+\varphi\right)\lambda^g}.$$

Proof of Transfer Multiplier (18)

Under the stated assumptions $y_t = 0$ for $t \geq 1$. At time 0,

$$y_0 = \left(1-\beta\omega\right)\left(\lambda^g y_0 + \epsilon_0\right) - \left(1-\beta\omega\right)\beta\omega\lambda^p t_1,$$

and since $t_1 = \beta^{-1}\epsilon_0$, then

$$y_0 = \frac{\left(1-\beta\omega\right)\epsilon_0 - \left(1-\beta\omega\right)\beta\omega\lambda^p\beta^{-1}\epsilon_0}{1-\left(1-\beta\omega\right)\lambda^g} = \frac{\left(1-\beta\omega\right)}{1-\left(1-\beta\omega\right)\lambda^g}\left[1-\lambda^p\omega\right]\epsilon_0$$

Therefore,

$$\frac{dy_0}{d\epsilon_0} = \frac{\left(1-\beta\omega\right)}{1-\left(1-\beta\omega\right)\lambda^g}\left[1-\lambda^p\omega\right].$$

C Appendix to Section 6

This Appendix complements our analysis of HANK models in Section 6. We start by illustrating the derivation of the log-linearized aggregate consumption function of a canonical, incomplete-markets, model. We next derive our generalized IKC and prove Propositions 6 and 7. We finally show how to accommodate a general feedback rule for monetary policy.

The aggregate consumption function

Consider a standard incomplete-markets consumption block, which will constitute the household sector of the canonical HANK model, as in Auclert, Rognlie, and Straub (2024). The economy is populated by a continuum of infinitely lived households indexed by $i \in [0,1]$. At each date t, household i chooses consumption $c_{i,t}$ and forms a plan over future consumption to maximize expected lifetime utility:

$$E_{i,t}\left[\sum_{h=0}^{\infty}\beta^h u\left(c_{i,t+h}\right)\right], \tag{36}$$

where β denotes the subjective discount factor, $u(c)$ is the utility function which satisfies the standard properties, and $E_{i,t}[\cdot]$ denotes household i's expectations at time t.

Idiosyncratic labor productivity follows a finite-state Markov process $\{z_{i,t}\}$ with transition matrix Π, and let $\pi(z)$ denote the stationary distribution over productivity states. The productivity process is normalized so that $\sum_z \pi(z)z = 1$. At time t, household i enters the period with financial assets $a_{i,t}$ and earns labor income $z_{i,t}Y_t$, where Y_t denotes aggregate output (or equivalently, aggregate labor income).

The household's period budget constraint is given by

$$c_{i,t+h} + \frac{a_{i,t+h}}{1+r_t} = z_{i,t+h}Y_{t+h} + a_{i,t+h-1}$$

subject to a borrowing constraint $a_{i,t+h} \geq \underline{a}$, where r_t denotes the real interest rate and $\underline{a}$ the exogenous borrowing limit. For simplicity, we assume a simple ad-hoc borrowing limit, but we could allow for more general, endogenous, borrowing constraints.

Throughout the analysis, we assume that households observe the current values of $z_{i,t}$ and $x_t = (Y_t, r_t)$, ensuring they can always satisfy their borrowing constraints (Carroll et al., 2020). We further assume that $z_{i,t}$, the household's idiosyncratic state, and

$$b_{i,t} \equiv \left(x_t, \{E_{i,t}[x_{t+h}]\}_{h=1}^{\infty}\right) = \left(\{E_{i,t}[x_{t+h}]\}_{h=0}^{\infty}\right),$$

the realized history of the aggregate outcomes and her forecasts about their future path, are sufficient statistics for her entire subjective beliefs about the future. This means, in particular, that there is no independent variation in higher moments of subjective beliefs beyond that spanned by $b_{i,t}$. The solution to the household's problem can then be expressed in terms of the following policy functions for consumption and saving:

$$c_{i,t} = c\left(a_{i,t-1}; b_{i,t}; z_{i,t}\right), \quad \text{and } a_{i,t} = a\left(a_{i,t-1}; b_{i,t}; z_{i,t}\right).$$

Recursive substitution of past assets reveals that consumption at date t depends on the history of idiosyncratic shocks and belief sequences:

$$c_{i,t} = c^*\left(b_i^t; z_i^t\right),$$

where $b_i^t = \left\{b_{i,s}\right\}_{s=-\infty}^t$ $z_i^t = \left\{z_{i,s}\right\}_{s=-\infty}^t$. The dependence on the initial asset $a_{i,-\infty}$ vanishes under standard ergodicity assumptions. To focus on the first-order aggregate dynamics of the economy, we linearize individual consumption decisions with respect to b_i^t (but not with respect to z_i^t) around a steady state characterized by $x_t = E_{i,t}\left[x_{t+h}\right] = x^{ss}$ for all t and h. The first-order approximation takes the form:

$$c_{i,t} = \alpha\left(z_i^t\right) + \sum_{s=-\infty}^{0}\sum_{h=0}^{\infty}\gamma_{s,h}^{Y}\left(z_i^t\right)E_{i,t+s}\left[Y_{t+s+h}\right] + \sum_{s=-\infty}^{0}\sum_{h=0}^{\infty}\gamma_{s,h}^{r}\left(z_i^t\right)E_{i,t+s}\left[r_{t+s+h}\right]. \tag{37}$$

Notice that the consumption function is not linearized with respect the history of idiosyncratic shocks, thus preserving the effect of cross-sectional heterogeneity.

Aggregate consumption is given by integrating over the continuum of agents:

$$C_t = \int_0^1 \alpha\left(z_i^t\right)di + \sum_{s=-\infty}^{0}\sum_{h=0}^{\infty}\int_0^1 \gamma_{s,h}^{Y}\left(z_i^t\right)\cdot E_{i,t+s}\left[Y_{t+s+h}\right]di + \sum_{s=-\infty}^{0}\sum_{h=0}^{\infty}\int_0^1 \gamma_{s,h}^{r}\left(z_i^t\right)\cdot E_{i,t+s}\left[r_{t+s+h}\right]di. \tag{38}$$

In general, this aggregate consumption function diverges from that in (21), due to its dependence on the higher-order moments of the belief distribution. Specifically, for aggregate income we have:

$$\int_0^1 \gamma_{s,h}^{Y}\left(z_i^t\right)\cdot E_{i,t+s}\left[Y_{t+s+h}\right]di = \overline{\gamma}_{s,h}^{Y}\cdot\overline{E}_{t+s}\left[Y_{t+s+h}\right] + \text{Cov}\left(\gamma_{s,h}^{Y}\left(z_i^t\right), E_{i,t+s}\left[Y_{t+s+h}\right]\right),$$

where $\overline{\gamma}_{s,h}^{y} = \int_0^1 \gamma_{s,h}^{y}\left(z_i^t\right)di$ and $\overline{E}_{t+s}\left[Y_{t+s+h}\right] = \int_0^1 E_{i,t+s}\left[Y_{t+s+h}\right]$. A similar property applies for real interest rates. The covariance term captures the dimension of belief heterogeneity that is relevant for aggregate dynamics—namely, the extent to which cross-sectional variation in beliefs aligns with the marginal propensities that govern individual behavioral responses.

To simplify the analysis, we assume that beliefs are orthogonal to idiosyncratic shocks, so that all covariance terms are zero. It follows that in this case:

$$C_t = \overline{\alpha} + \sum_{s=-\infty}^{0} \sum_{h=0}^{\infty} \overline{\gamma}^{Y}_{s,h} \cdot \overline{E}_{t+s}\left[Y_{t+s+h}\right] + \sum_{s=-\infty}^{0} \sum_{h=0}^{\infty} \overline{\gamma}^{r}_{s,h} \cdot \overline{E}_{t+s}\left[r_{t+s+h}\right], \tag{39}$$

where $\overline{\alpha} \equiv \int_0^1 \alpha\left(z_i^t\right) di$. Clearly, this not only fits assumption (21), but also reveals how heterogeneity influences the aggregate dynamics by regulating the "loads" $\overline{\gamma}^{Y}_{s,h}$ and $\overline{\gamma}^{r}_{s,h}$ with which the consensus forecasts enter aggregate consumption.

The derivation above proceeds under the assumption that individual consumption functions are differentiable with respect to aggregate variables, thereby allowing for the linearization in equation (37). However, this condition is stronger than necessary. For the purpose of characterizing aggregate consumption dynamics, it suffices that the *aggregate* consumption function is differentiable with respect to the sequence of aggregates.

Proof of Proposition 6

First, note that, we can express expectations as a cumulative sum of forecast revisions:

$$\overline{E}_t\left[x_{t+h}\right] = \overline{E}_0\left[x_{t+h}\right] + \sum_{\tau=1}^{t} \overline{FR}_\tau\left[x_{t+h}\right], \quad \text{for } t = 0, 1, 2, \ldots$$

So, aggregate demand can be written as

$$C_t = \mathscr{C}\left(\left\{\left\{\overline{E}_0\left[x_{t+s+h}\right] + \sum_{\tau=1}^{t+s} \overline{FR}_\tau\left[x_{t+s+h}\right]\right\}_{h=0}^{\infty}\right\}_{s=-t}^{0}\right),$$

$$C_t = \hat{\mathscr{C}}\left(\left\{\overline{E}_0\left[x_s\right]\right\}_{s=0}^{\infty}, \left\{\left\{\overline{FR}_\tau\left[x_s\right]\right\}_{s=\tau}^{\infty}\right\}_{\tau=0}^{t}\right) = \mathscr{C}\left(\left\{\left\{\overline{E}_0\left[x_s\right] + \sum_{\tau=1}^{k} \overline{FR}_\tau\left[x_s\right]\right\}_{s=k}^{\infty}\right\}_{k=0}^{t}\right),$$

where we have suppressed the notation for realizations and expectations prior to date 0 since these are held constant at steady state.

Note that

$$\frac{\partial C_t}{\partial \overline{E}_0\left[Y_s\right]} = \frac{\partial \hat{\mathscr{C}}\left(\left\{\overline{E}_0\left[x_s\right]\right\}_{s=0}^{\infty}, \left\{\left\{\overline{FR}_\tau\left[x_s\right]\right\}_{s=\tau}^{\infty}\right\}_{\tau=0}^{t}\right)}{\partial \overline{E}_0\left[Y_s\right]} = \sum_{k=0}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[Y_s\right]} = M_{t,s}$$

and

$$\frac{\partial C_t}{\partial \overline{E}_0\left[r_s\right]} = \frac{\partial \hat{\mathscr{C}}\left(\left\{\overline{E}_0\left[x_s\right]\right\}_{s=0}^{\infty}, \left\{\left\{\overline{FR}_\tau\left[x_s\right]\right\}_{s=\tau}^{\infty}\right\}_{\tau=0}^{t}\right)}{\partial \overline{E}_0\left[r_s\right]} = \sum_{k=0}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[r_s\right]} = M_{t,s}^{r},$$

where the derivatives are evaluated around steady state.

Furthermore, for $s \geq \tau$,

$$\frac{\partial C_t}{\partial \overline{FR}_\tau\left[Y_s\right]} = \frac{\partial \hat{\mathscr{C}}\left(\left\{\overline{E}_0\left[x_s\right]\right\}_{s=0}^{\infty}, \left\{\left\{\overline{FR}_\tau\left[x_s\right]\right\}_{s=\tau}^{\infty}\right\}_{\tau=0}^{t}\right)}{\partial \overline{FR}_\tau\left[Y_s\right]} = \begin{cases} 0 & \text{if } t < \tau \\ \sum_{k=\tau}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[Y_s\right]} & \text{if } t \geq \tau \end{cases}$$

and

$$\frac{\partial C_t}{\partial \overline{FR}_\tau\left[r_s\right]} = \frac{\partial \hat{\mathscr{C}}\left(\left\{\overline{E}_0\left[x_s\right]\right\}_{s=0}^{\infty}, \left\{\left\{\overline{FR}_\tau\left[x_s\right]\right\}_{s=\tau}^{\infty}\right\}_{\tau=0}^{t}\right)}{\partial \overline{FR}_\tau\left[r_s\right]} = \begin{cases} 0 & \text{if } t < \tau \\ \sum_{k=\tau}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[r_s\right]} & \text{if } t \geq \tau. \end{cases}$$

Now note that

$$\begin{aligned} \frac{\partial C_t}{\partial \overline{FR}_\tau\left[Y_s\right]} &= \sum_{k=\tau}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[Y_s\right]} = \sum_{\hat{k}=0}^{\min\{t-\tau,s-\tau\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_{\hat{k}+\tau}\left[Y_s\right]} \\ &= \sum_{\hat{k}=0}^{\min\{t-\tau,s-\tau\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t-\tau}\right)}{\partial \overline{E}_{\hat{k}}\left[Y_{s-\tau}\right]} = M_{t-\tau,s-\tau} \end{aligned}$$

and similarly

$$\begin{aligned} \frac{\partial C_t}{\partial \overline{FR}_\tau\left[r_s\right]} &= \sum_{k=\tau}^{\min\{t,s\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_k\left[r_s\right]} = \sum_{k=0}^{\min\{t-\tau,s-\tau\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t}\right)}{\partial \overline{E}_{k+\tau}\left[r_s\right]} \\ &= \sum_{k=0}^{\min\{t-\tau,s-\tau\}} \frac{\partial \mathscr{C}\left(\left\{\left\{\overline{E}_{\hat{k}}\left[x_{\hat{s}}\right]\right\}_{\hat{s}=\hat{k}}^{\infty}\right\}_{\hat{k}=0}^{t-\tau}\right)}{\partial \overline{E}_k\left[r_{s-\tau}\right]} = M_{t-\tau,s-\tau}. \end{aligned}$$

Now, note that the response of the vector of aggregate demand $\boldsymbol{C} = \left[C_t\right]$ to a change in time zero expectations is given by

$$d\boldsymbol{C} = \boldsymbol{M} \cdot \overline{E}_0\left[d\boldsymbol{Y}\right] + \boldsymbol{M} \cdot \overline{E}_0\left[d\boldsymbol{r}\right].$$

The change in aggregate demand to forecast revisions at time 1 is given by

$$d\boldsymbol{C} = \begin{bmatrix} 0 & \mathbf{0}_{1\times\infty} \\ \mathbf{0}_{\infty\times 1} & \boldsymbol{M} \end{bmatrix} \cdot \overline{FR}_1 [d\boldsymbol{Y}] + \begin{bmatrix} 0 & \mathbf{0}_{1\times\infty} \\ \mathbf{0}_{\infty\times 1} & \boldsymbol{M}^2 \end{bmatrix} \cdot \overline{FR}_1 [d\boldsymbol{r}]$$

and generally to forecast revisions at time h

$$d\boldsymbol{C} = \begin{bmatrix} \mathbf{0}_{h\times h} & \mathbf{0}_{h\times\infty} \\ \mathbf{0}_{\infty\times h} & \boldsymbol{M} \end{bmatrix} \cdot \overline{FR}_h [d\boldsymbol{Y}] + \begin{bmatrix} \mathbf{0}_{h\times h} & \mathbf{0}_{h\times\infty} \\ \mathbf{0}_{\infty\times h} & \boldsymbol{M}^2 \end{bmatrix} \cdot \overline{FR}_h [d\boldsymbol{r}].$$

Grouping together terms implies (24).

The Generalized IKC with Constant Expectations

In the special case where agents do not revise their expectations over time, the generalized IKC expression simplifies considerably. Specifically, suppose that agents form a fixed view of the future at time zero and maintain these expectations thereafter. That is, for all $t \geq 0$ and $h > 0$, forecast revisions satisfy $\overline{FR}_t [dx_{t+h}] = 0$. Under this restriction, the model reduces to a setting in which expectations are constant over time and all deviations from realized outcomes manifest as forecast errors.

Corollary C.1 (Generalized IKC under Fixed Expectations)**.** *Suppose that consumers do not update their expectations of future aggregates after time zero:*

$$\{E_t [dx_{t+h}]\}_{h=0}^{\infty} = \{E [dx_{t+h}]\}_{h=0}^{\infty}, \quad \forall t \geq 0.$$

Then, the generalized IKC representation simplifies to:

$$d\boldsymbol{C} = \boldsymbol{M} \cdot \overline{E} [d\boldsymbol{Y}] + \mathscr{E} \cdot \left\{ d\boldsymbol{Y} - \overline{E} [d\boldsymbol{Y}] \right\} + \boldsymbol{M}^r \cdot \overline{E} [d\boldsymbol{r}] + \mathscr{E}^r \cdot \left\{ d\boldsymbol{r} - \overline{E} [d\boldsymbol{r}] \right\}, \tag{40}$$

where the ***forecast-error Jacobians*** $\mathscr{E}$ *and* $\mathscr{E}^r$ *are given by:*

$$\mathscr{E} = \begin{bmatrix} M_{0,0} & 0 & 0 & \dots \\ M_{1,0} & M_{0,0} & 0 & \dots \\ M_{2,0} & M_{1,0} & M_{0,0} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \text{ and } \mathscr{E}^r = \begin{bmatrix} M^r_{0,0} & 0 & 0 & \dots \\ M^r_{1,0} & M^r_{0,0} & 0 & \dots \\ M^r_{2,0} & M^r_{1,0} & M^r_{0,0} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

This result highlights that, under fixed expectations, all deviations from anticipated paths enter through contemporaneously realized forecast errors. The resulting consumption dynamics can be captured by a convolution of these errors with the first column of the original Jacobian matrices.

Proof of Proposition 7

First, note that under assumption 4, people have constant expectations about the future $E_{i,t}\left[dY_{t+h}\right] = E_{i,0}\left[dY_{t+h}\right]$ and $E_{i,t}\left[dr_{t+h}\right] = E_{i,0}\left[dr_{t+h}\right]$. It follows from Proposition 6, that we can write the consumption of the informed, $d\boldsymbol{C}^{inf}$, and of the uninformed, $d\boldsymbol{C}^{0}$, as follows:

$$d\boldsymbol{C}^{inf} = \boldsymbol{M}E^{inf}\left[d\boldsymbol{Y}\right] + \mathscr{E}\cdot\left(d\boldsymbol{Y} - E^{inf}\left[d\boldsymbol{Y}\right]\right) + \boldsymbol{M}^{r}\cdot d\boldsymbol{r} \tag{41}$$

and

$$d\boldsymbol{C}^{0} = \mathscr{E}\cdot d\boldsymbol{Y} + \mathscr{E}^{r}\cdot d\boldsymbol{r}, \tag{42}$$

respectively. The equilibrium is given by

$$d\boldsymbol{Y} = \lambda^{p} d\boldsymbol{C}^{inf} + \left(1-\lambda^{p}\right) d\boldsymbol{C}^{0} \tag{43}$$

and the expectations of the informed solve:

$$E^{inf}\left[d\boldsymbol{Y}\right] = \lambda^{g} E^{inf}\left[d\boldsymbol{C}^{inf}\right] + \left(1-\lambda^{g}\right) E^{inf}\left[d\boldsymbol{C}^{0}\right]. \tag{44}$$

Combining (41) and (42) with (43), we obtain

$$d\boldsymbol{Y} = \mathscr{E}\cdot d\boldsymbol{Y} + \mathscr{E}^{r}\cdot d\boldsymbol{r} + \lambda^{p}\left\{\left(\boldsymbol{M}-\mathscr{E}\right)E^{inf}\left[d\boldsymbol{Y}\right] + \left(\boldsymbol{M}^{r}-\mathscr{E}^{r}\right)\cdot d\boldsymbol{r}\right\}, \tag{45}$$

and combining (41) and (42) with (44), we obtain

$$E^{inf}\left[d\boldsymbol{Y}\right] = \mathscr{E}\cdot E^{inf}\left[d\boldsymbol{Y}\right] + \mathscr{E}^{r}\cdot d\boldsymbol{r} + \lambda^{g}\left\{\left(\boldsymbol{M}-\mathscr{E}\right)E^{inf}\left[d\boldsymbol{Y}\right] + \left(\boldsymbol{M}^{r}-\mathscr{E}^{r}\right)\cdot d\boldsymbol{r}\right\}. \tag{46}$$

Note that, combining (45) and (46), now implies that

$$\frac{\left(I-\mathscr{E}\right)E^{inf}\left[d\boldsymbol{Y}\right] - \mathscr{E}^{r}\cdot d\boldsymbol{r}}{\lambda^{g}} = \frac{\left(I-\mathscr{E}\right)d\boldsymbol{Y} - \mathscr{E}^{r}\cdot d\boldsymbol{r}}{\lambda^{p}}$$

$$\Leftrightarrow \lambda^{p}E^{inf}\left[d\boldsymbol{Y}\right] = \lambda^{g}d\boldsymbol{Y} + \mathscr{M}^{0}\mathscr{E}^{r}\cdot\left(\lambda^{p}-\lambda^{g}\right)d\boldsymbol{r},$$

where we assume that $(I-\mathscr{E})$ possesses an inverse $\mathscr{M}^0$ (see conditions for this in Auclert, Rognlie, and Straub, 2024). Using this expression to replace $\lambda^p E^{inf}[d\boldsymbol{Y}]$ in 45, obtains

$$d\boldsymbol{Y} = \mathscr{E}\cdot d\boldsymbol{Y} + \mathscr{E}^r\cdot d\boldsymbol{r} + (\boldsymbol{M}-\mathscr{E})\left[\lambda^g d\boldsymbol{Y} + \mathscr{M}^0\mathscr{E}^r\cdot\left(\lambda^p-\lambda^g\right)d\boldsymbol{r}\right] + \lambda^p\left(\boldsymbol{M}^r-\mathscr{E}^r\right)\cdot d\boldsymbol{r}.$$

Finally, collecting terms implies that

$$d\boldsymbol{Y} = \left\{\lambda^g\boldsymbol{M} + \left(1-\lambda^g\right)\mathscr{E}\right\}d\boldsymbol{Y} + \left\{\lambda^p\boldsymbol{M}^r + \left(1-\lambda^p\right)\mathscr{E}^r\right\}d\boldsymbol{r} + (\boldsymbol{M}-\mathscr{E})\,\mathscr{M}^0\mathscr{E}^r\cdot\left(\lambda^p-\lambda^g\right)d\boldsymbol{r}.$$

Beyond a constant real interest rate

In the baseline, we have assumed that the central bank directly sets the real interest rate. However, the analytical framework developed above readily extends to accommodate a more conventional monetary policy specification in the form of a generalized Taylor rule. Specifically, suppose the central bank sets the nominal interest rate $\boldsymbol{i} = \{i_t\}_{t=0}^{\infty}$ according to:

$$d\boldsymbol{i} = \Phi_\pi\cdot d\boldsymbol{\pi} + \Phi_Y\cdot \boldsymbol{dY} + d\boldsymbol{m},$$

where $\boldsymbol{\pi} = \{\pi_t\}_{t=0}^{\infty}$ denotes the inflation rate at time t and $d\boldsymbol{m}$ is an exogenous monetary policy shock. The policy matrices Φ_π and Φ_Y allow for general dependence of the nominal interest rate on current, lagged, or anticipated values of inflation and output. Such dependence naturally accommodates interest rate smoothing and forward guidance within the Taylor rule formulation.

For tractability, we assume that price setters operate under full-information rational expectations and that inflation dynamics are governed by a generalized Phillips curve of the form:

$$d\boldsymbol{\pi} = K\cdot d\boldsymbol{Y},$$

where K is the reduced-form slope matrix capturing the sensitivity of inflation to output fluctuations.

Given the nominal rate path, the real interest rate can be expressed as:

$$d\boldsymbol{r} = d\boldsymbol{i} - F\cdot d\boldsymbol{\pi},$$

where F is the forward (diagonal-shifting) operator. Substituting the Taylor rule and Phillips curve yields:

$$d\boldsymbol{r} = \Phi\cdot \boldsymbol{dY} + d\boldsymbol{m}$$

where the generalized policy transmission matrix is defined as $\Phi \equiv [\Phi_\pi - F]\cdot K + \Phi_Y$. Substituting this expression into the original perfect-foresight IKC representation for aggregate consumption gives:

$$d\boldsymbol{C} = \boldsymbol{M}^{\text{new}} \cdot d\boldsymbol{Y} + \boldsymbol{M}^{r} \cdot d\boldsymbol{m},$$

where the modified demand sensitivity matrix is given by $\boldsymbol{M}^{new} \equiv \boldsymbol{M} + \boldsymbol{M}^{r} \cdot \Phi$.

This augmented formulation preserves the linearity and tractability of the original IKC framework, while allowing for a richer transmission mechanism from monetary policy to real activity. Using this representation, one can proceed analogously to derive the generalized IKC under various deviations from rational expectations, now applied to beliefs over $d\boldsymbol{Y}$ and $d\boldsymbol{m}$. Implicit in this extension is the assumption that households possess first-order knowledge of the structural relationships encoded in the Taylor rule and the Phillips curve.