Determinacy without the Taylor Principle*

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

September 27, 2021

Abstract

Our understanding of how monetary policy works is complicated by an equilibrium-selection conundrum: because the same path for the nominal interest rate can be associated with multiple equilibrium paths for inflation and output, there is a long-lasting debate about what the right equilibrium selection is. We offer a potential resolution by showing that a small friction in memory and intertemporal coordination can remove the indeterminacy. The unique surviving equilibrium is the same as that conventionally selected by the Taylor principle, but it no more relies on it. By the same token, no space is left for equilibrium selection by the Fiscal Theory of the Price Level, even when monetary policy is passive.

*For helpful comments and suggestions, we are grateful to Manuel Amador, Andy Atkeson, Marco Bassetto, Francesco Bianchi, Florin Bilbiie, V.V. Chari, Lawrence Christiano, John Cochrane, Martin Eichenbaum, Yuriy Gorodnichenko, Patrick Kehoe, Narayana Kocherlakota, Jennifer La'O, Stephen Morris, Emi Nakamura, Tom Sargent, Jon Steinsson, Christian Wolf, and seminar participants at Oxford University, UC Berkeley, Hitotsubashi University, SED, the Minnesota Conference in Macroeconomic Theory, the NBER Summer Institute (EFCE), and VEAMS.

†MIT and NBER; angelet@mit.edu.

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

1 Introduction

Are inflation and output gaps pinned down by monetary policy, as in the conventional solution of the New Keynesian model, or by government debt and deficits, as in an alternative solution of the *same* model that is consistent with the Fiscal Theory of the Price Level (FTPL)? Can the ZLB trigger a self-fulfilling deflationary spiral, or does inflation remain anchored?

The answers to such questions are complicated by the fact that, at least in the New Keynesian model, the same path for the nominal interest rate can be consistent with multiple equilibrium paths for output and inflation.[1] To avoid this indeterminacy, the convention is to assume that monetary policy satisfies the Taylor principle (Taylor, 1993), or that it is "active" (Leeper, 1991). This is often described as follows: in response to inflationary pressures, adjust the interest rate with enough ferocity to bring the economy back on track. But this description confounds the equilibrium selection and stabilization functions of feedback policies. Once these functions are separated (King, 2000; Atkeson et al., 2010), it becomes clear that the Taylor principle regards only the former—and this opens the pandora box of what the "right" equilibrium selection is.[2]

Cochrane (2011) has argued that the Taylor principle amounts to an off-equilibrium threat to "blow up" the economy, in the sense of triggering an explosion in inflation and the output gap; and he has pushed for the FTPL, originally articulated by Sims (1994) and Woodford (1995), as a superior alternative. But this theory, too, can be equated to a blow-up threat, now in the sense of violating the government's intertemporal budget constraint. And because off-equilibrium assumptions cannot be refuted by data, the debate is "fundamentally a religious, not scientific, issue" (Kocherlakota and Phelan, 1999, p.22).[3]

We offer a way out of this conundrum. We highlight how the multiplicity of equilibria under interest-rate pegs, or more generally under "passive" monetary policy, hinges on strong assumptions about memory and intertemporal coordination. Once we perturb these assumptions appropriately, this multiplicity disappears; the only surviving equilibrium is that known as the model's fundamental or minimum state variable (MSV) solution (McCallum, 1983, 2009).[4] This leaves little space for the FTPL, at least as currently formulated.

[1] This goes back to Sargent and Wallace (1975), who first highlighted this kind of indeterminacy, albeit in a flexible-price model. We discuss how nominal rigidity matters, or does not matter, in due course.

[2] To be precise, the Taylor principle is used to guarantee determinacy of *bounded* equilibria, which is our focus here. Unbounded equilibria, such as self-fulfilling hyper-inflations (Obstfeld and Rogoff, 1983, 2021; Cochrane, 2011) and self-fulfilling liquidity traps (Benhabib et al., 2002), are implicitly or explicitly ruled out by appropriate "exit clauses." See Atkeson et al. (2010), and especially their section on "hybrid" rules, for a careful treatment of this issue.

[3] Bassetto (2002) and Cochrane (2005) object to the blow-up interpretation of the FTPL and articulate ways around it. Atkeson et al. (2010) identify monetary policies that, too, avoid the blow-up criticism. And the debate goes on.

[4] The MSV concept for macroeconomic models is a close cousin of Markov Perfect Equilibrium for games.

Preview of results. A New Keynesian economy can be understood as a game among the consumers. How so? Because a consumer's optimal spending depends on others' spending via three GE channels: the feedback for aggregate spending to income (the Keynesian cross); the feedback for aggregate spending to inflation (the Phillips curve); and the response of monetary policy (the Taylor rule). The first two channels contribute to strategic complementarity, opening the door to equilibrium multiplicity; the third pulls in the opposite direction.[5]

In Sections 2 and 3, we formalize this prism as simply and transparently as possible, and use it to translate the Taylor principle to the following requirement: let the third channel be strong enough so as to guarantee a unique equilibrium when consumers can perfectly coordinate their behavior over time (which is what the representative-agent benchmark means for our purposes). In the rest of the paper, we then proceed to accommodate a friction in such coordination and to show how this can guarantee a unique equilibrium regardless of monetary policy.

For our main result, developed in Section 4, we model the friction as follows. In each period, a consumer learns perfectly the concurrent shocks, fundamental or otherwise; with probability $\lambda \in [0,1)$, she knows nothing else; and with the remaining probability, she inherits the information of another, randomly selected, player from the previous period. This lets λ parameterize the "depth" of social memory: for any t, the fraction of the population who "remembers" and can condition their actions on the shocks realized at any $\tau \leq t$ is $(1-\lambda)^{t-\tau}$.

The standard, representative-agent, case is nested with $\lambda = 0$; it translates to common knowledge of the economy's history (which defines what "perfect" coordination means for us); and it admits a continuum of sunspot and backward-looking equilibria whenever the Taylor principle is violated. Proposition 2 shows that, as soon as $\lambda > 0$, all these equilibria unravel. Indeed, only the fundamental/MSV solution survives, regardless of whether policy is active or passive.

Strictly speaking, this result precludes direct observation of the actions of others, or of endogenous outcomes such as inflation and output. But because such outcomes are functions of the underlying shocks, in the limit as $\lambda \to 0$ nearly all consumers are nearly perfectly, albeit heterogeneously, informed about arbitrarily long histories of *both* shocks and outcomes. From this perspective, our result illustrates the fragility of the model's non-fundamental solutions to the introduction of small, idiosyncratic noise in the knowledge of the economy's history.

We corroborate this message in Section 5 with two additional results, both of which are motivated by recursive representations of full-information equilibria. In such representations, finite memory of appropriately chosen endogenous variables helps replicate infinite memory of past

[5] The second channel is shut off with rigid prices and the third one is shut off with interest rate pegs. But the first channel is always there—whether hidden behind the Euler condition of the representative consumer in the textbook New Keynesian model, or salient in the "intertemporal Keynesian cross" of HANK models (Auclert et al., 2018).

shocks. For instance, pure sunspot equilibria are replicated by having agents keep track merely of today's sunspot and yesterday's inflation or output. Proposition 5 shows that this replication is itself fragile: it breaks as soon as we let the agents' observation of past inflation or output be contaminated with arbitrarily small idiosyncratic noise. Proposition 6 adds that a similar fragility can be present even when past outcomes are *perfectly* observed, provided that, for every period t, there is a small shock to fundamentals that is known at t but is "forgotten" at $t+1$.

The common thread between our results can be explained as follows. All the model's sunspot and backward-looking solutions are sustained by the following infinite chain: in any given period, current consumers are responding to a variable that has no intrinsic effect on either themselves or future agents (e.g., the current sunspot or the past rate of inflation) because they expect to be "rewarded" appropriately by future agents, who in turn are expected to act on the basis of a similar expectation about the behavior further into the future, and so on. Because such purely self-fulfilling chains have no anchor on fundamentals, they are exceedingly sensitive to perturbations of common knowledge. This echoes the literature on global games, subject to the following twist: in our context, the most relevant coordination is that of behavior over time, which in turn explains why the relevant perturbations relate, one way or another, to aggregate memory.

Interpreting our contribution. Our paper is subject to a similar qualification as the applied literature on global games: our results may or may not extend to other perturbations of common knowledge. Still, we view our results as a reinforcement of the logical foundations of the conventional practice, which takes for granted that the New Keynesian model's MSV solution is the "right" lens for interpreting the data and for guiding policy.

By the same token, our results leave no space for the FTPL, at least within the New Keynesian model: regardless of whether monetary policy is active or passive, a non-Ricardian fiscal policy implies non-existence of equilibrium. To paraphrase Kocherlakota and Phelan (1999), rejection of the non-Ricardian assumption turns from a religious choice to a logical necessity.[6]

Finally, we should note that sunspot-like volatility can obtain even when the equilibrium is unique, from overreaction to noisy public news (Morris and Shin, 2002) or shocks to higher-order beliefs (Angeletos and La'O, 2013). From this perspective, our contribution is not to rule out "animal spirits" altogether but rather to facilitate the study of their policy implications free of the equilibrium-selection conundrum. To put it differently, our results help recast policies that lean against animal spirits as a form of stabilization as opposed to equilibrium selection.

[6] Let us be clear that fiscal policy may still matter, but only insofar as it enters the model's MSV solution. This could be the case, for example, when taxes and deficits enter the IS curve because consumers are liquidity constrained, or when the debt burden enters the Taylor rule because the monetary authority internalizes fiscal objectives.

Related literature. Although Cochrane (2011, 2017, 2018) has been the most vocal advocate of the FTPL recently, this theory and the associated debate on whether equilibrium is selected by an "active" monetary policy or a "non-Ricardian" fiscal policy go back to Leeper (1991), Sims (1994) and Woodford (1995). Canzoneri, Cumby, and Diba (2010) review the debate and discuss how it fits in the broader context of the fiscal-monetary interaction.

As already alluded to, the debate has morphed in different forms over time. In response to correlated criticisms by Kocherlakota and Phelan (1999), McCallum (2001), Buiter (2002), Niepelt (2004) and others, Cochrane (2005) proposes that the government's intertemporal budget should be read as a valuation equation for government bonds, akin to the valuation equation for a company's stock; and he goes on to argue that the blow-up interpretation of the FTPL makes no sense under this prism. Bassetto (2002) offers a detailed, micro-structure foundation of the FTPL that, too, bypasses the blow-up criticism. Atkeson, Chari, and Kehoe (2010), on the other hand, propose a class of "sophisticated" monetary policies that avoid the corresponding criticism of the Taylor principle. The bottom line is that, although the debate has morphed in different forms over time, it has never ended. This is because the underlying core issue has always been the same one: the indeterminacy implied by interest-rate pegs (Sargent and Wallace, 1975).

What distinguishes our contribution is the attempt to resolve this indeterminacy—and, by extension, escape the endless disagreement on the meaning and sensibility of the non-Ricardian assumption—by introducing a friction in coordination. Our first result (Proposition 2), in particular, brings to mind Morris and Shin (1998, 2003) and Abreu and Brunnermeier (2003). Although the application and the formal argument are different, there is a close resemblance in terms of the discontinuity of equilibria to perturbations of common knowledge and the role of higher-order beliefs. Our third result (Proposition 6), on the other hand, is more closely connected to Bhaskar (1998) and Bhaskar, Mailath, and Morris (2012), which show that only Markov Perfect Equilibria survive in a class of games when a purification in payoffs is combined with certain restrictions in social memory. Together, our results hint at deep connections between seemingly disparate literatures, which deserve further exploration.

A large literature has already incorporated information/coordination frictions in the New Keynesian model (Mankiw and Reis, 2002; Woodford, 2003; Maćkowiak and Wiederholt, 2009; Angeletos and Lian, 2018). But it has *not* addressed the determinacy issue. Instead, it has focused on how information shapes the model's MSV solution and has assumed away all other solutions by invoking, implicitly or explicitly, the Taylor principle. We do the exact opposite: our perturbations remove all other solutions without necessarily affecting the MSV solution itself.

A different literature has attempted to refine the model's solutions by requiring that they are E-stable (McCallum, 2007; Christiano et al., 2018). This approach relies on specific assumptions about what it means for an equilibrium to be "learnable" and has had mixed success on the topic of interest.[7] Still, we view this approach and ours as complements in that they both contribute towards reinforcing the logical foundations of the standard, Keynesian, approach.

Although we commit on REE, both the indeterminacy problem and our resolution of it extend to a larger class of solution concepts, including cognitive discounting (Gabaix, 2020), diagnostic expectations (Bordalo et al., 2018), and Bayesian equilibrium with mis-specified priors about one another's knowledge or rationality (Angeletos and Sastry, 2021). Relative to REE, these concepts relax the perfect coincidence of subjective beliefs and objective distributions and, in so doing, can potentially relax the Taylor principle. But they do not escape the indeterminacy problem, because they preserve a fixed-point relation between beliefs and behavior.

By contrast, Level-K Thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019) produces a unique solution precisely because it shuts down this fixed point relation. But whenever the environment admits multiple REE, the Level-K solution becomes infinitely sensitive to the assumed Level-0 behavior as the depth of reasoning gets larger. In this sense, this concept does not "really" resolve the indeterminacy issue; it only translates one free variable (animal spirits or equilibrium selection) to another free variable (the analyst's choice of Level-0 behavior).

2 A Simplified New Keynesian Model

In this section we introduce our version of the New Keynesian model. This contains two unusual assumptions: a specific OLG structure for the consumers and an ad hoc Phillips curve. These assumptions ease the exposition, especially once we perturb common knowledge of the economy's history; but as discussed in Section 5, they do not drive the results.

An intertemporal Keynesian cross (aka a Dynamic IS equation)

Time is discrete and is indexed by t. There are overlapping generations of consumers, each living two periods. A consumer born at t has preferences given by

$$E_{i,t}\left[u(C^1_{i,t}) + \beta u(C^2_{i,t+1})e^{-\varrho_t}\right],$$

where $C^1_{i,t}$ and $C^2_{i,t+1}$ are consumption when young and old, respectively, $u(C) \equiv \frac{1}{1-1/\sigma}C^{1-1/\sigma}$, $\beta \in (0,1)$ is a fixed scalar, ϱ_t is an intertemporal preference shock (the usual proxy for aggregate demand shocks), and $E_{i,t}$ is the consumer's expectation. Young and old consumers earn the same

[7] For example, sunspot equilibria can be E-stable if the interest rate rule is written as a function of expected inflation (Honkapohja and Mitra, 2004). And there is a debate on how the E-stability of backward-looking solutions depends on the observability of shocks (Cochrane, 2011; Evans and McGough, 2018).

income. Young consumers can borrow or save using the single asset traded in the economy, a one-period nominal bond; old consumers pay out any outstanding debt, or eat their savings, before they die. The budget constraint of a consumer born at t are therefore given by $C^1_{i,t} + B_{i,t} = Y_t$ and $C^2_{i,t+1} = Y_{t+1} + \frac{I_t}{\Pi_{t+1}} B_{i,t}$, where where $B_{i,t}$ is her saving/borrowing in the first period, I_t is the (gross) nominal interest rate between t and $t+1$, and Π_{t+1} is the corresponding inflation rate.

Old consumers are "robots:" they face no optimizing margin, their consumption mechanically adjusts to meet their end-of-life budget. Young consumers, instead, are "strategic:" they optimally choose consumption and saving/borrowing, given their available information. After the usual log-linearization,[8] this translates to the following consumption function:

$$c^1_{i,t} = E_{i,t}\left[\tfrac{1}{1+\beta} y_t + \tfrac{\beta}{1+\beta} y_{t+1} - \tfrac{\beta}{1+\beta}\sigma(i_t - \pi_{t+1} - \varrho_t)\right]. \tag{1}$$

This is basically the Permanent Income Hypothesis. The only subtlety is that we have allowed young consumers to be imperfectly informed about, or inattentive to, current income and current interest rates—which explains why y_t and i_t appear inside the expectation operator.

Pick any t. Because the average saving/borrowing of the young has to be zero, $\int c^1_{i,t} di = y_t$; and because the average net wealth of old has to be zero as well, $\int c^2_{i,t} di = y_t$. Combining, we infer that the two groups consume the same—or equivalently that aggregate consumption, c_t, coincides with the average consumption of the young. Computing the latter from (1), and imposing $y_t = c_t$, we infer that, for any process of interest rate and inflation, the process for aggregate spending must satisfy the following equation:

$$c_t = \bar{E}_t\left[\tfrac{1}{1+\beta} c_t + \tfrac{\beta}{1+\beta} c_{t+1} - \tfrac{\beta}{1+\beta}\sigma(i_t - \pi_{t+1} - \varrho_t)\right], \tag{2}$$

where $\bar{E}_t[\cdot] = \int E_{i,t}[\cdot]\, di$ is the average expectations of the young.

As evident from its derivation, this equation combines consumer optimality with market clearing; and it encapsulates the positive feedback between aggregate spending and income, holding constant the real interest rate. This equation can thus be read as a special case of the "intertemporal Keynesian cross" (Auclert et al., 2018), or as a Dynamic IS equation.

Connection to standard New Keynesian model

Although equation (2), our version of the Dynamic IS equation, looks different from the familiar textbook counterpart, it actually nests it when there is full information. Indeed, in this benchmark $\bar{E}_t$ can be replaced by $\mathbb{E}_t$, which henceforth denotes the full-information rational expectation; and because full information implies knowledge of concurrent outcomes in any rational

[8]Throughout, we log-linearize around the steady state in which $\varrho_t = 0$, $\Pi_t = 1$, and $I_t = \beta^{-1}$; and we use lower-case variables to denote log-deviations from steady state.

expectations equilibrium, equation (2) reduces in this case to

$$c_t = \frac{1}{1+\beta} c_t + \frac{\beta}{1+\beta} \mathbb{E}_t[c_{t+1}] - \frac{\beta}{1+\beta} \sigma (i_t - \mathbb{E}_t[\pi_{t+1}] - \varrho_t),$$

or equivalently

$$c_t = \mathbb{E}_t[c_{t+1}] - \sigma(i_t - \mathbb{E}_t[\pi_{t+1}] - \varrho_t).$$

Clearly, this the same as the Euler condition of a representative, infinitely-lived consumer.

This clarifies the dual role of the adopted micro-foundations. With full information, they let our model translate to the standard, representative-agent, New Keynesian model. And away from this benchmark, they ease the exposition by letting the intertemporal Keynesian cross take a particularly simple form and by equating the players in our upcoming game representation to the young consumers. These simplifications are relaxed in Section 5, without changing the essence.

A Phillips curve and a Taylor rule

For the main analysis, we abstract from optimal price-setting behavior (firms are "robots") and impose the following, ad hoc Phillips curve:

$$\pi_t = \kappa (y_t + \xi_t), \tag{3}$$

where $\kappa \geq 0$ is a fixed scalar and ξ_t is a "supply" or "cost-push" shock. As discussed in Section 5, our results are robust to replacing (3) with the fully micro-founded, forward-looking, New Keynesian Phillips curve; these same is true if we employ a Neoclassical Phillips curve a la Lucas (1972). In all cases, the essence (for our purposes) is that there is a positive GE feedback from aggregate output to inflation. Equation (3) merely stylizes this feedback in a convenient form.

We finally assume that monetary policy follows a Taylor rule:

$$i_t = z_t + \phi \pi_t, \tag{4}$$

where z_t is a random variable, possibly correlated with ϱ_t and ξ_t, and $\phi \geq 0$ is a fixed scalar that parameterizes how aggressively the monetary authority raises the interest rate in response to inflationary pressures. As is well known and will be reviewed shortly, $\phi > 1$ is necessary and sufficient for uniqueness of bounded equilibrium in the standard paradigm—but *not* under our perturbations. Our results will indeed apply even if $\phi = 0$, which nests interest rate pegs.[9]

[9]Note that an interest rate peg can be state-contingent, via z_t; and that the latter can be correlated, possibly in an optimal way, with ϱ_t and ξ_t. Similarly to King (2000) and Atkeson et al. (2010), this allows to disentangle the stabilization and equilibrium selection functions of Taylor rules: the former is served by the design of z_t, the latter by the restriction $\phi > 1$.

The model in one equation—and the economy as a game

From (3) and (4), we can readily solve for π_t and i_t as simple functions of y_t, which itself equals c_t. Replacing into (2), we conclude that the model reduces to the following equation:

$$c_t = \bar{E}_t \left[(1-\delta_0)\theta_t + \delta_0 c_t + \delta_1 c_{t+1}\right] \tag{5}$$

where δ_0, δ_1 are fixed scalars and θ_t is a random variable, defined by

$$\delta_0 \equiv \frac{1-\beta\sigma\phi\kappa}{1+\beta} < 1, \quad \delta_1 \equiv \frac{\beta+\beta\sigma\kappa}{1+\beta} > 0, \quad \theta_t \equiv -\frac{1}{1+\phi\kappa\sigma}\left(\sigma z_t - \sigma\varrho_t + \sigma\phi\kappa\xi_t - \sigma\kappa\mathbb{E}_t[\xi_{t+1}]\right).$$

By construction, equation (5) summarizes private sector behavior and market clearing, for any information structure and any monetary policy. Different information structures change the properties of $\bar{E}_t$ but do not change the equation itself. Similarly, different monetary policies map to different values for δ_0 or different stochastic processes for θ_t, via the choice of, respectively, a value for ϕ or a stochastic process for z_t. But for any given monetary policy, we can understand equilibrium in the private sector by studying equation (5) alone.

Equation (5) and the micro-foundations behind it also facilitate the interpretation of the economy as a certain infinite-horizon game. In this game, the only players acting at t are the young consumers of that period (old consumers, firms, and the monetary authority are "robots," in the sense already explained) and their best responses are obtained by combining their optimal consumption functions with first-order knowledge of market clearing, the Phillips curve and the Taylor rule. This gives the *individual* best response at t as

$$c_{i,t} = E_{i,t}\left[(1-\delta_0)\theta_t + \delta_0 c_t + \delta_1 c_{t+1}\right], \tag{6}$$

and recasts (5) as the period-t average best response function. Under this prism, δ_0 and δ_1 parameterize, respectively, the intra-temporal and the inter-temporal degree of strategic complementarity, while θ_t identifies the game's fundamental (i.e., the only payoff-relevant exogenous random variable). Finally, by regulating the strength of the underlying GE feedbacks, different values for β, κ, and ϕ map to different degrees of strategic complementarity.

This game-theoretic prism is not strictly needed for proving our results, which work directly with (5). But it allows a one-to-one mapping between the Rational Expectations Equilibria of our economy and the Perfect Bayesian Equilibria of the game described above; it helps build insightful connections to the literatures on global games and beauty contests; and, once we add fiscal policy to the model (Section 8), it helps clear some of the confusion that the existing literature on the FTPL has created about how the non-Ricardian assumption works and what constitutes a "fundamental" in the New Keynesian model.

Fundamentals, sunspots, and the equilibrium concept

Aggregate uncertainty is of two sources: fundamentals and sunspots. As already mentioned, the former are herein conveniently summarized in θ_t.[10] We assume that this variable is a stationary, zero-mean, Gaussian process, admitting a finite-state representation.

Assumption 1 (***Fundamentals***)**.** *The fundamental θ_t admits the following representation:*

$$\theta_t = q' x_t \qquad with \qquad x_t = R x_{t-1} + \varepsilon_t^x, \tag{7}$$

where $q \in \mathbb{R}^n$ is a vector, R is an $n \times n$ matrix of which all the eigenvalues are within the unit circle (to guarantee stationarity), $\varepsilon_t^x \sim \mathcal{N}(\mathbf{0}, \Sigma_\varepsilon)$, and Σ_ε is a positive definite matrix.

This directly nests the case in which (ϱ_t, ξ_t, z_t) follows a VARMA of any finite length. It also allows x_t to contain "news shocks," or forward guidance about future monetary policy. We henceforth refer to x_t as the *fundamental state.*

We next introduce a sunspot variable:

Assumption 2 (***Sunspots***)**.** *The only source of aggregate uncertainty other than that behind x_t is a sunspot. This is given by a random variable $\eta_t \sim \mathcal{N}(0,1)$, which is independent of past, current and future fundamentals and is distributed independently and identically over time.*[11]

Let h^t capture the history of all shocks, fundamental or not, up to and including period t. To simplify the exposition, we assume that histories are infinite and, accordingly, focus on stationary equilibria. More precisely, we let $h^t \equiv \{x_{t-k}, \eta_{t-k}\}_{k=0}^{\infty}$ and we define an equilibrium as follows:

Definition 1 (**Equilibrium**)**.** *An equilibrium is any solution to equation (5) along which: expectations are rational, although potentially based on imperfect and heterogeneous information about h^t; the outcome is a stationary, linear function of the underlying shocks, or*

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \gamma_k' x_{t-k} \tag{8}$$

where $a_k \in \mathbb{R}$ and $\gamma_k \in \mathbb{R}^n$ are known coefficients for all k; and the outcome is bounded in the sense that $Var(c_t)$ is finite.

Recall that consumer optimality, firm behavior, and market clearing have already been embedded in equation (5). It follows that the above definition is the standard definition of a Rational Expectations Equilibrium (REE), except for the addition of three "auxiliary" restrictions:

[10] The fact that θ_t contains an expectation term does interfere with our results. For instance, it suffices to assume that the fundamental state x_t, introduced below, is a sufficient statistic $(z_t, \varrho_t, \xi_t, \bar{E}_t[\xi_{t+1}])$ and therefore also for θ_t.

[11] Although we are restricting η_t to be uncorrelated over time, we are not ruling out persistent sunspot fluctuations: such fluctuations are still possible insofar as agents condition their behavior on past sunspots. Furthermore, as discussed at the end of Section 4, our results are robust to letting η_t itself be persistent, except for a degenerate case.

stationarity, linearity, and boundedness. The stationarity restriction, which comes hand-in-hand with the assumption of infinite history, can readily be relaxed. The linearity restriction is strictly needed for tractability, but we do not have any reason to believe that it drives our results, plus it is commonplace in the literature. The last requirement is our version of "local determinacy" or "bounded equilibria."[12] This is herein treated as a primitive; but as usual, it can be justified by an "exit" strategy along the lines of Taylor (1993), Christiano and Rostagno (2001) and Atkeson et al. (2010), namely a commitment to switch from the Taylor rule to a money-growth-targeting regime, or to whatever it takes for keeping inflation (and the output gap) within some bounds.[13] Finally, and circling back to our game-theoretic prism, note that the following is true: because every agent is infinitesimal, one's deviations are of no consequence for others, there is hence no need to specify off-equilibrium beliefs, and our REE notion is basically the same as PBE.[14]

3 The Standard Paradigm and the Taylor Principle

In this section we consider the full-information version of our model, which is, in essence, the standard New Keynesian model. We first identify the model's fundamental/MSV solution; we next show how its determinacy hinges, under full information, on the Taylor principle; and we finally contextualize our departures from this benchmark.

Full information

Suppose that all consumers know the entire h^t, at all t. As shown earlier, it is then *as if* there is a representative, fully informed and infinitely lived, consumer—just as in the textbook case. Accordingly, equation (5), which summarizes equilibrium, reduces to the following:

$$c_t = \theta_t + \delta \mathbb{E}_t[c_{t+1}], \tag{9}$$

where $\mathbb{E}_t[\cdot] \equiv \mathbb{E}[\cdot|h_t]$ is the rational expectation conditional on full information and

$$\delta \equiv \frac{\delta_1}{1-\delta_0} = \frac{1+\kappa\sigma}{1+\phi\kappa\sigma} > 0.$$

[12] Note that $Var(c_t)$ can be finite only if there exists a scalar $M > 0$ such that $|a_k| \le M$ and $\|\gamma_k\|_1 \le M$ for all k, where $\|\cdot\|_1$ is the L^1-norm. Our upcoming result actually uses only this weaker form of boundedness.

[13] The credibility of such exit strategies, their precise formulation, and the subtlety of whether they amount to a threat of "equilibrium non-existence" (Cochrane, 2007) or a more "sophisticated" implementation (Atkeson et al., 2010), are important topics beyond the scope of our paper. The relevant observation for our purposes is, instead, the following: whereas the boundedness requirement must be combined with the Taylor principle in order to deliver global determinacy in the standard paradigm, it will alone do the job under our perturbations.

[14] Of course, the stationarity, linearity, and bounded restrictions embedded in our REE definition must be extended to its PBE counterpart for this equivalence to be exact.

Although δ is necessarily positive, it can be on either side of 1, depending on ϕ. We will see momentarily how this relates to equilibrium determinacy. Also note that the above is a single, first-order, difference equation in c_t alone. By contrast, the textbook New Keynesian model maps to a system of *two* such equations in the vector (c_t, π_t). What affords the present reduction in dimensionality is the omission of a forward-looking term in the Phillips curve. But as it will become clear in Section 5, this simplification is inessential. All we have done thus far is to reduce the standard model's determinacy question from a two-dimensional eigenvalue problem to the simpler question of whether δ, or equivalently the sum $\delta_0 + \delta_1$, is higher or lower than 1.

The fundamental/MSV Solution

Because equation (9) is purely forward looking and x_t is a sufficient statistic for both the concurrent θ_t and its expected future values, it is natural to look for a solution in which c_t is a function of x_t alone. Thus guess $c_t = \gamma' x_t$ for some $\gamma \in \mathbb{R}^n$; use this to compute $\mathbb{E}_t[c_{t+1}] = \gamma' R x_t$; and substitute into (9) to get $c_t = \theta_t + \delta\gamma' R x_t = [q' + \delta\gamma' R]x_t$. Clearly, the guess is verified if and only if γ' solves $\gamma' = q' + \delta\gamma' R$, which in turn is possible if and only if $I - \delta R$ is invertible (where I is the $n \times n$ identity matrix) and $\gamma' = q'(I - \delta R)^{-1}$.

This is known as the model's "fundamental" or "minimum state variable (MSV)" solution (McCallum, 1983). To guarantee its existence, we henceforth impose the following assumption:

Assumption 3. *The matrix $I - \delta R$ is invertible.*

And we write this solution as $c_t = c_t^F$, where

$$c_t^F \equiv q' (I - \delta R)^{-1} x_t. \tag{10}$$

As shown momentarily, other solutions are possible when, and only when, $\delta \geq 1$. But let us first note the following property of the MSV solution. Suppose that the infinite sum $\sum_{k=0}^{\infty} \delta^k R^k$ exists. Then, $(I - \delta R)^{-1} = \sum_{k=0}^{\infty} \delta^k R^k$ and

$$c_t^F = \sum_{k=0}^{\infty} \delta^k \mathbb{E}_t[\theta_{t+k}].$$

This illustrates that c_t^F can depend on the economy's history only insofar as this pins down the current θ_t or helps forecast its future values. And it verifies that c_t^F maps to what Blanchard (1979) calls the "forward-looking solution," namely the solution of iterating (9) forward.[15]

[15]What if $\sum_{k=0}^{\infty} \delta^k R^k$ does not exist (i.e., the sum fails to converge)? In this case, c_t^F remains an REE but is no more solvable by forward induction; and its correlation with θ_t can switch sign. This relates to whether the MSV solution can feature "neo-Fisherian" effects (Cochrane, 2017; García-Schmidt and Woodford, 2019), a question that is interesting but separate from that considered here. For our purposes, the relevant quality of the MSV solution is this: along it, history matters only insofar as it is part of x_t, the fundamental state variable. This contrasts with the model's other solutions, along which payoff-irrelevant histories serve as correlation devices.

Determinacy under full information and the Taylor Principle

We now turn attention to the question of whether there exist equilibria other than the MSV one. Let us first fix language:

Definition 2 (**Taylor principle**). *The Taylor principle is defined by the restriction* $\phi > 1$.

Note that $\phi > 1$ translates to $\delta_0+\delta_1 < 1$ and, equivalently, $\delta < 1$. The former can be read as "the overall degree of strategic complementarity is small to guarantee a unique equilibrium," the latter as "the dynamics are forward-stable." And conversely, $\phi < 1$ translates to "the complementarity is large enough to support multiple equilibria" ($\delta_0+\delta_1 > 1$) and the "dynamics are backward-stable" ($\delta > 1$). This underscores the tight connection between our way of thinking about determinacy (the size of the strategic complementarity) and the standard way (the size of the eigenvalue). The next proposition verifies this point and also characterizes the type of equilibria that emerge in addition to the MSV solution once the Taylor principle is violated.

Proposition 1 (**Full-information benchmark**). *Suppose that h^t is known to every i for all t, which means in effect that there is a representative, fully informed, agent. Then:*

(i) There always exist an equilibrium, given by the fundamental/MSV solution c_t^F, as in (10).

(ii) When the Taylor principle is satisfied, the above equilibrium is the unique one.

(iii) When this principle is violated, there exist a continuum of equilibria, given by

$$c_t = (1-b)c_t^F + bc_t^B + ac_t^\eta, \tag{11}$$

where $a, b \in \mathbb{R}$ are arbitrary scalars and c_t^B, c_t^η are given by

$$c_t^B \equiv -\sum_{k=1}^{\infty} \delta^{-k}\theta_{t-k} \qquad \text{and} \qquad c_t^\eta \equiv \sum_{k=0}^{\infty} \delta^{-k}\eta_{t-k}. \tag{12}$$

To understand the type of non-fundamental equilibria documented in part (iii) above, take equation (9), backshift it by one period, and rewrite it as follows:

$$\mathbb{E}_{t-1}[c_t] = \delta^{-1}(c_{t-1} - \theta_{t-1}). \tag{13}$$

Since η_t is unpredictable at $t-1$, the above is clearly satisfied with

$$c_t = \delta^{-1}(c_{t-1} - \theta_{t-1}) + a\eta_t, \tag{14}$$

for any $a \in \mathbb{R}$. As long as $\delta > 1$, we can iterate backwards to obtain

$$c_t = -\sum_{k=1}^{\infty} \delta^{-k}\theta_{t-k} + a\sum_{k=0}^{\infty} \delta^{-k}\eta_{t-k} = c_t^B + ac_t^\eta. \tag{15}$$

This is both bounded, thanks to $\delta > 1$, and a rational-expectations solution to (13), by construction. This verifies that $c_t^B + ac_t^\eta$ constitutes an equilibrium, for any $a \in \mathbb{R}$. Part (iii) of the Proposition adds that the same is true if we replace c_t^B with any mixture of it and the MSV solution.

When there are no fundamental shocks, $c_t^F = c_t^B = 0$ and the solution obtained above reduces to a pure sunspot equilibrium, of arbitrary aptitude a. Along it, agents respond to the current sunspot because and only because they expect future agents to keep reacting to it, in perpetuity.[16]

In the presence of fundamental shocks, the indeterminacy takes an additional, perhaps more disturbing, form: the same path for interest rates and other fundamentals can result to different paths for aggregate spending and inflation, even if we switch off the sunspots. Consider, for example, the solution given by $c_t = c_t^B$. Along it, the outcome is pinned down by past fundamentals and is invariant to both the current value of θ_t and any news about its future path—which is the exact opposite of what happens along c_t^F, the MSV solution.

The logic behind c_t^B is basically the same as that behind sunspot equilibria: agents respond to past shocks that are payoff-irrelevant looking forward, because and only because they expect future agents to keep doing the same, in perpetuity. This statement extends to any equilibrium of the form (11) for $b \neq 0$, and explains why all such equilibria can be thought of as both non-fundamental and backward-looking.[17]

Beyond the full-information benchmark: a challenge and the way forward

Consider conditions (14) and (15). Clearly, these are equivalent representations of the same equilibrium: the first is recursive, the second is sequential. This equivalence means that all the equilibria that can be supported by perfect knowledge of $h_t = \{x_{t-1}, \eta_{t-k}\}_{k=0}^{\infty}$ coincide with those that can be supported by perfect knowledge of (x_t, η_t) and (θ_{t-1}, c_{t-1}). But what if agents lack such perfect knowledge, as it is bound to the case in reality?

Regardless of what agents know or don't, one can *always* represent any equilibrium in a sequential form, or as in equation (8). This is simply because c_t *has* to be measurable in the history of exogenous aggregate shocks, fundamental or otherwise. But it is far from clear if and when there is an equivalent recursive representation. In fact, a finite-state recursive representation is generally impossible when agents observe noisy signals of endogenous outcomes, due to the infinite-regress problem first highlighted by Townsend (1983).[18]

This poses a challenge for what we want to do in this paper. On the one hand, we seek to highlight how fragile all non-fundamental solutions can be to perturbations of the aforementioned kinds of common-knowledge, or to small frictions in coordination. On the other hand, we need to make sure that these perturbations do not render the analysis intractable.

[16] This is the same as a traditional, rational-expectations bubble, except that it is not explosive, thanks to $\delta > 1$.

[17] Blanchard (1979) refers to the analogue of c_t^B in his analysis as a "backward-looking fundamental equilibrium;" but this is not *really* fundamental, in the sense we just explained.

[18] See Huo and Takayama (2021) for a detailed study of the issue.

To accomplish this dual goal, in the rest of the paper we follow three distinct but complementary strategies. Our main strategy, in Sections 4–5, takes off from (15), or the sequential representation; the other two strategies, in Section 7, circle back to (14), the recursive representation. All three strategies illustrate the fragility of non-fundamental equilibria, each one from a different angle. And they all keep the analysis tractable.

4 Uniqueness with Fading Social Memory

This section contains our main result. We introduce a friction in social memory and show how it yields a unique equilibrium no matter ϕ, or the size the strategic complementarity.

Main assumption

For the purposes of this and the next section, we replace the assumption of a representative, fully-informed agent with the following, incomplete-information variant:

Assumption 4 (**Social memory**)**.** *In every period t, a consumer's information set is given by*

$$I_{i,t} = \{(x_t, \eta_t), \cdots, (x_{t-s}, \eta_{t-s})\},$$

where $s \in \{0, 1, \cdots\}$ is drawn from a geometric distribution with parameter λ, for some $\lambda \in (0, 1]$.

To understand this assumption, note that herein s indexes the length of the history of shocks that the consumer knows. Next, recall that the geometric distribution means that $s = 0$ with probability λ, $s = 1$ with probability $(1-\lambda)\lambda$, and more generally $s = k$ with probability $(1-\lambda)^k\lambda$, for any $k \geq 0$. By the same token, the fraction of agents who know *at least* the past k realizations of shocks is given by $\mu_k \equiv (1-\lambda)^k$.

One can visualize this as follows. At every t, the typical player (young consumer) learns the concurrent shocks; with probability λ, she learns nothing more; and with the remaining probability, she inherits the information of another, randomly selected player from the previous period (a currently old consumer). In this sense, λ parameterizes the speed with social memory (or common-p belief of past shocks) fades over time.

Main result

The full-information benchmark can be nested with $\lambda = 0$, which translates to $I_{i,t} = h_t$ (perfect knowledge of the infinite history) for all i and t. But the question of interest is what happens for $\lambda > 0$, and in particular as $\lambda \to 0^+$. In this limit, the friction becomes vanishingly small, in the sense that almost every agent knows the history of shocks up to an arbitrarily distant point in the past. But the following is also true: no matter how small λ is, as long as it is not exactly zero, we have that $\lim_{k\to\infty}\mu_k = 0$, which means that shocks are expected to be "forgotten" in the very distant future. As shown next, this causes all non-fundamental equilibria to unravel.

Proposition 2 (**Determinacy without the Taylor principle**). *Suppose that social memory is imperfect in the sense of Assumption 4, for any $\lambda > 0$. Regardless of ϕ, or of δ_0 and δ_1, the equilibrium is unique and is given by the fundamental/MSV solution.*[19]

A detailed proof is provided in the Appendix. Here, we illustrate the main idea for the special case in which there are no fundamental disturbances, so the task reduces to checking for the existence of pure sunspot equilibria. That is, we specialize our equilibrium condition to

$$c_t = \delta_0 \bar{E}_t[c_t] + \delta_1 \bar{E}_t[c_{t+1}]; \tag{16}$$

we search for solutions of the form $c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k}$; and we verify that $a_k = 0$ for all k.

By Assumption 4, we have that, for all $k \geq 0$,

$$\bar{E}_t[\eta_{t-k}] = \mu_k \eta_{t-k}$$

where $\mu_k \equiv (1-\lambda)^k$ measures the fraction of the population at any given date that know, or remember, a sunspot realized k periods earlier. Future sunspots, on the other hand, are known to nobody. It follows that, along any candidate solution, average expectations satisfy

$$\bar{E}_t[c_t] = \bar{E}_t\left[\sum_{k=0}^{\infty} a_k \eta_{t-k}\right] = \sum_{k=0}^{+\infty} a_k \mu_k \eta_{t-k}$$

and similarly

$$\bar{E}_t[c_{t+1}] = \bar{E}_t\left[a_0\eta_{t+1} + \sum_{k=1}^{\infty} a_k \eta_{t+1-k}\right] = 0 + \sum_{k=0}^{+\infty} a_{k+1}\mu_k \eta_{t-k}.$$

By the same token, condition (16) rewrites as

$$\underbrace{\sum_{k=0}^{+\infty} a_k \eta_{t-k}}_{c_t} = \delta_0 \underbrace{\sum_{k=0}^{+\infty} a_k \mu_k \eta_{t-k}}_{\bar{E}_t[c_t]} + \delta_1 \underbrace{\sum_{k=0}^{+\infty} a_{k+1}\mu_k \eta_{t-k}}_{\bar{E}_t[c_{t+1}]}.$$

For this to be true for all sunspot realizations, it is necessary and sufficient that, for all $k \geq 0$,

$$a_k = \mu_k(\delta_0 a_k + \delta_1 a_{k+1}). \tag{17}$$

[19] Note that the fundamental/MSV solution remains the same as we move away from $\lambda = 0$ thanks to the assumption that $I_{i,t}$ contains x_t always. As mentioned in the Introduction, this helps isolate our contribution from the existing literature on informational frictions, which focuses on how the MSV solution is influenced by imperfect information about x_t but does not address the determinacy issue. Here, we do the exact opposite, but one could have it both ways: modify Assumption 4 so as to remove perfect information about x_t and reshape the MSV solution, while also preserving our argument for uniqueness.

Since $\delta_0 < 1$, $\delta_1 > 0$, and $\mu_k \in (0,1)$, the above is equivalent to

$$a_{k+1} = \frac{1-\delta_0\mu_k}{\delta_1\mu_k} a_k; \tag{18}$$

and because $\mu_k \to 0$ and hence $\frac{1-\delta_0\mu_k}{\delta_1\mu_k} \to \infty$ as $k \to \infty$, we have that $|a_k|$ explodes to infinity as $k \to \infty$ (and hence so does the variance of c_t), unless $a_0 = 0$. But $a_0 = 0$ implies $a_k = 0$ for all k. We conclude that the unique bounded equilibrium is $a_k = 0$ for all k, which herein corresponds to the MSV solution, since we have switched off the fundamental shocks. The proof in the Appendix extends the argument to the presence of such shocks.

Comparison to full information and the boundedness restriction

We expand on the intuition behind the above argument momentarily. But first, it is useful to repeat it for the knife-edge case with $\lambda = 0$. In this case, $\mu_k = 1$ for all k and condition (18) becomes

$$a_{k+1} = \delta^{-1} a_k,$$

where, recall, $\delta \equiv \frac{\delta_1}{1-\delta_0} = \frac{1+\kappa\sigma}{1+\phi\kappa\sigma}$. When $\delta < 1$ (equivalently $\phi > 1$), this still explodes as $k \to \infty$ unless $a_0 = 0$, which means that the unique bounded solution is once again $a_k = 0$ for all k. But when $\delta > 1$, the above remains bounded, and indeed converges to zero as $k \to \infty$, for arbitrary $a_0 = a \in \mathbb{R}$. This recovers the standard model's sunspot equilibria.

Note how *both* the standard argument with $\lambda = 0$ and our variant with $\lambda > 0$ use the boundedness assumption, namely that a_k does not explode. But whereas this assumption must complemented with the Taylor principle in order to rule out sunspot equilibria in the standard case, it *alone* does the job under our perturbation. The Taylor principle has become redundant.[20]

Intuition and additional remarks

Although the adopted micro-foundations pin down the values for δ_0 and δ_1 as specific functions of the underlying preference, technology and policy parameters, our argument did not rely at all on these restrictions. With this in mind, let us momentarily ignore these restrictions, set $\delta_0 = 0$ and $\delta_1 = \delta$ for arbitrary δ (possibly even negative), and simplify condition (17) to

$$a_k = \delta\mu_k a_{k+1}. \tag{19}$$

[20] Recall the standard justification of the boundedness assumption: the monetary authority commits to follow a Taylor rule as long as π_t, or c_t, stays within some bounds, and to switch from such an interest-rate-setting regime to a money-supply-setting regime, or to some other appropriately specified exit strategy, if inflation goes outside these bounds. As shown most clearly in Atkeson et al. (2010), in the standard paradigm such "hybrid" rules avoid the ad hoc boundedness assumption but continue to require the Taylor principle to guarantee determinacy. Our result, instead, suggests that such hybrid rules can do the job *without* the Taylor principle: it suffices to have a consensus that the monetary authority will "do whatever it takes" to keep inflation, or the output gap, within some bounds.

Focus now on the effects of the first-period sunspot and let $\{\frac{\partial c_t}{\partial \eta_0}\}_{t=0}^{\infty}$ stand for the corresponding impulse response function (IRF). We can then rewrite condition (19) as

$$\frac{\partial c_t}{\partial \eta_0} = \delta \mu_t \frac{\partial c_{t+1}}{\partial \eta_0}.$$

This is the same condition as that characterizing the IRF of c_t to η_0 in a "twin" representative-agent economy, in which condition (5) is modified as follows:

$$c_t = \tilde{\delta}_t \mathbb{E}_t[c_{t+1}], \quad \text{with} \quad \tilde{\delta}_t \equiv \delta \mu_t.$$

Under this prism, it is *as if* we are back to the standard New Keynesian model but the relevant eigenvalue, or the overall strategic complementarity, has become time-varying and has been reduced from δ to $\tilde{\delta}_t$. Furthermore, because $\mu_t \to 0$ as $t \to \infty$, we have that there is T large enough but finite so that $0 < \tilde{\delta}_t < 1$ for all $t \geq T$, regardless of δ. In other words, the twin economy's dynamic feedback becomes weak enough that c_t cannot depend on η_0 after T. By induction then, c_t cannot depend on η_0 before T either.[21]

This interpretation of our result must be clarified as follows. In the above argument, we studied the response of c_t to η_0. This means that our "twin" economy is defined from the perspective of period 0, and that $\tilde{\delta}_t = \mu_t \delta$ measures the feedback from $t+1$ to t in a very specific sense: as perceived from agents in period 0, when they contemplate whether to react to η_0. To put it differently, in this argument t indexes not the calendar time but rather the belief order, or how far into the future agents reason about the effects of an innovation today.

Let us explain. Because η_0 is payoff irrelevant in every single period, period-0 agents have an incentive to respond to it *if and only if* they are confident that period-1 agents will also respond to it, which can be true only if they are also confident that period-1 will themselves be confident that period-2 agents will do the same, and so on, ad infinitum. It is this kind of "infinite chain" that supports sunspot equilibria when $\lambda = 0$. And conversely, the friction we have introduced here amounts to the typical period-0 agent reasoning as follows:

> "I can see η_0. And I understand that it would make sense to react to it if I were confident that all future agents will keep conditioning their behavior on it *in perpetuity*. But I worry that future agents will fail to do so, either because they will be unaware of it, or because they may themselves worry that agents further into the future will not react to it. By induction, I am convinced that it makes sense not to react to η_0 myself."

Three remarks complete the picture. First, the reasoning articulated above, and the proof given earlier, can be understood as a chain of contagion effects from "remote types" (uninformed

[21] Although this argument assumed $\delta_0 = 0$, it readily extends to $\delta_0 \neq 0$. In this case, the twin economy has both δ_0 and δ_1 replaced by, respectively, $\mu_t \delta_0$ and $\mu_t \delta_1$. That is, both types of strategic complementarity are attenuated.

agents in the far future) to "nearby types" (informed agents in the near future) and thereby to present behavior. This underscores the high-level connection between our approach and the global games literature (Morris and Shin, 1998, 2003). Second, the aforementioned worries don't have to be "real" (objectively true). That is, we can reinterpret Assumption 4 as follows: agents don't necessarily forget themselves but believe that others will forget.[22] Finally, consider how such worries influence the response to a persistent innovation in the fundamental. Even if all future agents fail to react to it, current agents have an incentive to react to it, because it has a direct effect on their own payoffs. This highlights the following point: although all full-information equilibria, including the MSV solution, embed perfect coordination, the MSV solution is not as fragile as all other equilibria to the friction under consideration.

Persistent sunspots—and endogenous state variables

Let us now revisit the assumption that the sunspot η_t is uncorrelated over time. Proposition 2 readily extends to an arbitrary ARMA process for the sunspot, except for one knife-edge case: when η_t follows an AR(1) process with autocorrelation *exactly* equal to δ^{-1}. In this case, $c_t = c_t^F + a\eta_t$ is an equilibrium for any a and is supported by knowledge of (x_t, η_t) alone. Of course, such a situation seems exceedingly unlikely if the sunspot is an exogenous random variable. But could it be that an *endogenous* variable, such as c_{t-1}, can serve the same function?

We will return to this question in Section 7, but we offer a preliminary answer here:

Proposition 3 (**Nearly perfect knowledge of past outcomes**)**.** *Under Assumption 4, almost all agents become arbitrarily well informed about arbitrarily long histories of c_t as $\lambda \to 0$: for any mapping from h^t to c_t as in Definition 1, any $K < \infty$ arbitrarily large but finite, and any $\epsilon, \epsilon' > 0$ arbitrarily small but positive, there exists $\hat{\lambda} > 0$ such that, whenever $\lambda \in (0, \hat{\lambda})$, $Var\left(E_t^i[c_{t-k}] - c_{t-k}\right) \leq \epsilon$ for all $k \leq K$, for at least a fraction $1 - \epsilon'$ of agents, and for every period t.*

In this sense, our uniqueness result is compatible with indirect but almost perfect knowledge of current and past outcomes: it is *as if* agents have received arbitrarily precise signals about $\{c_t, c_{t-1}, ..., c_{t-K}\}$, and by extension for $\{\pi_t, \pi_{t-1}, ..., \pi_{t-K}\}$ and $\{i_t, i_{t-1}, ..., i_{t-K}\}$, too, for arbitrarily large K. In Section 7, we will show that uniqueness is compatible even with *perfect* knowledge of current and past outcomes, provided that we consider another perturbation.

[22] Strictly speaking, this requires a modification of the solution concept: from REE to PBE with misspecified priors about one another's knowledge, along the lines of Angeletos and Sastry (2021). But the essence is the same.

5 A Generalization

In this section we extend Proposition 2 to a more flexible class of games, featuring rich forward-looking behavior; we next explain how this allows us to nest a more standard New Keynesian model, where consumers are long lived and firms set prices optimally; and we finally discuss the key limitation of Assumption 4 and some ways around it.

A generalization

At this point it should be clear that the micro-foundations of equation (5) played no essential role in our argument. This suggest the following generalization. Maintain our assumptions about stochasticity and information but replace equation (5) with the following:

$$c_t = \bar{E}_t \left[\theta_t + \sum_{k=0}^{+\infty} \delta_k c_{t+k} \right] \tag{20}$$

for some scalars $\{\delta_k\}_{k=0}^{\infty}$, with $\delta_0 < 1$ and $\Delta \equiv \delta_0 + \sum_{k=1}^{\infty} |\delta_k| < \infty$. And interpret this equation as the average best response of a dynamic game in which (i) a continuum of players acts in each period, (ii) a player's optimal strategy is given by $c_{it} = E_{it}\left[\theta_t + \sum_{k=0}^{+\infty} \delta_k c_{t+k}\right]$, for any strategy played by other players now and in the future, and (iii) the coefficient δ_k identifies the slope of an agent's best response with respect to the average action k periods later.

This generalization allows outcomes to depend on expectations of outcomes in the entire infinite future, not just next period; and this dependence could be of arbitrary size and sign.[23] The analogue of the sum $\delta_0 + \delta_1$ from our main analysis, a measure of the overall strength of strategic interdependence, is now given by Δ. With $\Delta > 1$, multiple self-fulfilling equilibria can be supported under full information, in a similar fashion as in Section 3. But they unravel once $\lambda > 0$, because this again breaks the "infinite chain" behind them.

We verify this claim below. The proof is more convoluted than that of Proposition 4, and is delegated to the Appendix. But the basic logic is the same.

Proposition 4 (Generalized result)**.** *Consider the generalization described above, specify information as in Assumption 4, and let* $\lambda > 0$. *Whenever an equilibrium exists, it is unique and is given by the MSV solution.*[24]

[23] We are only restricting $\delta_0 < 1$. This is necessarily true in the (extended) New Keynesian model we describe next as long as $\phi \geq 0$ and it means that multiplicity can originate only from the dynamic feedback between c_t and $\{\bar{E}_t[c_{t+1}]\}_{k=1}^{\infty}$, as opposed to the static feedback between c_t and $\bar{E}_t[c_t]$. More succinctly, there is a unique "temporary" equilibrium for given expectations of the future, even though there may be multiple "dynamic" equilibria because of the endogeneity of these expectations.

[24] The MSV solution is now given by $c_t^F = \gamma' x_t$, where γ solves $\gamma' = q' + \delta_0 \gamma' + \sum_{k=0}^{\infty} \delta_k R^k \gamma'$. Clearly, this solution exists if and only if $(1-\delta_0)I - \sum_{k=0}^{\infty} \delta_k R^k$ exists and is invertible, which is the present analogue of Assumption 3.

Nesting a more conventional New Keynesian model

We now sketch how the above generalization of our uniqueness result helps accommodate a more "canonical" New Keynesian model. Consider, in particular, an OLG version, where the typical consumer dies with probability $\omega \in [0,1)$ in each period. Similarly to Farhi and Werning (2019) and Angeletos and Huo (2021), mortality risk lets the model accommodate a large MPC, building a bridge to HANK. More importantly for our purposes, death provides a concrete way to think about the decay in social memory. In particular, if we let consumers have perfect information about the aggregate shocks realized over their lifetime, and no information about those realized before their birth, we basically get a version of Assumption 4 with $\lambda = \omega$. But we could also let $\lambda > \omega$ or $\lambda < \omega$, to proxy for situations where consumers are, respectively, more or less informed.

Let us spell out the details. Preferences are standard, given by expected life-time utility. As long as she is alive, a consumer can use actuarily fair annuities to insure against the risk of dying with non-zero wealth. If she dies, she gets replaced by a newborn consumer with zero wealth. Each consumer may be subject to uninsurable, idiosyncratic shocks in the income she receives, the prices of the particular commodities she consumes, and the interest rate at which she can borrow or save. Finally, each consumer is rational and chooses consumption-saving optimally; but she may have to do so on the basis of imperfect information about, or with inattention to, income, wealth and other economic conditions.

This leads to the following optimal consumption function, after log-linearization:

$$c_{i,t} = E_{it}\left[(1-\beta\omega)\,w_{i,t} - \beta\omega\sigma \sum_{k=0}^{+\infty} (\beta\omega)^k \left(i_{i,t+k} - \pi_{i,t+k+1}\right) + \left(1-\beta\omega\right) \sum_{k=0}^{+\infty} (\beta\omega)^k\, y_{i,t+k}\right], \tag{21}$$

where $w_{i,t}$, $y_{i,t}$, $i_{i,t}$, $\pi_{i,t}$ are the individual's wealth, income, interest rate, and inflation rate. The above generalizes (1), the consumption function in our baseline model. The first term inside the expectation operator captures wealth, the second term captures intertemporal substitution, and the last term captures permanent income. $\beta\omega$ is the effective discount factor (inclusive of mortality risk) and $1-\beta\omega$ is the MPC (out of financial wealth and permanent income alike).

We now wish to express aggregate consumption as a function of the average expectations of interest rates, inflation and income. This task is complicated by the fact that aggregation of (21) yields the average expectations of *individual* variables. To go from them to the average expectations of the corresponding *aggregate* variables, we impose the following assumption, whose precise meaning we discuss right after:

Assumption 5 (No misperceptions). *Every consumer's information satisfies $I_{i,t} = I_{i,t}^{agg} \cup I_{i,t}^{idio}$, where $I_{i,t}^{agg} \subseteqq h_t$ and where $I_{i,t}^{idio}$ is such that, for any variable $x \in \{y, i, \pi, w\}$, $E_{it}[x_{it} - x_t] = f^{x,t}(I_{i,t})$ for some linear function $f^{x,t}$. Furthermore, $\int I_{i,t}^{idio} di$ is invariant to h_t.*

In words, a consumer's information is "orthogonalized" in two components: one capturing information about aggregate shocks ($I_{i,t}^{agg}$), and another capturing information about idiosyncratic shocks ($I_{i,t}^{idio}$). And the latter is invariant to h_t, at least on average across consumers.

What does this assumption rule out, what does it allow, and what does it buys us? It rules out misperception of aggregate shocks for idiosyncratic shocks—a possibility that may be realistic but seems largely orthogonal to the issue we are studying in this paper. It allows for otherwise unrestricted information—including arbitrarily precise information about the underlying sunspot and aggregate fundamentals. And it buys exactly what we want: for every $x \in \{y, i, \pi, w\}$, we have that $\int E_{it}[x_{it} - x_t]di = 0$, or equivalently $\int E_{it}[x_{it}]di = \bar{E}_t[x_t]$.[25] Also note that this assumption is entirely consistent with the representative-agent benchmark: in this benchmark, the assumption is trivially satisfied with $I_{i,t}^{idio} = \emptyset$ and $I_{i,t}^{agg} = h_t$. In a nutshell, this assumption affords a flexible specification of the information about all aggregate shocks, including sunspots, while avoiding the "rabbit hole" of the possible confusion between idiosyncratic and aggregate shocks.

We are now in business. Aggregating (21) across i, using the above property (that $\int E_{it}[x_{it}]di = \bar{E}_t[x_t]$), and using also the fact that aggregate wealth is zero (because there is no capital), we get

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

Next, since the goods market must clear in all periods (and the consumers know this fact), we can replace $y_{t+k} = c_{t+k}$ in the above to obtain the following equilibrium restriction:

$$c_t = -\beta\omega\sigma \left\{ \sum_{k=0}^{+\infty} (\beta\omega)^k \bar{E}_t [i_{t+k} - \pi_{t+k+1}] \right\} + (1-\beta\omega) \left\{ \sum_{k=0}^{+\infty} (\beta\omega)^k \bar{E}_t [c_{t+k}] \right\}. \tag{22}$$

This is the present analogue of condition (2) from our baseline model, or another example of the intertemporal Keynesian cross.

Turning to the production side of the economy, suppose, for simplicity, that firms are perfectly informed and exactly the same as in the New Keynesian model. This means that the ad hoc, static Phillips curve employed in the main analysis is now replaced by the micro-founded, forward-looking, New Keynesian Phillips curve:

$$\pi_t = \kappa c_t + \beta \mathbb{E}_t [\pi_{t+1}] + \kappa \xi_t, \tag{23}$$

for some $\kappa \geq 0$. Finally, let the Taylor rule be

$$i_t = z_t + \phi_c c_t + \phi_\pi \pi_t, \tag{24}$$

for some $\phi_c, \phi_\pi \geq 0$.

[25] To see this, note that the first part of the assumption gives $\int E_{it}[x_{it} - x_t]di = f^{x,t}\left(\int I_{i,t}\right)$. The second part then says that this object is a constant (invariant to h_t). Finally, because the unconditional expectation satisfies $E\left[\int E_{it}[x_{it} - x_t]di\right] = E[x_{it} - x_t] = 0$, this constant has to be zero. That is, $\int E_{it}[x_{it} - x_t]di = 0$, as claimed above.

Solving (23) and (24) for inflation and the interest rate, and replacing these solutions into (22), we can obtain c_t as a function of $\bar{E}_t[c_{t+k}]$ for all $k \geq 0$. That is, we can summarize all the model's equilibrium conditions to a special case of condition (20). In particular, the strategic complementarity coefficients are given by

$$\delta_k \equiv \left(1-\beta\omega-\beta\omega\sigma\phi_c\right)\left(\beta\omega\right)^k + \omega\sigma\kappa\left(-\beta\phi_\pi + \left(1-\beta\omega\phi_\pi\right)\frac{1-\omega^k}{1-\omega}\right)\beta^k,$$

for all k; and they readily satisfy $\delta_0 < 1$ and $\Delta < \infty$, for any $\beta, \omega \in (0,1)$ and $\sigma, \kappa, \phi_\pi, \phi_c \geq 0$. By the same token, Proposition 4 holds *exactly* in the present economy, provided that we impose Assumption 4 with $I_{i,t}^{agg}$ in place of $I_{i,t}$. That is, the assumption restricts what consumes know about aggregate shocks, but not necessarily what they know about idiosyncratic shocks.

To recap, our insights do not depend on the specifics consumer and firm behavior, or the precise structure of the intertemporal Keynesian cross and the Phillips curve. Instead, the only relevant question is the validity and meaning of Assumptions 4 and 5.

Markets, endogenous information, and endogenous coordination devices

We think that the aforementioned assumptions are the "right" ones for our purposes: the one rules out misperceptions, the other formalizes the friction of interest. But in a realistic market context, the available information may naturally confound different shocks (Lucas, 1972), plus one's information is bound to be endogenous to others' choices. This creates a tension between the micro-foundations of interest and the assumptions behind our uniqueness result.

In our view, this tension is moderated by the fact that we have concentrated on the limit as $\lambda \to 0^+$, which, as explained at the end of the previous section, translates to nearly perfect, private knowledge of arbitrarily long histories of aggregate outcomes, or the average actions of others. But let us be clear: so far we have established equilibrium uniqueness for a sequence of *exogenous*-information economies that converges to the standard, full-information benchmark as $\lambda \to 0$, and it is not clear how exactly this maps to *endogenous*-information economies.

This consideration, along with our earlier discussion about recursive equilibria, motivates the analysis of Section 7. There, we show how our message goes through if replace Assumption 4 with two other assumptions, which allow for direct signals of the aggregate outcomes (and hence also for endogenous coordination devices). But before exhausting the readers' patience with additional theoretical exercises, we invite them to think of our $\lambda \to 0^+$ limit as, at the very least, a refinement of full-information equilibria; and we discuss what this means for applied purposes.

6 Applied Lessons: FTPL and stabilization

In this section, we extend the analysis to allow for fiscal policy and explain in detail what our result means vis-a-vis the FTPL. We also offer a refined take on the equilibrium selection and stabilization functions of monetary policy,

On the Fiscal Theory of the Price Level

Let us momentarily go back to the basics: the textbook, three-equation, New Keynesian model. Add a fourth equation, the government's intertemporal budget constraint, written compactly (and in levels) as follows:

$$\frac{B_{t-1}}{P_t} = PVS_t, \tag{25}$$

where B_{t-1} denotes the outstanding nominal debt, P_t denotes the nominal price level, and PVS_t denotes the present discounted value of primary surpluses. Does the incorporation of this equation make a difference for the model's predictions about inflation and output?

The conventional approach says no by assuming that fiscal policy is Ricardian, in the sense that it adjusts to make sure that (25) holds along the MSV solution, which itself is pinned down by the model's other three equations alone. The FTPL argues the opposite by letting (25) be satisfied for a *different* solution, and by letting that solution identify the model's overall equilibrium.

To illustrate, consider a negative shock to tax revenue. Because such a shock does not enter the New Keynesian model's three famous equations, it does not change its MSV solution. But it of course reduces PVS_t, other things equal. Thus suppose that the following is true after the shock:

$$\frac{B_{t-1}^{nominal}}{P_t^{MSV}} > PVS_t,$$

where P_t^{MSV} denotes the price level predicted by the MSV solution. How is equilibrium restored? One possibility (the Ricardian regime) is that government raises taxes and/or cuts spending so as to make sure that (25) is satisfied along the MSV solution. But another possibility (the non-Ricardian regime) is that the government fails to do so, and yet the government remains solvent, because a different solution obtains and along it the price level adjusts to

$$P_t = P_t^{FTPL} \equiv \frac{B_{t-1}}{PVS_t} > P_t^{MSV}.$$

In a nutshell, the FTPL selects a particular sunspot equilibrium and uses it to clear the government's intertemporal budget.

As mentioned in the Introduction, Kocherlakota and Phelan (1999), Buiter (2002) and others have argued that a non-Ricardian policy amounts to a threat by the government to "blow up" its budget, and induce equilibrium non-existence, unless the right sunspot equilibrium is selected. But Bassetto (2002) and Cochrane (2005, 2018) provide alternative interpretations, which not only bypass the blow-up criticism but also blur the sharp separation between (25) and the model's other equations, making one wonder if it is misleading to think of the solution picked by the FTPL as a sunspot equilibrium. After all, aren't fiscal conditions part of an economy' fundamentals?

Given the space constraints, we can't give full justice to either side of the debate. But it seems fair to say that the literature has failed to converge, that some insiders have agreed to disagree, and that some outsiders remain confused about what exactly the theory says. We now explain how exactly our result fits in this debate, and how it helps remove some of the existing confusion.

To start with, revisit our characterization of optimal consumption. Relative to what we did in the previous section, there are exactly three changes: first, we let $\omega = 0$ so that consumers are infinitely lived and fiscal policy does not redistribute wealth across generations (a possibility that is empirically plausible but orthogonal to the FTPL); second, aggregate disposable income is $y_t - \tau_t$ instead of y_t, where τ_t are the taxes; and finally, the consumers' aggregate financial wealth is $b_{t-1} - p_t$ instead of 0, where $b_{t-1} - p_t$ is the real value of the outstanding nominal debt. Accordingly, aggregate consumption can now be expressed as follows:

$$c_t = \bar{E}_t\left[(1-\beta)(b_{t-1}-p_t) - \beta\sigma\sum_{k=0}^{+\infty}(i_{t+k}-\pi_{t+k+1}) + \left(1-\beta\right)\sum_{k=0}^{+\infty}\beta^k\left(y_{t+k}-\tau_{t+k}\right)\right]. \tag{26}$$

As in the previous section, the key step for obtaining (26) from the corresponding individual consumption functions is to impose Assumption 5. In the present context, this translates to the following: at least on average, consumers do not misperceive a change in fiscal policy as an improvement or deterioration of their *idiosyncratic* fortunes. (Such confusion, rational or otherwise, may be empirically plausible; but it is *not* what the existing versions of the FTPL are about.)

Next, consider the government. The per-period budget constraints together with a non-Ponzi condition give the following intertemporal restriction (after log-linearization):

$$b_{t-1} - p_t = \sum_k \beta^k(\tau_{t+k} - g_{t+k}). \tag{27}$$

Note that this has to hold for every realization of future uncertainty, not just on average, because debt is non-contingent. But if this holds with probability 1, it must also hold under the expectation of every agent, and hence under the average expectation as well, provided of course that agents are rational. Using this fact along with $y_t = c_t + g_t$ (from market clearing), we can reduce (26) to the following equation:

$$c_t = -\beta\sigma\left\{\sum_{k=0}^{+\infty}\beta^k\bar{E}_t\left[i_{t+k}-\pi_{t+k+1}\right]\right\} + \left(1-\beta\right)\left\{\sum_{k=0}^{+\infty}\beta^k\bar{E}_t\left[c_{t+k}\right]\right\}, \tag{28}$$

which is *exactly* the same condition as that obtained in the previous section.

The description of the economy is completed by the Taylor rule and by the NKPC. The first remains exactly the same as before. The second now includes g_t in the real marginal cost, as usual, but this is formally the same as re-defining ξ_t to contain g_t. Importantly, the outstanding level of debt and the path of deficits do not enter these equations, just as they do not enter (28). It follows that the analysis of the previous section remains intact. That is, the equilibrium process for c_t is still pinned down by equation (20); the equilibrium process for π_t and i_t are then pinned down by the Phillips curve and the Taylor rule; and the fiscal authority's policy rule, F, does not enter the determination of any of these objects.

Corollary. *Whenever an equilibrium exists, it corresponds to the MSV solution and has the following property: inflation and output are entirely invariant to the level of government debt and the fiscal policy rule F, regardless of whether monetary policy is active or passive.*

This statement starts with the qualification "whenever an equilibrium exists" to account for the following, obvious, possibility: even though equation (20) is well defined for *every F* and is invariant to it, the government may of course be insolvent for *some F*. This translates as follows: a fiscal policy *has* to be "Ricardian," or else it is incompatible with equilibrium existence.

More on the FTPL

Although our main point about the FTPL has already been made, it is worth expanding on the derivation of (28) and on the precise meaning of (20) in the present context. To reach (28), we started with the optimal consumption functions, not just with the Euler conditions; this avoided Cochrane's criticism that the Dynamic IS curve of the New Keynesian model does not fully capture consumer optimality. We next combined consumer optimality with market clearing and government solvency, or more precisely with the consumers' first-order knowledge of these properties, and showed that this recovers exactly the same "intertemporal Keynesian cross" as when there is no government. Finally, to go from (28) to (20), we used the consumers' first-order knowledge of the NKPC and of the Taylor rule.

Throughout these steps, we did *not* rely on Assumption 4 or on *any* other assumption about what information consumers may have about the underlying aggregate shocks. In fact, equations (28) and (20) would remain valid even if we had allowed consumers to have mis-specified priors about these shocks, the policy rule followed by the fiscal authority, and the information or the rationality of *other* consumers. For, as long as consumers are *themselves* rational, they can infer that public debt is not real wealth, and this suffices for their best responses to be invariant to both the outstanding level of public debt and the rule F. More succinctly, these objects do not appear anywhere in the game that summarizes the optimal individual behavior and all the GE interaction of the consumers—and this is true even in the full-information benchmark. Assumption 4 was ultimately used to guarantee that this game has a unique equilibrium, but the property that this game is invariant to fiscal policy applies more generally.

Why are we emphasizing these points? Because they help not only understand more precisely what we have shown but also avoid some of the confusion in the literature. To see what we have in mind, go back to the representative-agent, full-information case and to the particular example consider in the onset of this section, where PVS_t was treated as exogenous and

$$\frac{B_{t-1}}{P_t^{MSV}} > PVS_t = \frac{B_{t-1}}{P_t^{FTPL}}.$$

Cochrane (2005, Section 2.3.1) reads the above as follows: the MSV solution corresponds to an off-equilibrium situation in which the market value of government debt is too high, causing consumers to feel more wealthy than they ought to be in equilibrium. Because of this wealth effect, Cochrane argues, the representative consumer will try to buy more consumption than what possible in equilibrium. There will thus be excess aggregate demand, which will push the price level up, from P_t^{MSV} to P_t^{FTPL}. In his words, "We will then see 'aggregate demand' giving 'inflationary pressures,' as the Fed loves to say."

This narrative sounds plausible and may even be empirically relevant. But in the light of our analysis, it requires either an aggregate misperception or a departure from rationality. Why so? Because our derivation above showed that, as soon as consumers are optimizing, understand the economy's structure, and do not confound fiscal policy with idiosyncratic shocks, public debt and the rule F drop out of the picture, in the sense made precise above.[26]

To recap: Under our approach, the MSV solution emerges as the only possible equilibrium, regardless of whether monetary policy is active or passive. By the same token, there is no space for fiscal policy to be non-Ricardian and no reason to pontificate whether (25) is a true constraint or a valuation equation. But of course there is room for fiscal considerations—such as seigniorage or the real debt burden—to enter the monetary authority's choice of $\{z_t\}$ and thereby the MSV solution. In other words, this solution *itself* is logically consistent with the "unpleasant arithmetic" of Sargent and Wallace (1981), the evidence in Sargent (1982), the Ramsey literature on how monetary policy can substitute for fiscal policy and/or ease tax distortions (e.g., Chari et al., 1994; Benigno and Woodford, 2003; Correia et al., 2008), and some topical discussions. Perhaps this is what the FTPL is *meant* to be about, once freed up from the equilibrium selection conundrum.

[26]Plus, we think that our approach can offer more clarity even if one wishes to modify our information assumptions. At a high level, this is because we have move beyond the Walrasian perspective and have instead represented the economy as a game, where the consumers' best responses are well-defined even if their expectations about one another's behavior are off equilibrium.

Feedback rules, and equilibrium selection vs stabilization

We now comment on another applied lesson of our paper, a refined take on the equilibrium selection and stabilization functions of monetary policy.

Go back to the textbook New Keynesian model. Let $\{i_t^o, \pi_t^o, c_t^o\}$ denote the optimal path for the interest rate, inflation, and output, as a function of the underlying demand and supply shocks. And ask the following question: what does it take for this to be implemented as the *unique* bounded equilibrium?

The textbook answer goes as follows. If the monetary authority observes the fundamental shocks driving the Ramsey optimum, this optimum can be implemented as an equilibrium, and in particular as the model's MSV solution, by setting the interest rate as follows:

$$i_t = i_t^o.$$

But such a policy (an interest rate peg) is subject to a possible coordination failure: agents may end up playing a different equilibrium that the MSV solution. To avoid this risk, the monetary authority must instead commit to the following feedback rule:

$$i_t = i_t^o + \phi(\pi_t - \pi_t^o),$$

for any $\phi > 1$.[27] Note then that the value of ϕ, or the feedback from π_t to i_t, does not affect the properties of the optimum; it merely makes sure that the optimum is uniquely implemented, in the sense that no other bounded equilibrium is possible.[28]

But what if the monetary authority does not observe the underlying shocks? Feedback rules may then be useful for the purpose of replicating the optimal contingency of interest rates on shocks, or for optimal stabilization. And, in general, this function could be at odds with that of equilibrium selection. See Galí (2008, p.101) for an illustration with cost-push shocks, and Loisel (2021) for a general formulation. Seen from this perspective, our results help ease the potential conflict between equilibrium selection and stabilization: because feedback rules are no more needed for equilibrium selection, they are "free" to be used for stabilization.

At the same time, our results pave the way for recasting the *spirit* of the Taylor principle as a form of stabilization instead of a form of equilibrium selection. By this we mean the following. When the equilibrium is unique (whether thanks to our perturbations or otherwise), sunspot-like volatility may still obtain from overreaction to noisy public news (Morris and Shin, 2002) or shocks to higher-order beliefs (Angeletos and La'O, 2013). In particular, suppose that we relax As-

[27] This is nested in (4) with $z_t = i_t^o - \phi\pi_t^o$, which is indeed feasible as long as the monetary authority observes the shocks that drive the Ramsey optimum.

[28] While some textbook treatments stop here, the most careful ones combine the Taylor principle with escape clauses that take care of unbounded equilibria. See, e.g., Atkeson, Chari, and Kehoe (2010).

sumption 4 in our main analysis so as to remove common knowledge of the fundamental state, x_t, and accommodate independent shocks to higher-order beliefs of future monetary policy or other fundamentals. Then, we can maintain the MSV solution as the economy's unique equilibrium but also let this solution fluctuate in response to these shocks. In the eyes of an outside observer the economy may appear to be ridden with "animal spirits." And a policy that "leans against the wind" may well help contain the effects of such animal spirits basically in the same as it does with other, less exotic, demand and supply shocks.

7 Robustness: Observing Past Outcomes

In the end of Section 4, we highlighted that, although our key assumption excluded direct observation of the past endogenous outcomes, such as output and inflation, it allowed agents to face arbitrarily little uncertainty about them, in the sense of Proposition 3. We now push this argument further, by showing how uniqueness can obtain if we let agents have direct, and possibly even perfect, knowledge of the past outcomes. This also circles back to our discussion of sequential and recursive representations of full-information equilibria.

Recursive sunspot equilibria: an example of fragility

Consider our baseline model, where consumption choices must solve

$$c_{i,t} = \mathbb{E}_{i,t}[\delta_0 c_t + \delta_1 c_{t+1}]. \tag{29}$$

In the full-information case, this boils down to

$$c_t = \theta_t + \delta \mathbb{E}_t[c_{t+1}],$$

where $\delta = \frac{\delta_1}{1-\delta_0}$. Let us momentarily shut down the fundamentals, assume $\delta > 1$, and focus on the following, pure sunspot equilibrium:

$$c_t = c_t^{\eta} \equiv \sum_{k=0}^{\infty} \delta^k \eta_{t-k}. \tag{30}$$

As noted earlier, this can be represented in recursive form as

$$c_t = \eta_t + \delta^{-1} c_{t-1}. \tag{31}$$

It follows that perfect knowledge of yesterday's outcome can readily substitute for perfect knowledge of the infinite history of past sunspots. Intuitively, c_{t-1} serves as a sufficient statistic of the infinite history of sunspots.

Taken at face value, this challenges our message that multiplicity hinges on "infinite" memory: it suffices that agents have very short memories, provided that they keep track of the past average action, here c_{t-1} (or π_{t-1}). But as shown next, this logic itself is fragile.

Abstract from shocks to the fundamentals, let sunspots be Gaussian, and let information sets be given by

$$I_{i,t} = \{\eta_t, s_{i,t}\},$$

where

$$s_{i,t} = c_{t-1} + u_t + \varepsilon_{i,t}$$

is a signal of the past aggregate outcome, $u_t \sim \mathcal{N}(0, v)$ is aggregate noise, $\varepsilon_{i,t} \sim \mathcal{N}(0,\sigma)$ is idiosyncratic noise, and $v, \sigma \geq 0$are fixed parameters. One can interpret $s_t \equiv c_{t-1} + u_t$ as a publicly available macroeconomic statistic about past output (or inflation), u_t as the measurement error in this statistic, and $\varepsilon_{i,t}$ as idiosyncratic noise in i's observation of this statistic (or "cognitive noise" due to rational inattention). When $\sigma = 0$, every agent observes the *exact* same signal and this fact is common knowledge. In this case, multiplicity survives, regardless of how large v, the measurement error, is. When instead $\sigma > 0$ but arbitrarily small, agents knowledge of the macroeconomic statistic is only slightly blurred by idiosyncratic noise. As shown next, this causes sunspot equilibria to unravel.

Proposition 5. *Consider the economy described above. For any $\sigma > 0$, not matter how small, and regardless of δ_0 and δ_1, there is a unique equilibrium and it corresponds to the MSV solution.*

The proof is simple and instructive. We work out the special case in which the measurement error is absent ($v = 0$), and leave the more general case to the Appendix. Since information sets are given by $I_{i,t} = \{\eta_t, s_{i,t}\}$, any (stationary) strategy can be expressed as

$$c_{i,t} = a\eta_t + bs_{i,t},$$

for some coefficients a and b. Then, $c_{t+1} = a\eta_{t+1} + bc_t$; and since agents have no information about the *future* sunspot, $\mathbb{E}_{i,t}[c_{t+1}] = b\mathbb{E}_{i,t}[c_t]$. Next, note that $\mathbb{E}_{i,t}[c_t] = a\eta_t + b\chi s_{it}$, where

$$\chi \equiv \frac{Var(c_{t-1})}{Var(c_{t-1}) + \sigma^2} \in (0,1].$$

Combining these facts, we infer that condition (29), the individual best response, reduces to

$$c_{i,t} = \mathbb{E}_{i,t}[\delta_0 c_t + \delta_1 c_{t+1}] = (\delta_0 + \delta_1 b)\mathbb{E}_{i,t}[c_t] = (\delta_0 + \delta_1 b)\left\{a\eta_t + b\chi s_{it}\right\}.$$

It follows that a strategy is a best response to itself if and only if

$$a = (\delta_0 + \delta_1 b)a \qquad \text{and} \qquad b = (\delta_0 + \delta_1 b)b\chi. \tag{32}$$

Clearly, $a = b = 0$ is always an equilibrium, and it corresponds to the MSV solution. To have a sunspot equilibrium, on the other had, it must be that $a \neq 0$ (and also that $|b| < 1$, for it to be bounded). From the first part of condition (32), we see that this $a \neq 0$ if and only if $\delta_0 + \delta_1 b = 1$, which is equivalent to $b = \delta^{-1}$. But then the second part of this condition reduces to $1 = \chi$, which in turn is possible if and only if $\sigma = 0$ (since $Var(c_{t-1}) > 0$ whenever $a \neq 0$).

This circles back to our discussion of sequential and recursive equilibria. In the main analysis, we took off from the sequential form (30), represented information in terms of direct signals of the infinite history of shocks, and showed how a small perturbation in this domain results in a unique equilibrium. Here, we took off from the recursive form (31), represented information in terms of the relevant endogenous state variable, and showed how a small perturbation in this domain results, once again, in a unique equilibrium.

At this point one may raise the following question: could it be that multiple equilibria are supported by noisy idiosyncratic observations of *longer* histories of the endogenous outcome (or equivalently the average actions of others)? For example, what if $I_{i,t} = \{\eta_t\} \cup \{s_{i,t}^k\}_{k=1}^K$, where $s_{i,t}^k$ is a noisy private signal of c_{t-k}, for k running from 1 up to some finite $K \geq 1$? Or what if we replace η_t with noisy private signals of η_t itself and of its past values? We cannot handle such information structures in full generality, because of the complexities we alluded to earlier (signal-extraction and infinite regress). But we offer a complementary approach, which lets past outcomes be observed *perfectly*—and nevertheless obtains uniqueness.

Breaking the infinite chain even when past outcomes are perfectly observed

In the above exercise we focused on pure sunspot equilibria. Let us no bring back the fundamental shocks and consider any of the equilibria of the form $c_t^B + ac_t^\eta$, which, recall, were obtained by "solving the model backwards." These are replicated by letting each consumer play the following recursive strategy:

$$c_{i,t} = \delta^{-1}(c_{t-1} - \theta_{t-1}) + a\eta_t. \tag{33}$$

Contrary to the strategy that supported the pure sunspot equilibrium, the above strategy requires that the agents at t know not only c_{t-1} but also θ_{t-1}. Why is knowledge of θ_{t-1} necessary? Because this is what it takes for agents at t to know how to undo the direct, intrinsic effect of θ_{t-1} on the incentives of the agents at $t-1$, or to "reward" them for not responding to their intrinsic impulses.

This suggests that the "infinite chain" that supports all backward-looking equilibria—and all sunspot equilibria, as well—can break if the agents at t do not know what exactly it takes to reward the agents at $t-1$. To make this point crisply, we proceed as follows.

First, we introduce a new fundamental disturbance, denoted by ζ_t, which can be arbitrarily small but is not observed by future agents. That is, we modify equation (5) to

$$c_{i,t} = \mathbb{E}_{i,t}[(1-\delta_0)(\theta_t + \zeta_t) + \delta_0 c_t + \delta_1 c_{t+1}]. \tag{34}$$

The introduction of ζ_t allows us to maintain perfect knowledge of θ_t (or x_t) itself, for symmetry with the earlier analysis, and at the same time parameterize the perturbation by the support of ζ_t. In particular, we let ζ_t be drawn independently over time, as well as independently of any other shock in the economy, from a uniform distribution with support $[-\varepsilon, +\varepsilon]$, where ε is positive but arbitrarily small.

Second, we abstract from informational heterogeneity *within* periods, that is, we let $I_{i,t} = I_t$ for all i and all t. This guarantees that $c_{it} = c_t$ for all i and t, and therefore that we can think of the economy as a sequence of representative agents, or a sequence of players, one for each period, whose best response is given by

$$c_t = \theta_t + \zeta_t + \delta \mathbb{E}[c_{t+1} | I_t]. \tag{35}$$

where $\mathbb{E}[\cdot|I_t]$ is the rational expectation conditional on I_t, the information set of the period-t representative agent. But unlike the conventional case, today's representative agent need not inherit all the information of yesterday's representative agent: I_t does not necessarily nest I_{t-1}.

Finally, we let I_t contain perfect knowledge of arbitrary long histories of the endogenous outcome, the sunspots, and the "main" fundamental; but we preclude knowledge of the past values of the payoff perturbation introduced above. Formally:

Assumption 6. *For each t, there is a representative agent whose information is given by*

$$I_t = \{\zeta_t\} \cup \{x_t, \cdots, x_{t-K_\theta}\} \cup \{\eta_t, \cdots, \eta_{t-K_\eta}\} \cup \{c_{t-1}, \cdots, c_{t-K_c}\}$$

for finite but possibly arbitrarily large K_η, K_c, and K_θ.

When $\varepsilon = 0$ (the ζ_t shock is absent), Assumption 6 allows replication of all sunspot and backward-looking equilibria with extremely short memory, i.e., with $K_\eta = 0$ and $K_\theta = K_c = 1$. This is precisely the recursive representation of these equilibria in the standard paradigm. But there is again a discontinuity: once $\varepsilon > 0$, all the non-fundamental equilibria unravel, no matter how long the memory of outcomes and all other shocks may be.

Proposition 6. *Suppose that Assumption 6 holds and $\varepsilon > 0$. Regardless of δ, there is unique equilibrium and is given by $c_t = c_t^F + \zeta_t$, where c_t^F is the same MSV solution as before.*

To further illustrate the logic behind this result, abstract from the θ_t shock (but of course keep the ζ_t shock) and let $I_t = \{\zeta_t, \eta_t, c_{t-1}\}$. In this case, "solving the model backwards," which literally means having the agents at $t+1$ create indifference for the agents t, requires that

$$\mathbb{E}[c_{t+1}|I_t] = \delta^{-1}(-\zeta_t + c_t).$$

Since the only "news" contained in I_{t+1} relative to I_t are η_{t+1} and ζ_{t+1}, the above is true if and only if c_{t+1} satisfies

$$c_{t+1} = a\eta_{t+1} + d\zeta_{t+1} + \delta^{-1}(-\zeta_t + c_t)$$

for some $a, d \in \mathbb{R}$. As noted before, the agents at $t+1$ may extract information about ζ_t from their knowledge of c_t. But since ζ_t is not *directly* known and c_{t+1} has to be measurable in $I_{t+1} = \{\zeta_{t+1}, \eta_{t+1}, c_t\}$, the above condition can hold only if c_t itself is measurable in ζ_t and not in any other shock, such as the sunspots realized at t or earlier. In short, because the agents at $t+1$ does not know a (small) component of the "preferences" of the agents at t, it is impossible to support the aforementioned chain of indifference. The proof in the Appendix shows that an extension of this logic rules out all equilibria but the MSV solution.[29]

8 Additional Discussion

In this section, we expand on the connection of our paper to four literatures: the one on global games; the on on purification and Markov perfect equilibria; the one on discounted Euler conditions; and the one on Level-k Thinking.

High-level connections between our results, and between them and the literature

Our last result, Proposition 6, is closely connected to Bhaskar (1998) and Bhaskar, Mailath, and Morris (2012). These works have shown that only Markov Perfect Equilibria (which in our context translate to the MSV solution) survive in a certain class of games when a purification in payoffs is combined with "finite" social memory (the latter being defined in a manner analogous to Assumption 6 here). Even though our environment is different, Proposition 6 is a close cousin of these earlier results. But this is not the case for our first result, Proposition 2 or 4. There, the key assumption was of different kind (contrast Assumption 4 to Assumption 6), and so was the formal argument: there was a contagion from "remote" type (agents in the far future) to "nearby" types (agent in the near future), akin to those found in the global games literature.

[29] Note that $c_t = c_t^F + \zeta_t$ is MSV solution of the perturbed model. This differs from c_t^F, the original MSV solution, because the relevant fundamentals now include ζ_t. But as $\varepsilon \to 0$, the new solution converges to the old one. Also note that the argument given above goes through even if the ζ_t shock occurs only every, say, 10 periods rather than every single period, because once there is a chance that the chain will break at some future date the whole thing unravels. Finally, the argument goes through even that the agents at $t+1$ know ζ_t perfectly, provided that agents at t are (incorrectly) worried that this may not be the case.

A broad lesson of this literature is that determinacy ultimately hinges on whether information is private versus public (where "public" means not merely publicly available but common knowledge, or at least high common-p belief). The results of Mailath and Morris (2002) and Pęski (2012), which like the aforementioned works by Bhaskar (1998) and Bhaskar, Mailath, and Morris (2012) shift the focus to Markov Perfect equilibria in dynamic games, seem consistent with this logic: Mailath and Morris (2002) relies on "almost public monitoring" to support multiple, non-Markovian equilibria, and Pęski (2012) goes in the opposite direction to rule them out. But the precise relation between these literatures remains unclear, at least to us. Furthermore, our Proposition 5 offered an example that looked like almost public monitoring (everybody observed the past aggregate action only tiny idiosyncratic noise) and nevertheless obtained uniqueness. This is all to say that there are not only deep connections but also subtle differences, all of which deserve further study.

Finally, the results of Weinstein and Yildiz (2007) suggest that, although multiple equilibria may be "degenerate" in an appropriate topology, this statement by itself can be vacuous: with enough freedom in choosing priors and information structures, one can recast equilibrium indeterminacy as strategic uncertainty along a unique equilibrium. Under this prism, a key task for theory is to understand how a model's determinacy and its predictions more generally depend on strong common knowledge assumptions, and what are plausible relaxations thereof. We hope that our paper has made some progress in this direction.

Discounted Euler equations

Suppose we replace our IS equation (2) with the following variant:

$$c_t = -m_i i_t + m_\pi \bar{E}_t[\pi_{t+1}] + m_c \bar{E}_t[c_{t+1}] + \varrho_t, \tag{36}$$

for some positive scalars m_i, m_π, m_c. When $m_c < 1$, this nests the "discounted" Euler equations generated by liquidity constraints in McKay et al. (2017) and by cognitive discounting in Gabaix (2020). The opposite case, $m_c > 1$, is consistent with the broader HANK literature (Werning, 2015; Bilbiie, 2020), as well as with over-extrapolation or "cognitive hyperopia". Finally, $m_i \neq m_\pi$ could capture differential attention to (or salience of) nominal interest rates and inflation.

With these modifications, the entire analysis goes through modulo the following adjustment in the definition of δ:

$$\delta = \frac{m_\pi \sigma\kappa + m_c}{1 + m_i \sigma\phi\kappa}$$

The Taylor principle is still the same in the δ space, but of course changes in the ϕ space: we now have that $|\delta| < 1$ if and only if $\phi \in (-\infty, \underline{\phi}) \cup (\overline{\phi}, +\infty)$, where

$$\underline{\phi} \equiv -\frac{m_\pi}{m_i} - \frac{1+m_c}{\sigma\kappa m_i} \qquad \text{and} \qquad \overline{\phi} \equiv \frac{m_\pi}{m_i} + \frac{m_c - 1}{\sigma\kappa m_i}$$

Depending on the m's, these thresholds can be either smaller or larger than the ones in the main analysis. In this sense, the model's region of indeterminacy may either shrink or expand by the above modifications. For instance, Gabaix (2020) assumes $m_i = m_\pi$ and $m_c < 1$, obtains $\overline{\phi} < 1$, and uses this to argue that cognitive discounting relaxes the Taylor principle and, thereby, eases the potential conflict between the stabilization and equilibrium selection functions of monetary policy. From this perspective, that paper and ours are complements. But none of these enrichments changes the fact that indeterminacy remains for sufficiently "passive" monetary policy, and this is where our approach offers a potential way out.

Alternative Solution Concepts

Throughout, we have preserved Rational Expectations Equilibrium (REE), relaxing only the assumption of perfect information about the past. REE is defined by the requirement that the agents' subjective model of the economy *exactly* coincides with the true model generated by their behavior. One can capture bounded rationality by allowing a discrepancy between the former and the latter. But as long as one allows for a two-way feedback between them, the kind of indeterminacy we have studied here remains possible, and so does our resolution to it.

This circles back our earlier discussion of Gabaix (2020): the solution concept in that paper allows the objective model to feed into the subjective model, albeit with a distortion relative to REE. The same is true for Diagnostic Expectations (Bordalo et al., 2018); for Perfect Bayesian Equilibrium with mis-specified priors (Angeletos and Sastry, 2021); and for Woodford (2019)'s model of "finite planning horizons," at least once learning is allowed (Xie, 2019). All these concepts are close cousins of REE in the sense that they preserve the two-way feedback between beliefs and outcomes, thus also preserving the indeterminacy problem we have addressed in this paper.

Contrast this class of concepts with Level-K Thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019). The latter pins down a unique solution by shutting down the feedback from objective truth to subjective beliefs. But this begs the question of how agents adjust their behavior over time, in the light of repeated, systematic discrepancies between what they expect to happen and what actually happens. Accordingly, we believe that Level-K Thinking is more appropriate for unprecedented experiences (e.g., the recent ZLB experience) than for the kind of stationary environments we are concerned with in this paper.

Furthermore, one may argue that Level-K Thinking does not "really" resolve the indeterminacy problem and, instead, only translates it to a different dimension: whenever $|\delta| > 1$, the level-k outcome becomes *infinitely* sensitive to the arbitrary level-0 outcome as $k \to \infty$. In this sense, one free variable (the sunspot) is replaced by another free variable (the level-0 outcome).[30] By contrast, our approach leaves neither kind of freedom in specifying beliefs.

This is not to say that our approach is "better." After all, the value of getting uniqueness within the REE paradigm depends on how much one wants to commit to REE in the first place. Furthermore, the two approaches are complementary in two regards: highlighting the role of higher-order beliefs; and solidifying the logical foundations of the MSV solution. We hope that the above discussion clarifies the differences in the two approaches, but perhaps their common ground is what matters the most for applied purposes.

9 Conclusion

In this paper we revisited the indeterminacy issue of the New Keynesian model. We highlighted how all sunspot and backward-looking equilibria hinge on an infinite chain between current and future behavior. And we showed how "easy" it can be to break this chain by relaxing the model's assumptions about memory and intertemporal coordination.

To keep the analysis tractable, we followed a few stark but complementary approaches. Proposition 2 allowed rich information heterogeneity within each period at the expense of abstracting from direct observation of past aggregate outcomes. Proposition 6 considered the opposite extreme. Proposition 5 was somewhere in the middle. We discussed the limits of these approaches and speculated that, in general, determinacy is likely to depend on the subtler question of how much common knowledge is afforded both within and across time.

Notwithstanding the last point, our results left no space for the FTPL and lend support to the practice of focusing on its fundamental/MSV solution regardless of whether monetary policy is "active" or "passive." To put it differently, our results provided a rationale for why equilibrium can be determinate even with interest rate pegs—or why monetary policy may be able to regulate aggregate demand without a strict reliance on off-equilibrium threats of the kind embedded in

[30] To clarify this point, consider what Level-K Thinking means in our setting. First, level-0 behavior is exogenously specified, by a random process $\{c_t^0\}$. Level-1 behavior is then defined as the best response to the belief that others play according to level-0 behavior, that is, $c_t^1 \equiv \theta_t + \delta \mathbb{E}_t[c_{t+1}^0]$,where $\mathbb{E}_t$ is the full-information expectation operator. This amounts to using the "wrong" beliefs about what other players do but the "correct" beliefs about the random variables θ_t and c_{t+1}^0. Iterating K times, for any finite K, gives the level-K outcome as $c_t^K \equiv \sum_{k=0}^{K} \delta^k \mathbb{E}_t[\theta_{t+k}] + \delta^K \mathbb{E}_t[c_{t+K}^0]$. The solution concept says that actual behavior is given by $c_t = c_t^K$ for all periods and states of nature, where both K and $\{c_t^0\}$ are free variables for the modeler to choose. Clearly, $\{c_t^K\}$ is uniquely determined for any given K and any given $\{c_t^0\}$. But because $\{c_t^0\}$ is a free variable, the original indeterminacy issue is effectively transformed to the modeler's (or the reader's) uncertainty about $\{c_t^0\}$. Furthermore, the bite of this uncertainty is most severe precisely when the indeterminacy issue is present: whenever $|\delta| > 1$, the sensitivity of $\{c_t^K\}$ to $\{c_t^0\}$ explodes to infinity as $K \to \infty$.

the Taylor principle.

Let us close with a comment on the strategic interaction between the monetary and the fiscal authority. Specifying a proper game between the two authorities requires a unique mapping from those player's actions—interest rates and government deficits, respectively—to their payoffs. Such a unique mapping is missing in the standard paradigm, because the same paths for interest rates and government deficits can be associated with multiple equilibria within the private sector. By fixing this "bug," our paper allows one to study the monetary-fiscal interaction as a proper game, for example as a game of attrition between the two authorities.

Appendix: Proofs

As discussed after Definition 1, our proofs use a weaker boundedness criterion than the requirement of a finite $Var(c_t)$. The next lemma verifies that that the latter implies the former. The rest of the Appendix provides the proofs for all the results.

Lemma 1. *Consider any candidate equilibrium, defined as in Definition 1. There exist a finite scalar $M > 0$ such that $|a_k| \leq M$ and $\|\gamma_k\|_1 \leq M$ for all k, where $\|\cdot\|_1$ is the L^1-norm.*

Substituting (7) into (8), we have that any candidate equilibrium can be rewritten as

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \Gamma_k' \varepsilon_{t-k}^x, \tag{37}$$

where, for all $k \geq 0$,

$$\Gamma_k' \equiv \sum_{l=0}^{k} \gamma_{k-l}' R^l. \tag{38}$$

Since η_t and ε_t^x are independent of each other as well as independent over time, we have

$$Var(c_t) = \sum_{k=0}^{\infty} \left(a_k^2 + \Gamma_k' \Sigma_\varepsilon \Gamma_k \right).$$

This can be finite only if $\lim_{k\to+\infty} |a_k| = 0$ and $\lim_{k\to+\infty} \|\Gamma_k\|_1 = 0$.[31] From (38), $\gamma_k' = \Gamma_k' - \Gamma_{k-1}' R$ for all $k \geq 1$. It follows that $\lim_{k\to+\infty} \|\gamma_k\|_1 = 0$ as well. We conclude that there exist a scalar $M > 0$, large enough but finite, such that $|a_k| \leq M$ and $\|\gamma_k\|_1 \leq M$ for all k.

Proof of Proposition 1

Part (i) follows directly from the fact that $c_t^F \equiv q' (I - \delta R)^{-1} x_t$ satisfies (9).

Consider part (ii). Let $\{c_t\}$ be any equilibrium and define $\hat{c}_t = c_t - c_t^F$. From (9),

$$\hat{c}_t = \delta \mathbb{E}_t [\hat{c}_{t+1}]. \tag{39}$$

From Definition 1,

$$\hat{c}_t = \sum_{k=0}^{\infty} \hat{a}_k \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_k' x_{t-k},$$

with $|\hat{a}_k| \leq \hat{M}$ and $\|\hat{\gamma}_k'\|_1 \leq \hat{M}$ for all k, for some finite $\hat{M} > 0$. From Assumptions 1–2, we have

$$\mathbb{E}_t[\hat{c}_{t+1}] = \sum_{k=0}^{+\infty} \hat{a}_{k+1} \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_{k+1}' x_{t-k} + \hat{\gamma}_0' R x_t.$$

[31] To prove the latter statement, note that, because Σ_ε is positive definite, there exists an invertible L such that $\Sigma_\varepsilon = L'L$ by Cholesky decomposition. The finiteness of $Var(c_t)$ then implies $\lim_{k\to+\infty} \|L\Gamma_k\|_1 = 0$, which implies $\lim_{k\to+\infty} \|\Gamma_k\|_1 = 0$.

The equilibrium condition (39) can thus be rewritten as

$$\sum_{k=0}^{\infty} \hat{a}_k \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}'_k x_{t-k} = \delta \left(\sum_{k=0}^{+\infty} \hat{a}_{k+1} \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}'_{k+1} x_{t-k} + \hat{\gamma}'_0 R x_t \right).$$

For this to be true for all t and all states of nature, the following restrictions on coefficients are necessary and sufficient:

$$\hat{a}_k = \delta \hat{a}_{k+1} \ \forall k \geq 0, \qquad \hat{\gamma}'_0 = \delta \hat{\gamma}'_1 + \delta \hat{\gamma}'_0 R \qquad \text{and} \qquad \hat{\gamma}'_k = \delta \hat{\gamma}'_{k+1} \ \forall k \geq 1.$$

When the Taylor principle is satisfied ($|\delta| < 1$), $\hat{a}_k$ and $\hat{\gamma}_k$ explodes unless $\hat{a}_0 = 0$ and $\hat{\gamma}'_1 = 0$. Since $I - \delta R$ is invertible from Assumption 3, $\hat{\gamma}'_0 = 0$ too. We know that the only bounded solution of (39) is $\hat{c}_t = 0$. As a result, c_t^F is the unique equilibrium.

Finally, consider part (iii). $c_t^B \equiv -\sum_{k=1}^{\infty} \delta^{-k} \theta_{t-k}$ and $c_t^\eta \equiv \sum_{k=0}^{\infty} \delta^{-k} \eta_{t-k}$ are bounded (the infinite sums converge) when the Taylor principle is violated ($|\delta| > 1$). c_t^B satisfies (9). So does $c_t = (1-b) c_t^F + b c_t^B + a c_t^\eta$ for arbitrary $b, a \in \mathbb{R}$.

Proof of Proposition 2

Since the sunspots $\{\eta_{t-k}\}_{k=0}^{\infty}$ are orthogonal to the fundamental states $\{x_{t-k}\}_{k=0}^{\infty}$, the argument in the main text proves that $a_k = 0$ for all k. We can thus focus on solutions of the following form:

$$c_t = \sum_{k=0}^{\infty} \gamma'_k x_{t-k}. \tag{40}$$

And the remaining task is to show that $\gamma'_0 = q'(I - \delta R)^{-1}$ and $\gamma'_k = 0$ for all $k \geq 1$, which is to say that only the MSV solution survives.

To start with, note that, since x_t is a stationary Gaussian vector given by (7), the following projections apply for all $k \geq s \geq 0$:

$$\mathbb{E}\left[x_{t-k} | I_t^s\right] = W_{k,s} x_{t-s},$$

where $I_t^s \equiv \{x_t, ..., x_{t-s}\}$ is the period-t information set of an agent born s periods before and

$$W_{k,s} = \mathbb{E}\left[x_{t-k} x'_{t-s}\right] \mathbb{E}\left[x_t x'_t\right]^{-1} = \mathbb{E}\left[x_t x'_t\right] \left(R'\right)^{k-s} \mathbb{E}\left[x_t x'_t\right]^{-1}$$

is an $n \times n$ matrix capturing the relevant projection coefficients.

Next, note that

$$\| W_{k,s} \|_1 \leq \| \mathbb{E}\left[x_t x'_t\right] \|_1 \| \left(R'\right)^{k-s} \|_1 \| \mathbb{E}\left[x_t x'_t\right]^{-1} \|_1, \tag{41}$$

where $\| \cdot \|_1$ is the 1-norm. Since all the eigenvalues of R are within the unit circle, we know the spectral radius $\rho(R) = \rho\left(R'\right) < 1$. From Gelfand's formula, we know that there exists $\bar{\Lambda} \in (0,1)$ and $M_1 > 0$ such that

$$\| \left(R'\right)^{k-s} \|_1 \leq M_1 \bar{\Lambda}^{k-s},$$

for all $k \geq s \geq 0$. Together with the fact that $E\left[x_t x_t'\right]$ is invertible (because Σ_ε is positive definite and $\rho(R) < 1$), we know that there exists $M_2 > 0$ such that

$$\|W_{k,s}\|_1 \leq M_2 \bar{\Lambda}^{k-s}. \tag{42}$$

Now, from Assumption 4, we know

$$\bar{E}_t\left[x_{t-k}\right] = (1-\lambda)^k x_{t-k} + \sum_{s=0}^{k-1} \lambda (1-\lambda)^s \mathbb{E}\left[x_{t-k} | I_t^s\right] \equiv \sum_{s=0}^{k} V_{k,s} x_{t-s}, \tag{43}$$

where, for all $k \geq s \geq 0$,

$$V_{k,k} = (1-\lambda)^k I_{n\times n} \quad \text{and} \quad V_{k,s} = \lambda (1-\lambda)^s W_{k,s}.$$

Together with (42), we know that there exits $M_3 > 0$ and $\Lambda = \max\left\{1-\lambda, \bar{\Lambda}\right\} \in (0,1)$ such that for all $k \geq s \geq 0$,

$$\|V_{k,s}\|_1 \leq M_3 \Lambda^k. \tag{44}$$

Now consider an equilibrium in the form of (40). From equilibrium condition (5), we know

$$\begin{aligned}
\sum_{k=0}^{+\infty} \gamma_k' x_{t-k} &= (1-\delta_0)\theta_t + \delta_0 \bar{E}_t\left[\sum_{k=0}^{+\infty} \gamma_k' x_{t-k}\right] + \delta_1 \bar{E}_t\left[\sum_{k=0}^{+\infty} \gamma_k' x_{t+1-k}\right] \\
&= \left((1-\delta_0) q' + \delta_0 + \delta_1 \gamma_0' R + \delta_1 \gamma_1'\right) x_t + \bar{E}_t\left[\sum_{k=1}^{+\infty} \left(\delta_0 \gamma_k' + \delta_1 \gamma_{k+1}'\right) x_{t-k}\right] \\
&= \left((1-\delta_0) q' + \delta_0 + \delta_1 \gamma_0' R + \delta_1 \gamma_1'\right) x_t + \sum_{k=1}^{+\infty} \left(\delta_0 \gamma_k' + \delta_1 \gamma_{k+1}'\right)\left(\sum_{s=0}^{k} V_{k,s} x_{t-s}\right),
\end{aligned}$$

where we use the fact that all agents at t know the values of the fundamental state x_t.

For this to be true for all states of nature, we can compare coefficients on each x_{t-k}, we have

$$\begin{aligned}
\gamma_0' &= (1-\delta_0) q' + \delta_0 \gamma_0' + \delta_1 \gamma_0' R + \delta_1 \gamma_1' \\
\gamma_k' &= \sum_{l=k}^{+\infty} \left(\delta_0 \gamma_l' + \delta_1 \gamma_{l+1}'\right) V_{l,k} \quad \forall k \geq 1.
\end{aligned} \tag{45}$$

From Definition 1, we know that there is a scalar $M > 0$ such that $\|\gamma_k'\|_1 \leq M$ for all $k \geq 0$, where $\|\cdot\|_1$ is the 1-norm. From (44) and (45), we know that, for all $k \geq 1$,

$$\|\gamma_k'\|_1 \leq (|\delta_0| + |\delta_1|) \sum_{l=k}^{+\infty} \|V_{l,k}\|_1 M \leq (|\delta_0| + |\delta_1|) M_3 \frac{\Lambda^k}{1-\Lambda} M. \tag{46}$$

Because $\lim_{k\to\infty} \Lambda^k = 0$, there necessarily exists an $\hat{k}$ finite but large enough $(|\delta_0| + |\delta_1|) M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} < 1$. So we know that, for all $k \geq \hat{k}$,

$$\|\gamma_k'\|_1 \leq (|\delta_0| + |\delta_1|) M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} M.$$

Now, we can use the above formula and (45) to provide a tighter bound of $\|\gamma'_k\|_1$: for all $k \geq \hat{k}$,

$$\|\gamma'_k\|_1 \leq \left((|\delta_0| + |\delta_1|)\, M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} \right)^2 M.$$

We can keep iterating. For for all $k \geq \hat{k}$ and $l \geq 0$,

$$\|\gamma'_k\|_1 \leq \left((|\delta_0| + |\delta_1|)\, M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} \right)^l M.$$

Since $(|\delta_0| + |\delta_1|)\, M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} < 1$, we then have $\gamma'_k = 0$ for all $k \geq \hat{k}$. Using (45) and doing backward induction, we then know $\gamma'_k = 0$ for all $k \geq 1$ and

$$\gamma'_0 = (1-\delta_0)\, q' + \delta_0 \gamma'_0 + \delta_1 \gamma'_0 R,$$

which means $\gamma'_0 = q' \left(I - \frac{\delta_1}{1-\delta_0} R \right)^{-1} = q' (I - \delta R)^{-1}$, where I use $\delta_0 < 1$. Together, this means that the equilibrium is unique and is given by $c_t = c^F_t$, where c^F_t is defined in (10).

Proof of Proposition 3

Consider a candidate equilibrium c_t in Definition 1. We first notice that, for the period-t agent born s periods ago, her information set I^s_t in Assumption 4 can be written equivalently as

$$I^s_t = \{\eta_{t-s}, ..., \eta_t, x_{t-s}, \varepsilon^x_{t-s+1}, \cdots, \varepsilon^x_t\},$$

where ε^x_t is the innovation in x_t in Assumption 1. From (37) in the proof of Lemma 1, we know that c_t can be written as

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \Gamma'_k \varepsilon^x_{t-k},$$

where Γ'_k is given by (38). From the law of total variances, we have

$$Var\left(E_t\left[c_t | I^s_t\right] - c_t\right) \leq Var\left(\sum_{k=s+1}^{\infty} a_k \eta_{t-k} + \sum_{k=s}^{\infty} \Gamma'_k \varepsilon^x_{t-k} \right).$$

Since η_t and ε^x_t are independent of each other as well as independent over time, the finiteness of $Var(c_t)$ implies that

$$\lim_{s\to+\infty} Var\left(\sum_{k=s+1}^{\infty} a_k \eta_{t-k} + \sum_{k=s}^{\infty} \Gamma'_k \varepsilon^x_{t-k} \right) = 0.$$

As a result, for any $\epsilon > 0$ arbitrarily small but positive, there exists $\hat{s}_0$, such that

$$Var\left(E_t\left[c_t | I^s_t\right] - c_t\right) \leq \epsilon$$

for all $s \geq \hat{s}_0$ and every t. Similarly, for each $k \leq K$, there exists $\hat{s}_k$, such that

$$Var\left(E_t\left[c_{t-k} | I^s_t\right] - c_{t-k}\right) \leq \epsilon$$

for all $s \geq \hat{s}_k$ and every t. Now, for any $\epsilon' > 0$ arbitrarily small but positive, we can find $\hat{\lambda} > 0$ such that $\left(1-\hat{\lambda}\right)^{\hat{s}_k} \geq 1-\epsilon'$ for all $k \in \{0, \cdots, K\}$. Together, this means that whenever $\lambda \in (0, \hat{\lambda})$, $Var\left(E_t^i[c_{t-k}] - c_{t-k}\right) \leq \epsilon$ for all $k \leq K$, for at least a fraction $1-\epsilon'$ of agents, and for every period t.

Proof of Proposition 4

For part (i), consider the following sunspot equilibrium c_t^η similar to (12)

$$c_t^\eta \equiv \sum_{k=0}^{\infty} x^k \eta_{t-k}.$$

For any equilibrium that satisfies the full-information rational expectations version of (20),

$$c_t = \mathbb{E}_t\left[\theta_t + \sum_{k=0}^{+\infty} \delta_k c_{t+k}\right], \tag{47}$$

we know that $c_t + c_t^\eta$ also satisfies (47) when

$$1 = \sum_{k=1}^{+\infty} \frac{\delta_k x^k}{1-\delta_0}. \tag{48}$$

When $\Delta > 1$. We know that (48) has a root $x \in (0,1)$

For part (ii), we first find the MSV solution $c_t^F = \gamma' x_t$ for some $\gamma \in \mathbb{R}^n$. From (5), we have

$$\gamma' = q' + \gamma'\left(\sum_{k=0}^{+\infty} \delta_k R^k\right).$$

Since $I - \sum_{k=0}^{+\infty} \delta_k R^k$ is invertible, the unique solution is $\gamma' = q'(I - \sum_{k=0}^{+\infty} \delta_k R^k)^{-1}$. We henceforth denote this solution as

$$c_t^F \equiv q'(I - \delta \sum_{k=0}^{+\infty} R^k)^{-1} x_t. \tag{49}$$

Consider an equilibrium taking the form of (8). We use (5):

$$\sum_{l=0}^{+\infty} a_l \eta_{t-l} + \sum_{l=0}^{\infty} \gamma_l' x_{t-l} = q' x_t + \bar{E}_t\left[\sum_{k=0}^{+\infty} \delta_k \left(\sum_{l=0}^{+\infty} a_l \eta_{t+k-l} + \sum_{l=0}^{\infty} \gamma_l' x_{t+k-l}\right)\right]. \tag{50}$$

We know

$$\bar{E}_t[\eta_{t-l}] = \begin{cases} \mu_l \eta_{t-l} & \text{if } l \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\mu_l = (1-\lambda)^l$ is the measure of agents who remember a sunspot realized l periods earlier as in the proof of Proposition 2. Comparing coefficient in front of η_{t-l} and using the facts that each sunspot is orthogonal to all fundamentals:

$$a_l = \mu_l \sum_{k=0}^{+\infty} \delta_k a_{k+l} \quad \forall l \geq 0. \tag{51}$$

Because $\lim_{l\to\infty}\mu_l = 0$, there necessarily exists an $\hat{l}$ finite but large enough $\mu_{\hat{l}}\sum_{k=0}^{\infty}|\delta_k| < 1$.

Since we are focusing bounded equilibria as in Definition 1, there exists a scalar $M > 0$, arbitrarily large but finite, such that $|a_l| \le M$ for all l. From (51), we then know that, for all $l \ge \hat{l}$,

$$|a_l| \le \mu_{\hat{l}} M \sum_{k=0}^{+\infty} |\delta_k|, \tag{52}$$

where we also use the fact that the sequence $\{\mu_l\}_{l=0}^{\infty}$ is decreasing. Now, we can use (51) and (52) to provide a tigehter bound of $|a_l|$. That is, for all $l \ge \hat{l}$,

$$|a_l| \le \left(\mu_{\hat{l}} \sum_{k=0}^{\infty} |\delta_k|\right)^2 M.$$

We can keep iterating. Since $\mu_{\hat{l}}\sum_{k=0}^{\infty}|\delta_k| < 1$, we then have $a_l = 0$ for all $l \ge \hat{l}$. Using (51) and doing backward induction, we then know $a_l = 0$ for all l.

Now, (50) can be simplified as

$$\begin{aligned}\sum_{l=0}^{\infty} \gamma_l' x_{t-l} &= q' x_t + \bar{E}_t \left[\sum_{k=0}^{+\infty} \delta_k \sum_{l=0}^{\infty} \gamma_l' x_{t+k-l}\right]. \\ &= q' x_t + \sum_{k=0}^{+\infty} \delta_k \sum_{l=0}^{k} \gamma_l' R^{k-l} x_t + \bar{E}_t \left[\sum_{l=1}^{+\infty} \left(\sum_{k=0}^{+\infty} \delta_k \gamma_{k+l}'\right) x_{t-l}\right].\end{aligned} \tag{53}$$

For this to be true for all states of nature, we can compare coefficients on each x_{t-l}:

$$\gamma_0' = q' + \sum_{k=0}^{+\infty} \delta_k \sum_{l=0}^{k} \gamma_l' R^{k-l} \tag{54}$$

$$\gamma_l' = \sum_{s=l}^{+\infty} \left(\sum_{k=0}^{+\infty} \delta_k \gamma_{k+s}'\right) V_{s,l} \quad \forall l \ge 1, \tag{55}$$

where $V_{s,l}$ is defined in (43). The above two equations can be re-written as:

$$\gamma_0' = \left(q' + \sum_{k=1}^{+\infty} \delta_k \sum_{l=1}^{k} \gamma_l' R^{k-l}\right) \left(I - \sum_{k=0}^{+\infty} \delta_k R^k\right)^{-1} \tag{56}$$

$$\gamma_l' = \left(\sum_{k=l+1}^{+\infty} \sum_{s=l}^{k} \delta_{k-s} \gamma_k' V_{s,l}\right) \left(I - \delta_0 V_{l,l}\right)^{-1} \quad \forall l \ge 1, \tag{57}$$

where, from (43), we know that $I - \delta_0 V_{l,l} = \left[1 - \delta_0 (1-\lambda)^l\right] I$ is invertible.

From Definition 1, we know that there is a scalar $M > 0$ such that $\|\gamma_l'\|_1 \le M$ for all $l \ge 0$, where $\|\cdot\|_1$ is the 1-norm. From (55), we know, for all $l \ge 1$

$$\|\gamma_l'\|_1 \le \left(\sum_{k=0}^{+\infty} |\delta_k|\right) \left(\sum_{s=l}^{+\infty} \|V_{s,l}\|_1\right) M \le \left(\sum_{k=0}^{+\infty} |\delta_k|\right) M_3 \frac{\Lambda^l}{1-\Lambda} M, \tag{58}$$

where M_3 and Λ are defined in (43) Because $\lim_{l\to\infty}\Lambda^l = 0$, there necessarily exists an $\hat{l}$ finite but large enough such that $\left(\sum_{k=0}^{+\infty}|\delta_k|\right) M_3 \frac{\Lambda^{\hat{l}}}{1-\Lambda} < 1$. So we know that, for all $l \geq \hat{l}$,

$$\|\gamma_l'\|_1 \leq \left(\sum_{k=0}^{+\infty}|\delta_k|\right) M_3 \frac{\Lambda^{\hat{l}}}{1-\Lambda} M.$$

Now, we can use the above formula and (55) to provide a tighter bound of $\|\gamma_l'\|_1$: for all $l \geq \hat{l}$,

$$\|\gamma_l'\|_1 \leq \left(\left(\sum_{k=0}^{+\infty}|\delta_k|\right) M_3 \frac{\Lambda^{\hat{l}}}{1-\Lambda}\right)^2 M.$$

We can keep iterating. Since $\left(\sum_{k=0}^{+\infty}|\delta_k|\right) M_3 \frac{\Lambda^{\hat{l}}}{1-\Lambda} < 1$, we then have $\gamma_l' = 0$ for all $l \geq \hat{l}$. Using (55) and doing backward induction, we then know $\gamma_l' = 0$ for all $l \geq 1$ and, from (54),

$$\gamma_0' = q' + \gamma_0' \left(\sum_{k=0}^{+\infty} \delta_k R^k\right),$$

which means $\gamma_0' = q'\left(I - \delta \sum_{k=0}^{+\infty} \delta_k R^k\right)^{-1}$. Together, this means that the equilibrium is unique and is given by $c_t = c_t^F$, where c_t^F is defined in (49). This proves the Proposition.

Proof of Proposition 6

Given Assumption 6, an possible equilibrium takes the form of

$$c_t = \sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k' x_{t-k} + \chi \zeta_t.$$

From (35), we have that

$$\begin{aligned}
\sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k' x_{t-k} + \chi \zeta_t &= \theta_t + \zeta_t + \delta \mathbb{E}[\sum_{k=0}^{K_\eta - 1} a_{k+1} \eta_{t-k} + \sum_{k=0}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_\theta - 1} \gamma_{k+1}' x_{t-k} | I_t] \\
&= q' x_t + \zeta_t + \delta \left[\sum_{k=0}^{K_\eta - 1} a_{k+1} \eta_{t-k} + \sum_{k=1}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_\theta - 1} \gamma_{k+1}' x_{t-k} + \gamma_0' R x_t\right] \\
&\quad + \delta \beta_1 \left[\sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k' x_{t-k} + \chi \zeta_t\right]
\end{aligned}$$

where we use Assumptions 1–2 and the fact that ζ_t is drawn independently over time. For this to be true for all states of nature, we can compare coefficients:

$$a_k = \delta a_{k+1} + \delta \beta_1 a_k \quad \forall k \in \{0, \cdots, K_\eta - 1\} \quad \text{and} \quad a_{K_\eta} = \delta \beta_1 a_{K_\eta} \tag{59}$$

$$\beta_k = \delta \beta_{k+1} + \delta \beta_1 \beta_k \quad \forall k \in \{1, \cdots, K_\beta - 1\} \quad \text{and} \quad \beta_{K_\beta} = \delta \beta_1 \beta_{K_\beta} \tag{60}$$

$$\gamma_k' = \delta \gamma_{k+1}' + \delta \beta_1 \gamma_k' \quad \forall k \in \{1, \cdots, K_\theta - 1\} \quad \text{and} \quad \gamma_{K_\theta}' = \delta \beta_1 \gamma_{K_\theta}' \tag{61}$$

$$\gamma_0' = q' + \delta \gamma_1' + \delta \beta_1 \gamma_0' + \gamma_0' R \quad \text{and} \quad \chi = 1 + \delta \beta_1 \chi. \tag{62}$$

First, from the second equation in (62), we know $\delta\beta_1 \neq 1$. Then, from the second parts of (59)–(61), we know $a_{K_\eta} = 0$, $\beta_{K_\beta} = 0$, and $\gamma'_{K_\theta} = 0$. From backward induction on (59)–(62), we know that all a, b, γ are zero except for the following:

$$\gamma'_0 = q' + \gamma'_0 R,$$

which means $\gamma'_0 = q' (I - R)^{-1}$. We also know that $\chi = 1$. We conclude that the unique solution is

$$c_t = c_t^F + \zeta_t,$$

where c_t^F is given by (10).

References

ABREU, D. AND M. K. BRUNNERMEIER (2003): "Bubbles and crashes," *Econometrica*, 71, 173–204.

ANGELETOS, G.-M. AND Z. HUO (2021): "Myopia and Anchoring," *American Economic Review*, 111, 1166–1200.

ANGELETOS, G.-M. AND J. LA'O (2013): "Sentiments," *Econometrica*, 81, 739–779.

ANGELETOS, G.-M. AND C. LIAN (2018): "Forward guidance without common knowledge," *American Economic Review*, 108, 2477–2512.

ANGELETOS, G.-M. AND K. A. SASTRY (2021): "Managing Expectations: Instruments vs. Targets," *Quarterly Journal of Economics*.

ATKESON, A., V. V. CHARI, AND P. J. KEHOE (2010): "Sophisticated monetary policies," *Quarterly Journal of Economics*, 125, 47–89.

AUCLERT, A., M. ROGNLIE, AND L. STRAUB (2018): "The intertemporal keynesian cross," .

BASSETTO, M. (2002): "A game–theoretic view of the fiscal theory of the price level," *Econometrica*, 70, 2167–2195.

BENHABIB, J., S. SCHMITT-GROHÉ, AND M. URIBE (2002): "Avoiding liquidity traps," *Journal of Political Economy*, 110, 535–563.

BENIGNO, P. AND M. WOODFORD (2003): "Optimal monetary and fiscal policy: A linear-quadratic approach," *NBER macroeconomics annual*, 18, 271–333.

BHASKAR, V. (1998): "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems," *The Review of Economic Studies*, 65, 135–149.

Bhaskar, V., G. J. Mailath, and S. Morris (2012): "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory," *The Review of Economic Studies*, 80, 925–948.

Bilbiie, F. O. (2020): "Monetary policy and heterogeneity: An analytical framework," .

Blanchard, O. J. (1979): "Backward and forward solutions for economies with rational expectations," *American Economic Review*, 69, 114–118.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018): "Diagnostic Expectations and Credit Cycles," *Journal of Finance*.

Buiter, W. H. (2002): "The Fiscal Theory of the Price Level: A Critique," *The Economic Journal*, 112, 459–480.

Canzoneri, M. B., R. E. Cumby, and B. T. Diba (2010): "The interaction between monetary and fiscal policy," *Handbook of monetary economics*, 3, 935–999.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994): "Optimal fiscal policy in a business cycle model," *Journal of política Economy*, 102, 617–652.

Christiano, L., M. Eichenbaum, and B. Johannsen (2018): "Does the New Keynesian Model Have a Uniqueness Problem?" .

Christiano, L. J. and M. Rostagno (2001): "Money growth monitoring and the Taylor rule," *National Bureau of Economic Research Working Paper*.

Cochrane, J. H. (2005): "Money as stock," *Journal of Monetary Economics*, 52, 501–528.

——— (2007): "Determinacy and Identification with Taylor Rules," .

——— (2011): "Determinacy and identification with Taylor rules," *Journal of Political economy*, 119, 565–615.

——— (2017): "The new-Keynesian liquidity trap," *Journal of Monetary Economics*, 92, 47–63.

——— (2018): "Michelson-Morley, Fisher, and Occam: The radical implications of stable quiet inflation at the zero bound," *NBER Macroeconomics Annual*, 32, 113–226.

Correia, I., J. P. Nicolini, and P. Teles (2008): "Optimal Fiscal and Monetary Policy: Equivalence Results," *Journal of Political Economy*, 116, 141–170.

Evans, G. W. and B. McGough (2018): "Equilibrium selection, observability and backward-stable solutions," *Journal of Monetary Economics*, 98, 1–10.

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

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

Galí, J. (2008): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press.

García-Schmidt, M. and M. Woodford (2019): "Are low interest rates deflationary? A paradox of perfect-foresight analysis," *American Economic Review*, 109, 86–120.

Honkapohja, S. and K. Mitra (2004): "Are non-fundamental equilibria learnable in models of monetary policy?" *Journal of Monetary Economics*, 51, 1743–1770.

Huo, Z. and N. Takayama (2021): "Rational Expectations Models with Higher Order Beliefs," *Yale University mimeo.*

King, R. G. (2000): "The new IS-LM model: language, logic, and limits," *FRB Richmond Economic Quarterly*, 86, 45–104.

Kocherlakota, N. and C. Phelan (1999): "Explaining the fiscal theory of the price level," *Federal Reserve Bank of Minneapolis Quarterly Review*, 23, 14–23.

Leeper, E. M. (1991): "Equilibria under "active" and "passive" monetary and fiscal policies," *Journal of monetary Economics*, 27, 129–147.

Loisel, O. (2021): "The implementation of stabilization policy," *Theoretical Economics*, 677–716.

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

Maćkowiak, B. and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769–803.

Mailath, G. J. and S. Morris (2002): "Repeated games with almost-public monitoring," *Journal of Economic theory*, 102, 189–228.

Mankiw, G. and R. Reis (2002): "Sticky Information Versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117, 1295–1328.

McCallum, B. T. (1983): "On non-uniqueness in rational expectations models: An attempt at perspective," *Journal of monetary Economics*, 11, 139–168.

——— (2001): "Indeterminacy, bubbles, and the fiscal theory of price level determination," *Journal of Monetary Economics*, 47, 19–30.

——— (2007): "E-stability vis-a-vis determinacy results for a broad class of linear rational expectations models," *Journal of Economic Dynamics and Control*, 31, 1376–1391.

——— (2009): "Inflation determination with Taylor rules: Is new-Keynesian analysis critically flawed?" *Journal of Monetary Economics*, 56, 1101–1108.

McKay, A., E. Nakamura, and J. Steinsson (2017): "The discounted euler equation: A note," *Economica*, 84, 820–831.

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

——— (2002): "Social Value of Public Information," *American Economic Review*, 92, 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.

NIEPELT, D. (2004): "The Fiscal Myth of the Price Level," *Quarterly Journal of Economics*, 119, 277–300.

OBSTFELD, M. AND K. ROGOFF (1983): "Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out?" *Journal of Political Economy*, 91, 675–87.

——— (2021): "Revisiting Speculative Hyperinflations in Monetary Models," *Review of Economic Dynamics*, 40, 1–11.

PĘSKI, M. (2012): "An anti-folk theorem for finite past equilibria in repeated games with private monitoring," *Theoretical Economics*, 7, 25–55.

SARGENT, T. (1982): "The Ends of Four Big Inflations," in *Inflation: Causes and Effects*, National Bureau of Economic Research, 41–98.

SARGENT, T. AND N. WALLACE (1981): "Some Unpleasant Monetarist Arithmetic," *Federal Reserve Bank of Minneapolis Quarterly Review*.

SARGENT, T. J. AND N. WALLACE (1975): ""Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy*, 83, 241–254.

SIMS, C. A. (1994): "A simple model for study of the determination of the price level and the interaction of monetary and fiscal policy," *Economic theory*, 4, 381–399.

TAYLOR, J. B. (1993): "Discretion versus policy rules in practice," in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 39, 195–214.

TOWNSEND, R. M. (1983): "Forecasting the forecasts of others," *Journal of Political Economy*, 91, 546–588.

WEINSTEIN, J. AND M. YILDIZ (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400.

WERNING, I. (2015): "Incomplete Markets and Aggregate Demand," *NBER Working Paper No. 21448*.

WOODFORD, M. (1995): "Price-level determinacy without control of a monetary aggregate," in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 43, 1–46.

——— (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*.

——— (2019): "Monetary policy analysis when planning horizons are finite," *NBER Macroeconomics Annual*, 33, 1–50.

XIE, Y. (2019): "Fiscal and Monetary Policy Interaction under Limited Foresight," *mimeo*.