Determinacy without the Taylor Principle*

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

February 24, 2022

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-standing debate about what the right equilibrium selection is. We offer a potential resolution by showing that small frictions in social memory and intertemporal coordination can remove the indeterminacy. Under our perturbations, the unique surviving equilibrium is the same as that selected by the Taylor principle, but it no more relies on it; monetary policy is left to play only a stabilization role; and fiscal policy needs to be Ricardian, 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 Hitotsubashi University, LSE, Oxford University, UC Berkeley, USC, Warwick, Yale, Pontifical Catholic University of Chile, SED, the Minnesota Conference in Macroeconomic Theory, the NBER Summer Institute (EFCE), and VEAMS. Chen Lian thanks the Alfred P. Sloan Foundation and the Clausen Center for International Business and Policy for financial support.

†MIT and NBER; angelet@mit.edu.

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

1 Introduction

Can monetary policy regulate inflation and aggregate demand? Does the ZLB trigger a deflationary spiral? Does Ricardian equivalence hold when taxation is non-distortionary, markets are complete, and consumers have rational expectations and long horizons? One may be inclined to answer "yes" to all these questions. But the right answer, at least within the dominant policy paradigm (the New Keynesian model), is that it depends on how equilibrium is selected.

The basic problem goes back to Sargent and Wallace (1975): the same path for the nominal interest rate is consistent with multiple equilibrium paths for inflation and, in the presence of nominal rigidity, for output too. The standard approach selects a specific equilibrium by assuming that monetary policy satisfies the Taylor principle (Taylor, 1993), or equivalently that it is "active" (Leeper, 1991). The model's three "famous" equations then admit a unique bounded solution, which is the one customarily used to interpret the data and guide policy.[1] An alternative approach, the Fiscal Theory of the Price Level (FTPL),[2] selects a different equilibrium by assuming that fiscal policy is "non-Ricardian." As emphasized by Cochrane (2017, 2018), this invalidates the answers provided above and elevates government debt and deficits to key determinants of inflation and output, even when these variables do not enter the model's three famous equations.

Because both approaches boil down to untestable assumptions about off-equilibrium strategies of the monetary and fiscal authorities, the debate about which approach is "right" has never been settled.[3] We offer a way out of this dead-end by demonstrating how the core issue hinges on delicate assumptions about social memory and intertemporal coordination. Once we perturb these assumptions, tinily but appropriately, the conventional solution of the New Keynesian model emerges as the unique rational expectations equilibrium regardless of whether monetary policy is active or passive. This reinforces the logical foundations upon which one can answer "yes" to the questions raised in the beginning. And it allows one to think about *both* the Taylor principle and the FTPL in new ways, liberated from the equilibrium-selection conundrum.

Preview of main results. A crucial stepping stone of our analysis is the translation of a New Keynesian economy to a dynamic coordination game among the consumers. The basic idea is

[1] Like the textbook treatment, we work with the linearized New Keynesian model and restrict attention to bounded equilibria, which amounts to studying the local determinacy question. As usual, this leaves outside the analysis self-fulfilling hyper-inflations (Obstfeld and Rogoff, 1983, 2021; Cochrane, 2011) and self-fulfilling liquidity traps (Benhabib et al., 2002), which can themselves be ruled out by appropriate escape clauses.

[2] See Leeper (1991), Sims (1994) and Woodford (1995) for the genesis of the FTPL, Cochrane (2005, 2017, 2018) for extensions and reinterpretations, and Bassetto (2002) for a careful game-theoretic perspective.

[3] Bassetto (2008) offers a concise and balanced review of the debate, and Canzoneri, Cumby, and Diba (2010) discuss how it fits in the broader context of the study of the fiscal-monetary interaction. For more thorough discussions of the role played by off-equilibrium policy threats, see Kocherlakota and Phelan (1999), King (2000), Bassetto (2002), Cochrane (2007), and Atkeson et al. (2010).

that one's optimal spending depends on others' spending via three GE channels: the feedback for aggregate spending to income (the Keynesian cross); the feedback from aggregate spending to inflation (the Phillips curve); and the response of monetary policy (the Taylor rule). The first two channels contribute to strategic complementarity, and in particular to a dynamic feedback strong enough to support multiple equilibria; the third pulls in the opposite direction.[4]

Under this prism, which we develop in Sections 2 and 3, the Taylor principle translates to the following requirement: let the third channel above be strong enough to guarantee a unique equilibrium when different generations of consumers can perfectly coordinate their behavior over time. In the rest of the paper, we instead accommodate a small but appropriate friction in such coordination and proceed to show how this can eliminate all equilibria except one, that known as the model's fundamental or minimum state variable (MSV) solution.

Our main result, developed in Section 4, models the friction as follows. In each period, a consumer learns the current shocks (payoff-relevant or not); with probability $\lambda \in [0,1)$, she knows nothing else; and with the remaining probability, she inherits the information of another, randomly selected, consumer from the previous period. This lets λ parameterize the speed at which social memory "fades" over time: for any t, the fraction of the population who "remembers" (i.e., can directly or indirectly condition their actions on) the shocks realized at any $\tau \leq t$ is $(1-\lambda)^{t-\tau}$.

The frictionless, 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. Our main result is that all these equilibria unravel as soon as $\lambda > 0$. Only the fundamental/MSV solution survives, regardless of whether the monetary 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 $\lambda \to 0$ consumers face vanishing uncertainty about the history of *both* shocks and outcomes, suggesting that our conclusions do not necessarily rest on assuming away either endogenous aggregation of information or endogenous coordination devices. We corroborate this point in Section 5 with two additional results, both of which allow for direct observation of endogenous outcomes. This requires an adjustment in the perturbation notion—in particular, the perturbation considered in Proposition 4 amounts to *immediate* forgetting of a small component of the fundamentals, whereas that considered in our main result amounts to *asymptotic* forgetting of the distant past—but the end result is the same.

[4]The second channel is shut off with rigid prices but the first channel is always there—whether hidden behind the representative consumer's Euler condition or made salient in the "intertemporal Keynesian cross" (Auclert et al., 2018). This, along with the central role of aggregate demand in Keynesian thinking, explains why we opt to represent the economy as a game among the consumers instead of one among the firms.

Interpreting our contribution. The logic behind our results echoes the literature on global games (Morris and Shin, 2002, 2003) and is subject to a similar qualification: indeterminacy may strike back if markets or other mechanisms facilitate enough coordination (Angeletos and Werning, 2006; Atkeson, 2000). That said, it is useful to recognize the following point. In our context, any equilibrium other than the MSV solution is sustained by a self-fulfilling infinite chain over different generations of players: today's consumers are responding to a payoff-irrelevant variable (e.g., the current sunspot or the past rate of inflation) because and only because they expect tomorrow's consumers to do the same on the basis of a similar expectation about behavior further into the future, and so on. This explains why the relevant perturbations relate, one way or another, to social memory. And it suggests that the requisite type of common knowledge might be harder to attain in our context than in the case of, say, a self-fulfilling bank run.

All in all, we therefore view our contribution not as a definite resolution of the New Keynesian model's indeterminacy problem but rather in the following terms: (i) as a new lens for understanding this problem; (ii) as a formal justification for treating it like a bug instead of a feature; and (iii) as an invitation to reconsider the meaning of *both* the Taylor principle and the FTPL. The first two points should be self-evident by now, so let us expand on the last.

Consider first the Taylor principle. Our uniqueness result removes the need for equilibrium selection but leaves room for sunspot-like fluctuations in the form of overreaction to noisy public news (Morris and Shin, 2002), shocks to higher-order beliefs (Angeletos and La'O, 2013; Benhabib et al., 2015), or related forms of bounded rationality (Angeletos and Sastry, 2021). This in turn lets the slope of the Taylor rule to play a new function: to regulate the macroeconomic complementarity and, thereby, the magnitude of such sunspot-like fluctuations along the unique equilibrium. Our contribution is therefore not to rule out "animal spirits" altogether but rather to recast the Taylor principle as a form of on-equilibrium stabilization instead of an off-equilibrium threat.

Consider next the FTPL. In Section 6, we show the following elementary result for a suitable extension of our model: as long as consumers are neither rationally confused nor plainly irrational, the economy reduces to the *same* game as that in the absence of fiscal policy. This formalizes the sense in which government debt and deficits are sunspots and directly implies that the equilibrium selected by the FTPL is not robust to our perturbations: fiscal policy *has* to be Ricardian even when monetary policy is passive.

Like our main result, this lesson is subject to the following qualification: at this point, it is anyone's guess whether the real world is better approximated by the full-information benchmark, our specific perturbations, or other plausible alternatives. Still, our analysis illustrates the logical fragility of the existing formulation of the FTPL and, by extension, the value of fresh takes on it. For instance, we would favor a reformulation in which the equilibrium-selection issue is moot (whether with the help of our perturbations or otherwise) and the fiscal-monetary interaction is modeled as a game of chicken between the two authorities.

Related literature. The title of our paper could have been shared by Atkeson et al. (2010): that paper, too, seeks to achieve determinacy without the Taylor principle. But this means something very different there. That paper replaces Taylor rules with a class of more sophisticated strategies, which, inter alia, avoid Cochrane (2011)'s criticism that active monetary policy amounts to a threat to "blow up" inflation. But Bassetto (2002, 2005) had previously shown that basically the same point holds for the FTPL: that theory can be extended to allow for more sophisticated strategies for the fiscal authority and to escape the corresponding criticism that the non-Ricardian assumption amounts to a threat to "blow up" the government budget (Kocherlakota and Phelan, 1999; Buiter, 2002). In short, these works do not address the equilibrium-selection conundrum; they only deepen it. By contrast, our paper shows how appropriate perturbations of the private sector's information/coordination can dispense with this conundrum altogether.

Our main result (Proposition 2) reminds Rubinstein (1989), Morris and Shin (1998, 2003), and Abreu and Brunnermeier (2003): certain equilibria unravel because of a series of contagion effects related to higher-order beliefs. Our second result (Proposition 3) has the flavor of rational inattention: agents observe an endogenous coordination device with idiosyncratic noise. Our third result (Proposition 4) connects to Bhaskar (1998) and Bhaskar, Mailath, and Morris (2012): it combines a purification in payoffs with finite social memory. The common thread is the relaxation of common knowledge and the resulting coordination friction. But the precise connections between our results and the related literatures 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: it has focused exclusively on how such frictions shape the model's MSV solution, while assuming 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, which seeks to formalize what it takes for a rational expectations equilibrium to be "learnable," and has had mixed success.[5] Still, we view this approach and ours as complements in that they both contribute towards reinforcing the logical foundations of the conventional, or "monetarist," approach.

[5]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).

The determinacy problem we are after extends from Rational Expectations Equilibrium (REE) to a larger class of solution concepts that relax the perfect coincidence between subjective beliefs and objective outcomes but preserve a fixed-point relation between the two, such as cognitive discounting (Gabaix, 2020) and diagnostic expectations (Bordalo et al., 2018). By contrast, Level-K Thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019) produces a unique solution because it shuts down completely the feedback from objective reality to subjective beliefs. This assumption seems too strong, especially in the context of stationary environments.

Finally, it seems a safe guess that our methods extend from our setting to a larger class of linear, forward-looking, rational-expectations models, like those studied in Blanchard (1979) and Blanchard and Kahn (1980). But they have be to be translated with caution in non-linear settings, especially when these feature multiple steady-state equilibria. In short, our paper speaks to the question of local determinacy but not necessarily to that of global determinacy.

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 6, 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

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

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 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

$$C^1_{i,t} + B_{i,t} = Y_t \qquad \text{and} \qquad C^2_{i,t+1} = Y_{t+1} + \frac{I_t}{\Pi_{t+1}}B_{i,t},$$

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. After the usual log-linearization, this translates to the following optimal consumption function:[6]

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

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 as Dynamic IS equation.

Connection to standard New Keynesian model

Although our version of the Dynamic IS equation looks different from its textbook counterpart, it actually nests it when there is full information. In this benchmark, $\bar{E}_t$ can be replaced by $\mathbb{E}_t$, which henceforth denotes the rational expectation conditional on full information about h_t. Along with the fact that c_t and i_t are measurable in h_t, this means that in this case equation (2) reduces to

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

or equivalently

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

which is evidently 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

[6] 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. Also, (1) is a version of the Permanent Income Hypothesis. The only subtlety is that we have allowed consumers to be inattentive to current income and current interest rates (which is why y_t and i_t appear inside the expectation operator). But as we explain below, such inattention can be vanishingly small under our main perturbation, and can be completely dispensed with under the variant perturbations.

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 6, 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 shown in Section 6, our arguments directly extend to the fully micro-founded, forward-looking, New Keynesian Phillips curve. We suspect the same is true for 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.[7] As is well known and will be reviewed shortly, $\phi > 1$ is necessary and sufficient for the 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.

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 single 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

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

[7] Similarly to King (2000) and Atkeson et al. (2010), letting z_t be correlated with ϱ_t and ξ_t helps disentangle the stabilization and equilibrium selection functions of Taylor rules in the standard paradigm: the former can be served by the design of z_t, the latter by the restriction $\phi > 1$.

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.

Parenthesis: Sticky vs flexible prices

The slope of the Phillips curve, κ, can be arbitrarily large but not literally infinite. Some degree of nominal rigidity is necessary in order to conceptualize the economy as a dynamic coordination game and, thereby, to introduce the relevant friction. We clarify this point in Appendix B. We also suspect that a version of our insights may apply to flexible-price models with a well-defined demand for money (e.g., models with money in the utility function). We will not explore this idea, but we will make clear that our formal arguments do not rely on the precise micro-foundations.

Fundamentals, sunspots, and the equilibrium concept

Aggregate uncertainty is of two sources: fundamentals and sunspots. The former are herein conveniently summarized in θ_t. The latter are represented by a random variable η_t that is independent of the current, past, and future values of θ_t. As explained in Section 5, our arguments extend to essentially arbitrary specifications of these variables. To ease the exposition, the main analysis makes the following simplification:

Assumption 1 (**Simplification**)**.** *Both the fundamental θ_t and the sunspot η_t are i.i.d. over time.*

Let h^t capture the history of all fundamentals and sunspots 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 $h^t \equiv \{\theta_{t-k}, \eta_{t-k}\}_{k=0}^{\infty}$ 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 information about h^t; the outcome is a stationary, linear function of the underlying shocks, namely*

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \gamma_k \theta_{t-k} \tag{7}$$

where $\{a_k, \gamma_k\}$ are known coefficients; and the outcome is bounded in the sense that $Var(c_t) < \infty$.[8]

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: stationarity, linearity, and boundedness. The first two restrictions are technical.[9] The third is of substance but has the usual interpretation: the exclusion of unbounded equilibria can be justified by appropriate "exit" strategies 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 money-growth targeting or whatever else it takes for keeping inflation within some bounds.

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, so there is no need to specify off-equilibrium beliefs. The economy's Rational Expectations Equilibria (REE) thus coincide with the corresponding game's Perfect Bayesian Equilibria (PBE).

3 The Standard Paradigm

In this section, we consider the full-information version of our model (which is, in essence, the standard New Keynesian model); we review its determinacy problem; and we finally contextualize our departures from this benchmark.

[8]Note that $Var(c_t)$ can be finite only if there exists a scalar $M > 0$ such that $|a_k| \leq M$ and $|\gamma_k| \leq M$ for all k.

[9]The stationarity restriction comes hand-in-hand with the assumption of infinite history and can readily be relaxed; see Appendix B for an illustration. The linearity restriction, on the other hand, is strictly needed for tractability; but we have no reason to believe that it drives our results, plus it is commonplace in the literature.

Full information, the MSV solution, and the Taylor principle

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:

$$c_t = \theta_t + \delta \mathbb{E}_t[c_{t+1}], \tag{8}$$

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

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

Note that δ is necessarily positive but can be on either side of 1, depending on ϕ.

Because equation (8) is purely forward looking and θ_t is i.i.d., $c_t = c_t^F \equiv \theta_t$ is necessarily an equilibrium. This is known as the model's "fundamental" or "minimum state variable (MSV)" solution (McCallum, 1983), and is the basis of the conventional understanding of how monetary policy works. For instance, if the central bank can adjust z_t in response to the underlying demand and supply shocks, she can guarantee $\theta_t = 0$. This directly translates to $c_t = 0$ ("closing the output gap") under the MSV solution—but not under others solutions.

To rule out other solutions and justify conventional policy predictions, the standard approach imposes the Taylor principle. In our context, just as in the textbook treatment, this principle is defined by the restriction $\phi > 1$. This in turn 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 discussion 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.[10]

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.

(ii) When the Taylor principle is satisfied ($\phi > 1$), the above equilibrium is the unique one.

[10]By restricting $\phi \geq 0$, we have guaranteed $\delta > 0$. If we allow $\delta < 0$, which is possible if ϕ is *sufficiently* negative, Proposition 1 and the discussion after it continue to hold, provided that we recast the Taylor principle as $\delta \in (-1, 1)$. As for our upcoming uniqueness result (Proposition 2), this readily extends to negative ϕ and δ.

(iii) When this principle is violated ($\phi < 1$), there exist a continuum of equilibria, given by

$$c_t = (1-b)c_t^F + bc_t^B + ac_t^\eta, \tag{9}$$

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

$$\underbrace{c_t^B \equiv -\sum_{k=1}^{\infty} \delta^{-k}\theta_{t-k}}_{\textit{backward-looking, pseudo-fundamental component}} \qquad \textit{and} \qquad \underbrace{c_t^\eta \equiv \sum_{k=0}^{\infty} \delta^{-k}\eta_{t-k}}_{\textit{pure sunspot component}}. \tag{10}$$

To understand the type of non-fundamental equilibria documented in part (iii) above, take equation (8), backshift it by one period, and rewrite it as follows:

$$\mathbb{E}_{t-1}[c_t] = \delta^{-1}(c_{t-1} - \theta_{t-1}). \tag{11}$$

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

$$c_t = \delta^{-1}(c_{t-1} - \theta_{t-1}) + a\eta_t, \tag{12}$$

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

$$c_t = -\sum_{k=1}^{\infty} \delta^{-k}\theta_{t-k} + a\sum_{k=0}^{\infty} \delta^{-k}\eta_{t-k} = c_t^B + ac_t^\eta. \tag{13}$$

This is both bounded, thanks to $\delta > 1$, and a rational-expectations solution to (11), by construction, which 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.

To illustrate what all these equilibria are, switch off momentarily the fundamental shocks. Then, $c_t^F = c_t^B = 0$ and (9) reduces to $c_t = ac_t^\eta$, which is a pure sunspot equilibrium of arbitrary aptitude. In this equilibrium, consumers respond to the current sunspot because and only because they expect future agents to keep reacting to it, in perpetuity.

Now let us switch off the sunspots and switch on the fundamentals. Multiplicity then takes the following form: the same path for interest rates or other fundamentals maps to a continuum of different paths for aggregate spending and inflation. Consider, for example, the solution given by $c_t = c_t^B$. Along it, aggregate spending is invariant to the current interest rate and *increases* with past interest rates. This may sound paradoxical but is sustained by basically the same self-fulling infinite chain as that described above: consumers spend more in response to higher interest rates because and only because they expect future consumers to do the same in perpetuity. The same is true for any equilibrium of the form (9) for $b \neq 0$, and explains why all such equilibria can be thought of as both non-fundamental and backward-looking.

All in all, the Taylor principle is therefore used not only to rule out sunspots but also to secure the logical foundations of the modern policy paradigm. The rest of our paper attempts to liberate these foundations from their strict reliance on the Taylor principle, or any substitute thereof.

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

Consider conditions (12) and (13). 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 $h_t = \{\theta_{t-k}, \eta_{t-k}\}_{k=0}^{\infty}$ coincide with those that can be supported by perfect knowledge of $(\theta_t, \eta_t; \theta_{t-1}, c_{t-1})$. 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 know, one can *always* represent any equilibrium in a sequential form, or as in equation (7). 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).

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.

To accomplish this dual goal, in the rest of the paper we follow two strategies. Our main one, in Section 4, takes off from (13), or the sequential representation. An alternative, in Section 5, circles back to (12), the recursive representation. Both strategies illustrate the fragility of non-fundamental equilibria, each one from a different angle.

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 regardless of monetary policy.

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

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

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 idiosyncratically random length of the history of shocks that an agent 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 at which social memory (or common-p belief of past shocks) fades over time.

Two remarks help complete the picture. First, although we have herein specified the available information about fundamentals in terms of the "summary" variable θ_t, this is only for expositional simplicity: we can readily replace θ_t in the statement of Assumption 2 with (ϱ_t, ξ_t, z_t), the vector of all the primitive payoff-relevant shocks. Second, when $\phi = 0$, knowledge of z_t translates to knowledge of i_t. That is, for the special case of interest rate pegs, Assumption 2 and our upcoming uniqueness result is consistent with *perfect* knowledge of the policy instrument. As for the more general case in which $\phi \neq 0$, Assumption 2 requires that consumers be uncertain about, or inattentive, to the current interest rate (and all other endogenous outcomes). But such uncertainty becomes vanishingly small in the limit as $\lambda \to 0^+$: in this limit, almost all consumers become nearly perfectly informed about nearly infinite histories of the exogenous shocks and, therefore, of the endogenous outcomes as well.[11]

Main result

The full-information benchmark is nested with $\lambda = 0$; this indeed translates to $I_{i,t} = h_t$ (perfect knowledge of the infinite history) for all i and t. 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.

[11] See Proposition 7 in Appendix B for a formalization of this point, as well as for a qualification. Also, let us emphasize that we can reconcile our upcoming uniqueness result with *perfect* knowledge of current interest rates and current income if (i) we abstract from the possibility that consumers extract information about the past sunspots from these variables; or (ii) we allow for such signal extraction but invoke a different perturbation argument, that developed in Subsection 5.

Proposition 2 (**Determinacy without the Taylor principle**). *Suppose that social memory is imperfect in the sense of Assumption 2, 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.*[12]

The result is proven in Appendix A for arbitrary δ_0 and δ_1. To illustrate the main argument as transparently as possible, here we set $\delta_0 = 0$ and $\delta_1 = \delta$, for arbitrary $\delta > 0$ (including $\delta > 1$). This zeroes in on the role of coordination across time. We also abstract from fundamentals and focus on ruling out pure sunspot equilibria. That is, we specialize equation (5) to

$$c_t = \delta \bar{E}_t[c_{t+1}]; \tag{14}$$

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 2, 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

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

By the same token, condition (14) rewrites as

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

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

$$a_k = \delta \mu_k a_{k+1}, \tag{15}$$

or equivalently

$$a_{k+1} = \frac{a_k}{\delta \mu_k}. \tag{16}$$

Because $\mu_k \to 0$ as $k \to \infty$, $|a_k|$ explodes to infinity, and hence a bounded solution does not exist, unless $a_0 = 0$. But $a_0 = 0$ implies $a_k = 0\ \forall k$. We conclude that the latter identifies the unique bounded solution. That is, all sunspot equilibria are ruled out and only the MSV solution survives.

[12] 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 2 so as to remove perfect information about x_t and reshape the MSV solution, while also preserving our argument for uniqueness.

Comparison to full information and the importance of $\lim_{k\to\infty}\mu_k = 0$

We will explain the essence of our result momentarily. But first, it is useful to repeat the above argument for the knife-edge case with $\lambda = 0$. In this case, $\mu_k = 1\ \forall k$ and condition (16) becomes

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

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 sunspot equilibria of Proposition 1.

Note next that the result does not depend on the assumption that memory decays at an exponential rate, but it depends on it vanishing asymptotically, i.e., on $\mu_k \to 0$ as $k \to \infty$. If instead $\mu_k \to \mu$ for some $\mu \in (0,1)$, multiplicity would have remained for $\delta > 1/\mu$; that is, the Taylor principle would have been relaxed but would not have been completely dispensed with. Notwithstanding this point, let us emphasize that key is not whether memory *actually* vanishes over time but rather how agents *reason* about the future. We expand on this next.

Intuition and the role of higher-order beliefs

To appreciate the essence of our result and what lies beneath it, focus 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 (15) 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.[13]

This interpretation of our result must be clarified as follows. Here we focused on the response

[13] 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.

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 further explain. Because η_0 is payoff irrelevant in every t, period-0 agents have an incentive to respond to it only if they are confident that period-1 agents will also respond to it, which in turn 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."

Two remarks help 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 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 2 as follows: agents don't forget themselves but worry that others will forget. 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: the fear that agents far in the future may fail to support a sunspot, or backward-looking, equilibrium causes any such equilibrium to unravel.

5 Robustness and Complementary Perturbations

In this section, we explain how our uniqueness result generalizes to more flexible specifications of the fundamentals and the sunspots, provided of course that Assumption 2 is maintained. We next replace this assumption with two variants, which accommodate direct observation of past outcomes and, thereby, endogenous coordination devices. As flagged in the Introduction, these variants may not fully address the question of realism. But they circle back to the discussion of extensive and recursive representations of REE; they allow our insights to translate from the one perspective to the other; and they ultimately reinforce our message that the indeterminacy problem can be treated as a bug instead of a feature. At the end of the section, we also comment on the distinction between local and global determinacy. Readers interested primarily in the applied content of our contribution may skip this section and jump to Section 6.

Persistent fundamentals

In the main analysis we assumed that the fundamental θ_t is uncorrelated over time. Relaxing this assumption changes the MSV solution but does not affect our determinacy result.

To illustrate, suppose that θ_t follows an AR(1) process: $\theta_t = \rho\theta_t + \varepsilon_t$, where $\rho \in (-1,1)$ is a fixed scalar and $\varepsilon_t \sim \mathcal{N}(0,1)$ is a serially uncorrelated innovation. As long as $\rho \neq 0$, an innovation affects payoffs not only today but also in the future. This naturally modifies the MSV solution. Indeed, if we guess that $c_t = \gamma\theta_t$ for some $\gamma \in \mathbb{R}$ and substitute this into (8), we infer that the guess is correct if and only if $\gamma = 1 + \delta\rho\gamma$. For this to admit a solution, it is necessary and sufficient that $\rho \neq \delta^{-1}$. Provided that this is the case, the MSV solution exists and is now given by $c_t^F = \frac{1}{1-\delta\rho}\theta_t$. Modulo this minor adjustment, Proposition 2 directly extends. This claim is verified in Appendix C, indeed for a more general specification of the fundamental uncertainty: such generality naturally modifies the MSV solution but does not interfere with our uniqueness argument.

Let us now zero in on the role of $\rho \neq \delta^{-1}$ in the above example. This restriction is used to guarantee the existence of the MSV solution. But it is not needed in our argument for ruling out any other solution. For the later purpose, it suffices to invoke Assumption 2 alone. Finally, note that the comparative statics of the MSV solution with respect to θ_t switch sign depending on whether ρ is lower or higher than δ^{-1}. In particular, when $\rho > \delta^{-1}$, the MSV solution exhibits the so-called neo-Fisherian property: a sufficiently persistent increase in the nominal interest rate triggers an *increase* in inflation and the output. This raises number of delicate questions, such as whether the neo-Fisherian property is realistic, whether the MSV solution can be obtained by forward induction, or even whether the New Keynesian model is mis-specified. But these questions are beyond the scope of our paper.

Persistent sunspots

Let us now revisit the assumption that the sunspot is serially uncorrelated. As in the case of fundamentals, this is assumption can readily be relaxed, except for one special 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 the concurrent θ_t and η_t alone. Social memory of the distant past is no more needed, because the exogenous sunspot happens to coincide with the *right* sufficient statistic of economy's infinite history.

This situation seems exceedingly unlikely insofar as the sunspot is an exogenous random variable. But what if agents can devise an *endogenous* sunspot? For instance, could it be that agents coordinate on an equilibrium that lets an endogenous outcome, such as perhaps c_t itself, replicate the requisite sunspot? This possibility still presumes significant intertemporal coordination, but the approach taken thus far was not designed to address it. We thus address it in the rest of this section, with the help of two variant perturbations.

Recursive sunspot equilibria: another example of fragility

Recall that, with full information, our model boils down to the following equation:

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

where $\delta \equiv \frac{\delta_1}{1-\delta_0}$ and $\mathbb{E}_t$ is the full-information rational expectation. Let us momentarily shut down the fundamentals, assume that $\delta > 1$, and focus on the set of all pure sunspot equilibria:

$$c_t = a \sum_{k=0}^{\infty} \delta^k \eta_{t-k}, \tag{17}$$

for arbitrary $a \neq 0$. As noted earlier, this can be represented in recursive form as

$$c_t = a\eta_t + \delta^{-1} c_{t-1}. \tag{18}$$

It follows that all sunspot equilibria can be supported with the following "minimal" information set: $I_{i,t} = \{\eta_t, c_{t-1}\}$. Intuitively, c_{t-1} *endogenously* serves the role of the knife-edge persistent sunspot discussed earlier.

Taken at face value, this challenges our message. But as shown next, this logic, too, can be fragile. Suppose that information is given by

$$I_{i,t} = \{\eta_t, s_{i,t}\}, \qquad \text{with} \qquad s_{i,t} = c_{t-1} + \varepsilon_{i,t}.$$

Here, $s_{i,t}$ is a private signal of the past aggregate outcome, $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2)$ is idiosyncratic noise, and $\sigma \geq 0$ is a fixed parameter. When $\sigma = 0$, we are back to the case studied above, and the entire set of sunspot equilibria is supported. When instead $\sigma > 0$ but arbitrarily small, agents' knowledge of the past outcome is only slightly blurred by idiosyncratic noise. As shown next, this causes all sunspot equilibria to unravel.

Proposition 3. *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 actually quite simple. But we prefer to delegate it to the Appendix, because the present example is still special in two regards: it rules out public signals of c_{t-1}; and it rules out information about longer histories.

The first limitation is easy to address: Proposition 3 readily generalizes to $s_{i,t} = c_{t-1} + v_t + \varepsilon_{i,t}$, where v_t is aggregate noise and $\varepsilon_{i,t}$ is idiosyncratic noise. This can be interpreted as a situation where a publicly available statistic is not only contaminated with measurement error but also observed with idiosyncratic noise due to rational inattention (Sims, 2003) or imperfect cognition (Woodford, 2019a). It is only in the knife-edge case in which the statistic is a "public signal" in the precise sense of common knowledge that multiplicity survives.[14]

The second limitation is more challenging, because it opens the pandora box of signal extraction and infinite regress. In the next subsection, we therefore offer a different approach, which manages to keep this box closed while accommodating direct—and indeed perfect—knowledge of long histories of aggregate output and inflation.

Breaking the infinite chain even when past outcomes are perfectly observed

In the above exercise we focused on pure sunspot equilibria. Let us now 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 can be replicated by letting $I_{i,t} \supseteq \{\eta_t, c_{t-1}, \theta_{t-1}\}$ and by having each consumer play the following recursive strategy:

$$c_{i,t} = \delta^{-1}(c_{t-1} - \theta_{t-1}) + a\eta_t. \tag{19}$$

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— breaks 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; we modify equation (5) to

$$c_{i,t} = E_{i,t}[(1-\delta_0)(\theta_t + \zeta_t) + \delta_0 c_t + \delta_1 c_{t+1}]; \tag{20}$$

and 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. This let us parameterize the payoff perturbation by ε, or the size of the support of ζ_t.

[14] We thank a referee from prompting us to clarify this subtlety.

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. Under the additional, simplifying assumption that I_t contains both θ_t and ζ_t, we can then write the best response of the period-t representative agent as

$$c_t = \theta_t + \zeta_t + \delta E[c_{t+1}|I_t]. \tag{21}$$

where $\delta \equiv \frac{\delta_1}{1-\delta_0}$, as always, and $E[\cdot|I_t]$ is the rational expectation conditional on I_t. This is similar to the standard, full-information benchmark, except that we have allowed for the possibility that today's representative agent does 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 3. *For each t, there is a representative agent whose information is given by*

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

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

When the ζ_t shock is absent, or $\varepsilon = 0$, Assumption 3 allows replication of all sunspot and backward-looking equilibria with extremely short memory, namely 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 may be.

Proposition 4. *Suppose that Assumption 3 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.*

How does this connect to our main result? Both results introduce a friction in social memory and intertemporal coordination, thus breaking the infinite chain that sustain all non-fundamental equilibria. But the exact friction is different: whereas it amounts to *asymptotic* forgetting of the distant past in our main result, here it amounts to *immediate* forgetting of a small component of the fundamentals. This also means a change in the formal arguments: whereas our main result echoes the global-game literature, the present one is more closely connected to Bhaskar (1998) and Bhaskar et al. (2012). The precise connections between our two results, as well as between the corresponding two literatures, deserve further study.

Local vs global determinacy

Throughout, we work with the linearized New Keynesian model and restricted equilibria to be bounded. As previously mentioned, this amounts to focusing on local determinacy around a given steady state (herein normalized to zero). But what about global determinacy?

Let us first address this question within the policy context of interest. To ensure global determinacy, the standard paradigm compliments the Taylor principle with an escape clause: to switch from interest-rate setting to a different policy regime, such as money-supply setting or even commodity-backed money, should inflation exit certain bounds.[15] Under the standard approach, the escape clause rules out all unbounded equilibria (i.e., self-fulfilling inflationary and deflationary spirals), while the Taylor principle rules out any bounded equilibrium other than the MSV solution. Under our approach, the Taylor principle becomes redundant but the escape clause—or a credible commitment to arrest explosive paths—is still needed.

Consider next other contexts, such as the OLG model of money by Samuelson (1958). This is a non-linear model and it admits two steady-state equilibria: an "autarchic" one, in which the old and the young consume their respective endowment and money is not traded; and a "bubbly" one, in which money facilitates Pareto-improving transfers between the young and the old. In addition, there is a continuum of bounded sunspot equilibria, all of which hover around the second steady state. In this context, we can't rule out either of the steady-state equilibria, because our methods maintain common knowledge of the steady state(s) themselves, nor can we say anything about global determinacy. But if we linearize that model around each steady state and apply our assumptions and results, we can guarantee local determinacy of *both* steady states, and can therefore rule out the aforementioned sunspot equilibria.[16]

This clarifies the scope of our theoretical contribution. It seems a safe guess that Proposition 5 extends to a general class of linear REE models, such as that considered in the classics by Blanchard (1979) and Blanchard and Kahn (1980). In non-linear settings, this is likely to translate to local determinacy. But our methods and results do not speak to the question of global determinacy—except for the specific policy context of interest and in the way explained above. With this qualification in mind, we next focus on our paper's applied contribution.

[15] See, inter alia, Wallace (1981), Obstfeld and Rogoff (1983, 2021), Christiano and Rostagno (2001), and the discussion of "hybrid" Taylor rules in Atkeson et al. (2010).

[16] We thank the editor for suggesting the link to Samuelson (1958) and a referee for suggesting a different non-linear example, which has the same flavor but is more directly comparable to our own setting. We use that example in Appendix D to further illustrate the issues discussed above.

6 Applied Lessons

In this section we translate our main result to two applied lessons: one regarding the FTPL, and another regarding the Taylor principle. To facilitate these translations, we first illustrate how our main result extends to a larger class of New Keynesian model than that employed thus far.

Nesting a larger class of New Keynesian economies

Borrowing insight from the HANK literature, let us bypass the micro-foundations of consumer behavior and instead assume directly that aggregate demand can be expressed as follows:

$$c_t = \mathscr{C}\left(\{\bar{E}_t[y_{t+k}]\}_{k=0}^{\infty}, \{\bar{E}_t[r_{t+k}]\}_{k=0}^{\infty}\right) + \varrho_t, \tag{22}$$

where $r_t \equiv i_t - \pi_{t+1}$ stands for the real interest rate, $\mathscr{C}$ is a linear function, and ϱ_t is an exogenous shock. This generalizes equation (1) from our baseline model, allowing aggregate consumption to depend on expectations about interest rates and income at all future periods, not just the next period. In Appendix E, we show how to obtain (22) from a perpetual-youth version of the New Keynesian model. This allows us to cast the decay in social memory as the byproduct of individual mortality. But this interpretation is not strictly needed. For the present purposes, we take equation (22) as given and think of it as a linear but otherwise flexible specification of the intertemporal Keynesian cross (Auclert et al., 2018).

Consider next the supply side. We now replace our baseline model's ad hoc, static Phillips with the standard, micro-founded, and forward-looking New Keynesian Phillips curve:

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

where $\kappa \geq 0$ and $\beta \in (0,1)$ are fixed scalars and ξ_t is, again, a cost-push shock. The micro-foundations of (22) are omitted because they are entirely standard: whenever given the opportunity by the "Calvo fairy," firms optimally reset their prices under rational expectations and with full information.[17] Finally, we let the Taylor rule be

$$i_t = z_t + \phi_y y_t + \phi_\pi \pi_t, \tag{24}$$

for some random variable z_t and some fixed scalars $\phi_c, \phi_\pi \geq 0$.[18]

[17] The assumption that firms, unlike consumers, have full information simplifies the exposition and maximizes proximity to the standard New Keynesian model, without affecting the essence. For, as long as the informational friction is present in the consumer side, it is not necessary to "double" it in the production side.

[18] We can readily accommodate forward-looking terms in the policy rule. This changes the exact values of the coefficients $\{\delta_k\}$ in the upcoming game representation, namely equation (25), but does not affect Proposition 5, because this holds for arbitrary such coefficients. What we cannot readily nest in (25) is a backward-looking Taylor rule, such as $i_t = z_t + \phi_\pi \pi_{t-1}$, or a backward-lookin Phillips curve. See, however, Appendix B for an illustration of why this does not upset our result, insofar as, of course, Assumption 2 is maintained.

The "famous" three equations are now given by (22), (23) and (24), along with $y_t = c_t$ (by market clearing). Solving (23) and (24) for inflation and the interest rate, and replacing these solutions into (22), we can obtain c_t as a linear function of $\{\bar{E}_t[y_{t+k}]\}_{k=0}^{\infty}$, or equivalently of $\{\bar{E}_t[c_{t+k}]\}_{k=0}^{\infty}$. We conclude that a process for c_t is part of an equilibrium if and only if it solves the following:

$$c_t = \bar{E}_t\left[(1-\delta_0)\theta_t + \sum_{k=0}^{+\infty}\delta_k c_{t+k}\right] \tag{25}$$

for some scalars $\{\delta_k\}_{k=0}^{\infty}$, with $\delta_0 < 1$ and $\Delta \equiv \delta_0 + \sum_{k=1}^{\infty}|\delta_k| < \infty$.

Similar to equation (5) in our baseline model, this equation conveniently summarizes all the underlying GE feedbacks.[19] These feedbacks are now more convoluted, and aggregate spending in any given period depends on expectations of the outcomes in all future periods as opposed to merely the next period, but the essence is similar. For our purposes, the key is that the economy translates to a game in which: (i) a continuum of players acts in each period; (ii) a player's optimal strategy is given by $c_{i,t} = E_{i,t}\left[\theta_t + \sum_{k=0}^{+\infty}\delta_k c_{t+k}\right]$ for any t, any realization of her information set $I_{i,t}$, and any strategy played by other players; and (iii) the coefficient δ_k identifies the slope of an agent's best response with respect to the average action k periods later.

Thus put aside the micro-foundations and focus on the game representation. The overall strategic interdependence, or the analogue of the sum $\delta_0 + \delta_1$ from our main analysis, 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 we introduce Assumption 2, because this again breaks the "infinite chain" behind them. We verify this claim below. The proof is more tedious than that of Proposition 5 and is delegated to Appendix A, but the basic logic is the same.

Proposition 5 (Generalized result). *Consider the above generalization, impose Assumption 2, and let $\lambda > 0$. Whenever an equilibrium exists, it is unique and is given by the MSV solution.*[20]

Feedback rules and Taylor principle: equilibrium selection or stabilization?

Go back to the textbook New Keynesian model. Let $\{i_t^o, \pi_t^o, c_t^o\}$ denote the optimal path for interest rates, inflation, and output, as a function of the underlying demand and supply shocks. And ask the following question: what does it take for the optimum to be implemented as the unique equilibrium? The textbook answer is that, as long as the monetary authority observes the

[19] Accordingly, the coefficients $\{\delta_k\}_{k=0}^{\infty}$ can be expressed as functions of the following "deeper" parameters, which regulate these feedbacks: the MPCs out of current and future income, $\{\frac{\partial C}{\partial y_k}\}_{k=0}^{\infty}$; the sensitivities of consumption to current and future real interest rates, $\{\frac{\partial C}{\partial r_k}\}_{k=0}^{\infty}$; the slope, κ, and the forward-lookingness, β, of the NKPC; and the policy coefficients, ϕ_π and ϕ_c.

[20] When θ_t is uncorrelated over time, the MSV solution is again given by $c_t^F = \theta_t$. More generally, it can be solved for in a similar way as in our earlier discussion of persistent fundamentals.

aforementioned shocks, it suffices to follow the following feedback rule, for any $\phi > 1$:

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

This is nested in (4) with $z_t = i_t^o - \phi\pi_t^o$, and is sometimes referred to as the "King rule" (after King, 2000). Note then that ϕ can take any value above 1, and this does not affect the properties of the optimum. That is, the feedback from π_t to i_t serves only the role of equilibrium selection; macroeconomic stabilization is instead achieved via the optimal design of z_t, and in particular via its correlation with the underlying demand and supply shocks.[21]

What if the monetary authority does not observe these shocks? Feedback rules may then help replicate the optimal contingency of interest rates on shocks. But 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, in effect turning upside down its conventional interpretation (Atkeson et al., 2010; King, 2000). By this, we mean the following. When the equilibrium is unique (whether thanks to our perturbations or otherwise) but strategic complementarity is sizable, sunspot-like volatility can obtain from overreaction to noisy public news (Morris and Shin, 2002), shocks to higher-order beliefs (Angeletos and La'O, 2013; Benhabib et al., 2015), or related forms of bounded rationality (Angeletos and Sastry, 2021). In this context, the slope of the Taylor rule admits a new function: by regulating the overall complementarity in the economy, it also regulates the magnitude of such sunspot-like fluctuations along the unique equilibrium. Our contribution is therefore not to rule out "animal spirits" altogether but rather to recast policies that lean against them as a type of on-equilibrium stabilization instead of an off-equilibrium policy threat.

To make this idea more concrete, suppose (i) that we preserve our informational assumptions about sunspots, (ii) we introduce correlated higher-order uncertainty about future fundamentals. By (i), we can maintain the MSV solution as the economy's unique equilibrium, while by (ii), we can let this solution fluctuate in response to correlated shocks in higher-order beliefs. In the eyes of an outside observer, or a policymaker, 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.

[21] While most textbook treatments stop here, a kosher analysis combines the Taylor principle with escape clauses that rule out cunbounded equilibria; see the related discussion at end of this section.

On the Fiscal Theory of the Price Level (FTPL)

We now turn to how our paper relates to the FTPL. To this goal, let us momentarily go back again to the basics: the textbook, three-equation, New Keynesian model. Add now a fourth equation, written compactly (and in levels) as follows:

$$\frac{B_{t-1}}{P_t} = PVS_t, \tag{26}$$

where B_{t-1} denotes the outstanding nominal debt, P_t denotes the nominal price level, and PVS_t denotes the real present discounted value of primary surpluses. Does the incorporation of this equation make a difference for the model's predictions about inflation and output?[22]

The standard approach says no by assuming that fiscal policy is "Ricardian," in the following sense: PVS_t is required to adjust so as to make sure that (26) holds no matter P_t. This allows prices and quantities to be determined by the MSV solution of the model's other three equations. The FTPL turns this upside down: it requires that P_t itself adjusts to make sure that (26) for any given PVS_t. This is a coherent theoretical alternative, provided that the price level is determined according to a *different* solution of the model's other three equations.

It should be intuitive at this point that, by removing all solutions other than the MSV one, our paper also removes the equilibrium selected by the FTPL. But our analysis and formal results have thus far abstracted from fiscal policy. Could it be that explicit incorporation of fiscal policy modifies the MSV solution or otherwise upsets the way we have thought about the issue? We now show how to fill in the hole, clarifying some subtleties on the way.

We start by assuming that consumers have infinite horizons, or are "dynasties" as in Barro (1974). This rules out inter-generational redistribution and makes our analysis directly comparable to the standard treatment of the FTPL. For simplicity, we also rule out idiosyncratic income or interest-rate shocks. But we allow, at least momentarily, for arbitrary information. We can then write the (log-linearized) individual consumption function as follows:

$$c_{i,t} = E_{i,t}\left[(1-\beta)\gamma w_{i,t} - \sigma\beta\sum_{k=0}^{+\infty}\beta^k\left(i_{t+k}-\pi_{t+k+1}\right) + (1-\beta)\sum_{k=0}^{+\infty}\beta^k\left(y_{t+k}-\tau_{t+k}\right)\right], \tag{27}$$

where $w_{i,t}$ is the household's real financial wealth in the beginning of period t, τ_{t+k} are the lump sum taxes she owns in period $t+k$, all other variables are as before, and γ is the steady-state ratio of aggregate private financial wealth to GDP (equivalently, that of public debt to GDP).[23]

[22] There is disagreement between the proponents and the opponents of the FTPL on whether (26) should be read as a "real" constraint on the fiscal authority, which must hold both on and off equilibrium, or merely as an equilibrium condition, the market's valuation of government debt. Here, we put aside this somewhat "philosophical" debate and focus instead on whether and how equation (26) matters, regardless of its "deeper" interpretation.

[23] Note that equation (27) accommodates not only arbitrary information about the future but also possible inattention to own wealth: $w_{i,t}$ is left inside the expectation operator. This is not strictly needed for any of the results

Equation (27) is basically the Permanent Income Hypothesis. To derive it, we only imposed individual optimality: we made no assumption about what a consumer knows about the economy, how she forms expectations about interest rates, taxes, etc., or how he reasons about the behavior of others. We now add the following "minimal" assumptions about such knowledge/reasoning:

Assumption 4. *Consumers are first-order rational, in the sense that they have first-order knowledge of equation (26), the Phillips curve, the Taylor rule, and market clearing.*

Assumption 5. *At least on average, consumers do not mis-perceive their idiosyncratic wealth, in the sense* $\int E_{i,t}[w_{i,t} - w_t]di = 0$.

Assumption 4 is implied by REE but is significantly weaker than it: rational expectations amounts to infinite-order knowledge of the facts stated in this assumption, as well as of others' rationality, whereas the assumption requires only first-order knowledge of the stated facts. More succinctly, we require that agents themselves understand that equation (26) must ultimately hold, but we do not necessarily require that they know that others know this fact, nor that they have rational expectations about others' beliefs and behavior.

Assumption 5, on the other hand, is trivially satisfied when there is a representative agent (in which case $w_{i,t} = w_t$ for all i), as well as when agents are heterogeneous but know both their own wealth and the aggregate wealth (in which case $\int E_{i,t}[w_{i,t} - w_t] = \int (w_{i,t} - w_t)di = 0$). More generally, this assumption rules out the possibility that consumers confuse aggregate changes in fiscal policy for idiosyncratic variation in wealth. Such confusion is possible in the presence of informational frictions (Lucas, 1972) and may even rationalize a failure of Ricardian equivalence. But this is clearly *not* what the existing formulation of the FTPL is about, so Assumption 5 seems fully appropriate for our purposes.

This assumption alone guarantees that we can aggregate equation (27) to get the following:

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

Next, equation (26) rewrites in log-linearized form as

$$b_{t-1} - p = \frac{1}{\gamma} E_{i,t} \left[\sum_{k=0}^{+\infty} \beta^k (\tau_{t+k} - g_{t+k}) \right]$$

By Assumption 4, consumers understand this equation, as well as the identities $w_t = b_{t-1} - p_t$ and $y_t = c_t + g_t$. We can thus use these facts in (28) to obtain the following equation:

stated below. But it helps reduce the tension between Assumption 2, which we invoke in Corollary 1, and the idea that consumers may learn about past, unobserved, sunspots from the observation of their own wealth.

$$c_t = \bar{E}_t\left[-\sigma\beta\sum_{k=0}^{+\infty}\beta^k\left(i_{t+k}-\pi_{t+k+1}\right)+\left(1-\beta\right)\sum_{k=0}^{+\infty}\beta^k c_{t+k}\right]. \tag{29}$$

This can be interpreted as a DIS equation. But regardless of interpretation, the key observation here is that this equation is independent of fiscal policy. Finally, using the consumers' knowledge of the Taylor rule and the Phillips curve, we can map this equation to a special case of equation (25). That is, *without* invoking Assumption 2 or any other assumption about social memory and coordination, we can reach the following elementary result.

Proposition 6. *Suppose that agents are first-order rational and do not mis-perceive their wealth, in the sense of Assumptions 4 and 5. Then, aggregate consumption satisfies equation (25), for some coefficients $\{\delta_k\}_{k=0}^{\infty}$ and some random variable θ_t. Furthermore, debt and deficits do not appear in this equation: δ_k is a function of $(\sigma,\beta,\kappa,\phi)$ for all k, and θ_t is a transformation of $(z_t,\varrho_t,\xi_t,g_t)$.*

This result contains two key messages. First, the economy admits a similar game representation as before. And second, government debt and deficits do not enter the payoffs/best responses of this game. More succinctly, this result formalizes the sense in which debt and deficit are "non-fundamental" and verifies that the MSV solution is invariant to them.[24]

As already flagged, this result itself is true regardless of whether social memory/intertemporal coordination is perfect or imperfect. But once we combine it with our main assumption, it allows us to translate Proposition 5 to the present context as follows:

Corollary 1. *Suppose expectations of aggregate outcomes are formed according to Assumption 2. Whenever an equilibrium exists, it corresponds to the MSV solution of equation (25) and is invariant to both the outstanding level of debt and to the fiscal rule F. To put it differently, fiscal policy has to be Ricardian, or else it leads to equilibrium non-existence.*

With Full Information		
	Fiscal Policy is	
	Ricardian	Non-Ricardian
Taylor Principle holds	Determinacy	No equilibrium
does not hold	Multiplicity	Determinacy

With Our Perturbations		
	Fiscal Policy is	
	Ricardian	Non-Ricardian
Taylor Principle holds	Determinacy	No equilibrium
does not hold	Determinacy	No equilibrium

Table 1: Standard Paradigm vs Our Approach

[24]The conventional justification of this idea is that public debt and deficits do not appear the representative consumer's Euler condition. Cochrane (2005) criticizes this view on the basis that it fails to take into account the consumer's budget constraint and transversality condition, and he seems to argue that this allows for government debt to have a wealth effect off equilibrium. Proposition 6 deals properly with this issue (using individual consumption functions instead of merely Euler conditions) and shows that the conventional view remains valid as long as consumers are "minimally" rational, in the sense we have made precise.

Table 1 helps position this lesson in the literature. The left panel, which is basically reproduced from Leeper (1991), summarizes the state of the art. According to it, the non-Ricardian assumption is consistent with equilibrium existence, and uniquely pins down inflation and output, when monetary policy is passive. The right panel summarizes our own take on the issue: the non-Ricardian assumption is equated to equilibrium non-existence regardless of whether monetary policy is active or passive. This explains the sense in which our approach transforms the rejection of the FTPL from a "religious choice" to a logical necessity—provided, of course, that one accommodates the type of informational/coordination friction we have formalized here.

We conclude with two important qualifications. First, while the offered lesson about the FTPL is valid given Assumption 2, one can of course question the latter's precise meaning and empirical plausibility. Note in particular that, in the present context, this assumption rules out *direct* private or public signals of b_{t-1}, thus also ruling out equilibria in which consumers condition their behavior on such signals. Nevertheless, Proposition 6, which does *not* depend on Assumption 2, makes clear that any such equilibrium is necessarily a non-fundamental one, along which public debt plays one and only one role: that of an endogenous sunspot. At a high level, this circles back to our discussion of endogenous sunspots in Section 5. But the endogenous sunspot is now of a different form, preventing applicability of the specific results developed in that section. This calls for further exploration of the informational assumptions that may or may not support the equilibrium selected by the FTPL. But it does not negate the essence of what we have shown: a precise formalization of the sense in which the equilibrium selected by the FTPL is both non-fundamental (Proposition 6) and fragile to certain perturbations (Corollary 1).

Second, our approach leaves ample room for debt and deficits, or expectations thereof, to drive inflation and output insofar as (i) these objects influence aggregate demand because of finite horizons, liquidity constraints, rational confusion, or even plain irrationality; or (ii) the monetary authority internalizes the fiscal ramifications of its policies. The first option lets b_{t-1} and $\tau_t - g_t$ enter directly our game representation, for given z_t and ϕ; the second one makes the monetary authority's choice of these objects endogenous to fiscal conditions. Both options thus allow government debt and deficits to drive the MSV solution, regardless of whether another solution exists or not. By the same token, our uniqueness result is logically consistent with the "unpleasant arithmetic" of Sargent and Wallace (1981), 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; Sims, 2022), any estimated link between fiscal conditions and inflation (e.g., Bianchi and Ilut, 2017; Bianchi et al., 2020; Chen et al., 2021), and the real-world concern that monetary policy may succumb to political pressure. Perhaps these are the issues that the spirit of the FTPL is meant to be about, once liberated from the equilibrium selection conundrum.

7 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 a delicate, infinite, self-fulfilling chain between current and future behavior. And we showed how to break this chain, and guarantee that the model's fundamental or MSV solution is the unique rational expectations equilibrium regardless of monetary or fiscal policy, by appropriately perturbing the model's assumptions about social memory and intertemporal coordination.

We thus 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 the Taylor principle or any other off-equilibrium threat. But we also discussed how one could accommodate sunspot-like volatility along the economy's unique equilibrium, and highlighted that a steeper Taylor rule could help regulate the size of such volatility in a continuous way. More succinctly, we first killed the Taylor principle as a form of equilibrium selection and then resurrected it as a form of macroeconomic stabilization.

We offered a similar two-sided approach to the FTPL. We first showed that, under our perturbations, the non-Ricardian assumption can be equated to equilibrium non-existence, regardless of whether monetary policy was active or passive. One may of course quibble with the realism of our perturbations. Still, by illustrating the potential fragility of the existing formulation of the FTLP, we not only lend support to the conventional use of the New Keynesian model but also paved the way for resurrecting the (appealing) spirit of the FTPL outside the (unappealing) equilibrium-selection conundrum.

To illustrate what we have in mind, consider the topical question of whether the large public debt in the US will trigger inflation by forcing the hands of the Fed towards more lax monetary policy, or the broader question of which authority is "dominant." In our view, such questions seem to call for modeling the interaction between the the fiscal and the monetary authorities as that of two players in a game, for example a game of chicken. But this requires in the first place the existence of a unique mapping from the players' actions—government deficits and interest rates, respectively—to their payoffs. Such a unique mapping is missing in the standard paradigm, because of the equilibrium determinacy problem: the same paths for government deficits and interest rates can be associated with multiple equilibria within the private sector, and thereby with multiple equilibrium payoffs for the two authorities. By providing a possible fix to this "bug," or at least a formal justification for bypassing it, our paper may open the way to new research on these important policy questions.

Appendix A: 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 $|\gamma_k| \leq M$ for all k.*

From Assumption 1 and Definition 1, we have

$$Var(c_t) = \left(\sum_{k=0}^{\infty} a_k^2\right) Var(\eta_t) + \left(\sum_{k=0}^{\infty} \gamma_k^2\right) Var(\theta_t).$$

This can be finite only if $\lim_{k\to+\infty} |a_k| = 0$ and $\lim_{k\to+\infty} |\gamma_k| = 0$. We conclude that there exist a scalar $M > 0$, large enough but finite, such that $|a_k| \leq M$ and $|\gamma_k| \leq M$ for all k.

Proof of Proposition 1

Part (i) follows directly from the fact that $c_t^F \equiv \theta_t$ satisfies (8).

Consider part (ii). Let $\{c_t\}$ be any equilibrium and define $\hat{c}_t = c_t - c_t^F$. From (8),

$$\hat{c}_t = \delta \mathbb{E}_t[\hat{c}_{t+1}]. \tag{30}$$

From Definition 1,

$$\hat{c}_t = \sum_{k=0}^{\infty} \hat{a}_k \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_k \theta_{t-k},$$

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

$$\mathbb{E}_t[\hat{c}_{t+1}] = \sum_{k=0}^{+\infty} \hat{a}_{k+1} \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_{k+1} \theta_{t-k}.$$

The equilibrium condition (30) can thus be rewritten as

$$\sum_{k=0}^{\infty} \hat{a}_k \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_k \theta_{t-k} = \delta\left(\sum_{k=0}^{+\infty} \hat{a}_{k+1} \eta_{t-k} + \sum_{k=0}^{\infty} \hat{\gamma}_{k+1} \theta_{t-k}\right).$$

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

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

When the Taylor principle is satisfied ($|\delta| < 1$), $\hat{a}_k$ and $\hat{\gamma}_k$ explodes unless $\hat{a}_0 = 0$ and $\hat{\gamma}_0 = 0$. We know that the only bounded solution of (30) 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 (8). So does

$c_t = (1-b)c_t^F + bc_t^B + ac_t^\eta$ 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 $\{\theta_{t-k}\}_{k=0}^{\infty}$, the argument in the main text proves that $a_k = 0$ for all k. We can thus focus on solutions of the following form:

$$c_t = \sum_{k=0}^{\infty} \gamma_k \theta_{t-k}. \tag{31}$$

And the remaining task is to show that $\gamma_0 = 1$ 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 θ_t is a stationary i.i.d. Gaussian variable from Assumption 1, the following projections apply for all $k \geq s \geq 0$:

$$\mathbb{E}\left[\theta_{t-k} | I_t^s\right] = 0,$$

where $I_t^s \equiv \{(\theta_t, \eta_t), \cdots, (\theta_{t-s}, \eta_{t-s})\}$ is the period-t information set of an agent with memory length s.

Now, from Assumption 2, we know

$$\bar{E}_t\left[\theta_{t-k}\right] = (1-\lambda)^k \theta_{t-k} + \sum_{s=0}^{k-1} \lambda (1-\lambda)^s \mathbb{E}\left[\theta_{t-k} | I_t^s\right] \equiv (1-\lambda)^k \theta_{t-k}. \tag{32}$$

Now consider an equilibrium in the form of (31). From equilibrium condition (5), we know

$$\begin{aligned}
\sum_{k=0}^{+\infty} \gamma_k \theta_{t-k} &= (1-\delta_0)\theta_t + \delta_0 \bar{E}_t\left[\sum_{k=0}^{+\infty} \gamma_k \theta_{t-k}\right] + \delta_1 \bar{E}_t\left[\sum_{k=0}^{+\infty} \gamma_k \theta_{t+1-k}\right] \\
&= \left((1-\delta_0) + \delta_0 + \delta_1 \gamma_1\right)\theta_t + \bar{E}_t\left[\sum_{k=1}^{+\infty} \left(\delta_0 \gamma_k + \delta_1 \gamma_{k+1}\right)\theta_{t-k}\right] \\
&= \left((1-\delta_0) + \delta_0 + \delta_1 \gamma_1\right)\theta_t + \sum_{k=1}^{+\infty} \left(\delta_0 \gamma_k + \delta_1 \gamma_{k+1}\right)(1-\lambda)^k \theta_{t-k},
\end{aligned}$$

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

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

$$\begin{aligned}
\gamma_0 &= (1-\delta_0) + \delta_0 \gamma_0 + \delta_1 \gamma_1 \\
\gamma_k &= \left(\delta_0 \gamma_k + \delta_1 \gamma_{k+1}\right)(1-\lambda)^k \quad \forall k \geq 1.
\end{aligned} \tag{33}$$

From Definition 1, we know that there is a scalar $M > 0$ such that $\left|\gamma_k\right| \leq M$ for all $k \geq 0$. From (33), we know that, for all $k \geq 1$,

$$\left|\gamma_k\right| \leq \left(|\delta_0| + |\delta_1|\right)(1-\lambda)^k M. \tag{34}$$

Because $\lambda > 0$, there necessarily exists an $\hat{k}$ finite but large enough $\left(|\delta_0| + |\delta_1|\right)(1-\lambda)^{\hat{k}} < 1$. We then know that, for all $k \geq \hat{k}$,

$$\left|\gamma_k\right| \leq (|\delta_0| + |\delta_1|)\,(1-\lambda)^{\hat{k}}\,M.$$

Now, we can use the above formula and (33) to provide a tighter bound of $\left|\gamma_k\right|$: for all $k \geq \hat{k}$,

$$\left|\gamma_k\right| \leq (|\delta_0| + |\delta_1|)^2\,(1-\lambda)^{2\hat{k}}\,M.$$

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

$$\left|\gamma_k\right| \leq (|\delta_0| + |\delta_1|)^l\,(1-\lambda)^{l\hat{k}}\,M.$$

Since $(|\delta_0| + |\delta_1|)\,(1-\lambda)^{\hat{k}} < 1$, we then have $\gamma_k = 0$ for all $k \geq \hat{k}$. Using (33) and doing backward induction, we then know $\gamma_k = 0$ for all $k \geq 1$ and

$$\gamma_0 = (1-\delta_0) + \delta_0\gamma_0,$$

which means $\gamma_0 = 1$, where I use $\delta_0 < 1$. Together, this means that the equilibrium is unique and is given by $c_t = c_t^F$, where $c_t^F = \theta_t$.

Proof of Proposition 3

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, $E_{i,t}[c_{t+1}] = bE_{i,t}[c_t]$. Next, note that $E_{i,t}[c_t] = a\eta_t + b\chi s_{it}$, where

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

Combining these facts, we infer that condition (6), the individual best response, reduces to

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

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

$$a = (\delta_0 + \delta_1 b)a \qquad \text{and} \qquad b = (\delta_0 + \delta_1 b)b\chi. \tag{35}$$

Clearly, $a = b = 0$ is always an equilibrium, and it corresponds to the MSV solution. To have a sunspot equilibrium, on the other hand, it must be that $a \neq 0$ (and also that $|b| < 1$, for it to be bounded). From the first part of condition (35), 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 $a \neq 0$).

Proof of Proposition 4

Given Assumption 3, an possible equilibrium takes the form of

$$c_t = \sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k \theta_{t-k} + \chi \zeta_t.$$

From (21), we have that

$$\begin{aligned}
\sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k x_{t-k} + \chi \zeta_t &= \theta_t + \zeta_t + \delta \mathbb{E}[\sum_{k=0}^{K_\eta -1} a_{k+1} \eta_{t-k} + \sum_{k=0}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_\theta -1} \gamma_{k+1} \theta_{t-k} | I_t] \\
&= \theta_t + \zeta_t + \delta \left[\sum_{k=0}^{K_\eta -1} a_{k+1} \eta_{t-k} + \sum_{k=1}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_\theta -1} \gamma_{k+1} \theta_{t-k} \right] \\
&+ \delta \beta_1 \left[\sum_{k=0}^{K_\eta} a_k \eta_{t-k} + \sum_{k=1}^{K_\beta} \beta_k c_{t-k} + \sum_{k=0}^{K_\theta} \gamma_k \theta_{t-k} + \chi \zeta_t \right]
\end{aligned}$$

where we use Assumption 1 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:

$$a_k = \delta a_{k+1} + \delta \beta_1 a_k \quad \forall k \in \left\{0, \cdots, K_\eta - 1\right\} \quad \text{and} \quad a_{K_\eta} = \delta \beta_1 a_{K_\eta} \tag{36}$$

$$\beta_k = \delta \beta_{k+1} + \delta \beta_1 \beta_k \quad \forall k \in \left\{1, \cdots, K_\beta - 1\right\} \quad \text{and} \quad \beta_{K_\beta} = \delta \beta_1 \beta_{K_\beta} \tag{37}$$

$$\gamma_k = \delta \gamma_{k+1} + \delta \beta_1 \gamma_k \quad \forall k \in \{1, \cdots, K_\theta - 1\} \quad \text{and} \quad \gamma_{K_\theta} = \delta \beta_1 \gamma_{K_\theta} \tag{38}$$

$$\gamma_0 = 1 + \delta \gamma_1 + \delta \beta_1 \gamma_0 \quad \text{and} \quad \chi = 1 + \delta \beta_1 \chi. \tag{39}$$

First, from the second equation in (39), we know $\delta \beta_1 \neq 1$. Then, from the second parts of (36)–(38), we know $a_{K_\eta} = 0$, $\beta_{K_\beta} = 0$, and $\gamma_{K_\theta} = 0$. From backward induction on (36)–(39), we know that all a, b, γ are zero except for the following:

$$\gamma_0 = 1.$$

We also know that $\chi = 1$. We conclude that the unique solution is

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

where $c_t^F = \theta_t$.

Proof of Proposition 5

We first note that the MSV solution of (25) is still given by $c_t^F = \theta_t$. Consider an equilibrium taking the form of (7). We use (5):

$$\sum_{l=0}^{+\infty} a_l \eta_{t-l} + \sum_{l=0}^{\infty} \gamma_l \theta_{t-l} = (1 - \delta_0) \theta_t + \bar{E}_t \left[\sum_{k=0}^{+\infty} \delta_k \left(\sum_{l=0}^{+\infty} a_l \eta_{t+k-l} + \sum_{l=0}^{\infty} \gamma_l \theta_{t+k-l} \right) \right]. \tag{40}$$

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:

$$a_l = \mu_l \sum_{k=0}^{+\infty} \delta_k a_{k+l} \quad \forall l \geq 0. \tag{41}$$

Because $\lim_{l\to\infty} \mu_l = 0$, there necessarily exists an $\hat{l}$ finite but large enough $\mu_{\hat{l}} \sum_{k=0}^{\infty} |\delta_k| < 1$.[25]

Since we are focusing bounded equilibria as in Definition 1, there exists a scalar $M > 0$, arbitrarily large but finite, such that $|a_l| \leq M$ for all l. From (41), we then know that, for all $l \geq \hat{l}$,

$$|a_l| \leq \mu_{\hat{l}} M \sum_{k=0}^{+\infty} |\delta_k|, \tag{42}$$

where we also use the fact that the sequence $\{\mu_l\}_{l=0}^{\infty}$ is decreasing. Now, we can use (41) and (42) to provide a tigehter bound of $|a_l|$. That is, for all $l \geq \hat{l}$,

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

We can keep iterating. Since $\mu_{\hat{l}} \sum_{k=0}^{\infty} |\delta_k| < 1$, we then have $a_l = 0$ for all $l \geq \hat{l}$. Using (41) and doing backward induction, we then know $a_l = 0$ for all l.

Now, (40) can be simplified as

$$\begin{aligned} \sum_{l=0}^{\infty} \gamma_l \theta_{t-l} &= (1-\delta_0)\theta_t + \bar{E}_t\left[\sum_{k=0}^{+\infty} \delta_k \sum_{l=0}^{\infty} \gamma_l \theta_{t+k-l}\right]. \\ &= (1-\delta_0)\theta_t + \sum_{k=0}^{+\infty} \delta_k \gamma_k \theta_t + \bar{E}_t\left[\sum_{l=1}^{+\infty}\left(\sum_{k=0}^{+\infty} \delta_k \gamma_{k+l}\right)\theta_{t-l}\right]. \end{aligned} \tag{43}$$

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

$$\gamma_0 = 1 - \delta_0 + \sum_{k=0}^{+\infty} \delta_k \gamma_k \tag{44}$$

$$\gamma_l = (1-\lambda)^l \sum_{k=0}^{+\infty} \delta_k \gamma_{k+l} \quad \forall l \geq 1. \tag{45}$$

[25] $\sum_{k=0}^{\infty} |\delta_k| < \infty$ because $\Delta < \infty$.

The above two equations can be re-written as:

$$\gamma_0 = (1-\delta_0)^{-1}\left(1-\delta_0+\sum_{k=1}^{+\infty}\delta_k\gamma_k\right) \tag{46}$$

$$\gamma_l = \left(1-(1-\lambda)^l\delta_0\right)^{-1}\left(\sum_{k=1}^{+\infty}\delta_k\gamma_{k+l}\right) \quad \forall l \geq 1, \tag{47}$$

where we use $\delta_0 < 1$.

From Definition 1, we know that there is a scalar $M > 0$ such that $\left|\gamma_l\right| \leq M$ for all $l \geq 0$. From (45), we know, for all $l \geq 1$

$$\left|\gamma_l\right| \leq (1-\lambda)^l\left(\sum_{k=0}^{+\infty}|\delta_k|\right)M. \tag{48}$$

Because $\lim_{l\to\infty}(1-\lambda)^l = 0$, there necessarily exists an $\hat{l}$ finite but large enough such that $\left(\sum_{k=0}^{+\infty}|\delta_k|\right)(1-\lambda)^{\hat{l}} < 1$. We then know that, for all $l \geq \hat{l}$,

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

Now, we can use the above formula and (45) to provide a tighter bound of $\left|\gamma_l\right|$: for all $l \geq \hat{l}$,

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

We can keep iterating. Since $\left(\sum_{k=0}^{+\infty}|\delta_k|\right)(1-\lambda)^{\hat{l}} < 1$, we then have $\gamma_l = 0$ for all $l \geq \hat{l}$. Using (47) and doing backward induction, we then know $\gamma_l = 0$ for all $l \geq 1$ and, from (46),

$$\gamma_0 = 1.$$

Together, this means that the equilibrium is unique and is given by $c_t = c_t^F = \theta_t$. This proves the Proposition.

Proof of Proposition 6 and Corollary 1.

Let us revisit our characterization of optimal consumption. Relative to what we did in (**??**), 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 $Y_t - T_t$ instead of Y_t, where Y_t are the taxes; third, the consumers' aggregate financial wealth is $W_t \equiv \int W_{i,t}di = B_{t-1}/P_t$ instead of 0, where B_{t-1}/P_t is the real value of the outstanding nominal debt. Accordingly, the consumer's budget constraint is given by

$$\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\frac{I_{t+j-1}}{\Pi_{t+j}}\right)^{-1}\right]C_{i,t+k}\right\} = W_{i,t}+\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\frac{I_{t+j-1}}{\Pi_{t+j}}\right)^{-1}\right](Y_{t+k}-T_{t+k})\right\} .$$

The government's budget in (26) can be written as

$$B_{t-1}/P_t = \sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\frac{I_{t+j-1}}{\Pi_{t+j}}\right)^{-1}\right](T_{t+k}-G_{t+k})\right\} . \tag{49}$$

Since a consumer understands that (49) holds, she understands that her budget can be written as

$$\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\frac{I_{t+j-1}}{\Pi_{t+j}}\right)^{-1}\right]C_{i,t+k}\right\} = W_{i,t}-W_t+\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\frac{I_{t+j-1}}{\Pi_{t+j}}\right)^{-1}\right](Y_{t+k}-G_{t+k})\right\} .$$

The consumer's optimal consumption function, in log-linearized form, can thus be written as follows:

$$c_{i,t} = E_{i,t}\left[(1-\beta)\gamma(w_{i,t}-w_t)-\sigma\beta\sum_{k=0}^{+\infty}\beta^k(i_{t+k}-\pi_{t+k+1})+(1-\beta)\sum_{k=0}^{+\infty}\beta^k(y_{t+k}-g_{t+k})\right], \tag{50}$$

where $\gamma = \frac{B^*}{Y^*}$ is the steady-state debt-to-GDP ratio and all lowercase variables represent log-deviations from the steady state.[26] Aggregating, and using Assumption 5, we arrive at

$$c_t = (1-\beta)\left\{\sum_{k=0}^{+\infty}\beta^k\bar{E}_t[y_{t+k}-g_{t+k}]\right\}-\beta\sigma\left\{\sum_{k=0}^{+\infty}\beta^k\bar{E}_t[i_{t+k}-\pi_{t+k+1}]\right\}. \tag{51}$$

This is the same as equation (**??**), except y_{t+k} is replaced with $y_{t+k}-g_{t+k}$, because consumers understand that the government absorbs part of the aggregate output. Crucially, neither the level of government debt nor the expected path of taxes shows up in this condition; and this is true despite the fact that no assumption has been made thus far about how consumers form expectations regarding one another's behavior or any aggregate variable. In other words, to reach condition (51) we have *not* used the full bite of REE; we have only assumed that consumers have first-order knowledge of condition (26) from Assumption 4.

From Assumption 4, consumers understand that the goods markets must clear; and second, consumers understand that inflation obeys the NKPC (23) and that monetary policy follows the Taylor rule (24). The first property allows us to replace the expectations of $\{y_{t+k}-g_{t+k}\}$ in condition (51) with those of $\{c_{t+k}\}$; the second allows us to do the same for expectations of $\{\pi_{t+k}\}$ and $\{i_{t+k}\}$.[27]

[26] The steady state is one in which $G_t = 0$, $Y_t = C_t = Y^*$, and $T^* = (1-\beta)B^* > 0$. Also, the following exception applies to the statement that all variables are in log-deviations: g_t is the ratio G_t/Y^*. This is a standard trick in the literature on fiscal multipliers (e.g., Woodford, 2011) and it simply takes care of the issue that the log-deviation of the government spending is not well defined when its steady-state value is 0.

[27] To be precise, although the expectations of $\{g_{t+k}\}$ drop out in the first step, they reemerge in the second step as long as $\kappa > 0$, because government spending enters the NKPC as a cost push shock. But this amounts to a redefinition of ξ_t, or θ_t, and is of no consequence for our purposes.

Putting everything together, we arrive at the *same* fixed-point relation between c_t and the average expectations of $\{c_{t+k}\}$, or the same "game" among the consumers, as when fiscal policy is absent. That is, the equilibrium process for c_t must still solve equation (25);[28] under our informational assumptions, the MSV solution of this equation continues to identify the unique possible equilibrium process for c_t; conditional on the latter, the processes for π_t and i_t are uniquely pinned down by the NKPC curve and the Taylor rule; and the fiscal authority's policy rule, F, does not enter the determination of any of these objects. This proves Proposition 6. Corollary 1 then follows from 5.

Appendix B: Additional Discussion

This Appendix Dorroborates various claims made in the main text. First, we explain why the simplification of infinite histories and stationary equilibria is non-essential. Second, we formalize the sense in which Assumption 2 is compatible with nearly perfect information of both exogenous shocks and endogenous outcomes.

Time 0 and non-stationary equilibria

In the preceding analysis, we let histories be infinite and restricted equilibria to be stationary. To understand what exactly this simplification does, abstract from fundamentals (this is without any loss), let calendar time start at $t = 0$, and modify (7) as follows:

$$c_t = b_t + \sum_{k=0}^{t} a_{t,k}\eta_{t-k},$$

where $\{a_{t,k}\}$ and $\{b_t\}$ are deterministic coefficients. Note that this allows for (i) a time-varying, non-zero deterministic intercept and (ii) the equilibrium load of a sunspot to be a function of not only its age (k) but also the calendar time.

It is straightforward to show that Assumption 2 continues to rule out sunspot fluctuations, that is, $a_{t,k} = 0$ for all t, k. But it does not immediately rule a deterministic, time-varying intercept.

[28]Minor qualification: g_t must now be included in the definition of θ_t, but this makes not difference for the argument made here.

In particular, c_t is now an equilibrium if and only if

$$c_t = b_t = \delta^{-t} b_0, \tag{52}$$

for arbitrary $b_0 \in \mathbb{R}$. At first glance, this appears to contradict our claim of equilibrium uniqueness. But this is only an artifact of introducing infinite social memory "through the back door."

Let us explain. Clearly, (52) is exactly the same as the following sunspot equilibrium:

$$c_t = \delta^{-t} \eta_0,$$

with the constant b_0 in place of the sunspot η_0. So all the "deterministic" equilibria obtained above are really sunspot equilibria in disguise. But by treating b_0 (equivalently, c_0) as a deterministic scalar instead of a random variable, we have artificially bypassed the friction of interest: we have effectively imposed that the initial sunspot can never be forgotten.

To sum up, insofar one wishes to remain true to the spirit of Assumption 2, one must treat any initial sunspot as a random variable rather than a deterministic constant. And provided that this is done, our result goes through.

Knowledge about endogenous outcomes

Although Assumption 2 excluded direct observation of endogenous aggregate outcomes, such as output and inflation, our main result can be said to compatible with nearly perfect knowledge of such outcomes, in the following sense:

Proposition 7 (Nearly perfect information about endogenous outcomes). *For any given 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 $\lambda \in (0, \hat{\lambda})$, $Var\left(E_t^i\left[c_{t-k}\right] - c_{t-k}\right) \le \epsilon$ for all $k \in \{0, 1, \cdots, K\}$, for at least a mass $1 - \epsilon'$ of agents and for every t. (And the same is true if we replace c_{t-k} with π_{t-k}, i_{t-k}, or any linear combination thereof.)*

Proof: Consider a candidate equilibrium c_t in Definition 1. We first use I_t^s to denote the information set of the period-t agent with memory length s:

$$I_t^s = \left\{\eta_{t-s}, \cdots, \eta_t, \theta_{t-s}, \cdots, \theta_t\right\}.$$

From Definition 1, we know that c_t can be written as

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \gamma_k \theta_{t-k}.$$

From the law of total variances, we have

$$Var\left(E_t\left[c_t | I_t^s\right] - c_t\right) \le Var\left(\sum_{k=s+1}^{\infty} a_k \eta_{t-k} + \sum_{k=s+1}^{\infty} \gamma_k \theta_{t-k}\right).$$

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

$$\lim_{s\to+\infty} Var\left(\sum_{k=s+1}^{\infty} a_k\eta_{t-k} + \sum_{k=s+1}^{\infty}\gamma_k\theta_{t-k}\right) = 0.$$

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

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

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

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

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 $\lambda \in (0,\hat{\lambda})$, $Var\left(E_t^i\left[c_{t-k}\right] - c_{t-k}\right) \leq \epsilon$ for all $k \leq K$, for at least a fraction $1-\epsilon'$ of agents, and for every period t.
□

The following important qualification, however, applies. The above result allows the mapping from h^t to c_t to be arbitrary but treats this mapping as fixed when λ is lowered towards 0. But the *equilibrium* mapping from h^t to c_t may well vary with λ, upsetting the result. In Section 5 we therefore present two alternative information structures, which allow for direct observation of past outcomes and properly deal with this endogeneity.

Alternative Monetary Policies

In the main analysis, we specify the monetary policy (4) where the nominal interest rate responds to *current* inflation. In the literature (e.g. Bullard and Mitra, 2002), variants of such rules have been proposed. One may wonder whether the alternative specifications change our lessons on determinacy. The answer is no.

For example, one specification is that the the nominal interest rate responds to forecasts of future inflation:

$$i_t = z_t + \phi\bar{E}_t\left[\pi_{t+1}\right], \tag{53}$$

where $\phi \geq 0$. A system consisting of (2), (3), and (53) can be nested by the general environment (25), and the determinacy result in Proposition 5 directly applies.

Another specification is that the nominal interest rate responds to lagged values of inflation:

$$i_t = z_t + \phi\pi_{t-1}, \tag{54}$$

where $\phi \geq 0$. Even though this case is not directly nested in Proposition 5, the result about how frictions in intertemporal coordination results in determinacy remains to hold. Specifically, consider the systems consisting of (2), (3), and (54). Finally, shut down fundamentals shocks $\varrho_t = \xi_t = z_t = 0$, so the MSV solution is $c_t = 0$. Proposition 2 can be recast as the follows:

Proposition 8 (**Alternative monetary policies**). *Suppose Assumption 2 holds, there are no shocks to fundamentals, and monetary policy takes the form of (54). The equilibrium is unique and is given by $c_t = 0$.*

Proof: From (2), (3), and (54), we have that any equilibrium must satisfy

$$c_t = \bar{E}_t\left[\tfrac{1}{1+\beta}c_t - \tfrac{\beta}{1+\beta}\sigma\phi\kappa c_{t-1} + \tfrac{\beta}{1+\beta}(1+\sigma\kappa)c_{t+1}\right]; \tag{55}$$

and since there are no shocks to fundamentals, we search for solutions of the form $c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k}$. The goal is to verify that $a_k = 0$ for all k.

By Assumption 2, 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

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

and similarly

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

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

For condition (19) to be true for all sunspot realizations, it is necessary and sufficient that,

$$a_0 = (1+\sigma\kappa)\, a_1,$$

and, for $k \geq 1$,

$$a_k = \mu_k\left(\frac{1}{1+\beta}a_k - \tfrac{\beta}{1+\beta}\sigma\phi\kappa a_{k-1} + \tfrac{\beta}{1+\beta}(1+\sigma\kappa)\, a_{k+1}\right).$$

We hence have, for $k \geq 1$,

$$a_{k+1} = \frac{\frac{1}{\mu_k} - \frac{1}{1+\beta}}{\frac{\beta}{1+\beta}(1+\sigma\kappa)}a_k + \frac{\sigma\phi\kappa}{1+\sigma\kappa}a_{k-1}. \tag{56}$$

Since $\frac{1}{\mu_k} - \frac{1}{1+\beta} > 0$, we know that, all $\{a_k\}_{k=0}^{+\infty}$ have the same sign if $a_0 \neq 0$. But because $\mu_k \to 0$, we have that $|a_k|$ explodes to infinity as $k \to \infty$ from 56 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. □

Sticky vs flexible prices

Our analysis has allowed the Phillips curve to have an arbitrary slope $\kappa \in [0, \infty)$. In this sense, our results do not depend on the degree of nominal rigidity, and they allow in particular the limit with nearly flexible prices ($\kappa \to \infty$). But what about the knife-edge case in which prices are *perfectly* flexible ("$\kappa = \infty$")?

To ease the exposition, let us address this question in our baseline model. Maintain our assumptions about consumers and monetary policy, but modify the production side so that prices are truly flexible and output is given by a fixed endowment (so that $c_t = y_t = 0$ in log-deviations). Clearly, our characterization of the individual optimal consumption function in (1) is still valid, and so does the intertemporal Keynesian cross obtained in condition (2), which we repeat below:

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

But now the assumption of a fixed endowment together with market clearing implies that $c_t = 0$, which in turns means that the above condition reduces to

$$\bar{E}_t \left[i_t - \pi_{t+1} - \varrho_t \right] = 0.$$

This is no other than the Fisher equation, only adapted to heterogeneous information.

For simplicity, switch off the discount rate shock, so that $\varrho_t = 0$, and let monetary policy peg the nominal interest rate at its steady-state value, so that $\phi = 0$ and $i_t = z_t = 0$. Recall that these restrictions are consistent with our main result, which guaranteed uniqueness for an arbitrary degree of nominal rigidity. But now that prices are perfectly flexible, these restrictions imply that the Fisher equation reduces to

$$\bar{E}_t [\pi_{t+1}] = 0.$$

Two properties are then evident. First, there is no feedback from expectations of future outcomes to current outcomes, or no intertemporal coordination of the type that has been at the core of our analysis thus far. And second, equilibrium pins down only the average expectation of inflation and not its precise realizations. In particular, $\pi_t = a\eta_t$, where η_t is the sunspot and $a \in \mathbb{R}$ is an arbitrary scalar, is an REE under our main assumption for every $\lambda > 0$ and, more generally, for every information structure such that $I_{i,t}$ merely contains η_t. In a nutshell, our uniqueness result does not apply and we are basically back to Sargent and Wallace (1981).

Although this clarifies the applicability of our result, we suspect that it ultimately speaks to an inherent "bug" of the baseline RBC model, or equivalently of the flexible-price core of the New Keynesian model. By design, this otherwise important conceptual benchmark is not well suited for understanding how nominal prices are determined: the nominal price level is both payoff-irrelevant and set by an "invisible hand." By contrast, the New Keynesian model ties the adjustment in nominal prices to the optimizing behavior of specific players, the firms, and allows one to recast the whole economy as a game between the firms and the consumers.[29] The presence of *some* nominal rigidity was essential for obtaining such a game in the first place. But once we got to this point, that is, once we properly accounted for both real output and nominal prices as the average actions of specific players, our analysis could proceed without any restriction on how large or small the nominal rigidity might be.

What about monetary models in which nominal rigidity is absent but nominal prices are otherwise payoff-relevant, such as models with money in the utility function? We suspect that a version of our results may be applicable in this case, but we leave this conjecture open for future research.

Discounted Euler Equations, and Beyond REE

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

$$c_t = -m_i i_t + m_\pi \bar{E}_t[\pi_{t+1}] + m_c \bar{E}_t[c_{t+1}] + \varrho_t, \tag{57}$$

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

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

[29] This point might have been blurred by our choice to solve out firm behavior and reduce the economy to a game among the consumers alone. But recall that this game embeded the best-responses of the firms, via the NKPC, which translates as follows: what we *really* did in this paper was to study the game played by both consumers and firms, for any given monetary policy rule.

Depending on the m's, these thresholds can be either smaller or larger than the textbook counterparts. 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. But the essence of the determinacy problem remains the same.

A similar point applies to Diagnostic Expectations as in Bordalo et al. (2018); Perfect Bayesian Equilibrium with mis-specified priors as in Angeletos and Sastry (2021); and Woodford (2019b)'s model of "finite planning horizons," at least once learning is allowed (Xie, 2019). All these concepts depart from REE by relaxing the exact coincidence between subjective beliefs and objective distribution; but they are close cousins to REE in 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 replaces the sunspot with another free variable, the analyst's specification of the level-1 belief.[30]

This explains the sense in which Level-K Thinking replaces one free variable in beliefs (the sunspot) with another free variable (the analyst's specification of the level-0 behavior). By contrast, our approach leaves neither kind of freedom in specifying beliefs.

This is not to say that our approach is "better." One may question the realism of both our

[30] Let us explain what we mean by this. Whenever $|\delta| > 1$, the level-k outcome becomes *infinitely* sensitive to the arbitrary level-0 outcome as $k \to \infty$. To see this, 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 $K \to \infty$.

main informational assumption and our approach's heavy reliance on REE. Furthermore, the two approaches are ultimately complementary in two regards: highlighting the role of higher-order beliefs; and solidifying the logical foundations of the MSV solution. Thus, while the above discussion clarifies the differences in the two approaches, perhaps their common ground is what matters the most for applied purposes.

Appendix C: General Fundamentals

In this Appendix we verify the claim that our main result extends to a more general specification for the fundamentals. In particular, we let θ_t variable be any stationary, zero-mean, Gaussian process, admitting a finite-state representation.

Assumption 6 (***Fundamentals***)**.** *The fundamental θ_t admits the following representation:*

$$\theta_t = q' x_t \qquad \text{with} \qquad x_t = R x_{t-1} + \varepsilon_t^x, \tag{58}$$

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.*

Definition 1 and Assumption 2 adapt to this generalization as follows.

Definition 2 (**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*

$$c_t = \sum_{k=0}^{\infty} a_k \eta_{t-k} + \sum_{k=0}^{\infty} \gamma_k' x_{t-k} \tag{59}$$

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.[31]

Assumption 7 (**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 $\lambda \in (0, 1]$.

[31] 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.

With these minor adjustments in place, we can readily extend our main result. As anticipated in the main text, the only subtlety regards the existence and characterization of the MSV solution. Let us explain.

Because equation (8) 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; this restriction indeed defines the MSV solution. 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 (8) 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 $\gamma' = q'(I-\delta R)^{-1}$. We conclude that the following assumption is necessary and sufficient for the existence of the MSV solution:

Assumption 8. *The matrix $I - \delta R$ is invertible.*

We can then reach the following result:

Proposition 9. *Proposition 2 continues to hold, modulo the following adjustment of the MSV solution:*

$$c_t^F \equiv q' \left(I - \delta R\right)^{-1} x_t. \tag{60}$$

Proof. Similarly to Lemma 1, the requirement $Var(c_t) < \infty$ implies a uniform bound on the coefficients a_k and γ_k : 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. We thus focus on candidate solutions that are bounded in this weaker sense.

Since the sunspots $\{\eta_{t-k}\}_{k=0}^{\infty}$ are orthogonal to the fundamental states $\{x_{t-k}\}_{k=0}^{\infty}$, the same argument as that used in Proposition 2 still proves that $a_k = 0$ for all k. We can thus focus on solutions of the following form:

$$c_t = \sum_{k=0}^{\infty} \gamma_k' x_{t-k}. \tag{61}$$

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 (58), 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 with memory length s and

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

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

Next, note that

$$\|W_{k,s}\|_1 \leq \|\mathbb{E}\left[x_t x_t'\right]\|_1 \|\left(R'\right)^{k-s}\|_1 \|\mathbb{E}\left[x_t x_t'\right]^{-1}\|_1, \tag{62}$$

where $\|\cdot\|_1$ is the 1-norm. Since all the eigenvalues of R are within the unit circle, we know its spectral radius is less than one: $\rho(R) = \rho\left(R'\right) < 1$. 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

$$\|W_{k,s}\|_1 \leq M_2 \bar{\Lambda}^{k-s}. \tag{63}$$

Now, from Assumption 7, we know that

$$\bar{E}_t\left[x_{t-k}\right] = (1-\lambda)^k x_{t-k} + \sum_{s=0}^{k-1} \lambda (1-\lambda)^s \mathbb{E}\left[x_{t-k} | I_t^s\right] \equiv \sum_{s=0}^{k} V_{k,s} x_{t-s}, \tag{64}$$

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

$$\|V_{k,s}\|_1 \leq M_3 \Lambda^k. \tag{65}$$

Now consider an equilibrium in the form of (61). From equilibrium condition (5), we know

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

For this to be true for all states of nature, it has to be that the load of x_{t-k} on the left hand side coincides with that on the right hand side, for all $k \geq 0$. That is, the $\{\gamma_k\}_{k=0}^{\infty}$ coefficients must solve the following system:

$$\begin{aligned}
\gamma_0' &= (1-\delta_0)\, q' + \delta_0 \gamma_0' + \delta_1 \gamma_0' R + \delta_1 \gamma_1' \\
\gamma_k' &= \sum_{l=k}^{+\infty} \left(\delta_0 \gamma_l' + \delta_1 \gamma_{l+1}'\right) V_{l,k} \quad \forall k \geq 1.
\end{aligned} \tag{66}$$

From the aforementioned boundedness property, 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. Using this fact along with (65) and (66), we can then infer that, for all $k \geq 1$,

$$\|\gamma'_k\|_1 \leq (|\delta_0|+|\delta_1|) \sum_{l=k}^{+\infty} \|V_{l,k}\|_1 M \leq (|\delta_0|+|\delta_1|)\, M_3 \frac{\Lambda^k}{1-\Lambda} M. \tag{67}$$

Because $\lim_{k\to\infty} \Lambda^k = 0$, there necessarily exists an $\hat{k}$ finite but large enough such that

$$(|\delta_0|+|\delta_1|)\, M_3 \frac{\Lambda^{\hat{k}}}{1-\Lambda} < 1. \tag{68}$$

It follows 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 (66) to provide a tighter bound for $\|\gamma'_k\|_1$: for all $k \geq \hat{k}$,

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

And then we can keep iterating the same argument to get the following: for all $k \geq \hat{k}$ and $l \geq 0$,

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

And since the term in the parenthesis is less than 1, we conclude that any non-zero value for γ'_k can be ruled out, for all $k \geq \hat{k}$. Using (66) and doing backward induction, we conclude that $\gamma'_k = 0$ for all $k \geq 1$.

We are then left with a single equation for γ'_0:

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

Under Assumption 8, the above reduces to $\gamma'_0 = q'(I-\delta R)^{-1}$, which corresponds to the MSV solution. And since we have already proved that $\gamma_k = 0$ for all $k \geq 1$ and $a_k = 0$ for all $k \geq 0$, we conclude that the MSV solution is the unique equilibrium. □

Appendix D: Non-linearities and Multiple Steady States

Here we use an example, suggested by a referee, to clarify that our result speaks only to local determinacy around a given steady state: global indeterminacy may still be possible, at least when non-linearities support multiple steady-state equilibria.

Suppose that an agent's best response is given by

$$c_{i,t} = \delta \mathbb{E}_{i,t}[c_{t+1}] - \omega \mathbb{E}_{i,t}[c_t^3], \tag{69}$$

for some scalars δ, ω. When $\omega = 0$, this reduces back to our baseline, linear model and our main result applies. The point here is to understand what happens when $\omega \neq 0$. Let us focus in particular on how ω matters when $\delta > 1$.

When $\omega \leq 0$, there is a unique steady state and is given by $c_{i,t} = 0$. When instead $\omega > 0$, (69) admits *three* steady states. These are given by

$$c_{i,t} = -\bar{c}, \qquad c_{i,t} = 0, \qquad \text{and} \qquad c_{i,t} = \bar{c},$$

where $\bar{c} \equiv \sqrt{\frac{\delta-1}{\omega}}$. If we linearize (69) around any of these steady states, we can apply our result to the corresponding linearized model. In this sense, our approach guarantees local determinacy around all three steady states and regardless of their eigenvalues. But our approach does not guarantee global determinacy.

This should not be totally surprising. In our baseline model, the unique steady steady, which is given by $c_{i,t} = 0$, serves as an anchor for expectations of future outcomes, in a similar way that the common prior serves as an anchor for higher-order beliefs in the static games of Morris and Shin (1998, 2002). When there are multiple steady states, each one of them can play this kind of anchoring role locally, helping guarantee local determinacy. But our approach is silent about global dynamics, such as jumps from one steady state to another.

To illustrate what we mean, consider the following example, which was proposed by a referee. Suppose there exists a sunspot following a two-state Markov chain with values $\eta_t \in \{-1, +1\}$ and transition probability π. Suppose next that all agents coordinate on playing the following strategy, which requires knowledge only of the concurrent sunspot realization:

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

for some $a \neq 0$. This means, more simply, that all agents coordinate on playing the same action, and that this action follows a two-state Markov chain with values $c_{i,t} \in \{-a, +a\}$ and transition probability π.

It is straightforward to check that this strategy constitutes an equilibrium if and only if $a = \sqrt{\frac{\delta(2\pi-1)-1}{\omega}}$, which in turn is well defined if and only if $\pi \in \left(\frac{1+\delta^{-1}}{2}, 1\right)$ Also, as $\pi \to 1$, we have that $a \to \bar{c}$, that is, this type of equilibrium translates to infrequent jumps across the two outer steady states. Finally, this type equilibrium is robust to imperfect knowledge of the distant past in the following sense: it suffices to have common knowledge of the current realization of the sunspot (which itself is persistent as long as $\pi \neq \frac{1}{2}$) and of the parameters π, a, and δ.

It is important to recognize that the equilibrium constructed above is *not* memoryless: the restriction $\pi > \frac{1+\delta^{-1}}{2}$ implies $\pi > \frac{1}{2}$, which means that the sunspot itself *has* to be persistent. This example therefore links to our discussion of persistent sunspots discussed in Section 5. But there is a key difference: whereas there was a unique value for the persistence parameter ρ that supported multiplicity in our linear setting, now there is a whole range of values for the corresponding parameter π that supports multiplicity in the present example.

Does this upset our main message? Not necessarily. First of all, we have been upfront that our paper is ultimately only about local determinacy, and from this perspective our result is still valid: if we linearize the present example around any of the three steady states, we still have local determinacy. Second, and related, the above example is not a "perturbation" of our original setting: for ω positive but small enough, the outer two steady states diverge to plus/minus infinity, and so do the values of c_t in the equilibrium constructed above. Last but not least, the above equilibrium still assumes a significant degree of dynamic coordination: to jump from one steady state to another, or more precisely between the two points of the Markov chain, agents must be confident not only that other agents will do the same today but also that future generations will stay at the new point with sufficient probability.

This begs the question of how sensitive the type of equilibrium constructed above is to perturbations of intertemporal common knowledge, albeit of a different from that those considered in this paper. But our methods are not equipped to answer this question. At the end of the day, we thus prefer to iterate our "real" take-home lesson: our contribution is not to argue that all kinds of dynamic indeterminacy are gone, but rather to shed new light on the (local) determinacy problem of the New Keynesian model, to provide a formal justification for treating this problem as bug, and to set the foundations for re-thinking both the Taylor principle and the FTPL.

References

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

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*.

ANGELETOS, G.-M. AND I. WERNING (2006): "Crises and prices: Information aggregation, multiplicity, and volatility," *american economic review*, 96, 1720–1736.

ATKESON, A. (2000): "Discussion of Morris and Shin's 'Rethinking multiple equilibria in macroeconomic modelling'," *NBER Macroeconomics Annual*, 15, 162–171.

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," .

BARRO, R. J. (1974): "Are government bonds net wealth?" *Journal of political economy*, 82, 1095–1117.

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

——— (2005): "Equilibrium and government commitment," *Journal of Economic Theory*, 124, 79–105.

——— (2008): "Fiscal theory of the price level," *The New Palgrave Dictionary of Economics*.

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

BENHABIB, J., P. WANG, AND Y. WEN (2015): "Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83, 549–585.

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.

BIANCHI, F., R. FACCINI, AND L. MELOSI (2020): "Monetary and fiscal policies in times of large debt: Unity is strength," *National Bureau of Economic Research Working Paper*.

BIANCHI, F. AND C. ILUT (2017): "Monetary/Fiscal Policy Mix and Agent's Beliefs," *Review of Economic Dynamics*, 26, 113–139.

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.

BLANCHARD, O. J. AND C. M. KAHN (1980): "The solution of linear difference models under rational expectations," *Econometrica*, 1305–1311.

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.

BULLARD, J. AND K. MITRA (2002): "Learning about monetary policy rules," *Journal of monetary economics*, 49, 1105–1129.

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 polítical Economy*, 102, 617–652.

CHEN, X., E. LEEPER, AND C. LEITH (2021): "Strategic Interactions in U.S. Monetary and Fiscal Policies," *Quantitative Economics*.

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.

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.

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.

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.

——— (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.

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.

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.

RUBINSTEIN, A. (1989): “The Electronic Mail Game: Strategic Behavior Under" Almost Common Knowledge",” *The American Economic Review*, 385–391.

SAMUELSON, P. A. (1958): “An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money,” *Journal of Political Economy*, 66, 467–482.

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.

——— (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

——— (2022): "Optimal fiscal and monetary policy with distortionary taxation," *Princeton University mimeo*.

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.

WALLACE, N. (1981): "A hybrid fiat–Commodity monetary system," *Journal of Economic Theory*, 25, 421–430.

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*.

——— (2011): "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 3, 1–35.

——— (2019a): "Modeling Imprecision in Perception, Valuation and Choice," *NBER Working Paper No. 26258*.

——— (2019b): "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*.