Determinacy without the Taylor Principle*

George-Marios Angeletos†  Chen Lian‡

March 21, 2022

Abstract

Our understanding of monetary policy is complicated by an equilibrium-selection conun-
drum: 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.

*We are grateful to the Editor and three anonymous referees for their extensive feedback and to Francesco Bianchi for his insightful discussion of our paper. For helpful comments and suggestions, we also thank 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, Pontifical Catholic University of Chile, Tsinghua University, UC Berkeley, USC, Warwick, Wharton, Yale, the Minnesota Conference in Macroeconomic Theory, the NBER Summer Institute (EFCE), the NBER EFG meeting, SED, 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 respond “yes” to all these questions. But the correct answer, at least within the dominant policy paradigm (the New Keynesian model), crucially 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 output. The standard approach selects a specific equilibrium by assuming that monetary policy satisfies the Taylor principle (Taylor, 1993), or equivalently that it is “active” in the sense of Leeper (1991). It is this selection that drives the model’s customary predictions, including “yes” to the aforementioned questions. But as stressed by Cochrane (2017, 2018), an alternative selection, based on the Fiscal Theory of the Price Level (FTPL), can lead to sharply different predictions. This approach elevates government debt and deficits to key drivers of inflation and output, even when these variables do not enter the model’s three “famous” equations.

Both approaches are equally coherent, at least in the sense of being consistent with rational expectations and the same micro-foundations. They are also hard to test, because they translate to different assumptions about off-equilibrium strategies of the monetary and fiscal authorities. As a result, the debate about which approach is “right” has never been settled.

We shed new light on this methodological issue, and offer a possible way out of the prevailing conundrum, by demonstrating how the indeterminacy problem of the New Keynesian model hinges on delicate assumptions about social memory and intertemporal coordination. Once we perturb these assumptions, appropriately but tinily, the model’s conventional solution 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.

The economy as a game. A crucial stepping stone of our analysis is the translation of a New Keynesian economy to a game among the consumers. The details are spelled out in Section 2

---

1There is, however, an important difference between the New Keynesian framework and flexible-price models, such as that of Sargent and Wallace (1975), which we explain in due course.


but basic idea is that optimal spending depends on expectations of others’ spending via three GE feedbacks: that from aggregate spending to income (the Keynesian cross); that from aggregate spending to inflation (the Phillips curve); and the response of monetary policy (the Taylor rule).

The last two feedbacks are shut off with, respectively, rigid prices and interest rate pegs. But the first one is always on, whether implicit behind the textbook model’s Euler equation or explicit in the “intertemporal Keynesian cross” (Auclert et al., 2018); and it directly translates to strategic complementarity among different generations of consumers. This explains both why our game representation extends to a wide class of Keynesian economies and why the determinacy question we are after is deeply connected to intertemporal coordination.

**Preview of results.** Under the above prism, the slope of the Taylor rule regulates the dynamic strategic complementarity among the consumers, and the Taylor principle translates to the following requirement: let this complementarity be small enough to guarantee a unique equilibrium when consumers can perfectly coordinate their behavior over time. Departing from this benchmark, we accommodate a small but appropriate friction in such coordination and proceed to show how this can guarantee determinacy even when the Taylor principle is violated.

Our main result, developed in Section 4, models the friction as follows. There are overlapping generations of finitely-lived consumers. Consumers know the shocks (payoff-relevant or not) realized during the lifetime, and may also inherit knowledge of past shocks from previous generations. But the transmission of knowledge need not be imperfect: for any \(t\), the fraction of the population who know, and can directly or indirectly condition their actions on, shocks realized at any \(\tau \leq t\) is \((1 - \lambda)^{t-\tau}\), where \(\lambda \in [0, 1)\) parameterizes the erosion of social memory over time.

The frictionless, representative-agent case is nested with \(\lambda = 0\); it guarantees that any shock remains common knowledge in perpetuity; 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 \(\lambda > 0\). Regardless of monetary policy, the only surviving equilibrium is the conventional one, known as the fundamental or minimum-state variable (MSV) solution.

Our main result abstracts from direct observation of past outcomes such as inflation and output, which could themselves play the role of coordination devices. But because such outcomes are functions of the underlying shocks, in the limit as \(\lambda \to 0\) consumers face vanishing uncertainty about them, suggesting that our conclusions do not necessarily rest on the aforementioned abstraction. We corroborate this intuition 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—but the end result is the same.
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 (Atkeson, 2000; Angeletos and Werning, 2006). Note, however, that our context’s multiplicity is sustained by a self-fulfilling infinite chain over different generations of players: today’s consumers are responding to a payoff-irrelevant variable 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 selecting the fundamental solution; and (iii) as an invitation to reconsider the applied 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 result removes the need for equilibrium selection but leaves room for sunspot-like fluctuations along the MSV solution in at least the following two forms: overreaction to noisy public news (Morris and Shin, 2002); and shocks to higher-order beliefs (Angeletos and La’O, 2013; Benhabib et al., 2015). This in turn lets the slope of the Taylor rule to play a new function: to regulate the macroeconomic complementarity and thereby the aforementioned kind of sunspot-like volatility. 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. This is the exact opposite of the prevailing theoretical approach but closer to how Taylor (1993) had originally thought about it.

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. More succinctly, fiscal policy has to be Ricardian even when monetary policy is passive.

Like our main result, this lesson is subject to an obvious but important qualification: it is debatable whether the real world is better approximated by the full-information benchmark or our specific perturbations. Still, our analysis illustrates a certain fragility in the existing formulation of the FTPL. This does not negate its basic spirit—namely the idea that debt and deficits may drive inflation—but calls for its reformulation outside the equilibrium selection conundrum.
Unbounded equilibria and flexible prices. Two additional clarifications are due. First, like most related literature, we work with the linearized New Keynesian model and restrict attention to bounded equilibria, which amounts to studying local determinacy around a given steady state. A separate question, outside our scope, is the design and credibility of the kind of “escape clauses” that help rule out self-fulfilling hyper-inflations and self-fulfilling liquidity traps.4

Second, our results are not sensitive to the degree of nominal rigidity, as long as there is some of it. But if prices are perfectly flexible, output and inflation are no longer demand determined, the economy can no longer be understood as a coordination game among the consumers, and our methods do not directly apply. This touches on a larger methodological question, whether flexible-price models are proper limits of sticky-price models (Kocherlakota, 2020). And it highlights that, contrary to conventional wisdom, the indeterminacy problem of the New Keynesian model is not the same as that of flexible-price models.

Related literature. Kocherlakota and Phelan (1999), Buiter (2002), Niepelt (2004) and others have interpreted the non-Ricardian assumption as an off-equilibrium threat to blow up the government budget. Cochrane (2005, 2011) has fired back by arguing not only that this interpretation is misguided but also that the Taylor principle itself amounts to a threat to blow up inflation and interest rates. While these arguments emphasize the subtlety of both approaches, they do not help resolve the conundrum: Bassetto (2002, 2005) and Atkeson, Chari, and Kehoe (2010) have shown that both approaches can be supported with more sophisticated policies, which avoid such incredible threats and guarantee a proper continuation equilibrium always. By contrast, our paper seeks to remove the need for equilibrium selection of either kind.5

Our main result, Proposition 2, reminds Rubinstein (1989) and the global-games literature (Morris and Shin, 1998, 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.


5 Adding to Cochrane (2005, 2011)’s critique, Neumeyer and Nicolini (2022) argue that the Taylor principle is time-inconsistent. While we abstract from the time-inconsistency question, our results suggest that this question should perhaps be redirected to the exit strategies that rule out “unbounded” equilibria (i.e., hyper-inflations and self-fulfilling liquidity traps). For it is only the latter kind of policy commitments, and not the Taylor principle, that matter for determinacy under our perturbations.
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).

A different literature has studied which of the model’s solutions are “learnable” in the sense of E-stability (McCallum, 2007; Christiano et al., 2018). This approach has produced mixed results. Still, this approach and ours shed complementary light on which solution seems most sensible.

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 them. This class includes cognitive discounting (Gabaix, 2020) and diagnostic expectations (Bordalo et al., 2018), but not Level-K Thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019). The latter produces a unique solution because it rules out entirely any feedback from objective reality to subjective beliefs. This may be reasonable in the context of unprecedented experiences but seems less appropriate in the context of stationary environments, which is our focus here.

2 A Simplified New Keynesian Model

Here we introduce our version of the New Keynesian model. As in the textbook version, output is demand-determined; there is no money, taxation, or government debt (these will be added in Section 6); and the monetary authority directly controls the nominal interest rate. Unlike the textbook version, the demand side consists of overlapping generations of consumers and the supply side is represented by an ad hoc Phillips curve. These twists, and a few other auxiliary assumptions that we make clear in due course, ease the exposition but 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. There is a measure one of consumers born at $t$, who has preferences

$$u(C_{i,t}^1) + \beta u(C_{i,t+1}^2)e^{-\rho t},$$

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).
where \( C_1^{i,t} \) and \( C_2^{i,t+1} \) are consumption when young and old, respectively, \( u(C) = \frac{1}{1-1/\sigma}C^{1-1/\sigma} \), \( \beta \in (0, 1) \) is a fixed scalar, and \( \rho_t \) is an intertemporal preference shock (the usual proxy for aggregate demand shocks). For simplicity, young and old consumers receive the same income and face the same interest rate and the same prices. The budget constraints in the first and second period of life of the consumer are therefore given by, respectively,

\[
C_1^{i,t} + B_{i,t} = Y_t \quad \text{and} \quad C_2^{i,t+1} = Y_{t+1} + \frac{I_t}{\Pi_{t+1}}B_{i,t},
\]

where \( Y_t \) is total income in period \( t \), \( B_{i,t} \) is the consumer’s saving/borrowing at \( t \) (when young), \( I_t \) is the nominal interest rate between \( t \) and \( t+1 \), and \( \Pi_{t+1} \) is the corresponding inflation rate. \(^7\)

Old consumers are “robots:” their consumption mechanically adjusts to meet the second-period budget. Young consumers are “strategic:” they optimally choose consumption and saving/borrowing. But they may have to do so subject to an informational friction.

The precise specification of this friction is the key to our results. For the time being, however, we take no stand on what the young consumers may or may not know when choosing their spending. We only require that this choice be optimal given their possibly arbitrary information. After the usual log-linearization,\(^8\) this translates to the following optimal consumption function:

\[
c_1^{i,t} = E_{i,t} \left[ \frac{1}{1+\beta}Y_t + \frac{\beta}{1+\beta}Y_{t+1} - \frac{\beta}{1+\beta}\sigma(i_t - \pi_{t+1} - \rho_t) \right], \quad (2)
\]

where \( E_{i,t} \) denotes the rational expectation conditional on the information of the young consumer, whatever that might happen to be.

Pick any \( t \). Since the old do not borrow or save (because they are about to die), and since there is no government borrowing or saving either, the average private saving of the young has to be zero in equilibrium, or equivalently \( \int c_1^{i,t} \, di = y_t \). But the average net wealth of the old has to be zero as well, or equivalently \( \int c_2^{i,t} \, di = y_t \). 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 (2), 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:

\[
c_t = \bar{E}_t \left[ \frac{1}{1+\beta}C_t + \frac{\beta}{1+\beta}C_{t+1} - \frac{\beta}{1+\beta}\sigma(i_t - \pi_{t+1} - \rho_t) \right], \quad (3)
\]

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

As evident from its derivation, equation (3) makes no assumption about monetary policy or the supply side. It only combines consumer optimality with market clearing (or, more precisely,

\(^7\)To ease the exposition, we side-step labor supply. The missing details are filled in Appendix B.2 but the basic point is this: because output is demand-determined, the specification of labor supply is inconsequential.

\(^8\)Throughout, we log-linearize around the steady state in which \( \rho_t = 0, \Pi_t = 1 \), and \( I_t = \beta^{-1} \); and we use lower-case variables to denote log-deviations from steady state.
with consumers’ understanding of market clearing). In so doing, this equation encapsulates a positive feedback between income and spending and can be read interchangeably as a special case of the “intertemporal Keynesian cross” (Auclert et al., 2018) or as a Dynamic IS equation.

**Connection to standard New Keynesian model**

Although 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 $E_t$, which henceforth denotes the rational expectation conditional on full information about $h_t$, the history of all exogenous shocks up to and inclusive of period $t$. Along with the fact that $c_t$ and $i_t$ must themselves be measurable in $h_t$, this means that in this case equation (3) reduces to

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

or equivalently

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

which is evidently the same as the Euler condition of a representative, infinitely-lived consumer.

This clarifies the dual role of the adopted OLG structure. With full information, it lets our model translate to the standard, representative-agent, New Keynesian model. And away from this benchmark, it eases 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:

$$\pi_t = \kappa(y_t + \xi_t),$$

where $\kappa \geq 0$ is a fixed scalar and $\xi_t$ is a “supply” or “cost-push” shock. The absence of a forward-looking term in (4) simplifies the exposition significantly, but does not drive the results: as shown in Section 6, our arguments directly extend to the fully micro-founded, forward-looking, New Keynesian Phillips curve. With either version of the Phillips curve, the essence (for our purposes) is that there is a positive GE feedback from aggregate output to inflation. Equation (4) merely stylizes this feedback in a convenient form.

We finally assume that monetary policy follows a Taylor rule of the following type:

$$i_t = z_t + \phi \pi_t,$$
where $z_t$ is a random variable and $\phi \geq 0$ is a fixed scalar. This readily nests $i_t = i^*_t + \phi(\pi_t - \pi^*_t) + \zeta_t$, where $i^*_t$ and $\pi^*_t$ are state-contingent “targets” and $\zeta_t$ is a pure monetary shock. Also, no restriction is imposed on how $z_t$ covaries with $\varrho_t$ and $\xi_t$; for instance, $z_t$ may track the natural rate of interest or lean against cost-push shocks. In the standard paradigm, this helps disentangle the stabilization and equilibrium selection functions of monetary policy: the former is served by the design of $z_t$, the latter by the restriction $\phi > 1$. \(^9\) Our perturbations will dispense with the latter function and guarantee determinacy even under interest-rate pegs (herein nested by $\phi = 0$).

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

From (4) and (5), we can readily solve for $\pi_t$ and $i_t$ as simple functions of $y_t$, which itself equals $c_t$. Replacing into (3), we conclude that the model reduces to the following single equation:

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

where $\delta_0, \delta_1$ are fixed scalars and $\theta_t$ is a random variable. \(^10\) These are given by

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

By construction, equation (6) 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 $\delta_0$ or different stochastic processes for $\theta_t$, via the choice of, respectively, a value for $\phi$ or a stochastic process for $z_t$. But for any given monetary policy, we can understand equilibrium in the private sector by studying equation (6) alone.

Equation (6) 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

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

and recasts (6) as the period-$t$ average best response function. Under this prism, $\delta_0$ and $\delta_1$ pa-

\(^9\)See King (2000) and Atkeson et al. (2010) for sharp articulations of this point. Also note that we are restricting $\phi \geq 0$. Letting $\phi < 0$ qualifies the Taylor principle (see footnote 12) but does not upset our own result. Finally, note that (5) has the monetary authority respond to current inflation. But as explained in Appendix B.5, our insights go through if the monetary authority responds to past inflation and/or expected future inflation.

\(^10\)For the time being, we take no stand how much is known about $\theta_t$ or its components, which is why $\theta_t$ appears inside the expectation operator in (6). Also, the fact that $\theta_t$ is multiplied by $1 - \delta_0$ is just a normalization.
rameterize, respectively, the intra-temporal and the inter-temporal degree of strategic complementarity, while $\theta_t$ 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 $\beta$, $\kappa$, and $\phi$ map to different degrees of strategic complementarity.

This game-theoretic prism is not strictly needed for proving our main result. But it helps translate the determinacy question from one about eigenvalues (Blanchard and Kahn, 1980) to one about intertemporal coordination.

**Fundamentals, sunspots, and the equilibrium concept**

Aggregate uncertainty is of two sources: fundamentals and sunspots. The former are herein conveniently summarized in $\theta_t$. The latter are represented by a random variable $\eta_t$ that is independent of the current, past, and future values of $\theta_t$. 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 $\theta_t$ and the sunspot $\eta_t$ 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 (6) along which: expectations are rational, although potentially based on imperfect information about $h^t$; and the outcome is given by

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

(9)

where $\{a_k, \gamma_k\}$ are known and uniformly bounded coefficients.

Recall that consumer optimality, firm behavior, and market clearing have already been embedded in equation (6). It follows that the above definition is the standard definition of a Rational Expectations Equilibrium (REE), except for the addition of three “auxiliary” restrictions embedded in (9): stationarity, linearity, and boundedness. The first two restrictions are technical. The third is of substance but has the usual interpretation: the exclusion of unbounded equilibria can be justified by appropriate “exit” strategies, namely a commitment to switch from the Taylor rule to money-growth targeting, fiat money, or whatever else it takes for keeping inflation within some bounds.11

11See the references in footnote 4. Also, consistent with the literature, we are imposing boundedness directly in the
Finally, and circling back to our game-theoretic prism, note that the following is true: because every consumer is infinitesimal, there is no need to specify off-equilibrium beliefs, and the economy’s REE 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.

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 (6), which summarizes equilibrium, reduces to the following:

$$c_t = \theta_t + \delta E_t[c_{t+1}],$$

where $E_t[\cdot] \equiv E[\cdot|h_t]$ 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 $\delta$ is necessarily positive but can be on either side of 1, depending on $\phi$.

Because equation (10) is purely forward looking and $\theta_t$ is i.i.d., $c_t = c_t^{E} \equiv \theta_t$ 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 determi-
nacy (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.\footnote{By restricting $\phi \geq 0$, we have restricted $\delta > 0$. If we allow $\delta < 0$, which is possible for $\phi$ sufficiently negative, Proposition 1 and the discussion after it continue to hold, provided that we recast the Taylor principle as $\delta \in (-1, 1)$, or equivalently as $\phi \in (-\infty, \phi) \cup (1, +\infty)$, where $\phi \equiv 1 - \frac{\theta}{\kappa \sigma} < -1$. This echoes Kerr and King (1996). More importantly, our own uniqueness result does not hinge on $\delta > 0$.}

**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 $c^F_t$.

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

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

$$c_t = (1 - b)c^F_t + bc^B_t + ac^H_t,$$

where $a, b \in \mathbb{R}$ are arbitrary scalars and $c^B_t, c^H_t$ are given by

$$c^B_t \equiv -\sum_{k=1}^{\infty} \delta^{-k} \theta_{t-k} \quad \text{and} \quad c^H_t \equiv \sum_{k=0}^{\infty} \delta^{-k} \eta_{t-k}.$$

To understand the type of non-fundamental equilibria documented in part (iii) above, take
equation (10), backshift it by one period, and rewrite it as follows:

$$E_{t-1}[c_t] = \delta^{-1}(c_{t-1} - \theta_{t-1}).$$

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

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

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

$$c_t = -\sum_{k=1}^{\infty} \delta^{-k} \theta_{t-k} + a \sum_{k=0}^{\infty} \delta^{-k} \eta_{t-k} = c^B_t + ac^H_t.$$

This is both bounded, thanks to $\delta > 1$, and a rational-expectations solution to (13), by con-
struction, which verifies that $c^B_t + ac^H_t$ constitutes an equilibrium, for any $a \in \mathbb{R}$. Part (iii) of the Propo-
sition adds that the same is true if we replace $c^B_t$ with any mixture of it and the MSV solution.

To illustrate what all these equilibria are, switch off momentarily the fundamental shocks.
Then, $c^F_t = c^B_t = 0$ and (11) reduces to $c_t = ac^H_t$, 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-fulfilling 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 (11) 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 (14) and (15). Clearly, these are equivalent representations of the same equilibrium: the first is recursive, the second is sequential. This equivalence means that all the equilibria that can be supported by perfect knowledge of $h_t = [\theta_{t-k}, \eta_{t-k}]_{k=0}^{\infty}$ 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 (9). 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 (15), or the sequential representation. An alternative, in Section 5, circles back to (14), 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

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

where $s_{i,t} \in \{0, 1, \cdots\}$ is an idiosyncratic random variable, drawn from a geometric distribution with parameter $\lambda$, for some $\lambda \in (0, 1]$.

To understand this assumption, note that herein $s_{i,t}$ indexes the random length of the history of shocks that a period-$t$ agent knows. Next, recall that the geometric distribution means that $s_{i,t} = 0$ with probability $\lambda$, $s_{i,t} = 1$ with probability $(1 - \lambda)\lambda$, and more generally $s_{i,t} = k$ 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 for sure; with probability $\lambda$, 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, $\lambda$ parameterizes the speed at which social memory (or common-p belief of past shocks) fades over time.

Finally, note that Assumption 2 rules out direct observation of endogenous outcomes, including current income and current interest rates. This is consistent with our characterization of optimal consumption in (2) and by extension with our game representation in (6), because both of them are valid for arbitrary information. But it also means that we must envision consumers choosing their spending under uncertainty about current income and current interest rates.

Such uncertainty can be motivated on its own right as the product of inattention, but is not strictly needed for our results. First, this uncertainty vanishes as $\lambda \to 0^+$, in a sense we qualify in Appendix B.4. Second, our analysis goes through if consumers observe perfectly their own
income and own interest rate, provided that we abstract from signal-extraction about payoff-irrelevant histories; see Appendix B.1. Finally, we can accommodate such signal-extraction if we adopt the variant perturbation of Section 5. We thus invite the reader to take Assumption 2 with an open mind: even though it may not the most realistic specification of information one can think of, it allows us to introduce a plausible perturbation away from common knowledge.

Main result

The full-information benchmark is nested with $\lambda = 0$; this indeed translates to $I_{i,t} = h_t$ for all $i, t,$ and $h_t$ (i.e., perfect and common knowledge of the infinite history at all times). The question of interest is what happens for $\lambda > 0$, and in particular as $\lambda \to 0^+$. In this limit, the friction vanishes in the following sense: almost every agent knows the history of shocks up to an arbitrarily distant point in the past. But the following is also true: as long as $\lambda$ is not exactly zero, we have that $\lim_{k \to \infty} \mu_k = 0$, which means that shocks are expected to be “forgotten” in the very distant future. As shown next, this causes all non-fundamental equilibria to unravel.

Proposition 2 (Determinacy without the Taylor principle). Suppose that social memory is imperfect in the sense of Assumption 2, for any $\lambda > 0$. Regardless of $\phi$, or of $\delta_0$ and $\delta_1$, the equilibrium is unique and is given by the fundamental/MSV solution.

The result is proven in Appendix A for arbitrary $\delta_0$ and $\delta_1$. 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 (6) to

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

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

\[\text{Note that the MSV solution itself is invariant to } \lambda \text{ thanks to the assumption that } I_{i,t} \text{ always contains } \theta_t \text{ (and this fact is common knowledge). As mentioned in the Introduction, this helps isolate our contribution from the existing literature on informational frictions, which allows imperfect information about } \theta_t \text{ but does not address the determinacy issue. Needless to say, one can have it both ways: if we modify Assumption } 2 \text{ to accommodate imperfect information about } \theta_t, \text{ we can improve the empirical properties of the MSV solution while also preserving our argument that this solution is the unique equilibrium regardless of monetary policy.}\]
nobody. It follows that, along any candidate solution, average expectations satisfy

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

By the same token, condition (16) rewrites as

$$\sum_{k=0}^{+\infty} a_k \eta_{t-k} = \delta \sum_{k=0}^{+\infty} a_{k+1} \mu_k \eta_{t-k},$$

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

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

or equivalently

$$a_{k+1} = \frac{a_k}{\delta \mu_k}.$$  \hspace{1cm} (18)

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 $a_k = 0 \forall k$, which proves that all sunspot equilibria are ruled out and only the MSV solution survives.\(^{14}\)

**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 (18) becomes

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

When $\delta < 1$ (equivalently $\phi > 1$), this still explodes as $k \to \infty$ unless $a_0 = 0$ and hence also $a_k = 0 \forall k$. 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 explains how $\lambda = 0$ 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.

\(^{14}\)Note that this argument does not depend on the sign of $\delta$, which basically verifies that our result extends to $\phi < 0.$
Intuition and the role of higher-order beliefs

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 (17) 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 $\eta_0$ in a “twin” representative-agent, full-information economy, in which condition (6) is modified as follows:

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

Under this prism, it is as if we are back to the standard New Keynesian model but the relevant eigenvalue, or the dynamic macroeconomic complementarity, has become time-varying and has been reduced from $\delta$ 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 $\delta$. In other words, the twin economy’s dynamic feedback becomes weak enough that $c_t$ cannot depend on $\eta_0$ after $T$. By backward induction, then, $c_t$ cannot depend on $\eta_0$ before $T$ either.\(^{15}\)

This interpretation of our result must be clarified as follows. Here we focused on the response of $c_t$ to $\eta_0$. 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 $\eta_0$. 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 $\eta_0$ 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 $\eta_0$. 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 $\eta_0$, or because they may themselves worry, like me, that agents further into the future will not react to it. This makes it iteratively optimal not to react to $\eta_0$.”

\(^{15}\)Although this argument assumed $\delta_0 = 0$, it readily extends to $\delta_0 \neq 0$. In this case, the twin economy has both $\delta_0$ and $\delta_1$ replaced by, respectively, $\mu_t \delta_0$ and $\mu_t \delta_1$. That is, both types of strategic complementarity are attenuated.
Three 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.

Last but not least, our argument, like the related arguments in the global games literature, relies on rational expectations (or more precisely on common knowledge of rationality, which itself is implied by REE). This cuts both ways. On the one hand, it lets our paper speak directly and precisely to the question of interest, namely the determinacy of rational expectations equilibria. On the other hand, it begs the question of how monetary policy should be designed if bounded rationality is itself the source of non-fundamental volatility. While this question is outside the scope of our paper, we touch again on it in Section 6.

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 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. We finally comment on two other subtleties: the distinction between local and global determinacy; and the role of nominal rigidity. Readers interested in our paper’s take-home lessons may skip this section and jump to Section 6.

Persistent fundamentals

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

To illustrate, suppose that $\theta_t$ follows an AR(1) process: $\theta_t = \rho \theta_{t-1} + \epsilon_t$, where $\rho \in (-1, 1)$ is a fixed scalar and $\epsilon_t \sim \mathcal{N}(0, 1)$ 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 (10), 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 \( \theta_t \) switch sign depending on whether \( \rho \) is lower or higher than \( \delta^{-1} \). 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 \( \eta_t \) follows an AR(1) process with autocorrelation exactly equal to \( \delta^{-1} \). 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 \( \theta_t \) and \( \eta_t \) 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: formally, the requisite sunspot is degenerate in the space of ARMA processes. 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 variable? We already hinted that such coordination, too, can be fragile: in the limit as \( \lambda \to 0^+ \), agents were arbitrarily well informed about exogenous shocks and endogenous outcomes alike, and yet uniqueness obtained. We now reinforce this message by showing how determinacy may remain with two variant information structures, which, unlike Assumption 2, allow for direct signals of endogenous outcomes.
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:

\[ c_t = a \sum_{k=0}^{\infty} \delta^k \eta_{t-k}, \]  

(19)

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

\[ c_t = a \eta_t + \delta^{-1} c_{t-1}. \]  

(20)

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

\[ I_{i,t} = \{\eta_t, s_{i,t}\}, \]  

with \( s_{i,t} = c_{t-1} + \varepsilon_{i,t} \).

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 \( \delta_0 \) and \( \delta_1 \), 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 Appendix A, 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, 2019). It is only in the knife-edge case in which the statistic is common knowledge that multiplicity survives.\(^{16}\)

\(^{16}\)See Section D. We thank a referee from prompting us to clarify this subtlety.
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 \(I_i \supseteq \{\eta_t, c_{t-1}, \theta_{t-1}\}\) and by having each consumer play the following recursive strategy:

\[
c_{i,t} = \delta^{-1}(c_{t-1} - \theta_{t-1}) + a\eta_t. \tag{21}
\]

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 \(\theta_{t-1}\). Why is knowledge of \(\theta_{t-1}\) necessary? Because this is what it takes for agents at \(t\) to know how to undo the direct, intrinsic effect of \(\theta_{t-1}\) 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 \(\zeta_t\); we modify equation (6) to

\[
c_{i,t} = E_{i,t}[1/(1-\delta_0)(\theta_t + \zeta_t) + \delta_0 c_t + \delta_1 c_{t+1}]; \tag{22}
\]

and we let \(\zeta_t\) be drawn independently over time, as well as independently of any other shock in the economy, from a uniform distribution with support \([-\epsilon, +\epsilon]\), where \(\epsilon\) is positive but arbitrarily small. This lets us parameterize the payoff perturbation by \(\epsilon\), or the size of the support of \(\zeta_t\).

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 \(c_{i,t} = c_t\) 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 \(\theta_t\) and \(\zeta_t\), we can then write the best response of the period-\(t\) representative agent as

\[
c_t = \theta_t + \zeta_t + \delta E[c_{t+1}|I_t]. \tag{23}
\]

where \(\delta \equiv \delta^{-1}_1(1-\delta_0)\), 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

\[
I_t = \{\zeta_t\} \cup \{\theta_t, \cdots, \theta_{t-K}\} \cup \{\eta_t, \cdots, \eta_{t-K}\} \cup \{c_{t-1}, \cdots, c_{t-K_c}\}
\]

for finite but possibly arbitrarily large \( K_\eta, K_c, \) and \( K_\theta \).

When the \( \zeta_t \) shock is absent, or \( \varepsilon = 0 \), Assumption 3 allows replication of all sunspot and backward-looking equilibria with 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 \( \delta \), there is unique equilibrium and is given by \( c_t = c^F_t + \zeta_t \), where \( c^F_t \) is the same MSV solution as before.

How does this connect to Proposition 2? Both results introduce a friction in social memory and intertemporal coordination, thus breaking the infinite chain behind all non-fundamental equilibria. But the exact friction is different: whereas in our main result it amounts to asymptotic forgetting of the distant past, here it amounts to immediate forgetting of a small component of the fundamentals. This also means a change in the formal argument: whereas our main result echoes the global games literature, the present one is more closely connected to Bhaskar (1998) and Bhaskar et al. (2012), which show how the combination of a payoff perturbation and finite social memory can rule out non-Markov perfect equilibria in a certain class of dynamic games. At a high level, related is also a literature that studies how multiplicity in repeated games depends on public versus private monitoring (e.g., Mailath and Morris, 2002; Pęski, 2012). The common thread is the role played by lack of common knowledge. But the precise connections between our two results and these literatures are elusive and deserve further study.

**Local vs global determinacy**

Throughout, we work with the linearized New Keynesian model and restrict 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 complements 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.\textsuperscript{17} 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 \textit{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 first
steady state. In this context, we cannot rule out either one of the steady-state equilibria, because
our methods presume common knowledge of any given steady state. By extension, we cannot
say anything about global determinacy either. 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.\textsuperscript{18}

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), provided that these can be recast as dynamic
coordination games along the lines we have illustrated here. 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 specific way explained
above.

\textbf{Sticky vs flexible prices}

Equation (8), the game representation of our baseline model, is valid for any value of \( \kappa \), the slope
of the Phillips curve. The same is true for equation (27), the generalization developed in this
section. This underscores that our game-theoretic prism and, by extension, our main result is
not unduly sensitive to the degree of price flexibility. But what if prices are \textit{literally} flexible, or
\textsuperscript{17}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).
\textsuperscript{18}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 B.6 to further illustrate the issues discussed above.
“κ = ∞”? In this case, aggregate demand ceases to matter for aggregate output and, as a result, the economy can no more be represented as a game among the consumers.

This begs the question of whether a version of insights applies to flexible-price models. While we will not address this question here, we wish to raise the following flag. In the existing literature, the real indeterminacy problem of the New Keynesian paradigm is treated as a direct translation of the nominal indeterminacy problem of flexible-price models. But our results suggest that the two problems may be fundamentally different: with any non-zero degree of nominal rigidity, output and inflation can be understood as the outcomes of a game among the consumers and our results go through; but this game ceases to be well defined when prices are “truly” flexible.

This touches on two delicate, larger methodological questions: whether flexible-price models are proper limits of models with nominal rigidity (Kocherlakota, 2020); and, conversely, whether the New Keynesian model itself is the “right” model for the questions of interest. These questions are beyond the scope of our paper. But the following discussion helps illustrate in simple terms (i.e. without our game-theoretic prism) why treating the aforementioned two determinacy problems as equal might be misleading.

Forget our game representation and our perturbations; go back to the textbook New Keynesian model; and let prices be entirely rigid (κ = 0). In this case, equilibrium boils down to the following joint restriction on aggregate output and the nominal interest rate:

\[ y_t = -\sigma i_t + E_t[y_{t+1}] \]

This is the model’s DIS equation, specialized in the case of rigid prices (⇒ E_t[π_{t+1}] = 0). The following properties are then evident: an interest rate peg produces indeterminacy; and the feedback rule \( i_t = \phi_y y_t \) induces a unique bounded equilibrium if and only if \( \phi_y > 0 \) or \( \phi_y \sigma < -2 \).

Next, consider the opposite extreme: an endowment economy with flexible prices, in the tradition of Sargent and Wallace (1981). In this case, equilibrium boils down to the following joint restriction on interest rates and inflation, which is simply the Fisher equation:

\[ i_t - E_t[\pi_{t+1}] = 0. \]

Once again, an interest rate peg produces indeterminacy; and the feedback rule \( i_t = \phi_\pi \pi_t \) induces a unique bounded equilibrium if and only if \( \phi_\pi > 1 \) or \( \phi_\pi < -1 \).

Clearly, the formal arguments are the same in both cases. But whereas the indeterminacy is exclusively a real phenomenon in the first case, it is exclusively a nominal one in the second.

What about the intermediate case, namely a general New Keynesian economy in which prices are sticky but not entirely rigid (0 < κ < ∞)? It is tempting to think that this blends the two kinds of indeterminacy; and the fact that determinacy can now be guaranteed with “hybrid” rules of
the form \( i_t = \phi_y y_t + \phi_\pi \pi_t \) may appear to corroborate this view. But the determinacy problem is still exclusively a real one in the following sense: once the path of \( y_t \) has been determined, the path of \( \pi_t \) is pinned down by the Phillips curve.

More succinctly, whereas flexible-price models feature nominal indeterminacy conditional on real allocations, the New Keynesian model features no such indeterminacy, regardless of the degree of the nominal rigidity.\(^{19}\) Consistent with this basic idea (which is not new but is worth emphasizing), our game representation and our formal arguments work equally well for any non-zero degree of nominal rigidity. And once our perturbations have guaranteed real determinacy, nominal determinacy follows for free.

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 models 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:

\[
\begin{align*}
    c_t = & \mathcal{C} \left( \left\{ \hat{E}_{t+k} [y_{t+k}] \right\}_{k=0}^{\infty}, \left\{ \hat{E}_{t+k} [r_{t+k}] \right\}_{k=0}^{\infty} \right) + \varrho_t, \\
    \end{align*}
\]

(24)

where \( r_t \equiv i_t - \pi_{t+1} \) stands for the real interest rate, \( \mathcal{C} \) is a linear function, and \( \varrho_t \) is an exogenous (and, for simplicity, perfectly observed) aggregate demand shock. This generalizes equation (2) 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 D, we show how to obtain a special case of (24) 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 (24) 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

\(^{19}\)Indeed, for any given output path \( \{y_t\}_{t=0}^{\infty} \), the inflation path \( \{\pi_t\}_{t=0}^{\infty} \) is pinned down by the Phillips curve; and since \( p_{-1} \) is historically predetermined, the price path \( \{p_t\}_{t=0}^{\infty} \) is also pinned down.
with the standard, micro-founded, and forward-looking New Keynesian Phillips curve:

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

(25)

where $\kappa \geq 0$ and $\beta \in (0, 1)$ are fixed scalars and $\xi_t$ is, again, a cost-push shock.$^{20}$ Finally, we let the Taylor rule be

$$i_t = z_t + \phi_y y_t + \phi_\pi \pi_t,$$

(26)

for some random variable $z_t$ and some fixed scalars $\phi_y, \phi_\pi \geq 0.$

The “famous” three equations are now given by (24), (25) and (26), along with $y_t = c_t$ (market clearing). Solving (25) and (26) for inflation and the interest rate, and replacing these solutions into (24), 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:

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

(27)

for some random variable $\theta_t$ that is a linear combination of the primitive shocks $(z_t, \xi_t, \varrho_t)$ and some coefficients $|\delta_k|_{k=0}^{\infty}$, with $\delta_0 < 1$ and $\Delta \equiv \delta_0 + \sum_{k=1}^{\infty} |\delta_k| < \infty$.

Similar to equation (6) in our baseline model, this equation helps translate the economy to a game among the consumers. Accordingly, the coefficients $|\delta_k|_{k=0}^{\infty}$ are transformations of deeper parameters that regulate the relevant GE feedbacks.$^{22}$ These feedbacks are now more complicated, and aggregate spending in any given period depends on expectations of economic activity in all future periods as opposed to merely the next period, but the essence is similar.

The overall strategic interdependence, or the analogue of the sum $\delta_0 + \delta_1$ from our main analysis, is now given by $\Delta$. With $\Delta > 1$, multiple self-fulfilling equilibria can be supported under full information, in a similar fashion as in Section 3. But they unravel under 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.

$^{20}$The micro-foundations of (24) are omitted because they are entirely standard. The only point worth mentioning is that (24) presumes that firms, unlike consumers, have full information. This 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.

$^{21}$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 (27), but does not affect Proposition 5, because this holds for arbitrary such coefficients. What we cannot readily nest in (27) is a backward-looking Taylor rule, such as $i_t = z_t + \phi_\pi \pi_{t-1}$, or a backward-looking Phillips curve. See, however, Appendix B.5 for an illustration of why this does not upset our result, insofar as, of course, Assumption 2 is maintained.

$^{22}$These parameters are: the MPCs out of current and future income, $|\frac{\partial C}{\partial y}|_{k=0}^{\infty}$; the sensitivities of consumption to current and future real interest rates, $|\frac{\partial C}{\partial r}|_{k=0}^{\infty}$; the slope, $\kappa$, and the forward-lookingness, $\beta$, of the NKPC; and the policy coefficients, $\phi_\pi$ and $\phi_c$. 
Proposition 5 (Generalized result). Consider the above generalization, impose Assumption 1 and 2, and let $\lambda > 0$. Whenever an equilibrium exists, it is unique and is given by the MSV solution.$^{23}$

Feedback rules and Taylor principle: equilibrium selection or stabilization?

Go back to the textbook New Keynesian model. Let $\{i^o_t, \pi^o_t, c^o_t\}$ 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 aforementioned shocks, it suffices to follow the following feedback rule, for any $\phi > 1$:

$$i_t = i^o_t + \phi(\pi_t - \pi^o_t).$$

This is nested in (5) with $z_t = i^o_t - \phi\pi^o_t$, and is sometimes referred to as the “King rule” (after King, 2000). Note then that $\phi$ can take any value above 1, and this does not affect the properties of the optimum. That is, the feedback from $\pi_t$ 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.

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

---

$^{23}$When $\theta_t$ is uncorrelated over time, the MSV solution is again given by $c^F_t = \theta_t$. More generally, it can be solved for in a similar way as in the extension of our baseline mode that adds persistent fundamentals (Appendix C).
off-equilibrium policy threat.\textsuperscript{24}

**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:

\[
\frac{B_{t-1}}{P_t} = PV S_t, \tag{28}
\]

where $B_{t-1}$ denotes the outstanding nominal debt, $P_t$ denotes the nominal price level, and $PV S_t$ denotes the real present discounted value of primary surpluses.\textsuperscript{25} Does the incorporation of this equation make a difference for the model's predictions about inflation and output?\textsuperscript{26}

The standard approach says no by assuming that fiscal policy is “Ricardian,” in the following sense: $PV S_t$ is required to adjust so as to make sure that (28) 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 logic upside down by letting fiscal policy be “non-Ricardian” in the following sense: it allows fiscal policy to choose $PV S_t$ as if (28) were not a constraint and requires that $P_t$ itself adjusts to make sure that (28) is satisfied for any given $PV S_t$. 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 perturbations also remove 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 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

\textsuperscript{24}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.

\textsuperscript{25}As before, we continue to work with the “cashless” New Keynesian framework, which explains why there is no seigniorage term in (28). But the arguments made below do not seem to depend on this simplification.

\textsuperscript{26}There is disagreement in the literature on whether (28) 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; see Kocherlakota and Phelan (1999) and Buiter (2002) for the first view, Cochrane (2005) for the second, and Bassetto (2002) for a hybrid with sharper game-theoretic foundations. Here, we put aside this debate and focus instead on the role played by consumer rationality, and in particular on the consumers’ understanding of the fact that, one way or another, equation (28) must hold.
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:

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

where \(w_{i,t}\) is the household’s real financial wealth in the beginning of period \(t\), \(\tau_{t+k}\) are the lump sum taxes she owns in period \(t + k\), all other variables are as before, and \(\gamma\) is the steady-state ratio of aggregate private financial wealth to GDP (equivalently, that of public debt to GDP).\(^{27}\)

Equation (29) 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 (28), 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 \(\int E_{i,t} [w_{i,t} - w_t] di = 0\), where \(w_t = \int w_{i,t} di\).

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 (28) 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 \(\int E_{i,t} [w_{i,t} - w_t] di = \int (w_{i,t} - w_t) di = 0\)). More generally, this assumption rules out the possibility that consumers confuse aggregate changes in fiscal policy for idiosyncratic variation in wealth. Such confusion can be fully rational 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 (29) to get the following:

\[ c_t = \bar{E}_t \left[ (1 - \beta) \gamma w_t - \sigma \beta \sum_{k=0}^{+\infty} \beta^k (i_{t+k} - \pi_{t+k+1}) + (1 - \beta) \sum_{k=0}^{+\infty} \beta^k (y_{t+k} - \tau_{t+k}) \right] . \] (30)

Next, equation (28) rewrites in log-linearized form as

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

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 (30) to obtain the following equation:

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

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 (27). 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 idiosyncratic wealth, in the sense of Assumptions 4 and 5. Then, aggregate consumption satisfies equation (27), for some coefficients \( \{\delta_k\}_{k=0}^{\infty} \) and some random variable \( \theta_t \). Furthermore, debt and deficits do not appear in this equation: \( \delta_k \) is a function of \( (\sigma, \beta, \kappa, \phi) \) for all \( k \), and \( \theta_t \) is a transformation of \( (z_t, \xi_t, g_t) \).

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.\(^{28}\)

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:

\(^{28}\)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.
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 (27) 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.

<table>
<thead>
<tr>
<th>With Full Information</th>
<th>With Our Perturbations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fiscal Policy is</strong></td>
<td><strong>Fiscal Policy is</strong></td>
</tr>
<tr>
<td><strong>Ricardian</strong></td>
<td><strong>Ricardian</strong></td>
</tr>
<tr>
<td><strong>Non-Ricardian</strong></td>
<td><strong>Non-Ricardian</strong></td>
</tr>
<tr>
<td><strong>Taylor Principle holds</strong></td>
<td><strong>Determinacy</strong></td>
</tr>
<tr>
<td><strong>Determinacy</strong></td>
<td><strong>No equilibrium</strong></td>
</tr>
<tr>
<td><strong>Multiplicity</strong></td>
<td><strong>Determinacy</strong></td>
</tr>
<tr>
<td><strong>No equilibrium</strong></td>
<td><strong>No equilibrium</strong></td>
</tr>
</tbody>
</table>

Table 1: Standard Paradigm vs Our Approach

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 three important qualifications. First, while Corollary 1 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 a non-fundamental one, along which public debt plays the role 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.

Second, our approach leaves 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, or rational confusion;29 or (ii) the monetary authority internalizes the

29The role of finite horizons and liquidity constraints for breaking Ricardian equivalence is well understood. Perhaps more interesting, Eusepi and Preston (2018) show how a plausible form of boundedly rational learning may produce outcomes akin to those predicted by the FTPL even if these features are absent and fiscal policy is Ricardian. We suspect that the form of rational confusion we have ruled out via Assumption 5 may play a similar role.
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 $\phi$; 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.\(^\text{30}\)

Finally, it is worth highlighting the following points. In line with standard Keynesian thinking, our game theoretic representation presumed that the monetary controls the returns to private saving. This in turn allowed us to abstract from an explicit modeling of the market for government bonds: it sufficed to study behavior in the market for goods. We don’t expect our lessons to be unduly sensitive to this abstraction, insofar as we remain within the New Keynesian framework. For instance, it seems straightforward to augment the model with a competitive banking sector, whose sole role is to arbitrage away any difference between three interest rates: that on central bank reserves, that on government bonds, and that on consumer saving/borrowing. But we must also flag, once again, the importance of nominal rigidity for the Keynesian, “demand-centric” view of the world and, by extension, for our game theoretic representation.\(^\text{31}\)

All in all, we therefore ask that the lessons offered in this section are not mis-read as an attack on the FTPL, and definitely not as a rejection of its core idea that debt and deficits may drive inflation. Rather, these lessons are meant to offer a new lens for understanding what it takes for this idea to make sense within the dominant policy paradigm. To put it differently, our own take is that the FTPL is “right” in the sense that it captures something very relevant in the real world but also that its current formulation is “distorted” by a heavy reliance on equilibrium selection.

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

\(^{30}\)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.

\(^{31}\)For instance, it may be valuable to bridge our work with that of Bassetto (2002). The latter focuses on a flexible-price economy, which, as already discussed, may behave quite differently from a Keynesian economy; but it digs deeper into the question of how the bond market works and how this relates to the FTPL.
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 fiscal and the monetary authorities as that of two players in a game, for example a game of chicken. But for such a game to be well defined, there must also exist a unique mapping from the two 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

Proof of Proposition 1

Part (i) follows directly from the fact that \( c_t^F = \theta_t \) satisfies (10).

Consider part (ii). Let \( \{c_t\} \) be any equilibrium and define \( \hat{c}_t = c_t - c_t^F \). From (10),

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

(32)

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 (32) 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:

\[
\hat{a}_k = \delta \hat{a}_{k+1} \quad \forall k \geq 0 \quad \text{and} \quad \hat{\gamma}_k = \delta \hat{\gamma}_{k+1} \quad \forall k \geq 0.
\]

When the Taylor principle is satisfied (\( \phi > 1 \) and \( \delta < 1 \)), \( \hat{a}_k \) and \( \hat{\gamma}_k \) explodes unless \( \hat{a}_0 = 0 \) and \( \hat{\gamma}_0 = 0 \). We know that the only bounded solution of (32) 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 (\( \phi \in [0,1) \) and \( \delta \in (0,1) \)). \( c_t^B \) satisfies (10). 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:

\[
c_t = \sum_{k=0}^{+\infty} \gamma_k \theta_{t-k}.
\]

(33)

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 \( \theta_t \) is a stationary i.i.d. Gaussian variable from Assumption 1, the
following projections apply for all \( k > s \geq 0 \):

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

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

Now, from Assumption 2, we know

\[
\tilde{E}_t \left[ \theta_{t-k} \right] = (1 - \lambda)^k \theta_{t-k} + \sum_{s=0}^{k-1} \lambda (1 - \lambda)^s \mathbb{E} \left[ \theta_{t-k} | I_t^s \right] = (1 - \lambda)^k \theta_{t-k}. \tag{34}
\]

Now consider an equilibrium in the form of (33). From equilibrium condition (6), we know

\[
\sum_{k=0}^{+\infty} \gamma_k \theta_{t-k} = (1 - \delta_0) \theta_t + \delta_0 \tilde{E}_t \left[ \sum_{k=0}^{+\infty} \gamma_k \theta_{t-k} \right] + \delta_1 \tilde{E}_t \left[ \sum_{k=0}^{+\infty} \gamma_k \theta_{t+1-k} \right]
\]

\[
= (1 - \delta_0) + \delta_0 \gamma_0 + \delta_1 \gamma_1 \theta_t + \delta_1 \tilde{E}_t \left[ \sum_{k=1}^{+\infty} (\delta_0 \gamma_k + \delta_1 \gamma_{k+1}) \theta_{t-k} \right]
\]

\[
= (1 - \delta_0) + \delta_0 \gamma_0 + \delta_1 \gamma_1 \theta_t + \sum_{k=1}^{+\infty} (\delta_0 \gamma_k + \delta_1 \gamma_{k+1}) (1 - \lambda)^k \theta_{t-k},
\]

where we use the fact that all agents at \( t \) know the values of the fundamental state \( \theta_t \).

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

\[
\gamma_0 = (1 - \delta_0) + \delta_0 \gamma_0 + \delta_1 \gamma_1
\]

\[
\gamma_k = (\delta_0 \gamma_k + \delta_1 \gamma_{k+1}) (1 - \lambda)^k \quad \forall k \geq 1. \tag{35}
\]

From Definition 1, we know that there is a scalar \( M > 0 \) such that \( |\gamma_k| \leq M \) for all \( k \geq 0 \). From (35), we know that, for all \( k \geq 1 \),

\[
|\gamma_k| \leq (|\delta_0| + |\delta_1|) (1 - \lambda)^k M. \tag{36}
\]

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

\[
|\gamma_k| \leq (|\delta_0| + |\delta_1|) (1 - \lambda) \hat{k} M.
\]

Now, we can use the above formula and (35) to provide a tighter bound of \( |\gamma_k| \): for all \( k \geq \hat{k} \),

\[
|\gamma_k| \leq (|\delta_0| + |\delta_1|)^2 (1 - \lambda)^{2\hat{k}} M.
\]

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

\[
|\gamma_k| \leq (|\delta_0| + |\delta_1|)^l (1 - \lambda)^{l\hat{k}} M.
\]

Since \((|\delta_0| + |\delta_1|) (1 - \lambda) \hat{k} < 1 \), we then have \( \gamma_k = 0 \) for all \( k \geq \hat{k} \). Using (35) 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 $I_{i,t} = \{\eta_{i,t}, s_{i,t}\}$, any (stationary) strategy can be expressed as
\[
c_{i,t} = a \eta_{i,t} + b s_{i,t},
\]
for some coefficients $a$ and $b$. Then, $c_{t+1} = a \eta_{t+1} + b c_t$, and since agents have no information about the future sunspot, $E_{i,t}[c_{t+1}] = b E_{i,t}[c_t]$. Next, note that $E_{i,t}[c_t] = a \eta_t + b \chi s_{i,t}$, where
\[
\chi = \frac{\text{Var}(c_{t-1})}{\text{Var}(c_{t-1}) + \sigma^2} \in (0, 1].
\]
Combining these facts, we infer that condition (8), the individual best response, reduces to
\[
c_{i,t} = E_{i,t}[\delta_0 c_t + \delta_1 c_{t+1}] = (\delta_0 + \delta_1 b) E_{i,t}[c_t] = (\delta_0 + \delta_1 b) \{a \eta_t + b \chi s_{i,t}\}.
\]
It follows that a strategy is a best response to itself if and only if
\[
a = (\delta_0 + \delta_1 b) a \quad \text{and} \quad b = (\delta_0 + \delta_1 b) b \chi.
\]
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 (37), we see that $a \neq 0$ 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 $\text{Var}(c_{t-1}) > 0$ 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 (23), we have that
\[
\sum_{k=0}^{K_0} a_k \eta_{l-k} + \sum_{k=1}^{K_1} \beta_k c_{t-k} + \sum_{k=0}^{K_0} \gamma_k x_{l-k} + \chi \zeta_t = \theta_t + \zeta_t + \delta E[\sum_{k=0}^{K_{\eta-1}} a_{k+1} \eta_{l-k} + \sum_{k=1}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_0} \gamma_{k+1} \theta_{l-k} | I_t]
\]

\[
= \theta_t + \zeta_t + \delta \left[ \sum_{k=0}^{K_0} a_{k+1} \eta_{l-k} + \sum_{k=1}^{K_{\beta-1}} \beta_{k+1} c_{t-k} + \sum_{k=0}^{K_0} \gamma_{k+1} \theta_{l-k} \right] + \delta \beta_1 \left[ \sum_{k=0}^{K_0} a_k \eta_{l-k} + \sum_{k=1}^{K_{\beta-1}} \beta_k c_{t-k} + \sum_{k=0}^{K_0} \gamma_k \theta_{l-k} + \chi \zeta_t \right]
\]

where we use Assumption 1 and the fact that \( \zeta_t \) is drawn independently over time. For this to be true for all states of nature, we can compare coefficients:

\[
a_k = \delta a_{k+1} + \delta \beta_1 a_k \quad \forall k \in \{0, \ldots, K_\eta-1\} \quad \text{and} \quad a_{K_\eta} = \delta \beta_1 a_{K_\eta} \quad (38)
\]

\[
\beta_k = \delta \beta_{k+1} + \delta \beta_1 \beta_k \quad \forall k \in \{1, \ldots, K_\beta-1\} \quad \text{and} \quad \beta_{K_\beta} = \delta \beta_1 \beta_{K_\beta} \quad (39)
\]

\[
\gamma_k = \delta \gamma_{k+1} + \delta \beta_1 \gamma_k \quad \forall k \in \{1, \ldots, K_\theta-1\} \quad \text{and} \quad \gamma_{K_\theta} = \delta \beta_1 \gamma_{K_\theta} \quad (40)
\]

\[
\gamma_0 = 1 + \delta \gamma_1 + \delta \beta_1 \gamma_0 \quad \text{and} \quad \chi = 1 + \delta \beta_1 \chi. \quad (41)
\]

First, from the second equation in (41), we know \( \delta \beta_1 \neq 1 \). Then, from the second parts of (38)–(40), we know \( a_{K_\eta} = 0, \beta_{K_\beta} = 0, \) and \( \gamma_{K_\theta} = 0 \). From backward induction on (38)–(41), we know that all \( a, b, \gamma \) are zero except for the following:

\[
\gamma_0 = 1.
\]

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

\[
c_t = c_t^E + \zeta_t,
\]

where \( c_t^E = \theta_t \).

**Proof of Proposition 5**

We first note that, with Assumption 1, the MSV solution of (27) is still given by \( c_t^E = \theta_t \). Consider an equilibrium taking the form of (9). We use (27):

\[
\sum_{l=0}^{+\infty} a_l \eta_{t-l} + \sum_{l=0}^{+\infty} \gamma_l \theta_{t-l} = (1 - \delta_0) \theta_t + \bar{E}_t \left[ \sum_{k=0}^{+\infty} \delta_k \left( \sum_{l=0}^{+\infty} a_l \eta_{t+k-l} + \sum_{l=0}^{+\infty} \gamma_l \theta_{t+k-l} \right) \right]. \quad (42)
\]

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 \( \eta_{t-l} \) and using the facts that each
sunspot is orthogonal to all fundamentals:

\[ a_l = \mu_l \sum_{k=0}^{+\infty} \delta_k a_{k+l} \quad \forall l \geq 0. \]  

(43)

Because \( \lim_{l \to \infty} \mu_l = 0 \), there necessarily exists an \( \hat{l} \) finite but large enough \( \mu_l \sum_{k=0}^{+\infty} |\delta_k| < 1 \).\(^{32}\)

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 (43), we then know that, for all \( l \geq \hat{l} \),

\[ |a_l| \leq \mu_l M \sum_{k=0}^{+\infty} |\delta_k|, \]

(44)

where we also use the fact that the sequence \( \{\mu_l\}_{l=0}^{+\infty} \) is decreasing. Now, we can use (43) and (44) to provide a tighter bound of \( |a_l| \). That is, for all \( l \geq \hat{l} \),

\[ |a_l| \leq \left( \mu_l \sum_{k=0}^{+\infty} |\delta_k| \right)^2 M. \]

We can keep iterating. Since \( \mu_l \sum_{k=0}^{+\infty} |\delta_k| < 1 \), we then have \( a_l = 0 \) for all \( l \geq \hat{l} \). Using (43) and doing backward induction, we then know \( a_l = 0 \) for all \( l \), where we use the fact that \( \delta_0 < 1 \).

Now, (42) can be simplified as

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

(45)

\[ = (1 - \delta_0) \theta_t + \sum_{k=0}^{+\infty} \delta_k \gamma_k \theta_t + \tilde{E}_t \left[ \sum_{l=1}^{+\infty} \left( \sum_{k=0}^{+\infty} \delta_k \gamma_{k+l} \right) \theta_{t-l} \right]. \]

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

\[ \gamma_0 = 1 - \delta_0 + \sum_{k=0}^{+\infty} \delta_k \gamma_k \]

(46)

\[ \gamma_l = \mu_l \sum_{k=0}^{+\infty} \delta_k \gamma_{k+l} \quad \forall l \geq 1. \]

(47)

The above two equations can be re-written as:

\[ \gamma_0 = (1 - \delta_0)^{-1} \left( 1 - \delta_0 + \sum_{k=1}^{+\infty} \delta_k \gamma_k \right) \]

(48)

\[ \gamma_l = (1 - \mu_l \delta_0)^{-1} \left( \sum_{k=1}^{+\infty} \delta_k \gamma_{k+l} \right) \quad \forall l \geq 1, \]

(49)

where we use \( \delta_0 < 1 \) and \( \mu_l < 1 \).

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

\(^{32}\)\( \sum_{k=0}^{+\infty} |\delta_k| < \infty \) because \( \Delta < \infty \).
\( (47) \), we know, for all \( l \geq 1 \)
\[
|\gamma_l| \leq \mu_l \left( \sum_{k=0}^{+\infty} |\delta_k| \right) M. \tag{50}
\]
Because \( \lim_{l \to \infty} \mu_l = 0 \), there necessarily exists an \( \hat{l} \) finite but large enough such that \( \left( \sum_{k=0}^{+\infty} |\delta_k| \right) \mu_{\hat{l}} < 1 \). We then know that, for all \( l \geq \hat{l} \),
\[
|\gamma_l| \leq \mu_{\hat{l}} \left( \sum_{k=0}^{+\infty} |\delta_k| \right) M.
\]
Now, we can use the above formula and \( (47) \) to provide a tighter bound of \( |\gamma_l| \): for all \( l \geq \hat{l} \),
\[
|\gamma_l| \leq \left( \mu_{\hat{l}} \right)^2 \left( \sum_{k=0}^{+\infty} |\delta_k| \right)^2 M.
\]
We can keep iterating. Since \( \left( \sum_{k=0}^{+\infty} |\delta_k| \right) \mu_i < 1 \), we then have \( \gamma_l = 0 \) for all \( l \geq \hat{l} \). Using \( (49) \) and doing backward induction, we then know \( \gamma_l = 0 \) for all \( l \geq 1 \) and, from \( (48) \),
\[
\gamma_0 = 1.
\]
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 \( (76) \), 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 \( W_t \equiv \int W_i,tdi = B_{t-1}/P_t \) 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
\[
\sum_{k=0}^{+\infty} \left\{ \prod_{j=1}^{k} \left( \frac{I_{t+j-1}}{\Pi_{t+j}} \right)^{-1} \right\} C_{i,t+k} = W_{i,t} + \sum_{k=0}^{+\infty} \left\{ \prod_{j=1}^{k} \left( \frac{I_{t+j-1}}{\Pi_{t+j}} \right)^{-1} \right\} (Y_{t+k} - T_{t+k}) \right\} .
\]
The government’s budget in \( (28) \) can be written as
\[
B_{t-1}/P_t = \sum_{k=0}^{+\infty} \left\{ \prod_{j=1}^{k} \left( \frac{I_{t+j-1}}{\Pi_{t+j}} \right)^{-1} \right\} (T_{t+k} - G_{t+k}) \right\} . \tag{51}
\]
Since a consumer understands that (51) holds from Assumption 5, she understands that her budget can be written as
\[\sum_{k=0}^{+\infty} \left\{ \prod_{j=1}^{k} \left( \frac{I_{t+j-1}}{I_{t+j}} \right)^{-1} \right\} C_{i,t+k} = W_{i,t} - W_{t} + \sum_{k=0}^{+\infty} \left\{ \prod_{j=1}^{k} \left( \frac{I_{t+j-1}}{I_{t+j}} \right)^{-1} \right\} (Y_{t+k} - G_{t+k})\] .

The consumer’s optimal consumption function, in log-linearized form, can thus be written as follows:
\[c_{i,t} = E_{i,t} \left[ (1 - \beta) \gamma \left(w_{i,t} - w_{t}\right) - \sigma \beta \sum_{k=0}^{+\infty} \beta^{k} (i_{t+k} - \pi_{t+k+1}) + (1 - \beta) \sum_{k=0}^{+\infty} \beta^{k} (y_{t+k} - g_{t+k}) \right],\] (52)
where \(\gamma = \frac{\beta^{*}}{\gamma^{*}}\) is the steady-state debt-to-GDP ratio and all lowercase variables represent log-deviations from the steady state.\(^{33}\) Aggregating, and using Assumption 5, we arrive at
\[c_{t} = (1 - \beta) \left\{ \sum_{k=0}^{+\infty} \beta^{k} \tilde{E}_{t} [y_{t+k} - g_{t+k}] \right\} - \beta \sigma \left\{ \sum_{k=0}^{+\infty} \beta^{k} \tilde{E}_{t} [i_{t+k} - \pi_{t+k+1}] \right\} .\] (53)

This is the same as equation (76), 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 (53) we have not used the full bite of REE; we have only assumed that consumers have first-order knowledge of condition (28) from Assumption 4.

From Assumption 4, consumers understand that the goods markets must clear; and second, consumers understand that inflation obeys the NKPC (25) and that monetary policy follows the Taylor rule (26). The first property allows us to replace the expectations of \([y_{t+k} - g_{t+k}]\) in condition (53) with those of \([c_{t+k}]\); the second allows us to do the same for expectations of \([\pi_{t+k}]\) and \([i_{t+k}]\).\(^{34}\)

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 (27);\(^{35}\) under our informational assumptions, the MSV solution of this equation continues to identify the unique possible

\(^{33}\)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.

\(^{34}\)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 \(\xi_{t}\), or \(\theta_{t}\), and is of no consequence for our purposes.

\(^{35}\)Minor qualification: \(g_{t}\) must now be included in the definition of \(\theta_{t}\), but this makes not difference for the argument made here.
equilibrium process for \( c_t \); conditional on the latter, the processes for \( \pi_t \) 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 Proposition 5.

B Appendix B: Additional Material for Sections 2-4

This Appendix corroborates various claims made in the main text. First, we formalize the sense in which Assumption 2 is compatible with nearly perfect information of both exogenous shocks and endogenous outcomes. Second, we show how to fill in the missing details about labor supply. Third, we explain why the simplification of infinite histories and stationary equilibria is non-essential. Fourth, we illustrate how our result extends to variants of such Taylor rules, whereby monetary policy responds to either past inflation or its expected future value. Finally, we discuss how our result may translate in non-linear settings featuring multiple steady states.

B.1 A Variant with Observation of Current Outcomes

Our baseline model abstracts from idiosyncratic shocks. It also lets consumers be inattentive to, or face uncertainty about, their current income and interest rates. We now relax these assumptions and show how to nest the economy in the same game form as that in the main text, modulo an inconsequential adjustment in the coefficients \( \delta_0 \) and \( \delta_1 \).

For any aggregate variable \( x_t \in \{y_t, i_t, \pi_t, \varrho_t\} \), let \( x_{i,t} \) be the corresponding individual-level variable. Suppose further that future idiosyncratic shocks are unpredictable, so that \( E_{i,t}[x_{i,t+1}] = E_{i,t}[x_{t+1}] \) for any such variable, and that consumer \( i \) observes \( (y_{i,t}, i_{i,t}, \pi_{i,t}, \varrho_{i,t}) \) when setting \( c_{i,t} \). Then, the optimal consumption function of a young consumer, equation (2) in the main text, is modified as follows:

\[
c^y_{i,t} = -\beta \sigma (i_{i,t} - \varrho_{i,t}) + (1 - \beta) y_{i,t} + \beta E_{i,t} y_{i,t+1}
\]

Aggregating the above and using the fact, explained in the main text, that aggregate consumption equals the average consumption of the young, we infer that

\[
c_t = -\frac{\beta}{1+\beta} \sigma (i_t - \varrho_t) + \frac{1}{1+\beta} y_t + \bar{E}_t \left[ \frac{\beta}{1+\beta} y_{t+1} + \frac{\beta}{1+\beta} \sigma \pi_{t+1} \right].
\]
Combining this with market clearing \((y_t = c_t \text{ and } y_{t+1} = c_{t+1})\), and solving out for \(c_t\) we get

\[ c_t = -\sigma \left( i_t - \bar{E}_t [\pi_{t+1}] - \varrho \right) + \bar{E}_t [c_{t+1}]. \]

That is, the DIS curve is now the same as in the representative-agent benchmark, modulo the replacement of that agent's full-information expectation with the average, incomplete-information expectation in the population. By the same token, once we substitute out the interest rate and inflation, our game representation becomes

\[ c_t = \theta_t + \delta \bar{E}_t [c_{t+1}]. \]

That is, the game representation is even simpler than that in the main text.

Clearly, Proposition 5 continues to hold, provided that consumers form expectations about future aggregate outcomes in the manner implied by Assumption 2. But now there is a tension between this assumption and the assumption made above that consumers observe their own income and interest rates. By invoking Assumption 2, we have effectively abstracted from the possibility that consumers extract information about payoff-irrelevant aggregate histories from their own individual wealth, income and interest rates. This seems realistic, especially given that the idiosyncratic fluctuations are much larger than the aggregate ones. But it also brings to the forefront the technical complications that our analysis has painstakingly tried to bypass, either by abstracting from signal extraction (here) or by allowing it but introducing different perturbations (in Section 5).

With signal extraction, there might exist non-fundamental equilibria in which the observation of own income and own interest rates may reveal information about past sunspots, and such endogenous information may not necessarily satisfy 2. Such signal extraction is bound to confound sunspots with idiosyncratic fundamentals, even if there are no aggregate fundamental shocks. Such confounding can itself be the source of multiple equilibria (Benhabib et al., 2015; Gaballo, 2017; Acharya et al., 2017), albeit of a different kind than those obtained in the full-information benchmark. All in all, we are unsure what it takes for our uniqueness result to be robust to such signal extraction—and we cannot really address the issue because of the severe technical complications introduced by signal-extraction and infinite-regress problems.

This circles back to our discussion of why our main approach treats information as exogenous. That said, it should be clear from the above that the observability of current income and interest rates is not relevant per se: if consumers observe these objects but their expectations of future aggregate outcomes continue to satisfy Assumption 2, then the MSV solution remains the unique equilibrium. Finally, if one insists that young consumers not only observe these objects but also freely condition their expectations of future outcomes on them, then our main argument
no more applies, but uniqueness can still be obtained via the alternative perturbation considered in Section 5.

B.2 Flexible Labor Supply

In the main text, we sidestepped labor supply and production. We now show how to fill in the missing details, without affecting our results.

Production is given by

$$Y_t = N_t,$$

(54)

where $N_t$ is total employment. Since output is demand-determined, labor demand is given by $N_t = Y_t = C_t$, where $C_t$ is aggregate spending. Conditional on $C_t$, the specification of labor supply therefore matters only in the determination of the real wage and the split of total income between labor income and firm profits. What needs to be shown, however, is that $C_t$ is determined in the same way as in the main text.

As in the main text, there are overlapping generations of consumers, each living two periods. But unlike the main text, consumers choose not only how much to spend but also how much to work. Accordingly, the complete preferences are given by

$$u(C_{1,i}^1, t) - v(N_{1,i}^1) + \beta e^{-\psi} u(C_{1,i+1}^2) - v(N_{1,i+1}^2),$$

and the complete budgets in the two periods of life are given by

$$C_{1,i}^1 + B_{i,t} = W_t N_{1,i,t} + D_{1,t}$$

and

$$C_{1,i+1}^2 = W_{t+1} N_{1,i+1} + D_{1,t+1} + \frac{I_{t}}{\Pi_{t+1}} B_{i,t},$$

where $v(N) \equiv \frac{1}{1+\psi} N^{1+\psi}$, $N_{1,i,t}$ and $N_{1,i+1}$ are the amounts of labor supplied when young and old, respectively, $W_t$ is the real wage, $D_{1,t}$ and $D_{1,t+1}$ the real firm profits distributed to young and old agents, and all other variables are the same as in Section 1.

To simplify, we assume that old consumers choose consumption and labor supply under full information.\(^{36}\) After the usual log-linearization, this translates to the following optimal rules for the old consumers:\(^{37}\)

$$n_{1,i,t}^2 = \frac{1}{\psi} \left( w_t - \frac{1}{\sigma} c_{i,t}^2 \right)$$

and

$$c_{i,t}^2 = \beta^{-1} b_{i,t-1} + \Omega \left( w_t + n_{i,t}^2 \right) + (1 - \Omega) d_{i,t}^2,$$

\(^{36}\)This simplification is not strictly needed; see for instance Section 6 and Appendix D for how to extend the analysis to a setting in which consumers of all ages are subject to the same informational friction. But it mirrors the main text’s treatment of the old consumers as “robots” and, as it will be shown below, helps reduce the economy to exactly the same game as that obtained in the main text.

\(^{37}\)In the log-linear version, we let $b_{i,t}$ be given by $B_{i,t}/C^*$, where $C^*$ is steady-state spending. This is standard in the literature and takes care of the issue that wealth $B^*$ = 0 in steady state.
where Ω is the ratio of labor income to total income in steady state. Young consumers, on the other hand, are subject to the informational friction of interest, so that their optimal rules are given by

\[ n_{i,t}^1 = \frac{1}{\psi} \left( E_{i,t}[w_t] - \frac{1}{\sigma} c_{i,t}^1 \right) \]

\[ c_{i,t}^1 = E_{i,t} \left[ \frac{1}{1+\beta} \left( \Omega \left( w_t + n_{i,t}^1 \right) + (1 - \Omega) d_t^1 \right) + \frac{\beta}{1+\beta} \left( \Omega \left( w_{t+1} + n_{i,t+1}^2 \right) + (1 - \Omega) d_t^2 \right) \right] - \frac{\beta}{1+\beta} \sigma (i_t - \pi_{t+1} - \varphi_t) \],

where \( E_{i,t} \) is the rational expectation conditional on a young consumer’s information set, whatever that might be.\(^{38}\)

So far, we have taken no stand on how firm profits are distributed between the young and the old. To map exactly to the analysis in Section 1, we henceforth let

\[ D_t^1 = Y_t - W_t \int N_{i,t}^1 d i \quad \text{and} \quad D_t^2 = Y_t - W_t \int N_{i,t}^2 d i. \]  

(55)

This can be justified by having the government tax all firm profits and redistribute them according to the above rule. Alternatively, we can assume that firms live for two periods; and that young (respectively, old) firms are owned exclusively by young (old) consumers and employ exclusively young (old) workers. Either way, the key is that the average income of the young is the same as that of the old (and hence they are both equal to \( Y_t \)), just as in Section 1. Relaxing this assumption complicates the game representation but does not change the essence.

Using the above, we infer that

\[ E_{i,t} \left[ \Omega \left( w_t + n_{i,t}^1 \right) + (1 - \Omega) d_t^1 \right] = E_{i,t} \left[ \Omega \left( w_t + n_{i,t}^1 \right) + (1 - \Omega) d_t^1 \right] = E_{i,t} \left[ y_t \right] \]

\[ E_{i,t} \left[ \Omega \left( w_{t+1} + n_{i,t+1}^2 \right) + (1 - \Omega) d_t^2 \right] = E_{i,t} \left[ \Omega \left( w_{t+1} + n_{i,t+1}^2 \right) + (1 - \Omega) d_t^2 \right] = E_{i,t} \left[ y_{t+1} \right], \]

where \( n_{i,t}^1 = \int n_{i,t}^1 d i \) and \( n_{i,t+1}^2 = \int n_{i,t+1}^2 d i \). As a result, the young consumer’s optimal consumption can be written as

\[ c_{i,t}^1 = E_{i,t} \left[ \frac{1}{1+\beta} y_t + \frac{\beta}{1+\beta} y_{t+1} - \frac{\beta}{1+\beta} \sigma (i_t - \pi_{t+1} - \varphi_t) \right]. \]

Aggregating the above equation, we get

\[ \int c_{i,t}^1 d i = \bar{E}_t \left[ \frac{1}{1+\beta} y_t + \frac{\beta}{1+\beta} y_{t+1} - \frac{\beta}{1+\beta} \sigma (i_t - \pi_{t+1} - \varphi_t) \right]. \]

Since the average private saving of the young has to be zero in equilibrium (\( \int b_{i,t} d i = 0 \)), similar to the main analysis, we still have that

\[ \int c_{i,t}^1 d i = \int c_{i,t}^2 d i = c_t = y_t. \]

\(^{38}\)Similarly to the main text, the above allows the young consumers to be uncertain about, or inattentive to, current income (here, wages and dividends) and current interest rates. But as discussed there, such inattention is vanishingly small when \( \lambda \to 0^+ \), and can be dispensed with along the lines spelled out in Appendix B.
Putting everything together, we arrive at the same DIS equation as in the main text:

\[
c_t = \hat{E}_t \left[ \frac{1}{1+\beta} c_t + \frac{\beta}{1+\beta} c_{t+1} - \frac{\beta}{1+\beta} \sigma (i_t - \pi_{t+1} - \rho_t) \right].
\]

By direct implication, the rest of the analysis remains the same as well.

**B.3 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 (9) 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. In particular, \( c_t \) is now an equilibrium if and only if

\[
c_t = b_t = \delta^{-t} b_0, \tag{56}
\]

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, (56) 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 \( \eta_0 \). That is, 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 remains 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.
B.4 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 \( \lambda \in (0, \hat{\lambda}) \), \( \text{Var} (E_t [c_{t-k}] - c_{t-k}) \leq \epsilon \) 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 \( \pi_{t-k}, i_{t-k}, \) or any linear combination thereof.)

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

\[
I_s^t = \{\eta_{t-s}, \eta_t, \theta_{t-s}, \cdots \theta_t\}.
\]

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

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

Since \( \eta_t \) and \( \theta_t \) are independent of each other as well as independent over time, the finiteness of \( \text{Var} (c_t) \) implies that

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

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

\[
\text{Var} (E_t [c_t | I_s^t] - c_t) \leq \epsilon
\]

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

\[
\text{Var} (E_t [c_{t-k} | I_s^t] - c_{t-k}) \leq \epsilon
\]

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 \( (1 - \hat{\lambda}) \hat{s}_k \geq 1 - \epsilon' \) for all \( k \in \{0, \cdots, K\} \). Together, this means that whenever \( \lambda \in (0, \hat{\lambda}) \), \( \text{Var} (E_t [c_{t-k}] - c_{t-k}) \leq \epsilon \) 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 \( \lambda \) is lowered towards 0. But the
equilibrium mapping from $h^t$ to $c_t$ may well vary with $\lambda$, 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.

### B.5 Alternative Monetary Policies

In the main analysis, we let monetary policy respond to the current rate of inflation. Here, we illustrate how our result extends to variants of such Taylor rules, whereby monetary policy responds to either past inflation or its expected future value.

Consider first the following forward-looking rule:

$$i_t = z_t + \phi \mathbb{E}_t [\pi_{t+1}], \quad (57)$$

where $\phi \geq 0$. In this case, the economy still reduces to a game as in (6), albeit for different values for $\delta_0$ and $\delta_1$. But since our result does not depend on the values of these coefficients, Proposition 5 directly extends.

Suppose next the following backward-looking rule:

$$i_t = z_t + \phi \pi_{t-1}, \quad (58)$$

where $\phi \geq 0$. Even though this case is not directly nested in (6), a version of our argument still goes through.

**Proposition 8 (Alternative monetary policies).** Suppose that Assumption 2 holds, that there are no shocks to fundamentals, and monetary policy takes the form of (58). The equilibrium is unique and is given by the MSV solution.

**Proof:** From (3), (4), and (58), we have that any equilibrium must satisfy

$$c_t = \mathbb{E}_t \left[ \frac{1}{1+\beta} c_t - \frac{\beta}{1+\beta} \sigma \phi \kappa c_{t-1} + \frac{\beta}{1+\beta} (1 + \sigma \kappa) c_{t+1} \right]; \quad (59)$$

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 $k \geq 0$,

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

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

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

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

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

For condition (21) to be true for all sunspot realizations, it is necessary and sufficient that,

\[ a_0 = (1 + \sigma \kappa) a_1, \]

and, for \( k \geq 1, \)

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

We hence have, for \( k \geq 1, \)

\[ a_{k+1} = \frac{\mu_k - \frac{1}{1 + \beta}}{\beta (1 + \sigma \kappa)} a_k + \frac{\sigma \phi \kappa}{1 + \sigma \kappa} a_{k-1}. \] (60)

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 60 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 \( k, \) or equivalently \( c_t = 0 \) for all \( t \) and \( h^t, \) which herein corresponds to the MSV solution. \( \square \)

**B.6 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

\[ c_{i,t} = \delta \tilde{E}_{i,t}[c_{t+1}] - \omega \tilde{E}_{i,t}[c_t^3], \] (61)

for some scalars \( \delta, \omega. \) 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 \( \omega \) matters when \( \delta > 1. \)

When \( \omega \leq 0, \) there is a unique steady state and is given by \( c_{i,t} = 0. \) When instead \( \omega > 0, \) (61) admits three steady states. These are given by

\[ c_{i,t} = -\tilde{c}, \quad c_{i,t} = 0, \quad \text{and} \quad c_{i,t} = \tilde{c}, \]

where \( \tilde{c} \equiv \sqrt{\frac{\delta - 1}{\omega}}. \) If we linearize (61) 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 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 \( \pi \). 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 \( \pi \).

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 \( a \to \tilde{c} \), that is, this type of equilibrium translates to infrequent jumps across the two outer steady states. Finally, this type of 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 \( \pi, a, \) and \( \delta \).

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 \( \rho \) that supported multiplicity in our linear setting, now there is a whole range of values for the corresponding parameter \( \pi \) 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 \( \omega \) 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 form from 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 a bug, and to set the foundations for re-thinking both the Taylor principle and the FTPL.

C 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 $\theta_t$ variable be any stationary, zero-mean, Gaussian process, admitting a finite-state representation.

**Assumption 6 (Fundamentals).** The fundamental $\theta_t$ admits the following representation:

$$\theta_t = q' x_t \quad \text{with} \quad x_t = Rx_{t-1} + \varepsilon_t^x,$$

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), $\varepsilon_t^x \sim \mathcal{N}(0, \Sigma_{\varepsilon})$, and $\Sigma_{\varepsilon}$ is a positive definite matrix.

This directly nests the case in which $(\varrho_t, \xi_t, z_t)$ 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 (Equilibrium).** An equilibrium is any solution to equation (6) along which: expectations are rational, although potentially based on imperfect and heterogeneous information about $h^t$; and the outcome is given by

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

where $a_k \in \mathbb{R}$ and $\gamma_k \in \mathbb{R}^n$ are known and uniformly bounded coefficients.\(^{39}\)

\(^{39}\)This means that there exists a 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.
Assumption 7 (Social memory). In every period $t$, a consumer’s information set is given by

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

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

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 (10) is purely forward looking and $x_t$ is a sufficient statistic for both the concurrent $\theta_t$ 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 $E_t[c_{t+1}] = \gamma' R x_t$; and substitute into (10) 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 $\gamma'$ 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:

$$c_t^F \equiv q' (I - \delta R)^{-1} x_t.$$  \hspace{1cm} (64)

Proof. 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:

$$c_t = \sum_{k=0}^\infty \gamma'_k x_{t-k}. \hspace{1cm} (65)$$

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 (62), the following projections apply for all $k \geq s \geq 0$:

$$E[x_{t-k} | I_t^s] = W_{k,s} x_{t-s},$$
where $I_t^s \equiv \{x_t, ..., x_{t-s}\}$ is the period-t information set of an agent with memory length s and

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

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

Next, note that

$$\|W_{k,s}\|_1 \leq \|\mathbb{E}[x_t x_t']\|_1 \| (R^t)^{k-s} \|_1 \| \mathbb{E}[x_t x_t']^{-1} \|_1,$$

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: $\rho(R) = \rho(R') < 1$. From Gelfand’s formula, we know that there exists $\tilde{\Lambda} \in (0,1)$ and $M_1 > 0$ such that

$$\| (R^t)^{k-s} \|_1 \leq M_1 \tilde{\Lambda}^{k-s},$$

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

$$\|W_{k,s}\|_1 \leq M_2 \tilde{\Lambda}^{k-s},$$

for all $k \geq s \geq 0$. Now, from Assumption 7, we know that

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

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

$$V_{k,k} \equiv (1 - \lambda)^k I_{n \times n} \quad \text{and} \quad V_{k,s} \equiv \lambda (1 - \lambda)^s W_{k,s}.$$ Together with (67), we know that there exists $M_3 > 0$ and $\Lambda = \max\{1 - \lambda, \tilde{\Lambda}\} \in (0,1)$ such that for all $k \geq s \geq 0$,

$$\|V_{k,s}\|_1 \leq M_3 \Lambda^k.$$ (69)

Now consider an equilibrium in the form of (65). From equilibrium condition (6), we know

$$\sum_{k=0}^{+\infty} \gamma'_k x_{t-k} = (1 - \delta_0) \theta_t + \delta_0 \tilde{E}_t \left[ \sum_{k=0}^{+\infty} \gamma'_k x_{t-k} \right] + \delta_1 \tilde{E}_t \left[ \sum_{k=0}^{+\infty} \gamma'_k x_{t+1-k} \right]$$

$$= \left( (1 - \delta_0) q' + \delta_0 \gamma'_0 + \delta_1 \gamma'_0 R + \delta_1 \gamma'_1 \right) x_t + \tilde{E}_t \left[ \sum_{k=1}^{+\infty} \delta_0 \gamma'_k + \delta_1 \gamma'_k x_{t-k} \right]$$

$$= \left( (1 - \delta_0) q' + \delta_0 \gamma'_0 + \delta_1 \gamma'_0 R + \delta_1 \gamma'_1 \right) x_t + \sum_{k=1}^{+\infty} \left( \delta_0 \gamma'_k + \delta_1 \gamma'_{k+1} \right) \sum_{s=0}^k V_{k,s} x_{t-s}.$$ (68)

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:

\[
\gamma_0' = (1 - \delta_0) q' + \delta_0 \gamma_0' + \delta_1 \gamma_0'R + \delta_1 \gamma_1'
\]

\[
\gamma_k' = \sum_{l=k}^{+\infty} (\delta_0 \gamma_l' + \delta_1 \gamma_{l+1}') V_{l,k} \quad \forall k \geq 1.
\]  \hspace{1cm} (70)

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 (69) and (70), we can then infer that, for all \(k \geq 1\),

\[
\|\gamma_k'\|_1 \leq (|\delta_0| + |\delta_1|) \sum_{l=k}^{+\infty} \|V_{l,k}\|_1 M \leq (|\delta_0| + |\delta_1|) M_3 \frac{\Lambda^k}{1 - \Lambda} M.
\]  \hspace{1cm} (71)

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

\[
(|\delta_0| + |\delta_1|) M_3 \frac{\Lambda^{\hat{k}}}{1 - \Lambda} < 1.
\]  \hspace{1cm} (72)

From (71), 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 (70) 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 \(\gamma_k'\) can be ruled out, for all \(k \geq \hat{k}\). Using (70) and doing backward induction, we conclude that \(\gamma_k' = 0\) for all \(k \geq 1\), where I use \(\delta_0 < 1\).

We are then left with a single equation for \(\gamma_0'\):

\[
\gamma_0' = (1 - \delta_0) q' + \delta_0 \gamma_0' + \delta_1 \gamma_0'R.
\]

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. \(\Box\)
D Appendix D: Additional Material for Section 6

In this Appendix we present two micro-foundations for (24), the intertemporal Keynesian cross used in Section 6, and thereby for (27), the corresponding game representation. The first micro-foundation abstracts from idiosyncratic shocks while also allowing for inattention to own income, interest rates, and wealth; the second relaxes these assumptions.

An extension to Proposition 3

Here, we generalize Proposition 3 by letting information be given by

\[ I_{i,t} = \{\eta_t, s_{i,t}\}, \quad \text{with} \quad s_{i,t} = c_{t-1} + v_t + \epsilon_{i,t}. \tag{73} \]

where \( v_t \sim \mathcal{N}(0, \sigma_v^2) \) is an aggregate noise and \( \epsilon_{i,t} \sim \mathcal{N}(0, \sigma^2) \) is idiosyncratic noise. They are independent of each other, other shocks, and across time. 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, 2019). The case studied in main text can be nested by letting \( \sigma_v^2 = 0 \).

Corollary 2. Proposition 3 can be extended to the case in which the information is given by (73).

Proof: 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 \( \eta_{t+1} \) and \( v_{t+1}, E_{i,t}[c_{t+1}] = bE_{i,t}[c_t] \). Next, note that \( E_{i,t}[c_t] = a\eta_t + b\chi s_{i,t} \), where

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

Combining these facts, we infer that condition (8), the individual best response, reduces to

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

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

\[ a = (\delta_0 + \delta_1 b) a \quad \text{and} \quad b = (\delta_0 + \delta_1 b) b\chi. \tag{74} \]

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 (74), we see that \( a \neq 0 \) 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 \( a \neq 0 \)). □
A micro-foundation for (24) and (27)

Consider a “perpetual youth,” overlapping generations economy. Preferences are standard, given by expected lifetime utility, and the survival rate is invariant to age, given by \( \omega \in (0, 1] \). When a consumer dies, she gets replaced by a newborn consumer, who has zero wealth. As in Blanchard (1985), consumers can trade actuarily fair annuities, whose return conditional on survival is given in equilibrium by the risk-free rate plus \( \omega \). Furthermore, consumers have perfect recall over their lifetime. But unlike Blanchard (1985), information is not necessarily transferred from dying consumers to newborn consumers. This allows us to think of the decay in social memory, namely Assumption 2, as a byproduct of mortality. But this interpretation is not strictly needed.

We can then write the (log-linearized) individual consumption function as follows:

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

where \( w_{i,t} \) is the household’s real financial wealth in the beginning of period \( t \), \( \varrho_{i,t} \) is an exogenous demand shock (which may have both an idiosyncratic and an aggregate component), \( \sigma > 0 \) is the elasticity of intertemporal substitution, \( \beta \in (0, 1) \) is the subjective discount factor, and \( \omega \in (0, 1) \) is the mortality rate. Note that equation (75) 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 helps reduce the tension between Assumption 2 and the idea that consumers may learn about past, unobserved, sunspots from the observation of their own wealth. But as discussed below, it is not strictly needed.

Similarly to Section 6, we next assume that consumers do not mis-perceive their idiosyncratic wealth, in the sense of Assumption 5. In the present context, this translates to \( \int E_{i,t}[w_{i,t}]di = 0 \), because aggregate wealth itself is zero. By aggregating (29) across \( i \), and using \( \int E_{i,t}[w_{i,t}]di = 0 \), we arrive at the following aggregate consumption function:

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

which is clearly nested in (24). Using market clearing together with (25) and (26), we can then substitute the average expectations of \( [i_{t+k}] \) and \( [\pi_{t+k}] \) as functions of the average expectations

---

40 The above interpretation restricts \( \lambda = 1 - \omega \), where \( 1 - \omega \) is the probability of death. But we could have \( \lambda < 1 - \omega \) if newborn consumers inherit some of the information of the dying consumers. And conversely, we could justify \( \lambda > 1 - \omega \) by letting consumers be altruistic towards future generations. For instance, if consumers are “dynasties” as in Barro (1974), they choose consumption as if they are infinitely lived (\( \omega = 0 \)), but we can still justify \( \lambda > 0 \) as the product of physical death. Last but not least, we can think of \( \lambda > 0 \) as the by-product of bounded recall within the lifecycle of an individual. This would add a behavioral flavor to our approach, which we welcome but do not strictly require.

41 In the log-linear version, we let \( w_{i,t} \) be given by \( W_{i,t}/C^* \), where \( C^* \) is steady-state spending. This is standard in the literature and takes care of the issue that wealth \( W^* = 0 \) in steady state.
of \{c_{t+k}\} and reach the following version of equation (27):\footnote{For simplicity, we let \(\varphi_t\) be commonly known at \(t\), so that it can be subsumed inside the expectation operator.}
\[
c_t = \tilde{E}_t \left[ (1 - \delta_0) \theta_t + \sum_{k=0}^{+\infty} \delta_k c_{t+k} \right],
\]
where, for all \(k \geq 0\),
\[
\delta_k \equiv (1 - \beta \omega - \beta \omega \sigma \phi_y) (\beta \omega)^k + \omega \sigma \kappa \left( -\beta \phi_y + (1 - \beta \omega \phi_y) \frac{1 - \omega^k}{1 - \omega} \right) \beta^k
\]
and \(\theta_t\) is a linear combination of \((\varrho_t, \xi_t, z_t)\). Note then that \(\delta_0 < 1\) and \(\Delta \equiv \delta_0 + \sum_{k=1}^{\infty} |\delta_k| < \infty\), which means that the only restrictions on these coefficients required by Proposition 5 are readily satisfied. It follows that, as long as Assumption 2 holds, Proposition 5 also holds: the equilibrium is unique and is given by the MSV solution regardless of monetary policy.

A variant with idiosyncratic shocks and perfect attention to own variables

In the above, we abstracted from idiosyncratic shocks and, at the same time, allowed consumers to be inattentive to, or face uncertainty about, their concurrent wealth, income, and interest rates. We now relax these assumptions and show how to derive a close cousin to condition (77). This mirrors the related exercise in Appendix B.1 for our baseline model and leads to basically the same conclusion: everything goes through insofar as one is willing to abstract from signal extraction and maintain Assumption 2.

For any aggregate variable \(x_t \in \{w_t, y_t, i_t, \pi_t, \varphi_t\}\), let \(x_{i,t}\) be the corresponding individual-level variable. Suppose further that idiosyncratic shocks are unpredictable and that consumer \(i\) observes \((w_{i,t}, y_{i,t}, i_{i,t}, \pi_{i,t}, \varphi_{i,t})\) when setting \(c_{i,t}\). Then, the individual consumption function (75) is modified as follows:
\[
c_{i,t} = (1 - \beta \omega) w_{i,t} + (1 - \beta \omega) y_{i,t} - \beta \omega \sigma (i_{i,t} - E_{i,t}[\pi_{i,t+1}]) + E_{i,t} \left[ -\beta \omega \sigma \sum_{k=1}^{+\infty} (\beta \omega)^k (i_{t+k} - \pi_{t+k+1}) + (1 - \beta \omega) \sum_{k=1}^{+\infty} (\beta \omega)^k y_{t+k} \right] + \varphi_{i,t}.
\]
If we aggregate this condition, impose market clearing in all periods, use the NKPC and the Taylor rule to substitute \([\pi_{t+k}]_{k=0}^{\infty}\) and \([i_{t+k}]_{k=0}^{\infty}\) as functions of \([c_{t+k}]_{k=0}^{\infty}\), we arrive at the following equation:
\[
c_t = (1 - \delta_0) \theta_t + \delta_0 c_t + \sum_{k=1}^{+\infty} \delta_k \tilde{E}_t[c_{t+k}],
\]
where the coefficients \(\{\delta_k\}\) are defined as before. That is, we have arrived at exactly the same equation as equation (77) above, except that now \(\tilde{E}_t[c_t]\) is replaced by \(c_t\) itself.

In effect, what we have done so far is to shut down the coordination friction within time
but preserve it across time: the consumers act, on average, as if they know what other consumers are doing today, but they still need to form expectations about how future consumption will be determined and, more specifically, how it may depend on payoff-irrelevant histories. By the same token, if we solve (78) for $c_t$, we can recast the equilibrium in the following game form:

$$c_t = \theta_t + \sum_{k=1}^{\infty} \frac{\delta_k}{1-\delta_0} \hat{E}_t[c_{t+k}],$$

which is again nested in (27). It follows that Proposition 5 continues to apply, provided that expectations of aggregate outcomes satisfy Assumption 2. Note, however, that this means that we have effectively abstracted from the possibility that consumers extract information about payoff-irrelevant aggregate histories from their own individual wealth, income and interest rates, as in Appendix B.1.

**References**


60