NBER WORKING PAPER SERIES

DIAGNOSTIC BUSINESS CYCLES

Francesco Bianchi
Cosmin L. Ilut
Hikaru Saijo

Working Paper 28604
http://www.nber.org/papers/w28604

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021, Revised August 2021

This paper supersedes an earlier version, titled "Implications of Diagnostic Expectations: Theory and Applications". We thank Daniele d'Arienzo, Spencer Kwon, Peter Maxted, Andrei Shleifer, and Stephen Terry for helpful comments and all seminar participants at the 2021 NBER Summer Institute, 2021 SED annual meeting, and Hitotsubashi University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Diagnostic Business Cycles
Francesco Bianchi, Cosmin L. Ilut, and Hikaru Saijo
NBER Working Paper No. 28604
March 2021, Revised August 2021
JEL No. D8,D9,E03,E3,E71

ABSTRACT

A large psychology literature argues that, due to selective memory recall, decisionmakers' forecasts of their future circumstances appear overly influenced by the new information embedded in their current circumstances. We adopt the diagnostic expectations (DE) paradigm (Bordalo et al. (2018)) to capture this feature of belief formation and develop the micro-foundations for applying DE to business cycle models, while demonstrating its empirical relevance for aggregate dynamics. First, we develop behavioral foundations to address the theoretical challenges associated with modeling the feedback between optimal actions and agents' DE beliefs in the presence of (i) endogenous variables and (ii) time-inconsistencies in those optimal actions due to memory recall based on distant past. Second, we build on our theory to propose a portable solution method to study DE in dynamic stochastic general equilibrium models, which we use to estimate a quantitative New Keynesian model augmented with DE. We uncover a critical role played by both endogenous states and distant memory recall under DE in successfully replicating the boom-bust economic cycle observed in the data in response to a monetary policy shock.

Francesco Bianchi
Social Sciences Building, 201B
Department of Economics
Duke University
Box 90097
Durham, NC 27708-0097
and CEPR
and also NBER
francesco.bianchi@duke.edu

Cosmin L. Ilut
Department of Economics
Duke University
223 Social Sciences Building
Box 90097
Durham, NC 27708
and NBER
cosmin.ilut@duke.edu

Hikaru Saijo
University of California at Santa Cruz
Economics Department
401 Engineering 2 Building
1156 High Street
Santa Cruz, CA 95064
hsaijo@ucsc.edu

1 Introduction

A large psychology and experimental literature documents that decision-makers' forecasts of their future circumstances appear overly influenced by the surprises embedded in their current circumstances. In economics, this critical feature of belief formation has been captured by the diagnostic expectations (DE) paradigm, formulated recently by Bordalo et al. (2018) and based on the representativeness heuristic of probabilistic judgments introduced by Kahneman and Tversky (1972). For example, according to this view, a high current level of financial resources that is 'unusual' when compared to her reference belief, i.e what she expected to see currently based on past information, triggers more vivid memories of good times for the agent. This selective memory recall then leads her to overly inflate the likelihood of her future resources being high with respect to the true distribution of future outcomes.

While promising in the breadth of its potential implications, so far the DE paradigm has been typically studied in environments where the extent to which a circumstance is 'surprising' is characterized by two properties: (i) those circumstances are determined exogenously and (ii) the surprise is perceived with respect to a reference belief based on the last period (or immediate past) information set. However, these two characteristics appear overly restrictive in applications because (i) in a large set of situations, decisions involve a feedback between agents' beliefs and endogenously determined economic states, and (ii) the type of selective memory recall that best accounts for the empirical evidence may be based on more distant information sets than just the last period.[1]

Motivated by these observations, our paper has three main contributions. First, we develop micro-foundations that allow us to jointly address the theoretical challenges associated with modeling (i) the feedback between optimal actions and agents' DE beliefs over both exogenous and *endogenous* variables, and (ii) the time-inconsistencies in those optimal actions that arise when selective memory recall is based on a more *distant* past, rather than just the immediate past. Second, we build on these foundations to propose a *portable solution method* to study DE in linear recursive macroeconomic models, which can thus accommodate large-scale dynamic stochastic general equilibrium models. Third, we leverage the tractability of our proposed method to incorporate DE into a quantitative New Keynesian model of the type widely used for policy analysis. We estimate a critical and novel role played by endogenous states and distant memory recall that through DE allow the model to replicate the empirically documented boom-bust cycle in response to a monetary policy shock.

[1] For example, Bordalo et al. (2020b) find that a reference belief based on the four quarters ago information set seems to account well for the empirical over-reaction observed in the surveys of professional forecasters, while Bordalo et al. (2019b) argue that the sluggishness in expected returns is best explained by a reference information set eleven quarters in the past.

In deriving a behavioral model of DE we build on the recent formulation of Gennaioli and Shleifer (2010) and Bordalo et al. (2018). In particular, under the tractable assumption of normality of the data generating process, DE distort current forecasts made under the true density (which we refer to as rational expectations, or RE) with a term that depends on the difference between current RE (the representative, or diagnostic group) and lagged RE (the reference, or comparison group). Thus, the size of the distortion is proportional to the revision in RE (or the representative information). In Bordalo et al. (2018), this idea is formalized in terms of two parameters. A parameter θ controls the severity of the distortion, while a parameter J controls the lag of those reference beliefs.

Applied theory contribution. In the first part of the paper we use two simple consumption-savings models to analyze the properties characterizing the feedback between DE beliefs and optimal actions in the presence of endogenous states. We start with a two-period consumption-savings problem where the agent only needs to forecast total future resources. These are given by the sum of accumulated savings (K_t) entering next period, and future stochastic income (Y_{t+1}). Critically, the DE beliefs and the agent's optimal response of the random variable K_t to the current income realization Y_t are to be determined jointly.

A first important property, that we label *endogenous predictability*, arises because a given non-zero response of K_t to the exogenous Y_t is a source of conditional predictability from the Y_t realization to the random variable that the agent is interested in forecasting (i.e. total future resources). By assuming that Y_{t+1} is iid, we make this point stark, as there is no further predictability coming from the exogenous stochastic component of future resources.[2]

In particular, following a current unusually high (low) income shock, and for a given positive response of savings K_t to this innovation, the agent correctly realizes that her future resources are more likely to be higher (lower) than usual. Due to her imperfect memory, an agent subject to the representativeness heuristic recalls more vividly state realizations that are representative in light of the new information contained in this unusual state of high (low) *expected* resources, and becomes *overly influenced* by her perception of this new information. Thus, following a current positive income surprise, the agent becomes more optimistic about future available resources, and importantly, more than under the true distribution, leading her ro make saving decisions under an "as if" optimistic view of future resources.

Under quadratic utility and iid income shocks, both DE and RE agents take optimal

[2]Thus, if the DE distortion would apply in isolation only to the exogenous income component of the random variable to be forecasted, the revision in conditional expectations under the true density would be zero and there would be no effects arising from DE. In addition, we note that, consistent with the analysis of Bordalo et al. (2018), for memory recall to be even activated and for DE to matter, a necessary condition is that the variance of the future income shocks is non-zero, i.e. that there is some 'residual uncertainty' (in the general language of Gennaioli and Shleifer (2010)) in forecasting future resources, given the new information.

actions to keep a flat *expected* consumption profile. Under RE beliefs, this amounts to saving half of her income in the first period. Under DE beliefs, given her over-reaction to the diagnostic information, she decides to consume more and save less today than the RE agent, with a marginal propensity to save that decreases with the representativeness parameter θ. Thus, when the income innovation in the first period is unusually high (low), the agent seems to save too little (much), compared to the RE agent. While puzzling from the perspective of an external RE observer, this behavior is optimal under DE. Thus, DE can rationalize the apparent lack of consumption smoothing documented by a large empirical literature that finds that in the data the marginal propensity to consume (MPC) is puzzlingly large, even for agents that are not financially constrained (see Jappelli and Pistaferri (2010) for a survey).

In extending the model to multiple periods, we need to confront a second important property of DE, namely that when the reference point for the DE distortion is not pinned down by the immediate past ($J > 1$), *the law of iterated expectations (LOIE) fails.* Intuitively, this occurs because when forming expectations about the future and J is large, the information set pinning down the DE distortion can be antecedent to the current information set. In a multi-period model, the failure of the LOIE critically matters because it leads to time-inconsistent choices, as optimal plans decided in the past become suboptimal upon re-evaluation as a result of the change in beliefs induced by imperfect memory.

We illustrate these implications by extending the two-period model to include a third period. To address the issue of time inconsistency, and study the resulting *interaction* of *endogenous predictability* and *failure of LOIE under distant memory*, we use insights from the microeconomic theory (e.g. O'Donoghue and Rabin (1999)) and consider two alternative assumptions regarding agents' beliefs about future selves' behavior. Under the first approach, coined in this literature as *naïveté,* the agent fails to take into account that her preferences are time-inconsistent and thinks that in the future she will make choices under perfect memory recall, or RE. However, when the future arrives, the agent ends up changing behavior and be again subject to her imperfect memory recall. The second approach to deal with time-inconsistency consists of assuming *sophistication.* In this case, when solving her current problem, the agent understands how imperfect memory recall changes her future preferences.

We use an analytical illustration based on $J = 2$ and show how under the naïve approach, DE beliefs lead to savings policy functions in period 1 and 2 characterized by (i) a novel, non-zero, response to expectations formed two periods ago; and (b) a muted response to actual available savings chosen one period ago. The critical reason behind these results is that the saving decision in period 1 is *not* a sufficient statistic for the comparison group pinning down the selective memory recall process.[3] Instead, the agent looks at expectations

[3] When $J = 1$, since savings chosen in the first period also pins down reference expectations in the second

formed two periods ago. Given that under naïveté these expectations were formed thinking that the agent was going to behave rationally, the agent will typically be surprised by the actual savings at her disposal and she will end up over-reacting to this surprise in terms of her saving decision in the second period. Thus, higher savings in period 1 will lead the agent to choose lower savings in period 2 by inducing optimism about future resources.

Under sophistication, the agent takes into account that at time 2, the future self would undertake a suboptimal choice from the time 1 perspective. Therefore, knowing the future DE policy, the agent takes into account the impact of the current saving decision on the future perceived suboptimal choice. This feature presents itself through two different channels. First, the agent takes into account that her saving choices will affect future resources available to a future self with distorted beliefs. Second, the agent realizes that she can also affect her own reference point, given that this depends on her past actions. Because of these channels, the solution under sophistication involves a significantly higher level of complexity.

The full characterization of the naïve and the sophistication solutions in the three period model allows us to see clearly the economic mechanisms at play. However, in extending this theoretical framework to more realistic and quantitatively relevant business cycle models we propose to focus specifically on the naïveté approach, as a coherent micro-founded model of beliefs and behavior that can also be easily characterized methodologically. More broadly, the required hyper-rationality behind sophistication arguably runs counter to the motivation of accounting for belief heuristics, since this is usually viewed as a cognitive, mental shortcut that allows agents to make judgments quickly and efficiently (Tversky and Kahneman (1975) and Kahneman (2011)). As such, the naïve approach is arguably psychologically more coherent and consistent with the underlying foundation of diagnostic beliefs as a heuristic reflecting a memory representation affected by imprecise, selective, and less than fully rational recall.

Methodological contribution and quantitative evaluation. In the second part of the paper we leverage our theoretical insights to first explain how to solve linear general equilibrium models in the presence of DE, and specifically under the naïveté approach, by using standard solution methods, such as Sims (2000). Intuitively, solving a model featuring DE recursively requires at each point in time forming DE beliefs based on state variables inherited from the DE economy but, by the characteristic of naïveté model of beliefs, having agents expect that future variables will follow a counterfactual RE law of motion (derived under perfect memory). Since it can be applied easily to large state space models, we emphasize that our solution method is *portable, tractable* and, importantly, also allows for

period, then (i) both the naïve and sophisticated problems lead to the same, time-consistent, optimal savings policies, and (ii) these DE policies feature the same response to the endogenous economic state as for the RE policy function, but (iii) as in the two period-model, a muted response to the current income innovation.

general forms of how memory recall loads on different past information sets.

We apply this solution method to incorporate DE into a New Keynesian model with nominal and real frictions that are typical of quantitative business cycle models used for policy analysis (eg. Christiano et al. (2005) and Smets and Wouters (2007)). Given our particular interest in the role played by distant memory recall, we model the reference distribution entering the representativeness heuristic in a flexible manner, as a weighted average of lagged RE expectations. The weights can activate various combinations of information sets over the last 32 quarters to form the comparison group. We parsimoniously model them using a two-parameter Beta distribution that we estimate.

We estimate the model using a Bayesian version of the impulse-response-function (IRF) matching method developed by Christiano et al. (2010), where the empirical IRFs are recovered using a local Jordà (2005) projection to a monetary policy shock (identified by the Romer and Romer (2004) approach and extended by Coibion et al. (2017)). The targeted moments are the empirical IRFs of aggregate consumption, hours, inflation, and the Federal Funds rate. As 'untargeted' moments that serve as external validation, we also compare the model-implied IRFs to a monetary policy shock of investment, GDP, and SPF inflation expectations against their empirical counterparts.

We find that the DE model reproduces the empirical IRFs to a monetary policy shock well, successfully generating, as in the data, a persistent and hump-shaped boom-bust cycle in consumption and hours. In contrast, a counterfactual RE model, where we set the diagnostic parameter $\theta = 0$ while holding fixed other estimated parameters, generates transitory and negligible responses, indicating that DE are a critical economic force in the estimated model. In turn, a re-estimated RE model also fails in delivering the boom-bust dynamics and the amplitude of the responses observed in the data. As a result, the marginal likelihood, a Bayesian measure of fit that penalizes models with more parameters, heavily favors the estimated DE over the re-estimated RE model. In addition, the DE model is also able to match remarkably well the other untargeted empirical responses.

The estimated memory weights are centered on expectations formed six quarters ago, with positive weights assigned to expectations formed between three and eleven quarters ago. If we counterfactually impose that only very recent memory ($J = 1$) or only two-period-ago expectations ($J = 2$) matter, both the frequency and the amplitude of the boom-bust cycles are significantly dampened, at odds with the data. We distill the key economic mechanism through which this otherwise rich DE model fits the IRF dynamics by focusing on the equilibrium connection between perceived consumption and inflation paths, as implied by the optimal intertemporal consumption smoothing under DE.

In particular, per standard consumption smoothing logic, the perception of a higher than usual future price level compared to the current price leads to an downward adjustment in the expected real consumption growth compared to steady state. We show that this perception of changes in nominal prices under DE can be decomposed into two terms: (i) the one-step-ahead expectation of inflation under DE and (ii) the perceived innovation in the current price level (a state variable), compared to the reference distribution. We label this second term the *perceived innovation in cumulative inflation*, because the surprise in today's price level reflects the cumulative inflation between the current period and the time at which reference expectations were formed. In terms of relative variation of the two components, we find that the movement in this second term is much more ample and persistent. Intuitively, since inflation is expected to have low persistence (under RE, per naïveté), applying DE in isolation only over the first component matters less. In contrast, the large effect of DE comes through the second component, the accumulated surprises in the price level compared to its reference distribution, which moves beliefs a lot by shifting the predicted price level path.

The specific boom-bust dynamics of this perceived innovation in cumulative inflation help rationalize the observed boom-bust cycle in consumption. This perception can be described by following the realized path of inflation. On impact, because of an increase in utilization, inflation declines. This determines a negative surprise in the price level and a lower than usual expected future price level that is consistent for the agent with a perceived acceleration in consumption. Inflation eventually starts picking up, leading first to a reduction in the negative surprises for the price level and then eventually to positive surprises. This path determines a *reversal* in the perceived innovation in cumulative inflation, which moves into the positive territory during the bust part of the cycle, when agents find the resulting high perceptions of future price level consistent with their pessimism about future consumption. This reversal helps to account for why the economic boom induced by an expansionary monetary policy shock does not simply slowly subside to converge back to steady state from above. Instead, as in the data, there is an inflection point (around period 15) where the boom turns into a bust and a general decline in economic activity.

More broadly, these findings highlight the two critical aspects of DE that we emphasize throughout this paper. First, the feedback from beliefs to actions, in the presence of endogenous states creates distortions in beliefs that extend well beyond the lag in reference expectations. When DE apply to exogenous variables, DE and RE naturally realign themselves in the IRF after J periods, as the initial shock becomes part of the information set of the comparison group and no further distortions are activated. However, when DE also apply to endogenous variables, past decisions affect current expectations, generating new and time-varying distortions that in turn feed into current decisions, creating waves of optimism and pessimism that generate boom-bust cycles - a form of Minsky (1977) moments.

Second, our quantitative analysis also further illustrates the importance of considering distant memory for a given level of DE distortions. When memory is more immediate, only shocks occurring in between the time of the reference expectations and today can lead to significant surprises. Instead, under a more distant memory agents expectations are constantly revised as the economy is quite far from where they expected it to be when they formed those reference expectations, leading typically to larger revisions and belief distortions. Consistent with our counterfactuals based on more immediate past, this explains why distant memory affects not only the lag at which reversals occurs but also the amplitude of fluctuations.

Our paper is closely related to some recent contributions that study DE in macro models. Bordalo et al. (2019a) analyzes DE about a TFP process to account for credit cycles, Maxted (2020) builds a He and Krishnamurthy (2019) style macro-finance model featuring DE, while d'Arienzo (2020) introduces DE into a term-structure model to study bond market puzzles. L'Huillier et al. (2021) further shares a similar interest with us in introducing DE into linear, dynamic general equilibrium models. As anticipated earlier, we contribute to the literature in two key ways. First, we address the conceptual challenges of modeling the role of endogenous states and distant memory recall in jointly affecting DE beliefs and optimal actions. A particular dimension here is that, compared to this existing work, we confront the problem of time inconsistency that arises in the empirically relevant case of distant memory by providing a behavioral foundation of naïvete and sophistication.[4] Second, in quantitative terms, we propose and use an easily portable solution method to estimate a New Keynesian model augmented with our DE structure of beliefs to show that the feedback between actions and those DE beliefs is critical in replicating the boom-bust cycle we recover from data.

2 A Two-Period Consumption-Savings Model

We start with a simple two-period consumption-savings model as a laboratory to study the feedback between Diagnostic Expectations (DE) beliefs and optimal actions in the presence of endogenous states. In particular, an agent born at a generic time 1 inherits beliefs from J periods ago and capital K_0 from last period. Her budget constraints in periods 1 and 2 are

$$C_1 + K_1^{\theta} = Y_1 + (1+r)K_0;\; C_2 = Y_2 + (1+r)K_1^{\theta}, \tag{1}$$

where C_1 and C_2 are her consumption choices, K_1^{θ} is her savings choice at time 1 and Y_1, Y_2 are the realizations of an exogenously given labor income process.

In particular, here we assume that the exogenous income Y has zero persistence: $Y_{t+1} =$

[4] In this context, L'Huillier et al. (2021) study the role of endogenous states in driving DE beliefs, but their analysis and solution method applies only when memory is based on the immediate past. d'Arienzo (2020) explores the LOIE failure as a mechanism for a maturity increasing overreactions of expectations to news. Here we connect this failure to time-inconsistency and study it in models with endogenous states.

$\overline{Y} + \varepsilon_{t+1}$, where ε_{t+1} are mean zero iid normal shocks with variance $\sigma^2 > 0$. In this way we isolate what we label an *endogenous predictability* mechanism under DE, in which the action K_1^θ alone induces persistence in the future resources available for consumption.

The two-period assumption greatly simplifies this problem, as $K_0 = 0$ and her optimal end-of-life $K_2 = 0$, since past and current agents are assumed not to care about offsprings. In this simple model, we also assume for simplicity a real interest rate $r = 0$, a discount factor $\beta = 1$, and a quadratic utility function $u(C) = bC - .5C^2$.[5]

Under DE the agent maximizes

$$\max_{K_1^\theta} \left[u(C_1) + \mathbb{E}_1^\theta u(C_2) \right], \tag{2}$$

subject to the budget constraints in (1), where $\mathbb{E}_1^\theta$ is formed under a distorted conditional density h^θ, to be specified below. Under this density, the first-order condition that characterizes the optimal choice in (2) is given by

$$u'(C_1) = \mathbb{E}_1^\theta \left[u'(C_2) \right]. \tag{3}$$

Since here the marginal utility is linear in consumption, the tradeoff simply involves

$$C_1 = \mathbb{E}_1^\theta \left[C_2 \right]. \tag{4}$$

2.1 Diagnostic beliefs given law of motion of endogenous state

The random C_2 is known by the time-2 budget constraint to equal $\overline{Y} + \varepsilon_2 + K_1^\theta$. We conjecture a response of the optimal K_1^θ to the current state $Y_1 = \overline{Y} + \varepsilon_1$, given by

$$K_1^\theta = \alpha^\theta \varepsilon_1.$$

Given this conjecture, C_2 then follows a conditionally normal distribution

$$C_2 \sim N(\mu_{2|1}^C, \sigma^2), \text{ where } \mu_{2|1}^C \equiv \overline{Y} + \alpha^\theta \varepsilon_1, \tag{5}$$

so that the conditional mean $\mu_{2|1}^C$ is proportional to the current ε_1 realization due to the (for now given) *response* α^θ of K_1. Equation (5) defines the true distribution, which we denote by $h(\tilde{C}_2 | \mu_{2|1}^C = \tilde{\mu}_{2|1}^C)$, associated with some realization $\tilde{C}_2$, for a given α^θ and current realization of $\tilde{\mu}_{2|1}^C$.[6] We use 'tildes' (when needed for a sharper formalism) to indicate the specific realization of any given random variables.

DE and the Representativeness Heuristic. In formulating the DE distortion, we build on the work of Gennaioli and Shleifer (2010) and Bordalo et al. (2018). The fundamental

[5] More specifically, $b > 0$ and $C < b$ so that utility is increasing in consumption in that region.

[6] We choose to present the analysis for now in terms of the conditional mean as a random variable to showcase the conceptual generality of the argument, even though the specific environment is purposefully kept simple for cleaner analytics with iid shocks. Note that if the exogenous labor income would be persistent, the conditional mean $\mu_{2|1}^C$ would also load, through that exogenous persistence, on the current income realization.

psychological first-principle basis for this belief model is that due to limited and selective memory retrieval, an agent's probability assessment is overweighted by event realizations that are "representative," in the sense of the Kahneman and Tversky (1972) representativeness heuristic of probabilistic judgments. This heuristic has been motivated and documented by a large psychology and experimental literature (see recently Bordalo et al. (2020a) and more broadly Bordalo et al. (2018)). The basic intuition behind this heuristic and the associated DE model is that the judged probability of an otherwise uncertain event partly reflects its "true," objective, frequency, as well as a subjective element that reflects the accessibility of that event in the agent's working memory. When new information arrives, the agent's memory process does not costlessly collect all past available data to form the probability judgment, conditional on the past and new data, but instead selectively recalls more (less) past events that are more (less) associated with, or representative of, the current news.

In particular, following this work means, for our context given above, modeling the distortion in beliefs arising from the representativeness heuristic as the density $h^{\theta}(\tilde{C}_2)$

$$h^{\theta}(\tilde{C}_2) = h(\tilde{C}_2|\mu^C_{2|1} = \tilde{\mu}^C_{2|1}) \left[\frac{h(\tilde{C}_2|\mu^C_{2|1} = \tilde{\mu}^C_{2|1})}{h(\tilde{C}_2|\mu^C_{2|1} = \mathbb{E}_{1-J}\mu^C_{2|1})} \right]^{\theta} \frac{1}{a}, \tag{6}$$

where a is an integration constant, ensuring that $h^{\theta}(\tilde{C}_2)$ integrates to one.

There are three important elements in this distorted distribution. First, as introduced above, $h(\tilde{C}_2|\mu^C_{2|1} = \tilde{\mu}^C_{2|1})$ is the true density. Second, $\mathbb{E}_{1-J}\mu^C_{2|1}$ is the *comparison group* for the random variable $\mu^C_{2|1}$, where $\mathbb{E}_{1-J}$ denotes the expectation operator under the true density conditional on the information set J periods ago from this generic time 1. This comparison group gives the state prevailing if there is no news, compared to the immediate ($J = 1$), or more distant past ($J > 1$). In our example above, due to the iid assumption of the income shocks this comparison group takes the simple form $\mathbb{E}_{1-J}\mu^C_{2|1} = \overline{Y}$. Third, here the parameter $\theta \geq 0$ measures the severity of the distortion. When $\theta = 0$, the agent's memory retrieval is perfect and beliefs collapse to the standard frictionless model. When $\theta > 0$, memory is limited and the agent's judgments are shaped by representativeness. As introduced intuitively above, this formulation captures the notion that the agent has the true distribution in the back of her mind, but selectively retrieves and overweighs realizations $\tilde{C}_2$ that are representative (or diagnostic) of the group consisting of $\left\{\mu^C_{2|1} = \tilde{\mu}^C_{2|1}\right\}$ relative to the comparison group consisting of $\left\{\mu^C_{2|1} = \mathbb{E}_{1-J}\mu^C_{2|1}\right\}$. Because $h^{\theta}(\tilde{C}_2)$ overweighs the most diagnostic future outcomes, Bordalo et al. (2018) call these expectations *diagnostic*.

While it is in general difficult to characterize analytically the distorted distribution, Bordalo et al. (2018) show how the *normality* assumption over the true density (in our case appearing in equation (5) from the normality of ε_2), leads to a tractable characterization of the conditional distribution $h^\theta(.)$. In particular, it remains normally distributed, with the same variance σ^2, but a distorted mean

$$\mathbb{E}_1^\theta(C_2) = \tilde{\mu}_{2|1}^C + \theta \left(\tilde{\mu}_{2|1}^C - \mathbb{E}_{1-J} \mu_{2|1}^C \right), \tag{7}$$

where the extra term $\theta(\tilde{\mu}_{2|1}^C - \mathbb{E}_{1-J}\mu_{2|1}^C)$ captures the *over-reaction* of the conditional mean to the new information.

2.2 Optimal savings choice given DE beliefs

We have thus characterized the DE beliefs, given a response α^θ of savings entering next period. The agent's problem in equation (2) is to optimally choose this response, given the resulting DE beliefs, a choice that we now characterize. In particular, under the conjectured α^θ we substitute in the tradeoff of equation (4) the resulting C_1 from the time 1 budget constraint, the distorted conditional expectation $\mathbb{E}_1^\theta(C_2)$ of equation (7), and the conditional mean $\mu_{2|1}^C$ specific to this environment (see equation (5)), to obtain

$$\varepsilon_1(1-\alpha^\theta) = \mathbb{E}_1 \left[\varepsilon_2 + \alpha^\theta \varepsilon_1 \right] + \theta \left[\mathbb{E}_1 \left(\varepsilon_2 + \alpha^\theta \varepsilon_1 \right) - \mathbb{E}_{1-J} \left(\varepsilon_2 + \alpha^\theta \varepsilon_1 \right) \right].$$

Since income innovations are unpredictable, the optimal response α^θ solves

$$\varepsilon_1(1-\alpha^\theta) = \alpha^\theta(1+\theta)\varepsilon_1, \tag{8}$$

which obtains the result summarized by Proposition 1 below.

Proposition 1. *The optimal marginal propensity α^θ to save out of a transitory income shock ε_1 under DE is lower than under the RE (i.e. perfect memory) policy and decreases with θ*

$$\alpha^\theta = \frac{1}{2+\theta} < \alpha^{RE} = \frac{1}{2}. \tag{9}$$

The optimal response under perfect memory (or 'rational expectations', RE), obtained in equation (8) in the limiting case of $\theta = 0$ is immediately equal to $\alpha^{RE} = 0.5$. Indeed, the RE agent saves half of the current income shock to achieve a perfectly flat expected consumption profile under RE ($C_1^{RE} = \mathbb{E}_1 C_2^{RE}$) since then $C_1^{RE} = \overline{Y} + 0.5\varepsilon_1$ and $C_2^{RE} = \overline{Y} + \varepsilon_2 + 0.5\varepsilon_1$.

Under DE, when the income innovation ε_t is unusually high (low), the agent seems to save too little (much), compared to the RE agent. While puzzling from the perspective of an outsider that evaluates the future under RE, this behavior is optimal under DE. Indeed, as in the RE case, the expected consumption profile achieved by K_1^θ and evaluated under DE, is perfectly flat, since $C_1^\theta = Y_1 - K_1^\theta$, and thus by Proposition 1

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta = \overline{Y} + \frac{1+\theta}{2+\theta}\varepsilon_1, \tag{10}$$

where we use the θ superscript to denote resulting choices under DE. The actual average consumption tomorrow however, under the true distribution and the chosen K_1^θ, is $\mathbb{E}_1 C_2^\theta =$

$\mathbb{E}_1\left(Y_2 + K_1^\theta\right) = \overline{Y} + (2+\theta)^{-1}\varepsilon_1$, responding less to the income shock ε_1 by a factor of $(1+\theta)$ than expected under DE in equation (10).

The intuition for this result is at the heart of the *endogenous predictability* mechanism that we emphasize in this section. Here $Y_2 + K_1^\theta = C_2^\theta$ is the random future financial resources available, and by the budget constrain also future consumption. In particular, given a current unusually high (low) income shock ε_1 and thus level of assets K_1^θ, the agent correctly realizes that her future available resources and consumption are more likely to be high (low) than usual, where the residual uncertainty about C_2^θ comes from the stochasticity in Y_2. In the case where Y_2 is iid, as analyzed here, this conditional predictability of future resources comes just from K_1^θ, which through the response α^θ induces that *endogenous* persistence from ε_1 to the random variable C_2^θ. An agent subject to the representativeness heuristic is then *overly influenced* by her perception of the new information contained in this unusual state of high (low) expected resources $\overline{Y} + K_1^\theta$. Due to her imperfect memory, she recalls more vividly state realizations that are representative in light of this new information. The over-influence of this new information contained in ε_1 and thus K_1^θ means that she inflates, compared to the true distribution, the likelihood of future resources $Y_2 + K_1^\theta$ to be high (low), while she deflates the likelihood of states characterized by low (high) future resources.

High MPC. Therefore, given high (low) assets today, the agent is more optimistic (pessimistic) than usual about future available resources, and importantly, more than under the true distribution. Thus, a larger current income than usual leads the agent to make saving decisions under an "as if" optimistic view of future resources. Given this view, the agent optimally consumes more and saves less today than the RE agent, resulting in a *high* marginal propensity to consume (MPC), i.e. $1-\alpha^\theta$, and a lack of consumption smoothing from the point of view of an external observer. Importantly, this apparent puzzling behavior of a lack of consumption smoothing has been well documented by a large empirical literature (eg. see Jappelli and Pistaferri (2010) for a survey), which finds that the MPC out of unexpected *temporary* income shocks is *puzzlingly large*, even for agents that are not financially constrained.[7]

2.3 Uncertainty and distorted beliefs

Having described the main mechanism behind the joint determination of DE beliefs and optimal actions in the presence of endogenous states, we now make a couple of general remarks on our approach, based on the observation that the distorted density h^θ formulated

[7] While liquidity frictions are used to account for the high MPC of rich but liquidity-constrained agents (eg. Kaplan and Violante (2014)), Kueng (2018), Fagereng et al. (2020) and McDowall (2020) provide evidence that even agents with high liquid wealth have significantly higher MPCs than implied by standard models.

in equation (6), has two joint appealing properties, as emphasized by Bordalo et al. (2018): (1) through its over-weighing of diagnostic information, this density captures the role of selective memory recall in affecting beliefs, and (2) this formulation is particularly convenient to employ when the processes over which it applies are conditionally normal.

Deterministic Processes and Residual Uncertainty. First, we reiterate that the representativeness heuristic and the DE belief is a model of imperfect probability judgments, or, in statistical terms, one of forecasting an otherwise uncertain event. Gennaioli and Shleifer (2010) describe this heuristic as one where a decision-maker's memory influences the likelihood judgment of possible scenarios (i.e. missing data) in light of some new data, but still with some *residual uncertainty* remaining about that otherwise missing data. In this context, it follows that if the agent is *only* interested in forecasting (or in statistical terms 'now-casting') at time t a predetermined variable like K_t, conditional on time t information, then that conditional belief is simply its current observed value.[8] In our example, this means

$$\mathbb{E}_1^{\theta}(K_1^{\theta}) = \mathbb{E}_1(K_1^{\theta}) = K_1^{\theta}. \tag{11}$$

Intuitively, the representativeness heuristic does not influence behavior in this case, since memory recall is not activated when the new data completely eliminates uncertainty over the variable to be forecasted (as it does in this case for K_1^{θ}).[9]

However, the sheer presence of income shocks, ε_2 even if just iid in our model, activates the need of memory recall in forecasting the stochastic future consumption C_2 so that

$$\mathbb{E}_1^{\theta}(K_1^{\theta} + Y_2) \neq \mathbb{E}_1(K_1^{\theta}) + \mathbb{E}_1^{\theta}(Y_2).$$

This inequality is the behavioral manifestation of DE beliefs over a sum of random variables of which one is predetermined. While this formal result has appeared already in Bordalo et al. (2018), our analysis brings it forward as a key implication of imperfect memory recall that will help us in incorporating DE beliefs in dynamic macroeconomic models, which feature endogenous and thus predetermined variables, like K_1^{θ}.

Primitives and Distorted Beliefs. Second, as illustrated by equation (7), we have made use of the tractability arising in a model with Gaussian shocks where perceived tradeoffs are linear, thus maintaining conditional normality. However, in more general cases, the marginal value inside the operator $\mathbb{E}_1^{\theta}$ in equation (3) will not be conditionally normal.

[8] In the language developed in Bordalo et al. (2018), to compute $\mathbb{E}_1^{\theta}(K_1^{\theta})$, its observed realization constitutes its infinitely representative state (see appendix in Bordalo et al. (2018) on Corollary 1).

[9] The same observation immediately implies that if $\sigma^2 = 0$ in the conditional distribution of equation (5), all residual uncertainty about C_2 would then be eliminated, as it is known to equal $C_2 = \overline{Y} + K_1^{\theta}$. Thus, in this case, similar to equation (11), $\mathbb{E}_1^{\theta}(C_2) = \mathbb{E}_1(C_2)$ and the resulting optimal $\alpha^{\theta} = \alpha^{RE}$. In contrast to this result, and general approach of how predetermined variables matter in forecasting, L'Huillier et al. (2021) assume that DE activates even absent any such residual uncertainty.

Indeed, in the class of models we analyze in Section 4, it is the *log-linearized* Euler equations that have this property, involving log-linear deviations (from steady state) of variables, such as future consumption or inflation, which linearly load on Gaussian shocks. We exploit the convenient formulation of the representativeness heuristic based on the density h^{θ} in equation (6) by applying it on those relevant Gaussian objects that enter into the log-linearized perceived tradeoffs, leading to distorted expectations that resemble equation (7).

Finally, note that based on the distorted beliefs underlying these perceived tradeoffs, the conditional utilities are also immediately evaluated under those densities. For example, in this model, the conditional belief $\mathbb{E}_1^{\theta} u(C_2)$ in (2) is influenced by the representativeness heuristic by being evaluated under the distorted conditional density of C_2, i.e. $h^{\theta}(\tilde{C}_2)$ defined in (6). Technically, this evaluation will generally differ from applying the formulation of h^{θ} in equation (6) to $u(C_2)$ itself, since the quadratic utility and normality of C_2 imply that $u(C_2)$ follows a χ^2-distribution, for which a tractable description similar to that in (7) is not readily available. Thus, while the representativeness heuristic clearly applies to both the conditional utilities and marginal tradeoffs, the modeling choice is where to leverage the convenient functional representation of equation (6). Our primitive approach, in line with what proposed in Bordalo et al. (2018), consists of emphasizing the role of the representativeness heuristic in distorting the perceptions of the marginal tradeoffs, and through that building the implied distribution for other objects of interest, such as $\mathbb{E}_1^{\theta} u(C_2)$. Overall, we thus find the direct modeling of perceptions of linearized marginal tradeoffs as distorted by the density h^{θ} as formulated in equation (6) appealing because: (a) in linearized models these perceptions guide actual (marginally driven) decisions, and (b) in standard Gaussian environments these tradeoffs can be tractably characterized, a feature that we leverage throughout the paper.

3 Dynamics and Imperfect Memory Recall

The two-period model analyzed above showcases how DE matter when the agent optimally chooses an action that creates endogenous predictability in the perceived future evolution of the relevant states. As we discuss in this section, a key challenge in taking the model to multi-periods is the *failure of the Law of Iterated Expectations (LOIE) under distant memory*. The reason that LOIE is important in dynamic models is that its failure generally leads to time-inconsistency, a property that we confront in this section.

3.1 Failure of the LOIE under Distant Memory

We first present the issue of the LOIE in isolation, for a *given* path of the forecasted random variable of interest, which we generically take here as consumption C_t at different dates.

Consider some arbitrary periods $t > J$, integers $m \geq 1, n \geq 1$, and some comparison group

$t-J$, where $J \geq 1$. The same formalism that lead to equation (7) can be extended to (as discussed in Bordalo et al. (2018), Corollary 1)

$$\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right] = \mathbb{E}_t^\theta \left[\mathbb{E}_{t+m} C_{t+m+n} + \theta \left(\mathbb{E}_{t+m} C_{t+m+n} - \mathbb{E}_{t+m-J} C_{t+m+n}\right)\right].$$

Applying the DE distortion at time t, the RHS further becomes

$$\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right] = (1+\theta) \left[\mathbb{E}_t C_{t+m+n} + \theta \left(\mathbb{E}_t C_{t+m+n} - \mathbb{E}_t \mathbb{E}_{t+m-J} C_{t+m+n}\right)\right] - \theta \mathbb{E}_{t-J} C_{t+m+n}.$$

We are then interested in establishing whether the LOIE holds, as implied by

$$\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right] = \mathbb{E}_t^\theta \left[C_{t+m+n}\right]. \tag{12}$$

Lemma 1. *For a given m, the LOIE holds under DE if and only $J \leq m$.*

To prove this, the key term in $\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right]$ that matters is the perceived surprise

$$\mathbb{E}_t C_{t+m+n} - \mathbb{E}_t \mathbb{E}_{t+m-J} C_{t+m+n}. \tag{13}$$

Consider first the case of $J \leq m$. Then the time t information set is a *subset of the future* time $(t+m-J)$ information set and we can apply LOIE under the true process, which holds given that $\mathbb{E}_t \mathbb{E}_{t+m-J} C_{t+m+n} = \mathbb{E}_t C_{t+m+n}$ for $J \leq m$. It follows that the surprise in equation (13) is zero and the LOIE holds under the DE operator:

$$\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right] = \mathbb{E}_t C_{t+m+n} + \theta \left(\mathbb{E}_t C_{t+m+n} - \mathbb{E}_{t-J} C_{t+m+n}\right) = \mathbb{E}_t^\theta \left[C_{l+m+n}\right]. \tag{14}$$

In contrast, suppose that $J > m$. In that case, the conditioning time t information set *includes the past* time $(t+m-J)$. Therefore, the perceived surprise (13) is not zero and constitutes an additional source of variation for $\mathbb{E}_t^\theta \left[\mathbb{E}_{t+m}^\theta C_{t+m+n}\right]$ in equation (14), which now becomes

$$(1+\theta) \left[\mathbb{E}_t C_{t+m+n} + \theta \left(\mathbb{E}_t C_{t+m+n} - \mathbb{E}_{t+m-J} C_{t+m+n}\right)\right] - \theta \mathbb{E}_{t-J} C_{t+m+n}$$

Thus, for the generic case of $\mathbb{E}_t C_{t+m+n} \neq \mathbb{E}_{t+m-J} C_{t+m+n}$, the LOIE as stated in equation (12) does not hold.

Intuitively, when the lag J of the reference distribution exceeds the forecast horizon m, taking the time t expectation over the $t+m$ DE forecast of C_{t+m+n} introduces an *additional* lagged forecast (here $\mathbb{E}_{t+m-J} C_{t+m+n}$) which would not be otherwise included in the time t DE forecast of C_{t+m+n} itself. This case of $J > m$ is not just a theoretical curiosity. For example Bordalo et al. (2020b) find that values of $J = 4$ quarters seem to account well for the empirical over-reaction observed in the surveys of professional forecasters, while Bordalo et al. (2019b) argue that $J = 11$ quarters explains well the sluggishness in expected returns.

The analysis above also clarifies the important role of agents' selective memory process in building the comparison group. In particular, we note that the LOIE holds under DE only when $J = 1$. Indeed, in that case, the term in equation (13) necessarily becomes zero, since $\mathbb{E}_t \mathbb{E}_{t+m-1} C_{t+m+n} = \mathbb{E}_t C_{t+m+n}$ for any $m \geq 1$, $n \geq 1$. Intuitively, the current DE forecast of any future conditional DE belief does not bring in any further lagged information than the time $t-1$ information, rendering it equivalent to the DE belief $\mathbb{E}_t^\theta C_{t+m+n}$.

3.2 A Three Period Consumption-Savings Problem

Our analysis so far has emphasized two generic DE properties (*endogenous predictability* and the *failure of the LOIE under distant memory*). We reiterate that the formalism behind these properties is not novel to our paper. It has been noticed and proposed as characterizing the DE operator appearing in equation (6) by previous work, such as Bordalo et al. (2018). Our key contribution here is thus to bring these properties forward as insightful and promising ways to study: (1) the role of DE beliefs over exogenous and endogenous variables in dynamic models, and (2) how the role of past memory introduces additional informational state variables that can alter significantly the model's dynamics. To do so, we now extend the consumption-savings model of Section 2 to include a third period, which allows us to capture the intuition on how endogenous states matter in jointly distorting beliefs and actions, with a particular emphasis on the role of distant memory.

In this extension, the time 1 problem is now to choose actual savings K_1^θ (as a function of K_0 and ε_1) and a contingent plan $K_2^{\theta,p}$ (as a function of K_1^θ and ε_2) so to maximize current utility and the expected discounted sum of future utilities (recall that $\beta = 1$)

$$\max_{K_1^\theta, K_2^{\theta,p}} \left\{ u(C_1^\theta) + \mathbb{E}_1^\theta \left[u(C_2^{\theta,p}) + u(C_3^{\theta,p}) \right] \right\}, \tag{15}$$

$$s.t.\ C_1^\theta = Y_1 + K_0 - K_1^\theta(K_0, \varepsilon_2)$$
$$C_2^{\theta,p} = Y_2 + K_1^\theta(K_0, \varepsilon_1) - K_2^{\theta,p}(K_1^\theta, \varepsilon_2);\ C_3^{\theta,p} = Y_3 + K_2^{\theta,p}(K_1^\theta, \varepsilon_2) - K_3,$$

where end-of-life savings K_3 is optimally set to zero.

The key source of possible time-inconsistency is that at time 2, conditional on K_1^θ and ε_2, the agent re-optimizes over her initially planned $K_2^{\theta,p}$, by looking for a K_2^θ that solves

$$\max_{K_2^\theta} \left[u(C_2^\theta) + \mathbb{E}_2^\theta u(C_3^\theta) \right], \tag{16}$$

where $C_2^\theta = Y_2 + K_1^\theta - K_2^\theta(K_1^\theta, \varepsilon_2)$ and $C_3^\theta = Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2) - K_3$.

LOIE and Perceived Tradeoffs. As we have detailed in section 3.1, the LOIE for the two-step-ahead expectation holds if and only if $J = 1$. We show below that this property is intimately linked to time-inconsistency between planned and actual future choices. In particular, for a given time 1 policy $K_1^\theta(K_0, \varepsilon_1)$ we can establish the following Proposition.[10]

Proposition 2. *The conditional time-2 optimal solution $K_2^\theta(K_1^\theta, \varepsilon_2)$ is identical ('time-consistent') to the time-1 optimal contingent plan $K_2^{\theta,p}(K_1^\theta, \varepsilon_2)$ if and only if $J = 1$.*

[10]Proofs for the formal results of the remaining Lemmas and Propositions are in the Appendix.

While the optimal time-1 plan $K_2^{\theta,p}$ in equation (15) is set such that $\mathbb{E}_1^\theta\left[C_2^{\theta,p} - C_3^{\theta,p}\right] = 0$, the conditional optimal K_2^θ solves the time-2 perceived tradeoff $C_2^\theta - \mathbb{E}_2^\theta C_3^\theta = 0$. Critically, we show, as part of the proof, that the time-1 perceived consumption smoothing between C_2 and C_3, under the given $K_1^\theta(K_0, \varepsilon_1)$ and that optimal policy $K_2^\theta(K_1^\theta, \varepsilon_2)$, is

$$\mathbb{E}_1^\theta\left[C_2^\theta - C_3^\theta\right] = (1+\theta)\theta\left[\mathbb{E}_1 C_3^\theta - \mathbb{E}_{2-J} C_3^\theta\right], \text{ if } J > 1, \tag{17}$$

and equals to zero only when $J = 1$. Thus, the conditional optimal K_2^θ implements exactly the time-1 desired consumption path under $K_2^{\theta,p}$ if and only if $J = 1$. When $J > 1$, the time-2 decision under DE is based on $\mathbb{E}_{2-J} C_3^\theta$, which similar to the LOIE property of Lemma 1, introduces a different information set than its time-1 forecast $\mathbb{E}_1 C_3^\theta$. By equation (17) and Proposition 2, this difference in information sets leads, when memory is based on more distant past, to a misalignment of intertemporal perceived tradeoffs and thus to time-inconsistency.

3.3 Time Consistency when Memory is Based on Immediate Past

When memory recall is based on the immediate past, i.e. $J = 1$, then the savings plan $K_2^{\theta,p}$ under DE is time-consistent, and thus equal to K_2^θ, per Proposition 2. We now characterize the resulting optimal DE saving functions K_1^θ and K_2^θ when $J = 1$.

In particular, we first conjecture that the optimal policy functions under RE are

$$K_1^{RE} = \alpha_{K_0}^{RE} K_0 + \alpha_{\varepsilon_1}^{RE}\varepsilon_1;\ K_2^{RE} = \alpha_{K_1}^{RE} K_1^{RE} + \alpha_{\varepsilon_2}^{RE}\varepsilon_2. \tag{18}$$

while the optimal policy functions K_1^θ and K_2^θ are given by

$$K_1^\theta = \alpha_{K_0}^\theta K_0 + \alpha_{\varepsilon_1}^\theta\varepsilon_1;\ K_2^\theta = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta\varepsilon_2. \tag{19}$$

Proposition 3. *When $J = 1$, compared to the RE policy functions K_1^{RE} and K_2^{RE}, the optimal policy functions K_1^θ and K_2^θ feature the same optimal response to the endogenous state but a muted response to the current income innovation, i.e.*

$$\alpha_{K_0}^\theta = \alpha_{K_0}^{RE} = \frac{2}{3}; \alpha_{K_1}^\theta = \alpha_{K_1}^{RE} = \frac{1}{2}.$$
$$\alpha_{\varepsilon_1}^\theta = \frac{2}{3+\theta} < \alpha_{\varepsilon_1}^{RE} = \alpha_{K_0}^{RE};\ \alpha_{\varepsilon_2}^\theta = \frac{1}{2+\theta} < \alpha_{\varepsilon_2}^{RE} = \alpha_{K_1}^{RE}.$$

To see the intuition, first consider the time 2 problem in (16), where the optimal K_2^θ solves the tradeoff $C_2^\theta = \mathbb{E}_2^\theta C_3^\theta$. When $J = 1$, for a given state K_1^θ and exogenous innovation ε_2, by using the time 2 and 3 budget constraint, this tradeoff amounts to

$$\varepsilon_2 + K_1^\theta - K_2^\theta = \mathbb{E}_2(\varepsilon_3 + K_2^\theta) + \theta\left[\mathbb{E}_2(\varepsilon_3 + K_2^\theta) - \mathbb{E}_1\left(\varepsilon_3 + K_2^\theta\right)\right]. \tag{20}$$

Under the conjecture in equation (19), the perceived surprise at time 2, $\mathbb{E}_2 C_3^\theta - \mathbb{E}_1 C_3^\theta$, just equals the (endogenous) exposure of K_2^θ to ε_2. The over-reaction of this new information affects the DE beliefs by a factor θ. By substituting the conjectured coefficients $\alpha_{K_1}^{\theta}$ and $\alpha_{\varepsilon_2}^{\theta}$ into the tradeoff (20), we obtain their values characterized in Proposition 3.

The key economic observation here is that when $J = 1$, the economic state K_1^{θ} also serves as the *necessary and sufficient* conditioning information to form the comparison group $\mathbb{E}_{2-J}(\varepsilon_3 + K_2^{\theta})$, i.e. the object that controls the agent's selective memory according to the representativeness heuristic. Therefore, the DE beliefs' over-reaction to the new information, $K_2^{\theta} - \mathbb{E}_{2-J}K_2^{\theta}$, only contains the current innovation ε_2 and not the endogenous state K_1^{θ}. This over-sensitivity of beliefs to ε_2 leads to a behavior where the response to the state is the same as for the RE solution ($\alpha_{K_1}^{\theta} = \alpha_{K_1}^{RE}$), while the response to the exogenous income shock is muted ($\alpha_{\varepsilon_2}^{\theta} < \alpha_{\varepsilon_2}^{RE}$). The reason for the latter effect is the same as in the two-period model (see equation (9) and the discussion around Proposition 1).

We now move back to the time 1 problem in (15), where K_1^{θ} solves

$$C_1 = \mathbb{E}_1^{\theta}\left[C_2^{\theta,p} + \frac{\partial K_2^{\theta,p}}{\partial K_1^{\theta}}\left(C_3^{\theta,p} - C_2^{\theta,p}\right)\right]. \tag{21}$$

Intuitively, the benefit of higher K_1^{θ} involves the direct effect of increasing consumption tomorrow and the indirect effect of affecting consumption smoothing between period 2 and 3 $(C_3^{\theta,p} - C_2^{\theta,p},)$ through the optimal plan $K_2^{\theta,p}(K_1^{\theta}, \varepsilon_2)$, which recall that here coincides with the actual choice at time 2, $K_2^{\theta}(K_1^{\theta}, \varepsilon_2)$. The tradeoff in equation (21) can be broken in[11]:

$$C_1 = \mathbb{E}_1^{\theta}C_2^{\theta,p} + \alpha_{K_1}^{\theta}\mathbb{E}_1^{\theta}\left[C_3^{\theta,p} - C_2^{\theta,p}\right].$$

By the time-consistency established in Proposition 2 when $J = 1$, the term $\mathbb{E}_1^{\theta}\left[C_3^{\theta,p} - C_2^{\theta,p}\right] = 0$, under both the plan and the anticipation of the future choice K_2^{θ}. Thus, the tradeoff in equation (21) becomes $C_1 = \mathbb{E}_1^{\theta}C_2^{\theta}$, where we have already characterized the K_2^{θ} policy. We can apply a similar logic and procedure as for finding K_2^{θ} above to show that the response of K_1^{θ} to the state K_0 is the same as for the RE solution ($\alpha_{K_0}^{\theta} = \alpha_{K_0}^{RE}$), while the response to the exogenous income shock is muted ($\alpha_{\varepsilon_1}^{\theta} < \alpha_{\varepsilon_1}^{RE}$) and decreasing with θ.

3.4 Beliefs over Future Actions

When $J > 1$ Proposition 2 shows the time-inconsistency between the planned K_2 and what the agent believes she will actually choose for K_2 once time 2 arrives. In looking for the agent's current optimal action we then need to model her current beliefs about her future actions when faced with this inherent time-inconsistency. To build a coherent model of belief formation, that allows us to study the interaction of *endogenous predictability* and the *failure*

[11]Since C_2 and C_3 are conditionally normal and both have residual uncertainty as of time 1, due to the normally distributed income shocks and the conjectured K_1^{θ} and K_2^{θ}, the DE operator is additive over these two random variables (see eg. Corollary 1 in Bordalo et al. (2018) for details on DE additivity).

of the LOIE under distant memory, we use insights from the microeconomic theory (e.g. the seminal work by Strotz (1955) and Pollak (1968)) that point to two different assumptions regarding agents' current belief about future selves' behavior.

Naïveté. The first approach, coined in this literature as *naïveté* (in the sense of O'Donoghue and Rabin (1999) and used for example in Akerlof (1991)), models an agent who does not forecast her future self's behavior to be governed by the representativeness heuristic. Her time 1 problem is now

$$\max_{K_1^{\theta,n}} \left\{ u(C_1^{\theta}) + \mathbb{E}_1^{\theta} \left[u(C_2^{RE}) + u(C_3^{RE}) \right] \right\} \tag{22}$$

where the agent at time 1 believes her time 2 future self will take the action K_2^{RE} so to

$$\max_{K_2^{RE}} \left[u(C_2^{RE}) + \mathbb{E}_2 u(C_3^{RE}) \right]. \tag{23}$$

The θ-superscript and RE-superscript on a time t variable signify choices that are made under a DE and RE policy function, respectively, taking as given the state variable entering that period. From the budget constraints, the (forecasted) consumption choices are therefore

$$C_1^{\theta} = Y_1 + K_0 - K_1^{\theta,n}(.);\ C_2^{RE} = Y_2 + K_1^{\theta,n}(.) - K_2^{RE}(.); C_3^{RE} = Y_3 + K_2^{RE}(.) - K_3^{RE}(.),$$

where $K_1^{\theta,n}(.)$ (and $K_2^{RE}(.)$) signify the choice resulting from a DE under naïveté (and RE, respectively) policy function that solve (22) (and (23), respectively) and trivially $K_3^{RE}(.) = 0$.[12]

While these are her beliefs at time 1 looking ahead, entering period 2 with the state realization K_1^{θ} and new information determined at time 2, her problem is actually influenced by the representativeness heuristic, so her conditionally optimal action is

$$\max_{K_2^{\theta,n}} \left[u(C_2^{\theta}) + \mathbb{E}_2^{\theta} u(C_3^{RE}) \right]. \tag{24}$$

where $C_2^{\theta} = Y_2 + K_1^{\theta} - K_2^{\theta,n}(.)$ and $C_3^{RE} = Y_3 + K_2^{\theta,n}(.) - K_3^{RE}(.)$.

The behavioral interpretation of equations (22), (23) and (24) is that, at time 1, the agent maximizes assuming that after time 2 she will not be subject to any heuristics driving her memory recall (i.e. she will act 'fully rationally'), even though at time 2 she ends up changing behavior and be in fact subject to her otherwise imperfect memory recall.

Sophistication. The second typical approach in modeling agent's beliefs over future behavior is to consider *sophistication* (eg. Laibson (1997)). Entering period 2, her problem is

$$\max_{K_2^{\theta,s}} \left[u(C_2^{\theta}) + \mathbb{E}_2^{\theta} u(C_3^{\theta}) \right] \tag{25}$$

[12] There is no material distinction between $K_3^{RE}(.)$ and $K_3^{\theta}(.)$ since they both equal zero. We refer to them separately to highlight their conceptual difference as being taken under different beliefs.

where now

$$C_2^\theta = Y_2 + K_1^{\theta,s} - K_2^{\theta,s}(.); C_3^\theta = Y_3 + K_2^{\theta,s}(.) - K_3^\theta(.). \tag{26}$$

Sophistication means that at time 1 the agent understands that her future action is dictated by equation (25) (as well as $K_3^\theta(.) = 0$). Thus, the sophisticated agent solves

$$\max_{K_1^{\theta,s}} \left\{ u(C_1^\theta) + \mathbb{E}_1^\theta \left[u(C_2^\theta) + u(C_3^\theta) \right] \right\}, \tag{27}$$

where current $C_1^\theta = Y_1 + K_0 - K_1^{\theta,s}(.)$, while C_2^θ and C_3^θ are determined as in (26).

Comparison Groups for Memory Retrieval. In forming the DE beliefs of the naïve and sophisticated agents we also have to formulate the comparison groups that enter those DE beliefs. Consider first a naïve agent and take for example period 2, when the agent forms the forecast over time 3 consumption, as $\mathbb{E}_2^\theta(Y_3 + K_2^{\theta,n})$. At any past date before time 2, the defining characteristic of the naïve agents is that they expect their future selves to act under the RE savings policy function. To be consistent with these beliefs, we assume that their counterpart naïveté comparison group for a variable like $K_2^{\theta,n}$ at time 2 is $\mathbb{E}_{2-J} K_2^{RE}$, i.e. the conditional expectation made by the former self of the naïve agent as of J periods ago of the RE savings choice at time 2, under the true density. Consider now a sophisticated agent that forms the forecast $\mathbb{E}_2^\theta(Y_3 + K_2^{\theta,s})$. In contrast to the naïvete case and to continue to maintain belief consistency across selves, we assume that the sophisticated comparison group for $K_2^{\theta,s}$ is $\mathbb{E}_{2-J} K_2^{\theta,s}$, i.e. the conditional expectation of the DE savings choice at time 2 made J periods ago by the former sophisticated self, under the true density.

Euler Equations. Per the objective functions in (24) and (25), conditional on arriving in period 2 with the corresponding inherited savings from time 1 and observing new information ε_2, the optimal savings $K_2^{\theta,n}$ and $K_2^{\theta,s}$ solve the respective consumption smoothing problems

$$\varepsilon_2 + K_1^{\theta,n} - K_2^{\theta,n} = \mathbb{E}_2^\theta[\varepsilon_3 + K_2^{\theta,n}]; \;\; \varepsilon_2 + K_1^{\theta,s} - K_2^{\theta,s} = \mathbb{E}_2^\theta[\varepsilon_3 + K_2^{\theta,s}], \tag{28}$$

where the respective comparison groups for the $\mathbb{E}_2^\theta$ belief are discussed above.

In turn, the optimal solution for $K_1^{\theta,n}$ under naïveté solves the intertemporal tradeoff

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^{RE} + \frac{\partial K_2^{RE}}{\partial K_1^{\theta,n}} \left(C_3^{RE} - C_2^{RE} \right) \right], \tag{29}$$

while under sophistication

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^\theta + \frac{\partial K_2^{\theta,s}}{\partial K_1^{\theta,s}} \left(C_3^\theta - C_2^\theta \right) \right]. \tag{30}$$

We can characterize the indirect effect of internalizing that the current choice affects the future problem and decision. First, consider the naïve agent.

Lemma 2. *Given K_0 and ε_1, the perceived naïve consumption smoothing between C_2 and C_3, under $K_1^{\theta,n}(K_0, \varepsilon_1)$ and policy K_2^{RE}, is $\mathbb{E}_1^{\theta}\left[C_2^{RE} - C_3^{RE}\right] = 0$ for $J \geq 1$. The optimal naïve choice $K_1^{\theta,n}$ in equation (29) thus solves the one-step tradeoff $C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{RE}$ for $J \geq 1$.*

Under naïveté the time 1 perceived behavior of the future self at time 2 is to optimally select K_2^{RE}, which conditional on the states entering that period achieves $\mathbb{E}_2 C_3^{RE} - C_2^{RE} = 0$. Thus, the consumption profile $C_3 - C_2$ as perceived at time 1 in equation (29), equals just the income innovation ε_3, unpredictable under $\mathbb{E}_1^{\theta}$. This induced unpredictability as of time 1 means that the marginal effect of the choice K_1^{θ} as a relevant state for future conditional optimal choices can be ignored, conditional on the agent believing that the future self implements the K_2^{RE} policy. Critically, by anticipating future actions taken under perfect memory recall, this result holds for any $J \geq 1$. The second, indirect, effect in equation (29) thus disappears, leading to the Euler equation $C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{RE}$ in Lemma (2).

In contrast, the intertemporal tradeoff under more distant memory is more complicated for the current sophisticated agent, as follows.

Lemma 3. *The tradeoff in equation (30) solved by the time-1 sophistication optimal solution $K_1^{\theta,s}(K_0, \varepsilon_1)$ simplifies to the one-step ahead tradeoff, $C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{\theta}$, if and only if $J = 1$.*

Intuitively, when $J > 1$ the agent understands that she will act under a future policy function $K_2^{\theta,s}(K_1^{\theta,s}, \varepsilon_2)$ leading to a future consumption path that is not perceived as optimal as of time 1. Therefore her current optimal choice $K_1^{\theta,s}$ tries to remedy this imbalance by affecting the state of her future action. This is the extra term in (30) that affects the current choice of the sophisticated agent when $J > 1$ and otherwise disappears when $J = 1$.

J = 1 Case Revisited. Having introduced the conceptual distinction between naïveté and sophistication, we now show that when $J = 1$ they both recover the same, time-consistent policy functions (in turn described by Proposition 3).

Proposition 4. *When $J = 1$, the naïveté $K_1^{\theta,n}$ and $K_2^{\theta,n}$ and sophistication policy functions $K_1^{\theta,s}$ and $K_2^{\theta,s}$ are the same and recover the DE optimal choices based on time-consistency.*

First, note that when $J = 1$, as shown in Proposition 3, K_1^{θ} and K_2^{θ} respond to the state variables in the same way as the RE policy functions K_1^{RE} and K_2^{RE}, respectively. In that sense, both types of policy functions are time-consistent with respect to the endogenous state.[13] Based on this property it follows that $\mathbb{E}_0 K_1^{RE} = \mathbb{E}_0 K_1^{\theta}$ and $\mathbb{E}_1 K_2^{RE} = \mathbb{E}_1 K_2^{\theta}$ so the comparisons groups of the naïve problem and the time-consistent problem are identical. Thus, given the same belief formation and optimality condition, the naïve solution for $K_2^{\theta,n}$ in

[13]The DE and RE policies differ in their response to the innovation ε_2, but since that is mean zero it does not systematically affect the current expectation of future tradeoffs and actions.

equation (28) must recover the time-consistent policy at time 2 from equation (20). This identical response to the endogenous variables also means that the tradeoff $C_1^\theta = \mathbb{E}_1^\theta C_2^{RE}$ implied by the naïve solution $K_1^{\theta,n}$ (see Lemma 2) recovers the same solution K_1^θ as that of Proposition 3, where the agent was planning in a time-consistent way to follow the policy K_2^θ.

Second, when $J = 1$, by Proposition 2 the time-1 tradeoff under sophistication in equation (30) is identical to the time-consistent one $C_1^\theta = \mathbb{E}_1^\theta C_2^\theta$. It follows immediately that $K_1^{\theta,s}$ and $K_2^{\theta,s}$ are equal to the time-consistent and naïveté counterparts, by further noting that the comparison groups are identical across these models of belief formation when $J = 1$.

3.5 Dynamics with Memory Recall of More Distant Past

Our discussion indicates that there are two fundamental ways in which the representativeness heuristic affects current choices differently when memory recall is based on more distant rather than the immediate past: (1) the role of that distant past in the construction of the comparison groups, and (2) the role of anticipating future actions. Our naïve and sophisticated approaches have offered two model-coherent ways to analyze these issues. We study them below in the three period model when the comparison group is based on $J = 2$.

3.5.1 Role of Informational States for Comparison Groups

We first study the period 2 problem, which shows transparently the role of comparison groups, since there is no meaningful continuation utility to compute there. In the next subsection, we move backwards to period 1 and study the role of anticipating future actions.

Proposition 5. *When $J = 2$ the time-2 naïveté and sophisticated policy functions are*

$$K_2^{\theta,n} = \alpha^\theta_{\mathbb{E}_0 K_1}\mathbb{E}_0 K_1^{RE} + \alpha^\theta_{K_1} K_1^{\theta,n} + \alpha^\theta_{\varepsilon_2}\varepsilon_2.$$
$$K_2^{\theta,s} = \alpha^\theta_{\mathbb{E}_0 K_1}\mathbb{E}_0 K_1^{\theta,s} + \alpha^\theta_{K_1} K_1^{\theta,s} + \alpha^\theta_{\varepsilon_2}\varepsilon_2.$$

Compared to the $J = 1$ case, the optimal coefficients are characterized by (i) a positive loading on the past informational state, (ii) a muted response to the current economic state K_1, and (iii) an identical (while still muted), response to the current innovation, as follows:

$$\alpha^\theta_{\mathbb{E}_0 K_1} = \frac{\theta}{2(2+\theta)};\ \alpha^\theta_{K_1} = \frac{1}{2+\theta};\ \alpha^\theta_{\varepsilon_2} = \frac{1}{2+\theta}.$$

Let us detail some of the formalism and intuition behind this important result. We focus on the naïveté case and then argue that the sophistication case is identical in nature. Thus, given inherited $K_1^{\theta,n}$, consider $K_2^{\theta,n}$ that solves the time 2 problem in equation (31) as

$$\varepsilon_2 + K_1^{\theta,n} - K_2^{\theta,n} = K_2^{\theta,n} + \theta\left(K_2^{\theta,n} - \mathbb{E}_{2-J} K_2^{RE}\right). \tag{31}$$

The key difference between $J = 1$ and $J = 2$ is how the comparison group, $\mathbb{E}_{2-J}K_2^{RE}$, affects conditional beliefs. We have characterized in Proposition 3 the RE laws of motion,

$K_2^{RE} = \alpha_{K_1}^{RE} K_1^{RE} + \alpha_{\varepsilon_2}^{RE}\varepsilon_2$. Thus, the comparison group for the current naïve self is

$$\mathbb{E}_0 K_2^{RE} = \alpha_{K_1}^{RE}\mathbb{E}_0 K_1^{RE}. \quad (32)$$

where $K_1^{RE} = \alpha_{K_0}^{RE} K_0 + \alpha_{\varepsilon_1}^{RE}\varepsilon_1$ describes the counterfactual evolution of K_1 expected as of time 0 by the naïve self. Therefore, under the conjectured policy function $K_2^{\theta,n}$ in Proposition 5, the DE beliefs over-react by a factor of θ to the new information comprised of

$$K_2^{\theta,n} - \mathbb{E}_0 K_2^{RE} = \alpha_{K_1}^{\theta} K_1^{\theta,n} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \left(\alpha_{\mathbb{E}_0 K_1}^{\theta} - \alpha_{K_1}^{RE}\right)\mathbb{E}_0 K_1^{RE}.$$

By substituting this over-reaction in equation (31), we recover the optimal $K_2^{\theta,n}$ coefficients.[14]

The critical reason behind these novel state dynamics is that the economic state $K_1^{\theta,n}$ is *not* a sufficient statistic anymore (as it was when $J = 1$) for the comparison group. With $J = 2$, the conditional expectation $\mathbb{E}_0 K_1^{RE}$ forms a separate *informational state* that affects time 2 choices. Indeed, the higher the $\mathbb{E}_0 K_1^{RE}$, the higher is the comparison group for $K_2^{\theta,n}$ (since $\alpha_{K_1}^{RE} > 0$ in equation (32)) and thus the more the DE agent is typically 'disappointed' by the perceived innovation in the conditional mean of future consumption, given by $K_2^{\theta,n} - \mathbb{E}_0 K_2^{RE}$. Over-reacting to this negative innovation, the agent perceives less future resources (a higher future marginal utility at time 3), and hence invests more in period 2, explaining why the loading $\alpha_{\mathbb{E}_0 K_1}^{\theta}$ on the informational state $\mathbb{E}_0 K_1^{RE}$ is positive for $\theta > 0$ in Proposition 5. Of course, this over-reaction caused by imperfect memory recall is absent in the RE case, where $\mathbb{E}_0 K_1^{RE}$ does not matter for the choice K_2^{RE}.

The other manifestation of the separate role of $K_1^{\theta,n}$ as an economic state (savings entering this period) and an information state (affecting memory formation for building $\mathbb{E}_{2-J} K_2^{RE}$), is that now the response $\alpha_{K_1}^{\theta}$ is muted compared to the $J = 1$ and RE cases (recall Proposition 3). This separate role can also be seen by rewriting the solution for $K_2^{\theta,n}$ in Proposition 5 as

$$K_2^{\theta,n} = \left(\alpha_{K_1}^{\theta} + \alpha_{\mathbb{E}_0 K_1}^{\theta}\right) K_1^{\theta,n} - \alpha_{\mathbb{E}_0 K_1}^{\theta}\left(K_1^{\theta,n} - \mathbb{E}_0 K_1^{RE}\right) + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2. \quad (33)$$

The first part captures the role of $K_1^{\theta,n}$ as an economic state, which influences the $K_2^{\theta,n}$ decision in the same as it does for the RE policy function, so that $\alpha_{K_1}^{\theta} + \alpha_{\mathbb{E}_0 K_1}^{\theta} = \alpha_{K_1}^{RE} = 1/2$.

The information role is captured by the second term $\left(K_1^{\theta,n} - \mathbb{E}_0 K_1^{RE}\right)$. Consider, for example, an increase in $K_1^{\theta,n}$ caused by a positive innovation in ε_1 (a conjecture verified in Section 3.5.2). A higher $K_1^{\theta,n}$ than expected at time 0 under the relevant comparison group, leads to a perceived positive innovation in $K_2^{\theta,n} - \mathbb{E}_0 K_2^{RE}$. Since agents are over-influenced by this surprise, they become over-optimistic about future resources and invest less. This

[14]The high MPC out of transitory income shocks, given by $1 - \alpha_{\varepsilon_2}^{\theta}$, is the same as in Proposition 1. This naïveté case also shows that the mechanism is different from two recent related approaches. Lian (2020) shows that (partial) sophistication is key for an agent to decide to save less today out of anticipations of future mistakes. In Ilut and Valchev (2020) agents are similarly naïvé as here, but have uncertainty over their optimal consumption functions, which endogenously leads to stable beliefs characterized by high MPC.

explains why the innovation $\left(K_1^{\theta,n} - \mathbb{E}_0 K_1^{RE}\right)$ enters with a negative sign in equation (33). The total effect of how $K_2^{\theta,n}$ responds to $K_1^{\theta,n}$ is then given by $\alpha_{K_1}^{RE} - \alpha_{\mathbb{E}_0 K_1}^{\theta} = \alpha_{K_1}^{\theta}$ and explains why when $J = 2$ there is a muted response of time 2 savings to $K_1^{\theta,n}$ compared to the RE and DE policy function based on the immediate past (i.e. $\alpha_{K_1}^{\theta} < \alpha_{K_1}^{RE}$).[15]

The solution for the sophisticated choice $K_2^{\theta,s}$ follows the same logic as for $K_2^{\theta,n}$, leading to the result in Proposition 5 that the optimal coefficients are the same. The subtle difference here is the comparison group formation. The naïveté solution can leverage the law of motion for K_1^{RE}, so that $\mathbb{E}_0 K_1^{RE}$ can be immediately plugged in the determination of $K_2^{\theta,n}$ as $\alpha_{K_0}^{RE} K_0$. In contrast, the corresponding $\mathbb{E}_0 K_1^{\theta,s}$ is more difficult to transparently assess because it requires computing a feedback effect between the (yet to be determined) $K_1^{\theta,s}$ chosen by the time 1 sophisticated DE agent, which in turn is a function of expectations about $K_2^{\theta,s}$.

Overall, our analysis brought forward the novel role of past endogenous states as informational variables that affect how memory forms 'benchmark' (or comparison) views of what is currently perceived as unusually high or low expected future resources. An agent acting the representativeness heuristic over-reacts to these perceptions. Thus, savings choices made in the more distant past (like K_0) have an independent and novel effect for decisions today.

3.5.2 Anticipating Future Actions when Distant Past Matters

In formulating the optimal current action at $t = 1$ the agent has to form beliefs over future actions.[16] Consider first the naïveté case. By Lemma 2 her anticipated future policy K_2^{RE} implements a perceived consumption path between time 2 and 3 that is on average flat, as expected of time 1 under DE, which is precisely what her time-1 optimal plan would dictate. Thus, the optimal $K_1^{\theta,n}$ only involves setting $C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{RE}$, even when $J > 1$. This tradeoff looks like the one solved by $K_2^{\theta,n}$ at time 2, i.e. $C_2^{\theta} = \mathbb{E}_2^{\theta} C_3^{RE}$, but lagged one period, so the resulting qualitative properties of the optimal $K_1^{\theta,n} = \alpha_{\mathbb{E}_{-1}K_0}^{\theta,n} \mathbb{E}_{-1} K_0 + \alpha_{K_0}^{\theta,n} K_0 + \alpha_{\varepsilon_1}^{\theta,n} \varepsilon_1$ resemble those of $K_2^{\theta,n}$ in Proposition 5, as we detail in Proposition 6 in the Appendix.

In contrast, the sophistication case is significantly more complicated because the agent at time 1 would choose a different plan for K_2 than what she anticipates is her optimal conditional action at time 2. Therefore, by equation (30), her optimal action $K_1^{\theta,s}$ aims to fix this misalignment by affecting the state of her anticipated policy function $K_2^{\theta,s}$, and solve

$$C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{\theta} + \alpha_{K_1}^{\theta} \mathbb{E}_1^{\theta} \left(C_3^{\theta} - C_2^{\theta} \right), \tag{34}$$

[15]Note that when $J = 1$, this innovation does not enter as an additional relevant state for $K_2^{\theta,n}$ (see equation (19)) because at time 2 the comparison group $\mathbb{E}_1 K_2^{RE}$ includes $K_1^{\theta,n}$ in the information set. In that case, the sole role of $K_1^{\theta,n}$ is as an economic state, re-affirming the intuition why here $\alpha_{K_1}^{\theta} + \alpha_{\mathbb{E}_0 K_1}^{\theta} = \alpha_{K_1}^{RE}$.

[16]To capture the novel informational state due to distant memory ($J = 2$), we allow here the exogenous $\mathbb{E}_{-1} K_0$ to enter the solution, given that this expectation could be different from the realized K_0. In a full infinite horizon model, like the one in section 4, this lagged expectation is part of the model solution.

where $\alpha^{\theta}_{K_1} = \partial K^{\theta,s}_2 / \partial K^{\theta,s}_1$ is given in Proposition 5.

As with the naïveté case we conjecture and verify that the optimal solution takes the form $K^{\theta,s}_1 = \alpha^{\theta,s}_{\mathbb{E}_{-1}K_0}\mathbb{E}_{-1}K_0 + \alpha^{\theta,s}_{K_0}K_0 + \alpha^{\theta,s}_{\varepsilon_1}\varepsilon_1$. There are three conceptual forces that affect these coefficients compared to their naïveté case, which we detail in Proposition 7 in the Appendix. First, the agent now anticipates that she will over-consume (relative to her naïve beliefs) at time 2 out of K_1, as the forecasted response of future savings out of capital entering period 2 is smaller than under naïveté, i.e. $\alpha^{\theta}_{K_1} < \alpha^{RE}_{K_1}$. This force alone, coming from the $\mathbb{E}^{\theta}_1 C^{\theta}_2$ term in (34), leads the agent to consume more today out of ε_1 to achieve consumption smoothing between period 1 and 2. Second, the misalignment of her perceived tradeoffs means that following a positive innovation ε_1, from the viewpoint of current self, the time 2 self will under-consume in period $t = 3$ relative to $t = 2$. This constitutes an indirect effect, i.e. the second term in (34), that leads to more savings. The race between these two forces is dominated here by the former, direct effect, as $\alpha^{\theta}_{K_1} < 0.5$, and thus the agent ends up saving less out of ε_1 than under naïveté, i.e. $\alpha^{\theta,s}_{\varepsilon_1} < \alpha^{\theta,n}_{\varepsilon_1}$.[17] Third, there is the conceptual difference of the comparison groups. With sophistication, the informational state $\mathbb{E}_0 K^{\theta,s}_1$ (a) matters for the $K^{\theta,s}_2$ solution in Proposition 5 but also (b) needs to be itself based on $K^{\theta,s}_1$, a choice that in turn is affected by $\mathbb{E}^{\theta}_1 K^{\theta,s}_2$ in equation (34). The effect of this fixed point consideration is less transparent, as it turns out to amplify or dampen, through a non-monotonic relationship with θ, the optimal responses of $K^{\theta,s}_1$ to K_0 and $\mathbb{E}_{-1}K_0$ compared to the $K^{\theta,n}_1$ case.

3.6 Naïveté and DE in Business Cycle Models

We have developed, in the context of a three-period consumption savings model, a theoretical framework to study the joint determination of actions and DE beliefs. In this process, we also emphasize the novel role played by imperfect memory retrieval based on more distant past and have introduced two formal representations of how the current self deals with the future selves' behavior. In this model we have fully characterized the naïve and the sophistication solutions in order to fully see the conceptual and economic mechanisms at play.

However, in extending this theoretical framework to more realistic and quantitatively relevant business cycle models, featuring a large state space, multiple decisions taken over an infinite horizon, and memory retrieval based on more distant past, we propose to focus specifically on the *naïveté* approach for the following conceptual and methodological reasons.

First, as noted by a large theory literature it is a coherent micro-founded model of beliefs and behavior. In our case, it also implies that the same approach recovers the time-consistent solution when memory recall is based on the immediate past. Second, our three-period model

[17] We also note that $\alpha^{\theta,s}_{\varepsilon_1} < \alpha^{RE}_{\varepsilon_1}$ for any $\alpha^{\theta}_{K_1} > 0$. This result also provides an illustration of the general analysis in Lian (2020), which shows how sophistication lowers the marginal propensity to save out of temporary income shocks compared to the current rational action due to the anticipation of future mistakes.

analysis highlights that the naïveté approach can be easily characterized methodologically, a property that we will exploit heavily as we build a portable solution method based on linearity in the next section. Third, the approach captures intuitive and rich dynamics implied by the formation of comparison groups in the representativeness heuristic: (i) different responses to the same endogenous and exogenous economics states that would matter in a fully rational model; as well as (ii) a novel response to additional states, which would not matter in the fully rational model, but do so here due to their informational role in memory formation.

In comparison, the sophistication approach may require an incredible amount of rationality and computational resources in such models. For example, in an infinite horizon version of the consumption-savings model the agent cannot solve for the optimal actions through backward iteration, but instead looks for a recursive policy function, where the continuation utilities are described recursively by an appropriate value function. The typical Euler equation for the current optimal savings choice would resemble equation (34), in that the lack of an envelope theorem-like argument means that the agents would take into account the impact of their current actions on their future perceived suboptimal choices. In general models, this entails computing the marginal life-time value of having extra savings both in terms of it being (a) an economic state and (b) an informational state. The latter effect means that the agent would have to internalize and evaluate the effect of the current savings in the formation of the comparison groups that will matter in the future selective memory recall of the past.

The computational complexities of the sophisticated solution therefore become very demanding, especially as informational states proliferate, both for the modelers, but presumably also for the economic agents. In this sense, the solution is not only difficult to characterize by the outside observers, but importantly this required hyper-rationality runs counter to the motivation of modeling agents' beliefs about their future circumstances as influenced by a heuristic, since this is usually viewed as a cognitive, mental shortcut that allows agents to make judgments quickly and efficiently (Tversky and Kahneman (1975) and Kahneman (2011)). As such, the naïve approach is arguably psychologically more coherent and consistent with the underlying foundation of diagnostic beliefs as a heuristic reflecting a memory representation affected by imprecise, selective, and less than fully rational recall.

Overall, our theoretical framework thus lays the conceptual ground for the naïve approach, compared to the sophisticated one, as being arguably a more realistic and computationally more efficient model of belief formation that captures the informational role of endogenous state variables under the representativeness heuristic. At the same time, we do not exclude that the sophisticated model may be more useful for some particular applications. Instead, by proof of concept, as we analyze in detail in the next section, we present the naïve approach as a 'portable extension of existing models' (as advocated by Rabin (2013)) that tractably incorporates the psychology foundation of the representativeness heuristic and the role of imperfect memory recall in standard business cycle models.

4 A Quantitative New Keynesian Model

In this section, we leverage the previous qualitative insights to incorporate DE into a *quantitative New Keynesian model* of the type widely used for policy analysis. We emphasize the critical role played by distant memory recall in this new class of models. Methodologically, we formally rely on the naïve approach to model beliefs, as argued earlier. This allows us to develop a *solution method* that tractably and recursively characterize equilibrium laws of motion when agents act under DE beliefs. We estimate our model and show that it replicates the empirical boom-bust cycle in response to a monetary policy shock.

4.1 The Model

The novelty of our analysis is to allow for DE in an otherwise standard economic environment: (i) the model features monopolistic competition in the labor and goods market, subject to (ii) adjustment costs in setting nominal prices, with (iii) consumption-investment decisions being influenced by real rigidities, in the form of habit formation and investment adjustment costs, and (iv) monetary policy follows a Taylor rule.

Household. A household j chooses capital K_t^θ, investment I_t^θ, capital utilization rate u_t^θ, bonds B_t^θ, consumption C_t^θ, labor $N_{j,t}^\theta$ and nominal wage $W_{j,t}^\theta$ to solve

$$\max_{K_t^\theta, I_t^\theta, u_t^\theta, B_t^\theta, C_t^\theta, N_{j,t}^\theta, W_{j,t}^\theta} \left[\ln(C_t^\theta - b\overline{C}_{t-1}^\theta) - \frac{(N_{j,t}^\theta)^{1+\eta}}{1+\eta} + \beta \mathbb{E}_t^\theta \mathcal{V}(\mathcal{S}_{t+1}^\theta) \right] \tag{35}$$

subject to the budget constraint

$$\begin{aligned} &P_t^\theta C_t^\theta + P_t^\theta I_t^\theta + P_t^{B,\theta} B_t^\theta + (\varphi_w/2)\left(W_{j,t}^\theta / W_{j,t-1}^\theta - \gamma \Pi\right)^2 W_t^\theta \\ &= B_{t-1}^\theta + P_t^\theta R_t^{k,\theta} u_t^\theta K_{t-1}^\theta + W_{j,t}^\theta N_{j,t}^\theta + \int_0^1 D_{i,t}^\theta di - P_t^\theta a(u_t^\theta) K_{t-1}^\theta. \end{aligned}$$

where P_t^θ is the price level, $R_t^{k,\theta}$ is the capital rental rate, and $\int_0^1 D_{i,t}^\theta di$ is the combined current nominal profits from intermediate firms, given below in the firms' profit maximization problem. $P_t^{B,\theta}$ is the price of bond that pays 1 unit of consumption at $t+1$ so $P_t^{B,\theta} = 1/R_t^\theta$, where R_t^θ is the gross nominal interest rate. Notice that we also allow for a capital utilization rate u_t^θ choice, subject to a resource cost specified as $a(u_t^\theta) = R^k(1+\tau)^{-1}\left((u_t^\theta)^{1+\tau} - 1\right)$.

Each household is monopolistically competitive in its labor supply. A perfectly competitive labor packer combines household labor and sells the composite labor N_t^θ to intermediate firms, described below, using the CES technology $N_t^\theta = \left[\int_0^1 (N_{j,t}^\theta)^{\frac{1}{\lambda_n}} dj\right]^{\lambda_n}$, where λ_n controls the steady-state wage markup. The packer's cost minimization leads to a standard demand curve taken by the household as an additional constraint in solving equation (35), namely $N^\theta_{j,t} = N^\theta_t \left[W^\theta_{j,t}/W^\theta_t\right]^{-\lambda_n/(\lambda_n-1)}$, where W^θ_t is the aggregate wage level.

As we detail below, our approach handles large state space models, which allows us to incorporate DE into a NK model with *nominal and real frictions* that are typical of such quantitative business cycle models (see eg. Christiano et al. (2005) and Smets and Wouters (2007)). In particular, the budget constraint describes how nominal wages are subject to an adjustment cost (as in Kim (2000)), governed by the parameter φ_w, where further γ is the rate of deterministic technological progress and Π is the steady-state inflation rate. On the preference side, note that in equation (35) we allow for habit formation where $\overline{C}^\theta_{t-1}$ is the average previous consumption and b is the external habit parameter.

Finally, the problem in equation (35) is further subject to the physical capital law of motion, which features a standard quadratic investment adjustment cost

$$K^\theta_t = (1-\delta)K^\theta_{t-1} + \left\{1 - (\kappa/2)\left((I^\theta_t/I^\theta_{t-1}) - \gamma\right)^2\right\} I^\theta_t,$$

where δ is the depreciation rate and κ is the adjustment cost parameter.

As explained in section 3.4, in this naïve approach, in evaluating the continuation value $\mathcal{V}(.)$ in equation (35), the household assumes that her and other agents' future conditional preferences and resulting conditionally optimal actions will be taken under perfect memory (or RE), given values of the states entering next period, collected in the vector $\mathcal{S}^\theta_{t+1}$. To construct that continuation value we thus set up a 'shadow' economy (indexed by RE) where the household problem is solved under perfect memory, conditional on inherited states:

$$\mathcal{V}(\mathcal{S}^\theta_t) = \max_{K^{RE}_t, I^{RE}_t, u^{RE}_t, B^{RE}_t, C^{RE}_t, N^{RE}_{j,t}, W^{RE}_{j,t}} \left[\ln(C^{RE}_t - b\overline{C}^\theta_{t-1}) - \frac{(N^{RE}_{j,t})^{1+\eta}}{1+\eta} + \beta \mathbb{E}_t \mathcal{V}(\mathcal{S}^{RE}_{t+1})\right],$$

subject to the budget constraint

$$\begin{aligned} &P^{RE}_t C^{RE}_t + P^{RE}_t I^{RE}_t + P^{B,RE}_t B^{RE}_t + \frac{\varphi_w}{2}\left(\frac{W^{RE}_{j,t}}{W^\theta_{j,t-1}} - \gamma\Pi\right)^2 W^{RE}_t \\ &= B^\theta_{t-1} + P^{RE}_t u^{RE}_t R^{k,RE}_t K^\theta_{t-1} + W^{RE}_{j,t} N^{RE}_{j,t} + \int_0^1 D^{RE}_{i,t}\, di - P^{RE}_t a(u^{RE}_t) K^\theta_{t-1}. \end{aligned}$$

The law of motion for capital is given by

$$K^{RE}_t = (1-\delta)K^\theta_{t-1} + \left\{1 - (\kappa/2)\left(I^{RE}_t/I^\theta_{t-1} - \gamma\right)^2\right\} I^{RE}_t,$$

while the labor demand curve is simply $N^{RE}_{j,t} = N^{RE}_t \left[W^{RE}_{j,t}/W^{RE}_t\right]^{-\lambda_n/(\lambda_n-1)}$.

Firms. The final output is produced by a perfectly competitive representative firm who combines a continuum of intermediate goods $Y^\theta_{i,t}$ using the CES technology:

$$Y^\theta_t = \left[\int_0^1 (Y^\theta_{i,t})^{\frac{1}{\lambda_f}} di\right]^{\lambda_f},$$

where λ_f controls the steady-state markup. Intermediate goods firms' production function is

$$Y_{i,t}^{\theta} = (u_{i,t}^{\theta} K_{i,t}^{\theta})^{\alpha} (\gamma^t N_{i,t}^{\theta})^{1-\alpha},$$

where $K_{i,t}^{\theta}$ and $N_{i,t}^{\theta}$ are the capital and labor employed by firm i. From the cost minimization problem, the real marginal cost is given by

$$MC_t^{\theta} = \frac{(R_t^{k,\theta})^{\alpha} (W_t^{\theta}/P_t^{\theta})^{1-\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha} u_{i,t}^{\theta} (\gamma^t)^{1-\alpha}}.$$

As with households, intermediate firms also face an adjustment cost (a-la Rotemberg (1982)) in changing their nominal price. Their problem is to choose $P_{i,t}^{\theta}$ to maximize

$$(C_t^{\theta} - bC_{t-1}^{\theta})^{-1} \left(P_{i,t}^{\theta} Y_{i,t}^{\theta} - P_t^{\theta} MC_t^{\theta} Y_{i,t}^{\theta} - \frac{\varphi_p}{2} \left(P_{i,t}^{\theta}/P_{i,t-1}^{\theta} - \Pi \right)^2 P_t^{\theta} Y_t^{\theta} \right) / P_t^{\theta} + \beta \mathbb{E}_t^{\theta} \mathcal{V}_f(P_{i,t}^{\theta}), \quad (36)$$

where φ_p is the price adjustment cost parameter. The continuation value $\mathcal{V}_f(P_{i,t-1}^{\theta})$ solves

$$\max_{P_{i,t}^{RE}} \left[(C_t^{RE} - bC_{t-1}^{\theta})^{-1} D_{i,t}^{RE}/P_t^{\theta} + \beta \mathbb{E}_t \mathcal{V}_f(P_{i,t}^{RE}) \right],$$

where $D_{i,t}^{RE} = \left(P_{i,t}^{RE} Y_{i,t}^{RE} - P_t^{RE} MC_t^{RE} Y_{i,t}^{RE} - .5\varphi_p \left(P_{i,t}^{RE}/P_{i,t-1}^{\theta} - \Pi \right)^2 P_t^{RE} Y_t^{RE} \right)$. Thus, in equation (36), firms' instantaneous payoff is given by current real profits and the continuation value is given by the discounted sum of real profits $V_f(P_{i,t}^{\theta})$. Under naivete, in computing continuation value, agents assume that firms inherit the chosen price $P_{i,t}^{\theta}$ (which is relevant for the adjustment cost) but future prices are set according to RE.

Market Clearing and Monetary Policy. The resource constraint is given by

$$C_t^{\theta} + I_t^{\theta} + \frac{\varphi_p}{2} \left(\Pi_t^{\theta} - \Pi \right)^2 Y_t^{\theta} + \frac{\varphi_w}{2} \left(\Pi_{w,t}^{\theta} - \gamma \Pi \right)^2 \frac{W_t^{\theta}}{P_t^{\theta}} + a(u_t^{\theta}) K_{t-1}^{\theta} = Y_t^{\theta},$$

where $\Pi_{w,t}^{\theta} \equiv W_t^{\theta}/W_{t-1}^{\theta}$ is the nominal wage inflation.

To close the model, we assume that the central bank follows a standard Taylor rule:

$$\frac{R_t^{\theta}}{R} = \left(\frac{R_{t-1}^{\theta}}{R} \right)^{\rho_R} \left\{ \left(\frac{\widetilde{\Pi}_t^{\theta}}{\Pi} \right)^{\phi_\pi} \left(\frac{Y_t^{G,\theta}}{\gamma Y_{t-1}^{G,\theta}} \right)^{\phi_Y} \right\}^{1-\rho_R} \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_R^2),$$

where $\widetilde{\Pi}_t^{\theta}$ is the annual inflation rate $\widetilde{\Pi}_t^{\theta} \equiv 0.25 \sum_{s=0}^{3} \Pi_{t-s}^{\theta}$ and ε_t is the iid monetary policy shock.[18] Finally, we provide the equilibrium conditions in Appendix B.

4.2 Solution Method

Our solution method exploits the fact that under DE, agents expect future actions to be taken under the RE policy function. Below we outline our solution method. We provide additional details and formulas in the online appendix C.

[18] We define GDP as $Y_t^{G,\theta} \equiv Y_t^{\theta} - 0.5\varphi_p \left(\Pi_t^{\theta} - \Pi \right)^2 Y_t^{\theta} - 0.5\varphi_w \left(\Pi_{w,t}^{\theta} - \gamma \Pi \right)^2 W_t^{\theta}/P_t^{\theta} - a(u_t^{\theta}) K_{t-1}^{\theta}$.

1. The first step of the solution algorithm consists of obtaining the shadow RE law of motion used by agents to form DE. We start from a linear RE system

$$\mathbf{\Gamma}_0 \mathbf{x}_t^{RE} = \mathbf{\Gamma}_1 \mathbf{x}_{t-1}^{RE} + \mathbf{\Psi}\varepsilon_t + \mathbf{\Pi}\eta_t^{RE}, \tag{37}$$

where $\mathbf{x}_t^{RE}$, ε_t and η_t^{RE} are vectors of endogenous variables, shocks, and expectation errors, respectively. This RE system is simply the RE version of the economy, with linear equilibrium conditions where DE ($\mathbb{E}_t^{\theta}$) is replaced with RE ($\mathbb{E}_t$).

A recursive law of motion can be obtained, using for example Sims (2000),

$$\mathbf{x}_t^{RE} = \mathbf{T}^{RE}\mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE}\varepsilon_t.$$

2. Consider a linear DE system

$$\mathbf{\Gamma}_0^{\theta} \mathbf{x}_t^{\theta} = \mathbf{\Gamma}_2^{\theta}\mathbb{E}_t^{\theta}\mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^{\theta}\mathbf{x}_{t-1}^{\theta} + \mathbf{\Psi}^{\theta}\varepsilon_t, \tag{38}$$

where we provide expressions for $\mathbf{\Gamma}_0^{\theta}$, $\mathbf{\Gamma}_2^{\theta}$, $\mathbf{\Gamma}_1^{\theta}$ and $\mathbf{\Psi}^{\theta}$ in the Appendix. Relative to the RE system (37), which implicitly defines expectations in $\mathbf{x}_t^{RE}$ by using expectation errors η_t^{RE}, the DE system (38) explicitly accommodates DE ($\mathbb{E}_t^{\theta}\mathbf{y}_{t+1}^{RE}$).

We can substitute the $\mathbb{E}_t^{\theta}\mathbf{y}_{t+1}^{RE}$ in the DE system (38) as

$$\mathbb{E}_t^{\theta}\mathbf{y}_{t+1}^{RE} = \mathbb{E}_t\mathbf{y}_{t+1}^{RE} + \theta(\mathbb{E}_t\mathbf{y}_{t+1}^{RE} - \mathbb{E}_t^{r}\mathbf{y}_{t+1}^{RE}), \tag{39}$$

where $\mathbb{E}_t^{r}\mathbf{y}_{t+1}^{RE}$ denotes the comparison group, or the *reference* distribution, characterizing the representativeness heuristic.

Our method allows for a general form of memory recall and thus of this comparison group. In particular, as we further explain below, we model this reference distribution in a flexible, yet parsimonious manner, as a weighted average of lagged RE expectations:

$$\mathbb{E}_t^{r}\mathbf{y}_{t+1}^{RE} = \sum_{j=1}^{J}\alpha_j\mathbb{E}_{t-j}\mathbf{y}_{t+1}^{RE}, \tag{40}$$

where $\{\alpha_j\}_{j=1}^{J}$ are weight parameters on lagged expectations (and thus $\sum_{j=1}^{J}\alpha_j = 1$).

Let $\mathbf{y}_t^{RE} = \mathbf{M}\mathbf{x}_t^{RE}$, where $\mathbf{M}$ is a selection matrix that selects variables from a vector $\mathbf{x}_t^{RE}$. Given the DE beliefs characterized by (39) and (40), the system (38) then becomes

$$\mathbf{\Gamma}_0^{\theta}\mathbf{x}_t^{\theta} = \mathbf{\Gamma}_2^{\theta}\left[(1+\theta)\,\mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^{\theta} - \sum_{j=1}^{J}\theta\alpha_j\mathbf{M}\left(\mathbf{T}^{RE}\right)^{j+1}\mathbf{x}_{t-j}^{\theta}\right] + \mathbf{\Gamma}_1^{\theta}\mathbf{x}_{t-1}^{\theta} + \mathbf{\Psi}^{\theta}\varepsilon_t, \tag{41}$$

which also makes clear that agents form DE based on state variables inherited from the DE economy, but under the assumption that in the future the economy will follow the RE law of motion.

3. Inverting matrices and rewriting (41) more compactly gives the DE law of motion:

$$\mathbf{z}_t^\theta = \mathbf{T}^\theta \mathbf{z}_{t-1}^\theta + \mathbf{R}^\theta \varepsilon_t, \tag{42}$$

where we provide expressions for $\mathbf{T}^\theta$ and $\mathbf{R}^\theta$ in Appendix C and note that $\mathbf{z}_t^\theta$ is a vector that includes not only $\mathbf{x}_t^\theta$ but its lags $\mathbf{x}_{t-1}^\theta, \ldots, \mathbf{x}_{t-J+1}^\theta$.

The key advantages of our solution method are thus its *portability and tractability*: a researcher can transform a standard linear dynamic equilibrium model (37) and compute the DE law of motion (42) with a few additional lines of code.

4.3 Estimation

Our aim is to demonstrate that DE matter in practical and policy-relevant settings. Thus we choose the estimation method that aligns with this goal. The starting point of our estimation is a local projection estimation of empirical impulse responses to a monetary policy shock using U.S. quarterly macroeconomic data over the sample period 1969Q1–2006Q4.[19] Specifically, we estimate the following regressions:

$$x_{t+h} = c^h + \tau^h t + \sum_{l=1}^{L} \alpha_l^h x_{t-l} + \sum_{i=0}^{I} \beta_i^h e_{t-i} + \varepsilon_{t+h}, \quad h = 0, \ldots, H \tag{43}$$

where x_t is the variable of interest and e_t is the Romer and Romer (2004) monetary policy shock, extended by Coibion et al. (2017). Our variables of interest are log real per capita GDP, log per capita hours worked, log real per capita consumption, log real per capita investment, log GDP deflator inflation and log Federal funds rate. The coefficients of interest are $\{\beta_0^h\}_{h=0}^H$. We set $L = I = 4$ and compute the impulse response for $H = 32$ horizons.

We estimate the model parameters using the Bayesian version of the impulse-response-matching method, developed by Christiano et al. (2010). In this method, the likelihood depends on how closely the model matches the empirical response to a shock. The likelihood is then combined with priors on the model parameters. In our empirical analysis below, we match the impulse responses of four variables: consumption, hours, inflation, and the Federal Funds rate. We then also use the implied responses of three other variables, namely investment, GDP, and SPF inflation expectations, as 'untargeted' moments that serve as external validation.[20]

[19] We do not include the period after 2007Q1 to avoid complications arising from the zero lower bound.

[20] To obtain real per capita GDP we divide real GDP by total population. Real per capita consumption is measured by the sum of personal consumption expenditure on nondurables and services divided by total population. Real per capita investment is the sum of gross private domestic investment and personal consumption expenditure on durables divided by total population. Per capita hours worked is the total hours in nonfarm business sector divided by total population.

We fix several parameters before the estimation. The deterministic growth rate γ and the steady-state inflation rate Π are set to 1.004 and 1.01, respectively, which imply a steady-state annual output growth rate of 1.6% and the annualized inflation rate of 4%. The capital share α, the discount factor β and the depreciation rate δ are set to 0.3, 0.99 and 0.025, respectively. We set λ_f and λ_n to 1.1, which imply steady-state price and wage markups of 10%. For parameters that are common in the New Keynesian literature, we center our priors around conventional values. For the diagnostic parameter θ, we center our prior around 1, in line with the estimates found in Bordalo et al. (2018) and Bordalo et al. (2019b). As explained above (see equation (40)), we allow for flexible reference expectations in memory recall and thus the comparison group is a weighted average of lagged expectations. To estimate the weights $\{\alpha_j\}_{j=1}^J$ on past memory, we consider a parsimonious parameterization. We set $J = 32$ and estimate the mean μ and the standard deviation σ of a Beta distribution. We then rescale and discretize the implied $Beta(\mu, \sigma^2)$ distribution to span the discrete interval [0,32] and obtain the weights $\widetilde{\alpha}_j$. We then apply the transformation $\alpha_j = \widetilde{\alpha}_j / (\sum_{j=1}^J \widetilde{\alpha}_j)$ so that $\{\alpha_j\}_{j=1}^J$ sum to one. We report the priors and all estimated parameters in Table 1 in the Online Appendix, while below we focus on the key parameters that control the effects of DE.

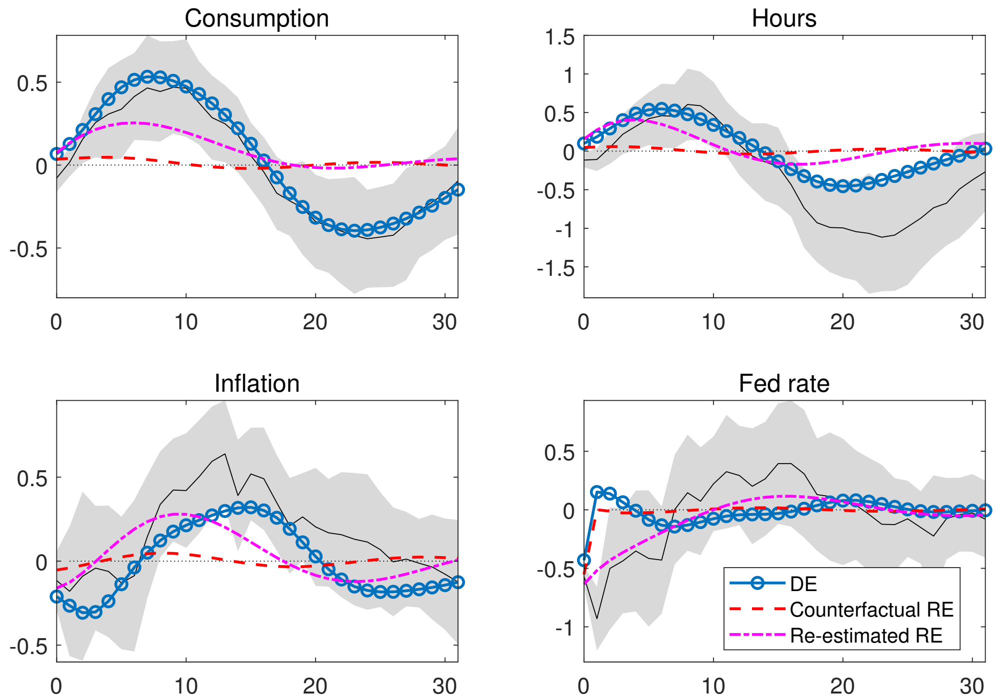
4.4 Results

Figure 1 presents the local projection impulse responses (black solid lines) to a one-standard-deviation expansionary monetary policy shock along with the 95% confidence bands. In response to a reduction in the Fed rate, real variables such as hours and consumption all increase in a hump-shaped manner, peaking around 10 quarters after the initial shock. These variables then undershoot below the steady states and reach their trough around 5 to 6 years after the shock, followed by a gradual recovery.[21] Inflation builds up slower and tends to peak at the end of the boom, followed by a slow return to the steady state.

The New Keynesian model with DE (blue lines with circles) reproduces the empirical impulse response functions (IRF) well, successfully generating, as in the data, the boom-bust cycle following the monetary policy shock. The counterfactual RE model, where we set the diagnostic parameter $\theta = 0$ while holding fixed other estimated parameters, generates transitory and negligible response. The difference between the DE IRF and the counterfactual RE IRF indicate that much of our success is due to the DE mechanism. This is confirmed by the results obtained when re-estimating the model under RE. The re-estimated RE model fails in delivering the boom-bust dynamics and the amplitude of the responses observed in the data. As a result, the marginal likelihood, a Bayesian measure of fit that penalizes models

[21] McKay and Wieland (2021) find a similar boom-bust pattern in their estimated responses to a monetary policy shock.

Figure 1: Impulse responses to a monetary policy shock: Fit for targeted responses



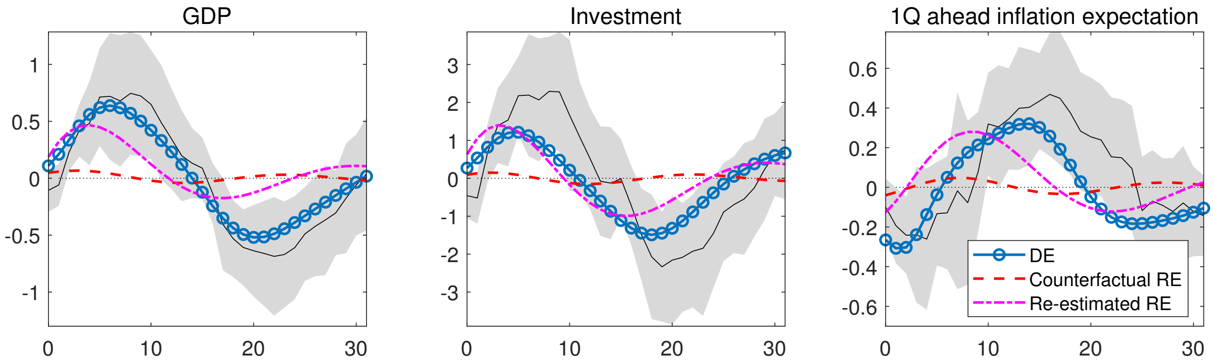
Notes: The black lines are the mean responses from the local projection and the shaded areas are the 95% confidence bands. The blue circled lines are IRFs from the baseline model with DE. The red dashed lines are IRFs from the counterfactual RE model where we set $\theta = 0$ while holding fixed other estimated parameters. The magenta dashed lines are IRFs from the re-estimated RE model. The consumption and hours responses are in percentage deviations from steady states while inflation and Fed rate are in annual percentage points.

with more parameters, is $(-464 - (-504) =)$ 40 log points higher in the DE model.

The model is also able to match remarkably well the IRF that were not targeted in the estimation. The first two panels of Figure 2 report the responses of GDP and investment to the monetary policy shock. The model delivers a good fit. The right panel of Figure 2 reports the impulse response of expected inflation.[22] The model generates inflation expectations that are very much in line with what observed in the data, even if we did not target inflation expectations in our estimation exercise. Figure 2 also shows that the re-estimated RE counterpart of the model does a worse job in accounting for this untargeted moments. Formally, we find that the root-mean-square error (RMSE) for the DE model is 0.54, while

[22] We measure inflation expectations using the median of the SPF survey responses of one-quarter-ahead inflation expectations. We assume that the model implied inflation expectations coincide with what an agent that knows the model would predict ($\mathbb{E}_t \widehat{\pi}^{\theta}_{t+1}$).

Figure 2: Impulse responses to a monetary policy shock: Fit for untargeted responses



Notes: The interpretation of the plotted lines follow their description for Figure 1. The responses of GDP and investment are in percentage deviations from the steady states while the inflation expectation is in annual percentage points.

for the RE re-estimated model is 0.75.[23]

We estimate $\theta = 1.91$ for the parameter controlling the severity of the DE distortion. The mean and standard deviation of the Beta distribution that controls the weights $\alpha_j's$ attached to each of the $J = 32$ lagged expectations entering the comparison group are 0.19 and 0.04, respectively. As shown in the left panel of Figure 3, these estimates imply that the weights are centered on the expectations formed six quarters ago, with positive weights assigned to expectations formed between three and eleven quarters ago. The right panel of Figure 3 shows how the impulse response for consumption changes as we vary the lag for the reference distribution. We consider the case in which only recent memory matters ($J = 1$) or when only two-period-ago expectations matter ($J = 2$). Reducing the lag impacts the frequency and the amplitude of the boom-bust cycles. As we will discuss below, when J increases the effects of past misperceptions accumulate, leading to larger fluctuations.

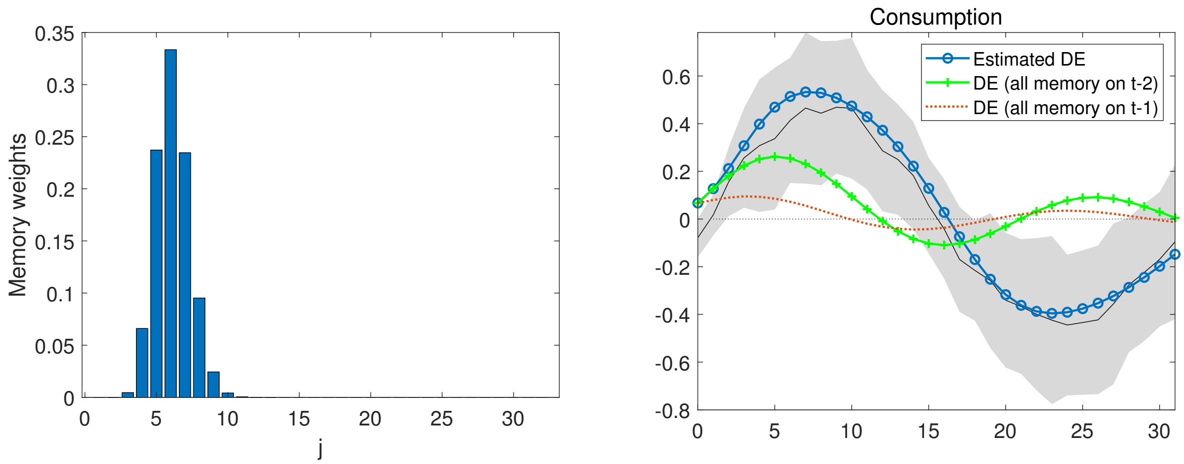
4.5 Mechanism: Perceived Consumption and Inflation Paths

As mentioned in Section 4.1 and detailed in the online appendix, due to its nominal and real frictions, the model features several intertemporal decisions (i.e. on optimal nominal wages, nominal prices, investment, capital and risk-free bonds), all of which are evaluated under DE. We can nevertheless distill the key mechanism through which this otherwise rich DE model fits the IRF dynamics by focusing on the DE Euler equation for bonds, which naturally connects the equilibrium perceived consumption and inflation paths.

[23]We compute $RMSE = \sqrt{\sum_{i=1}^{3}\sum_{t=1}^{T}(IRF^i_{data,t} - IRF^i_{model,t})^2/T}$, where $IRF^i_{data,t}$ and $IRF^i_{model,t}$ indicate the local projection IRF and model IRF, respectively, for GDP, investment and expected inflation.

Figure 3: Estimated selective memory

(a) Estimated memory weights α_j (b) Empirical role of distant memory



Notes: The left panel reports the estimated memory weights α_j. The right panel reports the consumption IRF in the estimated DE model (blue circled line) and the counterfactual model where only recent memory matters (orange dotted line) and when only two-period-ago memory matters (green line with plus signs).

Momentarily ignoring consumption habits, the Euler equation for bonds is

$$\left(C_t^\theta\right)^{-1} \frac{1}{P_t^\theta} = \beta R_t^\theta \mathbb{E}_t^\theta \left[\frac{\left(C_{t+1}^{RE}\right)^{-1}}{P_{t+1}^{RE}} \right] = \frac{\Pi}{R} R_t^\theta \mathbb{E}_t^\theta \left[\frac{\left(C_{t+1}^{RE}\right)^{-1}}{P_{t+1}^{RE}} \right]$$

In deviations from the steady state, we have:

$$-\widehat{c}_t^\theta = \widehat{r}_t^\theta + p_t^\theta + \pi - \mathbb{E}_t^\theta \left[\widehat{c}_{t+1}^{RE} + p_{t+1}^{RE} \right]$$

where hats denote log-deviations from the steady state, lowercase variables denote logs, and π denotes steady state net inflation. Here we can easily appeal to the additivity property under DE to separate terms inside expectations,[24] leading to

$$\mathbb{E}_t^\theta \widehat{c}_{t+1}^{RE} - \widehat{c}_t^\theta = \widehat{r}_t^\theta - \left[\mathbb{E}_t^\theta p_{t+1}^{RE} - p_t^\theta - \pi \right].$$

The perception of the future price level p_{t+1}^{RE} plays a critical role for the optimal intertemporal consumption smoothing decision. This price can be decomposed as the one-step ahead inflation (a jump variable) plus the current price level (a state variable entering next period):

$$p_{t+1}^{RE} = \pi_{t+1}^{RE} + p_t^\theta = \widehat{\pi}_{t+1}^{RE} + \pi + p_t^\theta$$

[24]This is possible as long as we maintain conditional normality of those individual random variables, a feature that our equilibrium objects satisfy.

Substituting this decomposition and applying the DE operator, we obtain:

$$\mathbb{E}_t^{\theta} p_{t+1}^{RE} - p_t^{\theta} - \pi = \mathbb{E}_t^{\theta} \widehat{\pi}_{t+1}^{RE} + \theta \underbrace{\left[p_t^{\theta} - \mathbb{E}_t^{r}\left(p_t^{RE}\right)\right]}_{\text{Surprise in price level}}.$$

DE act through two channels. The first term of the right hand side captures DE over future inflation, where future inflation is perceived as determined at $t+1$ under RE, given the value of the current nominal price p_t^{θ} entering as an endogenous state next period. The second component captures the role of memory recall in the perceived surprise in that same endogenous state variable with respect to the reference group, as emphasized in statistical and economic terms in the earlier sections.

As in equation (40) the reference group for p_t^{θ} is a weighted average of lagged expectations:

$$\mathbb{E}_t^{r}\left(p_t^{RE}\right) = \sum_{j=1}^{J} \alpha_j \mathbb{E}_{t-j}\left(p_t^{RE}\right),$$

which leads to a surprise in the current price level expressed as

$$p_t^{\theta} - \mathbb{E}_t^{r}\left(p_t^{RE}\right) = \underbrace{\sum_{j=1}^{J} \alpha_j \left(\widehat{\pi}_{t-j+1,t}^{\theta} - \mathbb{E}_{t-j} \widehat{\pi}_{t-j+1,t}^{RE}\right)}_{\text{perceived innovation in cumulative inflation}} \equiv \pi_{J,t}^{*},$$

where each term $\widehat{\pi}_{t-J+1,t} = \widehat{\pi}_{t-J} + \widehat{\pi}_{t-J+1} + ...\widehat{\pi}_t = p_t - p_{t-J} - \pi$ denotes the *cumulative* inflation between $t-J$ and t. For further reference, we denote that surprise, or *perceived innovation in cumulative inflation*, as the equilibrium object $\pi_{J,t}^{*}$.[25]

Putting everything together, the DE Euler equation becomes:

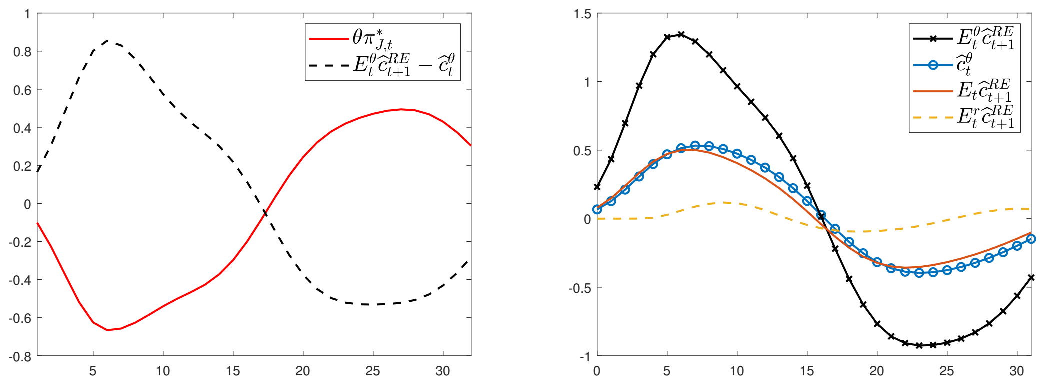
$$\mathbb{E}_t^{\theta} \widehat{c}_{t+1}^{RE} - \widehat{c}_t^{\theta} = \widehat{r}_t^{\theta} - \mathbb{E}_t^{\theta} \widehat{\pi}_{t+1}^{RE} - \theta \pi_{J,t}^{*}. \tag{44}$$

Intuitively, holding constant $\mathbb{E}_t^{\theta} \widehat{c}_{t+1}^{RE}$ and $\left(\widehat{r}_t^{\theta} - \mathbb{E}_t^{\theta} \widehat{\pi}_{t+1}^{RE}\right)$, a higher innovation $\pi_{J,t}^{*}$ makes the perceived expected future price relatively high, thus lowering the incentives to postpone consumption. As we discuss below, the equilibrium variation in this innovation turns out to be key in rationalizing the boom-bust dynamics.

The left panel of Figure 4 shows that the perceived innovation in cumulative inflation $\pi_{J,t}^{*}$ also exhibits a boom-bust pattern. Importantly, through the lenses of the DE Euler equation in (44), this response appears to be the mirror image of the DE expected consumption growth, further implying that the one period ahead component $\widehat{r}_t^{\theta} - \mathbb{E}_t^{\theta} \widehat{\pi}_{t+1}^{RE}$ is very stable compared to $\theta \pi_{J,t}^{*}$. The key observation in understanding why $\theta \pi_{J,t}^{*}$ matters much more in determining the perceived consumption smoothing tradeoffs than the one period component is that DE are computed with respect to the one-step-ahead price level, not with respect to one-step-ahead

[25] In the special case of $J = 1$, per our earlier analytical results, equilibrium variables under the RE law of motion respond to endogenous states in the same way as they do under the DE law of economy, making the equilibrium perceived innovation in cumulative inflation take the simpler but equivalent form $\pi_{1,t}^{*} = \widehat{\pi}_t^{\theta} - \mathbb{E}_{t-1} \widehat{\pi}_t^{\theta}$. This form recovers the nominal price surprise object that distorts consumption smoothing in the NK model of L'Huillier et al. (2021) who focus their analysis entirely on the $J = 1$ case.

Figure 4: Perceived innovation in cumulative inflation and consumption paths



Notes: The left panel shows the perceived innovation in cumulative inflation (multiplied by θ) and DE expected consumption growth. The right panel plots DE expected consumption ($\mathbb{E}^\theta_t \widehat{c}^{RE}_{t+1}$), realized equilibrium consumption ($\widehat{c}^\theta_t$), RE expected consumption ($\mathbb{E}_t \widehat{c}^{RE}_{t+1}$) and reference expectation ($\mathbb{E}^r_t \widehat{c}^{RE}_{t+1}$).

inflation. Thus even if inflation has relatively low persistence, the model still delivers large and persistent effects of DE from surprises in the price level, i.e. in cumulative inflation with respect to the reference expectation. In the estimated model, inflation under RE (which is what matters for DE beliefs, per the naïveté approach) has relatively low persistence, so applying DE in isolation only over it matters little in shaping the optimal consumption smoothing. However, surprises in the price level have large effects because they determine a shift in the whole perceived path of the price level.[26]

The left panel also plots the DE expected consumption growth in equation (44), given by

$$\mathbb{E}^\theta_t \widehat{c}^{RE}_{t+1} - \widehat{c}^\theta_t = \mathbb{E}_t \widehat{c}^{RE}_{t+1} + \theta \left(\mathbb{E}_t \widehat{c}^{RE}_{t+1} - \mathbb{E}^r_t \widehat{c}^{RE}_{t+1} \right) - \widehat{c}^\theta_t$$

where $\mathbb{E}^r_t \widehat{c}^{RE}_{t+1} = \sum_{j=1}^{J} \alpha_j \mathbb{E}_{t-j} \widehat{c}^{RE}_{t+1}$. The right panel of Figure 4 shows that, to accommodate this DE expected consumption growth, the realized equilibrium consumption $\widehat{c}^\theta_t$ follows a boom-bust movement, occurring with a corresponding dynamic of $\mathbb{E}^\theta_t \widehat{c}^{RE}_{t+1}$ that is even more pronounced. Per standard intuition, a surprise interest rate cut increases consumption $\widehat{c}^\theta_t$. Because of the relatively high estimated habit ($b = 0.8$), consumption moves sluggishly. Hence, the equilibrium conditional expectation under RE, $\mathbb{E}_t \widehat{c}^{RE}_{t+1}$, also rises by a similar amount. This increase causes DE agents' perception of future consumption to be overly influenced by the high consumption state and hence raises significantly the DE agents' perception $\mathbb{E}^\theta_t \widehat{c}^{RE}_{t+1}$ of future consumption. This over-reaction slowly subsides as the reference expectation $\mathbb{E}^r_t \widehat{c}^{RE}_{t+1}$

[26]In fact, even if inflation had no persistence in the shadow RE economy, surprises in inflation would still determine a change in the predicted value of the price level, activating DE.

starts to slowly rise, which catches up with the RE conditional expectation $\mathbb{E}_t \widehat{c}_{t+1}^{RE}$ around period 15. The reference expectation rises slowly because the estimated memory recall is fairly distant; recent events take time to sink into agent's mind.

We can then jointly rationalize the perceived consumption path and the perceived innovation in the cumulative inflation together, as required by the equilibrium condition in equation (44), as follows. On impact, because of an increase in utilization, inflation declines. This determines a negative surprise in the price level and a lower than usual expected future price level, which in equilibrium is consistent from the agent's perspective with a perceived acceleration in consumption.[27] Importantly, as shown in Figure 1 equilibrium inflation $\widehat{\pi}_t^{\theta}$ starts to rise above steady state during the economic boom (significantly so after period 8, as in the data). This rise first leads to a reduction in the negative surprises for the price level and then eventually to positive surprises, as indicated by the left panel of Figure 4. This path determines a *reversal* in the perceived innovation in cumulative inflation, which moves into the positive territory during the bust part of the cycle, when agents find the resulting high perceptions of future price levels consistent with their pessimism about future consumption.

The specific novel property of the model dynamics under DE is that this economic boom induced by an expansionary monetary policy shock does not simply slowly subside to converge back to the steady state from above. Instead, we observe that, as in the data, there is an inflection point (around period 15) where the boom turns into a bust and a general decline in economic activity. Moreover, the boom-bust cycle keeps generating misperceptions about the future as agents are surprised that the economy is away from their reference distributions. The feedback from beliefs to actions create further belief distortions that extend well beyond the lag in reference expectations. When DE apply to exogenous variables, DE and RE naturally realign themselves after J periods. Instead, when DE apply to endogenous variables, past decisions affect current expectations, generating new distortions that feed into current decisions, creating waves of optimism and pessimism - a form of Minsky (1977) moments.

The results presented above also illustrate the importance of considering distant memory for a given level of DE distortions. Distant memory creates larger revisions in expectations, leading to larger surprises, and larger belief distortions. This explains why the parameter J does not only affect the frequency of the boom-bust cycle but also the amplitude. When J is small, only shocks occurring in between the time of the reference expectations and today can lead to significant surprises. Instead, under distant memory agents expectations are constantly revised as the economy is quite far from where they expected it to be when they

[27]The online appendix shows that the perceived increase in consumption more than compensates for the habit stock. In other words, not only agents expect consumption to be higher in the future, but they also expect it to grow with respect to the habit stock, lowering the marginal utility.

formed the reference expectations.

Having explained in detail the consumption dynamics, we finally point out that in this economy there is positive co-movement between the key real aggregate variables (consumption, investment, hours). The economic channel is typical to the NK models. Intuitively, following the expansionary monetary policy shock, the demand for goods (consumption and investment) is stimulated. In this demand-driven economy, equilibrium is largely restored through a higher capacity utilization, which not only directly increases the supply of goods, but also leads to a larger labor productivity and thus stimulates firms' labor demand. While the intuition behind the co-movement between the key macroeconomic variables resembles qualitatively the standard NK mechanism, our DE model is remarkably more successful than its RE counterpart in delivering an ample, persistent, and hump-shaped boom-bust cycle. This is true for both the targeted and untargeted empirical IRFs.

5 Conclusion

In this paper, we build on the representativeness heuristic and DE paradigm proposed by Bordalo et al. (2018) to analyze the qualitative and quantitative implications of the joint determination of DE beliefs and optimal actions in the presence of (i) endogenous states and (ii) distant memory recall. These two characteristics are important because they fundamentally affect the extent to which the agent perceives a circumstance as 'surprising,' a notion central to the over-reacting distortion caused by the representativeness heuristic.

Accounting for these features in typical dynamic models require us to jointly confront two types of conceptual challenges. First, the presence of endogenous states means that in developing the agent's decision problem a form of endogenous predictability arises in forming DE, since the conditional predictability of future outcomes depends on the endogenous actions taken by agents. Second, when current DE are affected by memories formed in the distant past, the law of iterated expectations (LOIE) generally fails. This failure is intimately linked to time inconsistency in dynamic models because optimal plans decided in the past become suboptimal as a result of the change in beliefs induced by imperfect memory.

We use two simple consumption-savings models as a laboratory to provide behavioral micro-foundations that address these challenges in a psychologically and model-coherent way. In a two-period version, we isolate the role played by endogenous predictability for the feedback between DE beliefs and optimal savings choice. In a three-period extension, we tackle the issue of time-inconsistency by proposing and studying two models of belief formation about future behavior: naïveté or sophistication. We characterize how optimal actions respond differently to the same given set of variables that would matter in a fully rational model, but also to additional states, which would not matter in the fully rational model, but do so under DE due to their role in memory formation.

By leveraging our proposed theoretical foundations, we then provide a tractable solution method and develop a portable toolbox that can be used to enrich standard general equilibrium models with DE. In particular, we incorporate DE into a quantitative New Keynesian model of the type widely used for policy analysis. We uncover a critical and novel role played by endogenous states and distant memory recall, which allows the DE model to replicate the empirical boom-bust cycle dynamics in response to a monetary policy shock.

References

Akerlof, George A, "Procrastination and obedience," *The American Economic Review*, 1991, *81* (2), 1–19.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, Frederik Schwerter, and Andrei Shleifer, "Memory and representativeness," *Psychological Review*, 2020. forthcoming.

_ , Nicola Gennaioli, and Andrei Shleifer, "Diagnostic expectations and credit cycles," *The Journal of Finance*, 2018, *73* (1), 199–227.

_ , _ , _ , and Stephen J Terry, "Real credit cycles," 2019. Harvard, mimeo.

_ , _ , Rafael La Porta, and Andrei Shleifer, "Diagnostic expectations and stock returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

_ , _ , Yueran Ma, and Andrei Shleifer, "Overreaction in macroeconomic expectations," *American Economic Review*, 2020, *110* (9), 2748–82.

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

Christiano, Lawrence J., Mathias Trabandt, and Karel Walentin, "Involuntary Unemployment and the Business Cycle," 2010. NBER Working Paper.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silva, "Innocent Bystanders? Monetary Policy and Inequality," *Journal of Monetary Economics*, 2017, *88*, 70–89.

d'Arienzo, Daniele, "Maturity increasing overreaction and bond market puzzles," 2020. Bocconi University, Working Paper.

Fagereng, Andreas, Martin Blomhoff Holm, and Gisle James James Natvik, "MPC heterogeneity and household balance sheets," *AEJ: Macroeconomics*, 2020. forthcoming.

Gennaioli, Nicola and Andrei Shleifer, "What comes to mind," *The Quarterly Journal of Economics*, 2010, *125* (4), 1399–1433.

He, Zhiguo and Arvind Krishnamurthy, "A macroeconomic framework for quantifying systemic risk," *American Economic Journal: Macroeconomics*, 2019, *11* (4), 1–37.

Ilut, Cosmin L and Rosen Valchev, "Economic agents as imperfect problem solvers," 2020. NBER WP 27820.

Jappelli, Tullio and Luigi Pistaferri, "The consumption response to income changes," *Annual Review of Economics*, 2010, *2*, 479–506.

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, *95* (1), 161–182.

Kahneman, Daniel, *Thinking, fast and slow*, Macmillan, 2011.

_ **and Amos Tversky**, "Subjective probability: A judgment of representativeness," *Cognitive psychology*, 1972, *3* (3), 430–454.

Kaplan, Greg and Giovanni L Violante, "A model of the consumption response to fiscal stimulus payments," *Econometrica*, 2014, *82* (4), 1199–1239.

Kim, Jinill, "Constructing and estimating a realistic optimizing model of monetary policy," *Journal of Monetary Economics*, 2000, *45* (2), 329–359.

Kueng, Lorenz, "Excess sensitivity of high-income consumers," *The Quarterly Journal of Economics*, 2018, *133* (4), 1693–1751.

Laibson, David, "Golden eggs and hyperbolic discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

L'Huillier, Jean-Paul, Sanjay R Singh, and Donghoon Yoo, "Diagnostic Expectations and Macroeconomic Volatility," 2021. UC Davis, mimeo.

Lian, Chen, "Mistakes in future consumption, high MPCs now," 2020. MIT, Working Paper.

Maxted, Peter, "A macro-finance model with sentiment," 2020. Harvard, mimeo.

McDowall, Robert A, "Consumption behavior across the distribution of liquid assets," 2020. Working paper, NYU.

McKay, Alisdair and Johannes F. Wieland, "Lumpy durable consumption demand and the limited ammunition of monetary policy," *Econometrica*, 2021. forthcoming.

Minsky, Hyman P, "The financial instability hypothesis: An interpretation of Keynes and an alternative to "standard" theory," *Challenge*, 1977, *20* (1), 20–27.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, *89* (1), 103–124.

Pollak, Robert A, "Consistent planning," *The Review of Economic Studies*, 1968, *35* (2), 201–208.

Rabin, Matthew, "An approach to incorporating psychology into economics," *American Economic Review*, 2013, *103* (3), 617–22.

Romer, Christina D. and David H. Romer, "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 2004, *94* (4), 1055–1084.

Rotemberg, Julio J, "Monopolistic price adjustment and aggregate output," *The Review of Economic Studies*, 1982, *49* (4), 517–531.

Sims, Christopher A., "Solving linear rational expectations models," *Computational Economics*, 2000, *20* (1–2), 1–20.

Smets, F. and R. Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *The American Economic Review*, 2007, *97* (3), 586–606.

Strotz, Robert Henry, "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, 1955, *23* (3), 165–180.

Tversky, Amos and Daniel Kahneman, "Judgment under uncertainty: Heuristics and Biases," in "Utility, probability, and human decision making" 1975, pp. 141–162.

Appendices

A Omitted Proofs

A.1 Proof of Proposition 2

Compute $\mathbb{E}_1^\theta\left[C_2 - C_3\right]$ and replace $C_2 = \mathbb{E}_2^\theta C_3$ to obtain

$$\mathbb{E}_1^\theta\left[C_2 - C_3\right] = \mathbb{E}_1^\theta\left[\mathbb{E}_2^\theta C_3 - C_3\right]$$

The DE belief $\mathbb{E}_2^\theta C_3$ is

$$\mathbb{E}_2^\theta C_3 = \mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right)$$

therefore

$$\mathbb{E}_1^\theta\left[C_2 - C_3\right] = \mathbb{E}_1^\theta\left[\mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right) - C_3\right]$$

By applying the DE at time 1 this equals

$$(1+\theta)\mathbb{E}_1\left[\mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right) - C_3\right] - \theta\mathbb{E}_{1-J}\left[\mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right) - C_3\right]$$

The second term equals

$$\mathbb{E}_{1-J}\left[\mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right) - C_3\right] = 0$$

while the first term

$$\mathbb{E}_1\left[\mathbb{E}_2 C_3 + \theta\left(\mathbb{E}_2 C_3 - \mathbb{E}_{2-J} C_3\right) - C_3\right] = \theta\left(\mathbb{E}_1 C_3 - \mathbb{E}_1\mathbb{E}_{2-J} C_3\right)$$

If $J = 1$, this terms also equals zero and therefore

$$\mathbb{E}_1^\theta\left[C_2 - C_3\right] = 0 = C_2 - \mathbb{E}_2^\theta C_3$$

while if $J > 1$

$$\begin{aligned}\mathbb{E}_1^\theta\left[C_2 - C_3\right] &= (1+\theta)\theta\left(\mathbb{E}_1 C_3 - \mathbb{E}_{2-J} C_3\right) \\ &= (1+\theta)\theta\left[\mathbb{E}_1 K_2^\theta(K_1^\theta, \varepsilon_2) - \mathbb{E}_{2-J} K_2^\theta(K_1^\theta, \varepsilon_2)\right].\end{aligned}$$

Thus the conditionally optimal time-2 solution K_2^θ implements exactly time-1 plan $K_2^{\theta,p}$ if and only if $J = 1$.

A.2 Proof of Proposition 3

Time 2 policy. The general procedure is to work backwards from time 2. Let us immediately find the RE solution, which conditional on some K_1 entering period 2 solves

$$\varepsilon_2 + K_1 - K_2^{RE} = \mathbb{E}_2(\varepsilon_3 + K_2^{RE}), \tag{45}$$

which give us the RE coefficients for K_2^{RE} in Proposition 3.

For the DE agent, conditional on reaching period 2, the optimal K_2^θ solves the trade-off

$$C_2^\theta = \mathbb{E}_2^\theta C_3^\theta,$$

as implied by the problem in (16). When $J = 1$, for a given state K_1 and exogenous innovation ε_2, by using the time 2 and 3 budget constraint, this trade-off amounts to

$$\varepsilon_2 + K_1 - K_2^\theta = \mathbb{E}_2(\varepsilon_3 + K_2^\theta) + \theta \left[\mathbb{E}_2(\varepsilon_3 + K_2^\theta) - \mathbb{E}_1 \left(\varepsilon_3 + K_2^\theta\right)\right] \tag{46}$$

By substituting the conjectured coefficients $\alpha_{K_1}^\theta$ and $\alpha_{\varepsilon_2}^\theta$ into the trade-off (46), we obtain their values characterized in Proposition 3.

Time 1 policy. Moving backward, let us characterize the time 1 problem. By Corollary 2 the time-1 planned $K_2^{\theta,p}$ equals the policy function K_2^θ, chosen at time 2. In that case, the optimal solution for K_1^θ, solves the condition

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^\theta + \alpha_{K_1}^\theta \left(C_3^\theta - C_2^\theta\right)\right], \tag{47}$$

where the path for C_2^θ and C_3^θ are implied by the budget constraints. Technically, the DE operator over a sum of random variables satisfies the additivity property (see the proof of Corollary 1 in Bordalo et al. (2018) for details), so we can break the RHS of (47) into

$$\mathbb{E}_1^\theta(C_2^\theta) + \alpha_{K_1}^\theta \mathbb{E}_1^\theta \left(C_3^\theta - C_2^\theta\right) \tag{48}$$

which by Corollary 2 means

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta \tag{49}$$

since $\mathbb{E}_1^\theta \left(C_3^\theta - C_2^\theta\right)$. The RHS reflects the DE belief over C_2^θ, given the comparison group based on time 0 information, and as such equals

$$\mathbb{E}_1^\theta C_2^\theta = (1+\theta)\mathbb{E}_1 \left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^\theta) + K_1^\theta(1-\alpha_{K_1}^\theta)\right] - \theta\mathbb{E}_0 \left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^\theta) + K_1^\theta(1-\alpha_{K_1}^\theta)\right]$$

where we have substituted in $C_2^\theta = Y_2 + K_1^\theta - K_2^\theta$ the conjectured policy $K_2^\theta = \alpha_{K_1}^\theta K_1 + \alpha_{\varepsilon_2}^\theta \varepsilon_2$. Therefore, by using the unpredictability of income shocks, we have

$$\mathbb{E}_1^\theta C_2^\theta = \overline{Y} + (1-\alpha_{K_1}^\theta)\left[K_1^\theta + \theta\left(K_1^\theta - \mathbb{E}_0 K_1^\theta\right)\right].$$

Here $(1-\alpha_{K_1}^\theta)$ gives the conjectured exposure of C_2^θ to K_1^θ, which is its only source of endogenous persistence, and $\left(K_1^\theta - \mathbb{E}_0 K_1^\theta\right)$ is the new information about the conditional mean of C_2^θ. Under the conjectured solution for $K_1^\theta = \alpha_{K_0}^\theta K_0 + \alpha_{\varepsilon_1}^\theta \varepsilon_1$, this new information just equals $\alpha_{\varepsilon_1}^\theta \varepsilon_1$. Thus, the optimal K_1^θ solves

$$\varepsilon_1 + K_0 - K_1^\theta = (1-\alpha_{K_1}^\theta)\left[K_1^\theta + \theta\alpha_{\varepsilon_1}^\theta \varepsilon_1\right],$$

where we have $\alpha_{K_1}^\theta = 0.5$. This immediately recovers the optimal coefficients in Proposition 3. In the case of $\theta = 0$, this also solves for the corresponding RE coefficients.

A.3 Proof of Proposition 4

Policies under naïveté. Conjecture

$$K_1^{\theta,n} = \alpha_{K_0}^{\theta,n} K_0 + \alpha_{\varepsilon_1}^{\theta,n} \varepsilon_1; \quad K_2^{\theta,n} = \alpha_{K_1}^{\theta,n} K_1 + \alpha_{\varepsilon_2}^{\theta,n} \varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta C_3^{RE}$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_2^\theta C_3^{RE} &= (1+\theta)\mathbb{E}_2 \left[Y_3 + K_2^{\theta,n} \right] - \theta \mathbb{E}_1 \left[Y_3 + K_2^{RE} \right] \\
&= \overline{Y} + (1+\theta) K_2^{\theta,n} - \theta \mathbb{E}_1 K_2^{RE} \\
&= \overline{Y} + (1+\theta) K_2^{\theta,n} - \frac{1}{2}\theta K_1,
\end{aligned}$$

where we substituted in $\alpha_{K_1}^{RE} = 1/2$ in the third line. Connecting this with the LHS, we have

$$\varepsilon_2 + K_1 - K_2^{\theta,n} = (1+\theta) K_2^{\theta,n} - \frac{1}{2}\theta K_1.$$

Plugging in the conjectured solution $K_2^{\theta,n} = \alpha_{K_1}^{\theta,n} K_1 + \alpha_{\varepsilon_2}^{\theta,n} \varepsilon_2$ and equating coefficients give us $\alpha_{K_1}^{\theta,n} = 1/2 = \alpha_{K_1}^\theta$ and $\alpha_{\varepsilon_2}^{\theta,n} = 1/(2+\theta) = \alpha_{\varepsilon_2}^\theta$.

By Lemma 2 the time 1 trade-off is given by

$$C_1^\theta = \mathbb{E}_1^\theta C_2^{RE}.$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_1^\theta C_2^{RE} &= (1+\theta)\mathbb{E}_1 \left[Y_2 + K_1^{\theta,n} - K_2^{RE} \right] - \theta \mathbb{E}_0 \left[Y_2 + K_1^{RE} - K_2^{RE} \right] \\
&= (1+\theta)\mathbb{E}_1 \left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{RE}) + K_1^{\theta,n}(1-\alpha_{K_1}^{RE}) \right] - \theta \mathbb{E}_0 \left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{RE}) + K_1^{RE}(1-\alpha_{K_1}^{RE}) \right] \\
&= \overline{Y} + (1-\alpha_{K_1}^{RE}) \left[(1+\theta) K_1^{\theta,n} - \theta \mathbb{E}_0 K_1^{RE} \right] \\
&= \overline{Y} + \frac{1}{2} \left[(1+\theta) K_1^{\theta,n} - \frac{2}{3}\theta K_0 \right]
\end{aligned}$$

where we have substituted in the RE policy $K_2^{RE} = \alpha_{K_1}^{RE} K_1 + \alpha_{\varepsilon_2}^{RE} \varepsilon_2$ in the second line and substituted in $\alpha_{K_1}^{RE} = 1/2$ and $\alpha_{K_0}^{RE} = 2/3$ in the fourth line. Connecting this with the LHS, we have

$$\varepsilon_1 + K_0 - K_1^{\theta,n} = \frac{1}{2} \left[(1+\theta) K_1^{\theta,n} - \frac{2}{3}\theta K_0 \right].$$

Plugging in the conjectured solution $K_1^{\theta,n} = \alpha_{K_0}^{\theta,n} K_0 + \alpha_{\varepsilon_1}^{\theta,n} \varepsilon_1$ and equating coefficients give us $\alpha_{K_0}^{\theta,n} = 2/3 = \alpha_{K_0}^\theta$ and $\alpha_{\varepsilon_1}^{\theta,n} = 2/(3+\theta) = \alpha_{\varepsilon_1}^\theta$.

Policies under sophistication. Conjecture

$$K_1^{\theta,s} = \alpha_{K_0}^{\theta,s} K_0 + \alpha_{\varepsilon_1}^{\theta,s} \varepsilon_1; \quad K_2^{\theta,s} = \alpha_{K_1}^{\theta,s} K_1 + \alpha_{\varepsilon_2}^{\theta,s} \varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta C_3^\theta$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_2^\theta C_3^\theta &= (1+\theta)\mathbb{E}_2\left[Y_3 + K_2^{\theta,s}\right] - \theta\mathbb{E}_1\left[Y_3 + K_2^{\theta,s}\right] \\
&= \overline{Y} + (1+\theta)K_2^{\theta,s} - \theta\mathbb{E}_1 K_2^{\theta,s} \\
&= \overline{Y} + (1+\theta)K_2^{\theta,s} - \theta\alpha_{K_1}^{\theta,s}K_1.
\end{aligned}$$

Connecting this with the LHS, we have

$$\varepsilon_2 + K_1 - K_2^{\theta,s} = (1+\theta)K_2^{\theta,s} - \theta\alpha_{K_1}^{\theta,s}K_1.$$

Plugging in the conjectured solution $K_2^{\theta,s} = \alpha_{K_1}^{\theta,s}K_1 + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2$ and equating coefficients give us $\alpha_{K_1}^{\theta,s} = 1/2 = \alpha_{K_1}^\theta$ and $\alpha_{\varepsilon_2}^{\theta,s} = 1/(2+\theta) = \alpha_{\varepsilon_2}^\theta$.

By Corollary 3, the time 1 trade-off is given by

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta.$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_1^\theta C_2^\theta &= (1+\theta)\mathbb{E}_1\left[Y_2 + K_1^{\theta,s} - K_2^{\theta,s}\right] - \theta\mathbb{E}_0\left[Y_2 + K_1^{\theta,s} - K_2^{\theta,s}\right] \\
&= (1+\theta)\mathbb{E}_1\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{\theta,s}) + K_1^{\theta,s}(1-\alpha_{K_1}^{\theta,s})\right] - \theta\mathbb{E}_0\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{\theta,s}) + K_1^{\theta,s}(1-\alpha_{K_1}^{\theta,s})\right] \\
&= \overline{Y} + (1-\alpha_{K_1}^{\theta,s})\left[(1+\theta)K_1^{\theta,s} - \theta\mathbb{E}_0 K_1^{\theta,s}\right] \\
&= \overline{Y} + \frac{1}{2}\left[(1+\theta)K_1^{\theta,n} - \alpha_{K_0}^{\theta,s}\theta K_0\right]
\end{aligned}$$

where we have substituted in the DE policy $K_2^{\theta,s} = \alpha_{K_1}^{\theta,s}K_1 + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2$ in the second line and substituted in $\alpha_{K_1}^{\theta,s} = 1/2$ in the fourth line. Connecting this with the LHS, we have

$$\varepsilon_1 + K_0 - K_1^{\theta,s} = \frac{1}{2}\left[(1+\theta)K_1^{\theta,s} - \alpha_{K_0}^{\theta,s}\theta K_0\right].$$

Plugging in the conjectured solution $K_1^{\theta,s} = \alpha_{K_0}^{\theta,s}K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1$ and equating coefficients give us $\alpha_{K_0}^{\theta,s} = 2/3 = \alpha_{K_0}^\theta$ and $\alpha_{\varepsilon_1}^{\theta,s} = 2/(3+\theta) = \alpha_{\varepsilon_1}^\theta$.

A.4 Proof of Proposition 5

Time 2 policy under naïveté. Consider the conjecture

$$K_2^{\theta,n} = \alpha_{\mathbb{E}_0 K_1}^\theta \mathbb{E}_0 K_1^{RE} + \alpha_{K_1}^\theta K_1^{\theta,n} + \alpha_{\varepsilon_2}^\theta\varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta C_3^{RE}$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_2^\theta C_3^{RE} &= (1+\theta)\mathbb{E}_2\left[Y_3 + K_2^{\theta,n}\right] - \theta\mathbb{E}_0\left[Y_3 + K_2^{RE}\right] \\
&= \overline{Y} + (1+\theta)K_2^{\theta,n} - \theta\mathbb{E}_0 K_2^{RE} \\
&= \overline{Y} + (1+\theta)\left(\alpha_{K_0}^{\theta,n}K_0 + \alpha_{K_1}^{\theta,n}K_1 + \alpha_{\varepsilon_2}^{\theta,n}\varepsilon_2\right) - \theta\alpha_{K_1}^{RE}\mathbb{E}_0 K_1^{RE}
\end{aligned}$$

where we substituted in $\alpha_{K_0}^{RE} = 2/3$. Connecting this with the LHS, we have

$$\varepsilon_2 + K_1 - K_2^{\theta,n} = (1+\theta)\left(\alpha_{K_0}^{\theta,n}K_0 + \alpha_{K_1}^{\theta,n}K_1 + \alpha_{\varepsilon_2}^{\theta,n}\varepsilon_2\right) - \frac{2}{3}\theta\mathbb{E}_0 K_1^{RE}.$$

Plugging in the conjectured solution $K_2^{\theta,n} = \alpha_{\mathbb{E}_0 K_1}^{\theta}\mathbb{E}_0 K_1^{RE} + \alpha_{K_1}^{\theta}K_1^{\theta,n} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2$ and equating coefficients give us $\alpha_{\mathbb{E}_0 K_1}^{\theta} = \theta/[2(2+\theta)]$, $\alpha_{K_1}^{\theta,n} = 1/(2+\theta)$ and $\alpha_{\varepsilon_2}^{\theta,n} = 1/(2+\theta)$.

Time 2 policy under sophistication. Consider the conjecture

$$K_2^{\theta,s} = \alpha_{\mathbb{E}_0 K_1}^{\theta}\mathbb{E}_0 K_1^{\theta,s} + \alpha_{K_1}^{\theta}K_1^{\theta,s} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta C_3^\theta$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_2^\theta C_3^\theta &= (1+\theta)\mathbb{E}_2\left[Y_3 + K_2^{\theta,s}\right] - \theta\mathbb{E}_0\left[Y_3 + K_2^{\theta,s}\right] \\
&= \overline{Y} + K_2^{\theta,s} + \theta\left[K_2^{\theta,s} - \mathbb{E}_0 K_2^{\theta,s}\right] \\
&= \overline{Y} + \alpha_{\mathbb{E}_0 K_1}^{\theta}\mathbb{E}_0 K_1^{\theta,s} + \alpha_{K_1}^{\theta}K_1^{\theta,s} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \theta\left[\alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \alpha_{K_1}^{\theta}(K_1^{\theta,s} - \mathbb{E}_0 K_1^{\theta,s})\right].
\end{aligned}$$

Connecting this with the LHS, we have

$$\varepsilon_2 + K_1^{\theta,s} - K_2^{\theta,s} = \alpha_{\mathbb{E}_0 K_1}^{\theta}\mathbb{E}_0 K_1^{\theta,s} + \alpha_{K_1}^{\theta}K_1^{\theta,s} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \theta\left[\alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \alpha_{K_1}^{\theta}(K_1^{\theta,s} - \mathbb{E}_0 K_1^{\theta,s})\right].$$

Plugging in the conjectured solution $K_2^{\theta,s} = \alpha_{\mathbb{E}_0 K_1}^{\theta}\mathbb{E}_0 K_1^{\theta,s} + \alpha_{K_1}^{\theta}K_1^{\theta,s} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2$ and equating coefficients give us $\alpha_{\mathbb{E}_0 K_1}^{\theta} = 1/[2(2+\theta)]$, $\alpha_{K_1}^{\theta} = 1/(2+\theta)$ and $\alpha_{\varepsilon_2}^{\theta} = 1/(2+\theta)$.

A.5 Proof of Proposition 6

The Proposition below considers the time 1 savings policy under naïveté, referenced in section 3.5.2.

Proposition 6. *Compared to the $J = 1$ case, when $J = 2$ the naïveté policy function*

$$K_1^{\theta,n} = \alpha_{\mathbb{E}_{-1}K_0}^{\theta,n}\mathbb{E}_{-1}K_0 + \alpha_{K_0}^{\theta,n}K_0 + \alpha_{\varepsilon_1}^{\theta,n}\varepsilon_1$$

is characterized by (i) a positive loading on the past informational state $\mathbb{E}_{-1}K_0$, (ii) a muted response to the current economic state K_0, and (iii) an identical, muted, response to the *current innovation, as follows:*

$$\alpha^{\theta,n}_{\mathbb{E}_{-1}K_0} = \frac{2\theta}{3(3+\theta)};\ \alpha^{\theta,n}_{K_0} = \frac{2}{3+\theta};\ \alpha^{\theta,n}_{\varepsilon_1} = \frac{2}{3+\theta}.$$

To obtain the policy function, we start from the conjecture

$$K_1^{\theta,n} = \alpha^{\theta,n}_{\mathbb{E}_{-1}K_0}\mathbb{E}_{-1}K_0 + \alpha^{\theta,n}_{K_0}K_0 + \alpha^{\theta,n}_{\varepsilon_1}\varepsilon_1.$$

Proof. By Lemma 2 the time 1 tradeoff is given by

$$C_1^\theta = \mathbb{E}_1^\theta C_2^{RE}.$$

The RHS equals

$$\begin{aligned}
\mathbb{E}_1^\theta C_2^{RE} &= (1+\theta)\mathbb{E}_1\left[Y_2 + K_1^{\theta,n} - K_2^{RE}\right] - \theta\mathbb{E}_{-1}\left[Y_2 + K_1^{RE} - K_2^{RE}\right] \\
&= (1+\theta)\mathbb{E}_1\left[\overline{Y} + \varepsilon_2(1-\alpha^{RE}_{\varepsilon_2}) + K_1^{\theta,n}(1-\alpha^{RE}_{K_1})\right] - \theta\mathbb{E}_{-1}\left[\overline{Y} + \varepsilon_2(1-\alpha^{RE}_{\varepsilon_2}) + K_1^{RE}(1-\alpha^{RE}_{K_1})\right] \\
&= \overline{Y} + (1-\alpha^{RE}_{K_1})\left[(1+\theta)K_1^{\theta,n} - \theta\mathbb{E}_{-1}K_1^{RE}\right] \\
&= \overline{Y} + \frac{1}{2}\left[(1+\theta)K_1^{\theta,n} - \frac{2}{3}\theta\mathbb{E}_{-1}K_0\right]
\end{aligned}$$

where we have substituted in the RE policy $K_2^{RE} = \alpha^{RE}_{K_1}K_1 + \alpha^{RE}_{\varepsilon_2}\varepsilon_2$ in the second line and substituted in $\alpha^{RE}_{K_1} = 1/2$ and $\alpha^{RE}_{K_0} = 2/3$ in the fourth line. Connecting this with the LHS, we have

$$\varepsilon_1 + K_0 - K_1^{\theta,n} = \frac{1}{2}\left[(1+\theta)K_1^{\theta,n} - \frac{2}{3}\theta\mathbb{E}_{-1}K_0\right].$$

Plugging in the conjectured solution $K_1^{\theta,n} = \alpha^{\theta,n}_{\mathbb{E}_{-1}K_0}\mathbb{E}_{-1}K_0 + \alpha^{\theta,n}_{K_0}K_0 + \alpha^{\theta,n}_{\varepsilon_1}\varepsilon_1$ and equating coefficients give us $\alpha^{\theta,n}_{\mathbb{E}_{-1}K_0} = 2\theta/[3(3+\theta)]$, $\alpha^{\theta,n}_{K_0} = 2/(3+\theta)$ and $\alpha^{\theta,n}_{\varepsilon_1} = 2/(3+\theta)$. □

A.6 Proof of Proposition 7

The Proposition below considers the time 1 savings policy under sophistication, referenced in section 3.5.2.

Proposition 7. *When $J = 2$ we conjecture and verify the sophistication policy function*

$$K_1^{\theta,s} = \alpha^{\theta,s}_{\mathbb{E}_{-1}K_0}\mathbb{E}_{-1}K_0 + \alpha^{\theta,s}_{K_0}K_0 + \alpha^{\theta,s}_{\varepsilon_1}\varepsilon_1.$$

which compared to the naïveté policy function in Proposition 6 is characterized by the following properties (1) $\alpha^{\theta,s}_{\varepsilon_1} < \alpha^{\theta,n}_{\varepsilon_1}$; (2) $\alpha^{\theta,s}_{K_0} < \alpha^{\theta,n}_{K_0}$ if $\theta < 1$, and $\alpha^{\theta,s}_{K_0} > \alpha^{\theta,n}_{K_0}$ if $\theta > 1$; (3), $\alpha^{\theta,s}_{\mathbb{E}_{-1}K_0} > \alpha^{\theta,n}_{\mathbb{E}_{-1}K_0}$ if $\theta < 1$, and $\alpha^{\theta,s}_{\mathbb{E}_{-1}K_0} < \alpha^{\theta,n}_{\mathbb{E}_{-1}K_0}$ if $\theta > 1$.

Proof. Conjecture

$$K_1^{\theta,s} = \alpha^{\theta,s}_{\mathbb{E}_{-1}K_0}\mathbb{E}_{-1}K_0 + \alpha^{\theta,s}_{K_0}K_0 + \alpha^{\theta,s}_{\varepsilon_1}\varepsilon_1.$$

The time 1 tradeoff is given by

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^\theta + \alpha_{K_1}^\theta (\mathbb{E}_2 C_3^\theta - C_2^\theta) \right].$$

The RHS equals

$$\begin{aligned}
&\mathbb{E}_1^\theta \left[C_2^\theta + \alpha_{K_1}^\theta (\mathbb{E}_2 C_3^\theta - C_2^\theta) \right] = (1-\alpha_{K_1}^\theta)\mathbb{E}_1^\theta C_2^\theta + \alpha_{K_1}^\theta \mathbb{E}_1^\theta C_3^\theta \\
&= (1-\alpha_{K_1}^\theta)\left\{ (1+\theta)\mathbb{E}_1 \left[Y_2 + K_1^{\theta,s} - K_2^{\theta,s} \right] - \theta \mathbb{E}_{-1} \left[Y_2 + K_1^{\theta,s} - K_2^{\theta,s} \right] \right\} \\
&+ \alpha_{K_1}^\theta \left\{ (1+\theta)\mathbb{E}_1 \left[Y_3 + K_2^{\theta,s} \right] - \theta \mathbb{E}_{-1} \left[Y_3 + K_2^{\theta,s} \right] \right\}
\end{aligned}$$

After some algebra, we find that this equals

$$\begin{aligned}
&= \overline{Y} + (1-\alpha_{K_1}^\theta)(1+\theta) \left[\left(1 - \alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} - \alpha_{K_1}^\theta \right) \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} \mathbb{E}_{-1} K_0 + \alpha_{K_0}^{\theta,s} K_0 \right) + (1-\alpha_{K_1}^\theta)\alpha_{\varepsilon_1}^{\theta,s} \varepsilon_1 \right] \\
&- (1-\alpha_{K_1}^\theta)\theta \left(1 - \alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} - \alpha_{K_1}^\theta \right) \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} + \alpha_{K_0}^{\theta,s} \right) \mathbb{E}_{-1} K_0 \\
&+ \alpha_{K_1}^\theta (1+\theta) \left[\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} \mathbb{E}_{-1} K_0 + \alpha_{K_0}^{\theta,s} K_0 \right) + \alpha_{K_1}^\theta \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} \mathbb{E}_{-1} K_0 + \alpha_{K_0}^{\theta,s} K_0 + \alpha_{\varepsilon_1}^{\theta,s} \varepsilon_1 \right) \right] \\
&- \alpha_{K_1}^\theta \theta \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} + \alpha_{K_1}^\theta \right) \left(\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} + \alpha_{K_0}^{\theta,s} \right) \mathbb{E}_{-1} K_0
\end{aligned}$$

The LHS is given by

$$C_1^\theta = \overline{Y} + \varepsilon_1 + K_0 - K_1^{\theta,s}.$$

We then connect the LHS to the RHS and equate coefficients after substituting in the conjectured solution for $K_1^{\theta,s}$. Equating coefficients, we have

$$\begin{aligned}
\alpha_{\varepsilon_1}^{\theta,s} &= \frac{1}{1 + (1+\theta)\left[(1-\alpha_{K_1}^\theta)^2 + (\alpha_{K_1}^\theta)^2 \right]} = \frac{(2+\theta)^2}{(2+\theta)^2 + (1+\theta)\left[(1+\theta)^2 + 1 \right]} \\
\alpha_{K_0}^{\theta,s} &= \frac{1}{1 + (1+\theta)\left[(1-\alpha_{K_1}^\theta)(1 - \alpha_{\mathbb{E}_0 K_1}^\theta - \alpha_{K_1}^\theta) + \alpha_{K_1}^\theta (\alpha_{\mathbb{E}_0 K_1}^\theta + \alpha_{K_1}^\theta) \right]} \\
&= \frac{2(2+\theta)^2}{2(2+\theta)^2 + (1+\theta)\left[(1+\theta)(1+2\theta) + 3 \right]} \\
\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} &= \frac{\theta \left[(1-\alpha_{K_1}^\theta)(1 - \alpha_{\mathbb{E}_0 K_1}^\theta - \alpha_{K_1}^\theta) + \alpha_{K_1}^\theta (\alpha_{\mathbb{E}_0 K_1}^\theta + \alpha_{K_1}^\theta) \right]}{1 + (1-\alpha_{K_1}^\theta)(1 - \alpha_{\mathbb{E}_0 K_1}^\theta - \alpha_{K_1}^\theta) + \alpha_{K_1}^\theta (\alpha_{\mathbb{E}_0 K_1}^\theta + \alpha_{K_1}^\theta)} \alpha_{K_0}^{\theta,s} \\
&= \frac{\theta \left[(1+2\theta)(1+\theta) + 3 \right]}{2(2+\theta)^2 + (1+2\theta)(1+\theta) + 3} \alpha_{K_0}^{\theta,s}
\end{aligned}$$

which give the specific coefficients in Proposition 7. When we compare this sophistication solution to the naïveté one, we find the patterns stated in Proposition 7. □

B Equilibrium Conditions of the New Keynesian Model

- Capital Euler equation:

$$\mu_t^\theta = \beta \mathbb{E}_t^\theta \left[(C_{t+1}^{RE} - bC_t^\theta)^{-1} (R_{t+1}^{k,RE} u_{t+1}^{RE} - a(u_{t+1}^{RE})) + \mu_{t+1}^{RE}(1-\delta) \right],$$

where μ_t^θ is the Lagrangian multiplier on the capital accumulation equation.

- Utilization choice:

$$R_t^{k,\theta} = R^k (u_t^\theta)^\tau$$

- Investment first-order condition:

$$(C_t^\theta - bC_{t-1}^\theta)^{-1} = \mu_t^\theta \left\{1 - \frac{\kappa}{2}\left(\Delta I_t^\theta - \gamma\right)^2 - \kappa\left(\Delta I_t^\theta - \gamma\right)\Delta I_t^\theta\right\} + \beta \mathbb{E}_t^\theta \left[\mu_{t+1}^{RE} \kappa \left(\Delta I_{t+1}^{RE} - \gamma\right)\left(\Delta I_{t+1}^{RE}\right)^2\right]$$

- Investment growth:

$$\Delta I_t^\theta = I_t^\theta / I_{t-1}^\theta$$

- Consumption Euler equation:

$$Q_t^\theta = \frac{\beta R_t^\theta}{\Pi} \mathbb{E}_t^\theta \left[Q_{t+1}^{RE}\right]$$

- Definition of Q_t^θ :

$$\frac{Q_t^\theta}{Q_{t-1}^\theta} = \frac{\Pi}{\Pi_t^\theta}\left(\frac{C_t^\theta - bC_{t-1}^\theta}{C_{t-1}^\theta - bC_{t-2}^\theta}\right)^{-1}$$

- Capital accumulation:

$$K_t^\theta = (1-\delta)K_{t-1}^\theta + \left\{1 - \frac{\kappa}{2}\left(\frac{I_t^\theta}{I_{t-1}^\theta} - \gamma\right)^2\right\} I_t^\theta$$

- Real wage:

$$\widetilde{W}_t^\theta = MC_t^\theta (1-\alpha)\frac{Y_t^\theta}{N_t^\theta}$$

where $\widetilde{W}_t^\theta \equiv W_t^\theta / P_t^\theta$ is the real wage.

- Capital rental rate:

$$R_t^{k,\theta} = MC_t^\theta \alpha \frac{Y_t^\theta}{K_{t-1}^\theta}$$

- Production function:

$$Y_t^\theta = (u_t^\theta K_{t-1}^\theta)^\alpha (\gamma^t N_t^\theta)^{1-\alpha}$$

- Optimal price setting:

$$Q_t^\theta \left\{-\frac{1}{\lambda_f - 1} Y_t^\theta + \frac{\lambda_f}{\lambda_f - 1} MC_t^\theta Y_t^\theta - \varphi_p \left(\Pi_t^\theta - \Pi\right)\Pi_t^\theta Y_t^\theta\right\}$$
$$+ \frac{\beta \varphi_p}{\Pi} \mathbb{E}_t^\theta \left[Q_{t+1}^{RE}\left(\Pi_{t+1}^{RE} - \Pi\right)(\Pi_{t+1}^{RE})^2 Y_{t+1}^{RE}\right] = 0$$

- Optimal wage setting:

$$Q_t^\theta \left[\left(-\frac{1}{\lambda_n - 1} \right) N_t^\theta + (C_t^\theta - bC_{t-1}^\theta) \left(\frac{\lambda_n}{\lambda_n - 1} \right) \left(N_t^\theta \right)^{1+\eta} \frac{1}{\widetilde{W}_t^\theta} - \varphi_w \left(\Pi_{w,t}^\theta - \gamma \Pi \right) \Pi_{w,t}^\theta \right]$$
$$+ \frac{\beta \varphi_w}{\Pi} \mathbb{E}_t^\theta \left[Q_{t+1}^{RE} \left(\Pi_{w,t+1}^{RE} - \gamma \Pi \right) (\Pi_{w,t+1}^{RE})^2 \right] = 0$$

- Nominal wage inflation:

$$\Pi_{w,t}^\theta = \Pi_t^\theta \frac{\widetilde{W}_t^\theta}{\widetilde{W}_{t-1}^\theta}$$

- Resource constraint:

$$C_t^\theta + I_t^\theta + \frac{\varphi_p}{2} \left(\Pi_t^\theta - \Pi \right)^2 Y_t^\theta + \frac{\varphi_w}{2} \left(\Pi_{w,t}^\theta - \gamma \Pi \right)^2 \widetilde{W}_t^\theta + a(u_t^\theta) K_{t-1}^\theta = Y_t^\theta$$

- GDP:

$$Y_t^{G,\theta} = Y_t^\theta - \frac{\varphi_p}{2} \left(\Pi_t^\theta - \Pi \right)^2 Y_t^\theta - \frac{\varphi_w}{2} \left(\Pi_{w,t}^\theta - \gamma \Pi \right)^2 \widetilde{W}_t^\theta - a(u_t^\theta) K_{t-1}^\theta$$

- Taylor rule:

$$\frac{R_t^\theta}{R} = \left(\frac{R_{t-1}^\theta}{R} \right)^{\rho_R} \left\{ \left(\frac{\widetilde{\Pi}_t^\theta}{\Pi} \right)^{\phi_\pi} \left(\frac{Y_t^{G,\theta}}{\gamma Y_{t-1}^{G,\theta}} \right)^{\phi_Y} \right\}^{1-\rho_R} \varepsilon_t$$

C Solution Algorithm

We start from a linear RE system

$$\underset{n\times n}{\mathbf{\Gamma}_0} \underset{n\times 1}{\mathbf{x}_t^{RE}} = \underset{n\times n}{\mathbf{\Gamma}_1} \underset{n\times 1}{\mathbf{x}_{t-1}^{RE}} + \underset{n\times n_s}{\mathbf{\Psi}} \underset{n_s\times 1}{\varepsilon_t} + \underset{n\times n_e}{\mathbf{\Pi}} \underset{n_e\times 1}{\eta_t^{RE}}$$

where $\mathbf{x}_t^{RE}$, ε_t and η_t^{RE} are vectors of endogenous variables, shocks, and expectation errors, respectively. A recursive law of motion can be obtained, using for example Sims (2000), as:

$$\mathbf{x}_t^{RE} = \mathbf{T}^{RE} \mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE} \varepsilon_t.$$

Note that the solution can be divided based on the non-expectation $\left(\widetilde{\mathbf{x}}_t^{RE} \right)$ and expectation terms $\left(\mathbb{E}_t \mathbf{y}_{t+1}^{RE} \right)$:

$$\begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_t \mathbf{y}_{t+1}^{RE}} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{T}_{11}^{RE}} & \underset{(n-n_e)\times n_e}{\mathbf{T}_{12}^{RE}} \\ \underset{n_e\times(n-n_e)}{\mathbf{T}_{21}^{RE}} & \underset{n_e\times n_e}{\mathbf{T}_{22}^{RE}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_{t-1} \mathbf{y}_t^{RE}} \end{bmatrix} + \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{R}_1^{RE}} \\ \underset{n_e\times n_s}{\mathbf{R}_2^{RE}} \end{bmatrix} \varepsilon_t$$

where $\mathbf{y}_{t+1}^{RE}$ is a subset of $\widetilde{\mathbf{x}}_{t+1}^{RE}$.

Define:

$$\mathbf{x}_t^\theta = \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_t \mathbf{y}_{t+1}^{RE}\right)^\theta} \end{bmatrix}$$

Note that $\left(\mathbb{E}_t \mathbf{y}_{t+1}^{RE}\right)^\theta$ denotes the realized value for rational expectations, so it is different from $\mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE}$. We have:

$$\mathbb{E}_t \mathbf{y}_{t+1}^{RE} = \mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^\theta = \left(\mathbb{E}_t \mathbf{y}_{t+1}^{RE}\right)^\theta$$

where $\mathbf{M}$ is a matrix that extract the relevant elements from $\mathbf{T}^{RE}\mathbf{x}_t^\theta$. Note that the equation needs to be included to the system of equations for the DE model because it provides the law of motion for the realized expectations. To see this,

$$\begin{aligned} (\mathbb{E}_t \mathbf{y}_{t+1}^{RE})^\theta &= \underbrace{[\mathbf{M}_1 : \mathbf{0}]}_{\mathbf{M}} \begin{bmatrix} \mathbf{T}_{11}^{RE} & \mathbf{T}_{12}^{RE} \\ \mathbf{T}_{21}^{RE} & \mathbf{T}_{22}^{RE} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{x}}_t^\theta \\ (\mathbb{E}_t \mathbf{y}_{t+1}^{RE})^\theta \end{bmatrix} \\ &= \mathbf{M}_1\mathbf{T}_{11}^{RE}\widetilde{\mathbf{x}}_t^\theta + \mathbf{M}_1\mathbf{T}_{12}^{RE}(\mathbb{E}_t \mathbf{y}_{t+1}^{RE})^\theta \end{aligned}$$

so

$$-\mathbf{M}_1\mathbf{T}_{11}^{RE}\widetilde{\mathbf{x}}_t^\theta + (\mathbf{I} - \mathbf{M}_1\mathbf{T}_{12}^{RE})(\mathbb{E}_t \mathbf{y}_{t+1}^{RE})^\theta = 0.$$

It is useful to divide variables $\mathbf{x}_t^{RE}$ in the original gensys system into non-expectation terms and expectation terms:

$$\begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{0,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{0,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{\Gamma}_{0,21}} & \underset{n_e\times n_e}{\mathbf{\Gamma}_{0,22}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_t \mathbf{y}_{t+1}^{RE}} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{1,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{1,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{\Gamma}_{1,21}} & \underset{n_e\times n_e}{\mathbf{\Gamma}_{1,22}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}} \\ \underset{n_e\times 1}{\mathbb{E}_{t-1} \mathbf{y}_t^{RE}} \end{bmatrix} + \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{\Psi}_1} \\ \underset{n_e\times n_s}{\mathbf{\Psi}_2} \end{bmatrix} \underset{n_s\times 1}{\varepsilon_t} + \begin{bmatrix} \underset{(n-n_e)\times n_e}{\mathbf{\Pi}_1} \\ \underset{n_e\times n_e}{\mathbf{\Pi}_2} \end{bmatrix} \underset{n_e\times 1}{\eta_t^{RE}}$$

Then, the model under DE can be expressed using matrix notation as:

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t \tag{50}$$

where $\mathbf{\Gamma}_0^\theta$ includes the RE restrictions:

$$\begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{0,11}} & \underset{(n-n_e)\times n_e}{\mathbf{0}} \\ \underset{n_e\times(n-n_e)}{-\mathbf{M}_1\mathbf{T}_{11}^{RE}} & \underset{n_e\times n_e}{\mathbf{I} - \mathbf{M}_1\mathbf{T}_{12}^{RE}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_t \mathbf{y}_{t+1}^{RE}\right)^\theta} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times n_e}{-\mathbf{\Gamma}_{0,12}} \\ \underset{n_e\times n_e}{\mathbf{0}} \end{bmatrix} \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} + \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{1,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{1,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{0}} & \underset{n_e\times n_e}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_{t-1} \mathbf{y}_t^{RE}\right)^\theta} \end{bmatrix} + \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{\Psi}_1} \\ \underset{n_e\times n_s}{\mathbf{0}} \end{bmatrix} \varepsilon_t$$

Then:

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta E_t^\theta \mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t$$
$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \left[(1+\theta)\, \mathbb{E}_t \mathbf{y}_{t+1}^{RE} - \sum_{j=1}^{J} \theta \alpha_j \mathbb{E}_{t-j} \mathbf{y}_{t+1}^{RE} \right] + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t$$

Suppose that we do not need all elements in $\mathbf{x}_t^\theta$ to form expectations about the future.[28] In particular, we have

$$\mathbf{y}_t^{RE} = \mathbf{M}\mathbf{x}_t^{RE}$$
$$\mathbf{x}_t^{RE} = \mathbf{T}^{RE}\mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE}\varepsilon_t$$

but can be reduced to

$$\mathbf{y}_t^{RE} = \widetilde{\mathbf{M}}\widetilde{\mathbf{x}}_t^{RE}$$
$$\widetilde{\mathbf{x}}_t^{RE} = \widetilde{\mathbf{T}}^{RE}\widetilde{\mathbf{x}}_{t-1}^{RE} + \widetilde{\mathbf{R}}^{RE}\varepsilon_t$$

Then (50) becomes

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \left[(1+\theta)\, \mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^\theta - \sum_{j=1}^{J} \theta \alpha_j \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE}\right)^{j+1} \widetilde{\mathbf{x}}_{t-j}^\theta \right] + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t. \qquad (51)$$

This becomes:

$$\begin{aligned}\left[\mathbf{\Gamma}_0^\theta - \mathbf{\Gamma}_2^\theta (1+\theta)\, \mathbf{M}\mathbf{T}^{RE}\right] \mathbf{x}_t^\theta = {} & \left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta \theta \alpha_1 \mathbf{M}(\mathbf{T}^{RE})^2\right] \mathbf{x}_{t-1}^\theta \\ & - \mathbf{\Gamma}_2^\theta \theta \alpha_2 \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE}\right)^3 \widetilde{\mathbf{x}}_{t-2}^\theta \\ & \ldots \\ & - \mathbf{\Gamma}_2^\theta \theta \alpha_J \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE}\right)^{J+1} \widetilde{\mathbf{x}}_{t-J}^\theta \\ & + \mathbf{\Psi}^\theta \varepsilon_t.\end{aligned}$$

[28] The method can easily allow for the case where we need full elements in $\mathbf{x}_t^\theta$ to form expectations. The advantage of the current method is that its state space is smaller and hence is useful for a DSGE estimation, among other things.

The solution can be obtained inverting the LHS matrix:

$$\begin{aligned}\mathbf{x}_t^\theta =&(\mathbf{A}_0^\theta)^{-1}\left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta\theta\alpha_1\mathbf{M}(\mathbf{T}^{RE})^2\right]\mathbf{x}_{t-1}^\theta \\ &-(\mathbf{A}_0^\theta)^{-1}\mathbf{\Gamma}_2^\theta\theta\alpha_2\widetilde{\mathbf{M}}\left(\widetilde{\mathbf{T}}^{RE}\right)^3\widetilde{\mathbf{x}}_{t-2}^\theta \\ &\dots \\ &-(\mathbf{A}_0^\theta)^{-1}\mathbf{\Gamma}_2^\theta\theta\alpha_J\widetilde{\mathbf{M}}\left(\widetilde{\mathbf{T}}^{RE}\right)^{J+1}\widetilde{\mathbf{x}}_{t-J}^\theta \\ &+(\mathbf{A}_0^\theta)^{-1}\mathbf{\Psi}^\theta\varepsilon_t,\end{aligned}$$

where $\mathbf{A}_0^\theta \equiv \left[\mathbf{\Gamma}_0^\theta - \mathbf{\Gamma}_2^\theta(1+\theta)\mathbf{M}\mathbf{T}^{RE}\right]$.

Writing in a more compact form, we obtain

$$\underbrace{\begin{bmatrix}\mathbf{x}_t^\theta \\ \widetilde{\mathbf{x}}_{t-1}^\theta \\ \vdots \\ \widetilde{\mathbf{x}}_{t-J+1}^\theta\end{bmatrix}}_{\mathbf{z}_t^\theta} = \underbrace{\begin{bmatrix}\left(\mathbf{A}_0^\theta\right)^{-1}\left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta\theta\alpha_1\mathbf{M}(\mathbf{T}^{RE})^2\right] & -\left(\mathbf{A}_0^\theta\right)^{-1}\mathbf{\Gamma}_2^\theta\theta\alpha_2\widetilde{\mathbf{M}}\left(\widetilde{\mathbf{T}}^{RE}\right)^3 & \dots & -\left(\mathbf{A}_0^\theta\right)^{-1}\mathbf{\Gamma}_2^\theta\theta\alpha_J\widetilde{\mathbf{M}}\left(\widetilde{\mathbf{T}}^{RE}\right)^{J+1} \\ \mathbf{S} & \mathbf{0} & \dots & \mathbf{0} \\ & & \mathbf{I} & \\ \mathbf{0} & \dots & & \mathbf{0}\end{bmatrix}}_{\mathbf{T}^\theta}$$

$$\underbrace{\begin{bmatrix}\mathbf{x}_{t-1}^\theta \\ \widetilde{\mathbf{x}}_{t-2}^\theta \\ \vdots \\ \widetilde{\mathbf{x}}_{t-J}^\theta\end{bmatrix}}_{\mathbf{z}_{t-1}^\theta} + \underbrace{\begin{bmatrix}\left(\mathbf{A}_0^\theta\right)^{-1}\mathbf{\Psi}^\theta \\ \mathbf{0} \\ \vdots \\ \mathbf{0}\end{bmatrix}}_{\mathbf{R}^\theta}\varepsilon_t,$$

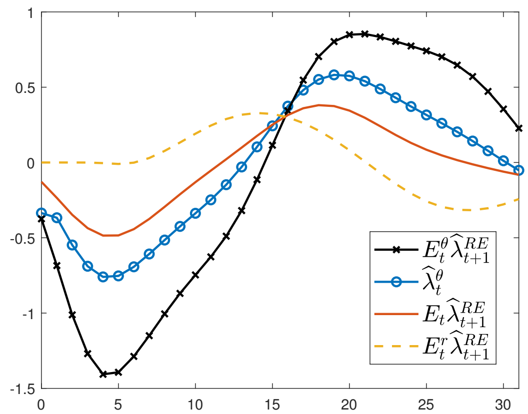
where $\mathbf{S}$ is a selection matrix that relates $\mathbf{x}_t^\theta$ to $\widetilde{\mathbf{x}}_t^\theta$:

$$\widetilde{\mathbf{x}}_t^\theta = \mathbf{S}\mathbf{x}_t^\theta.$$

Finally, we check that all variables over which we take DE present residual uncertainty. To do this, we define a vector $\mathbf{w}_t^{RE} = \mathbf{Q}\mathbf{x}_t^{RE}$ that extracts all relevant linear combinations from the vector $\mathbf{x}_t^{RE}$. This vector contains all and only the variables over which we compute DE. Then, for each element $w_{j,t}^{RE}$ of this vector we verify that the one-step-ahead conditional variance is positive:

$$Var_t\left(w_{j,t+1}^{RE}\right) = (\mathbf{Q}\mathbf{R}^{RE}\mathbf{\Sigma}(\mathbf{Q}\mathbf{R}^{RE})')_{j,j} > 0,$$

Figure 5: Impulse response of marginal utility



Notes: The Figure shows the DE marginal utility ($\mathbb{E}_t^\theta \widehat{\lambda}_{t+1}^{RE}$), realized equilibrium marginal utility ($\widehat{\lambda}_t^\theta$), RE marginal utility ($\mathbb{E}_t \widehat{\lambda}_{t+1}^{RE}$) and reference expectation of marginal utility ($\mathbb{E}_t^r \widehat{\lambda}_{t+1}^{RE}$).

where $\mathbf{\Sigma} \equiv \mathbb{E}_t[\varepsilon_{t+1}\varepsilon_{t+1}']$ and $(\cdot)_{j,j}$ indicates the j-th diagonal element of the matrix.

D Additional Results

In this appendix we report some additional results for the estimated DSGE model.

Table 1 reports the priors and the posterior mode for the model parameters of the DE model and RE re-estimated model. Standard deviations are reported in parentheses. The priors are symmetric across the two models and diffuse.

Figure 5 reports the impulse response of the marginal utility to an expensionary monetary policy shock, given that the estimated Euler Equation features habits:

$$-\mathbb{E}_t^\theta \left(\widehat{\lambda}_{t+1}^{RE}\right) + \widehat{\lambda}_t^\theta = \widehat{r}_t^\theta - \mathbb{E}_t^\theta \widehat{\pi}_{t+1}^{RE} - \theta \pi_{J,t}^* \tag{52}$$

where

$$\widehat{\lambda}_t^\theta = -\frac{\widehat{c}_t^\theta - b\gamma^{-1}\widehat{c}_{t-1}^\theta}{1 - b\gamma^{-1}}. \tag{53}$$

Marginal utility follows a symmetric pattern with respect to consumption, once controlling for habits. The initial increase in consumption is associated with low expected marginal utility that induces expectations of even lower marginal utility. Thus, agents expect consumption to increase even when controlling for the stock of habits. As the economy progresses in its response to the shock, consumption starts declining and marginal utility to increase. However, reference expectations for marginal utility also start increasing. This is because reference expectations were formed at a time of high consumption. Under RE, agents expect a fairly quick return to the steady state from above, implying consumption lower than the stock

Table 1: Estimated parameters

		Prior			Posterior mode	
		Type	Mean	Std	DE	RE
η	Inverse Frisch elasticity	G	2	0.3	2.44	1.74
					(0.30)	(0.30)
b	Consumption habit	B	0.5	0.2	0.80	0.93
					(0.01)	(0.01)
τ	Utilization cost	IG	1	1	0.22	0.26
					(0.01)	(0.01)
κ	Investment adjustment cost	G	2	0.2	3.30	5.24
					(0.20)	(0.32)
φ_p	Price adjustment cost	G	100	20	232.3	252.0
					(21.6))	(29.7)
φ_w	Wage adjustment cost	G	100	20	90.5	61.5
					(21.2)	(16.7)
ρ_R	Taylor rule smoothing	B	0.5	0.2	0.005	0.82
					(0.004)	(0.014)
ϕ_π	Taylor rule inflation	N	1.5	0.4	1.001	1.000
					(0.007)	(0.031)
ϕ_Y	Taylor rule output	N	0.1	0.05	0.67	0.20
					(0.02)	(0.05)
$100\sigma_R$	Monetary policy shock	IG	1	1	0.17	0.16
					(0.01)	(0.01)
θ	Diagnostic parameter	G	1	0.1	1.91	–
					(0.10)	
μ	Memory distribution mean	B	0.5	0.2	0.19	–
					(0.01)	
σ	Memory distribution stdev	G	0.2	0.05	0.04	–
					(0.004)	
Log marginal likelihood					**-464**	**-504**

Notes: 'DE' corresponds to the model with diagnostic expectations and 'RE' corresponds to the rational expectations version. B refers to the Beta distribution, N to the Normal distribution, G to the Gamma distribution, IG to the Inverse-gamma distribution. Posterior standard deviations are in parentheses and are obtained from draws using the random-walk Metropolis-Hasting algorithm. The marginal likelihood is calculated using Geweke's modified harmonic mean estimator.

of habits, leading to a positive RE marginal utility. However, under DE, the return to the steady state is slower than expected as agents remain overly optimistic for a while. Agents are still surprised by the high consumption, leading to a negative surprise in marginal utility, amplified by DE. Thus, past decisions feed into current beliefs, affecting the duration and amplitude of the cycle. It is only around 15 quarters that reference expectations catch up with the current marginal utility. As consumption moves below trend, agents start expecting a return to the steady from below, generating a negative reference expectation for marginal utility as consumption is expected to be higher than the stock of habits. In the bust phase, agents are surprised by the fact that consumption is still well below trend, leading to a positive surprise in marginal utility, that induces magnified DE of high marginal utility in the future.

How can we rationalize this behavior from the perspective of the Euler equation under DE in (44)? As mentioned in the paper, a key role is played by the surprise in cumulative inflation $\pi^*_{J,t}$ with respect to the reference expectations formed in the past. On impact, because of an increase in utilization, inflation declines. This determines a negative surprise in the price level that induces a misperception in the model relevant real interest rate that starts increasing. This perceived high real interest is, in the eyes of the agent, justified in light of a perceived acceleration in consumption that more than compensates for the habit stock. In other words, not only agents expect consumption to be higher in the future, but they also expect the marginal utility to be lower: $-\mathbb{E}^{\theta}_t\left(\widehat{\lambda}^{RE}_{t+1}\right)+\widehat{\lambda}^{\theta}_t>0$ implies $\mathbb{E}^{\theta}_t\left(\widehat{c}^{RE}_{t+1}-b\gamma^{-1}\widehat{c}^{RE}_t\right)-\left(\widehat{c}^{\theta}_t-b\gamma^{-1}\widehat{c}^{\theta}_{t-1}\right)>0$. Eventually, inflation starts picking up, leading first to a reduction in the negative surprises for the price level and then eventually to positive surprises. This determines a reversal in the model relevant real interest rate that moves into the negative territory during the bust part of the cycle, when agents find the perceived low real interest rate justified in light of their excessive pessimism. Now not only they expect consumption to decline, but also to do so in a way to increase the marginal utility.