NBER WORKING PAPER SERIES

IMPLICATIONS OF DIAGNOSTIC EXPECTATIONS: THEORY AND APPLICATIONS

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

We thank Daniele d'Arienzo, Spencer Kwon, Peter Maxted, Andrei Shleifer and Stephen Terry for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Implications of Diagnostic Expectations: Theory and Applications
Francesco Bianchi, Cosmin L. Ilut, and Hikaru Saijo
NBER Working Paper No. 28604
March 2021
JEL No. D8,D9,E03,E3,E71

ABSTRACT

A large psychology literature argues that decision-makers' forecasts of their future circumstances appear overly influenced by their perception of 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 a broad class of macroeconomic models. In this environment, DE apply to both exogenous and endogenous variables. We derive three theoretical properties of DE in the presence of endogenous variables: (i) endogenous predictability,(ii) endogenous non-stochasticity, and (iii) the failure of the law of iterated expectations under distant memory. We show that these properties imply (i) a joint determination of actions and DE; (ii) the possibility of silencing DE by policy actions; (iii) the possibility of time-inconsistency. We analyze two approaches to deal with the issue of time inconsistency: naivete and sophistication. We illustrate our analysis' relevance in two applications. First, we provide a portable solution algorithm to incorporate DE into recursive linear models. In an RBC model, DE generate rich and novel propagation dynamics and a boom-bust cycle. Second, a Fisherian model shows that policy makers' behavior has pervasive macroeconomic effects by activating or silencing DE.

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, an unusually high current level of financial resources triggers more vivid memories of good times for the agent, which 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 applications, so far the DE paradigm has been typically studied in environments where agents' economic circumstances are determined exogenously. However, in a large set of situations, decisions involve a feedback between agents' beliefs and endogenously determined economic states. Furthermore, in state-of-the-art macroeconomic models policymakers' interventions can alter economic outcomes, potentially affecting the variables that the agent is interested in forecasting. Motivated by these observations, this paper aims at modeling the feedback between actions and agents' beliefs over exogenous and endogenous variables, as affected by the psychologically-founded DE paradigm.

We develop the micro-foundations for applying the DE paradigm to both exogenous and endogenous variables and characterize the general equilibrium formulations for a large class of recursive macroeconomic models. In our analysis, we highlight the potential of the proposed approach to deliver micro and macro dynamics that fit the data better and document novel welfare implications. We also characterize optimal actions that 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 also provide tractable solution methods and develop a portable toolbox that can be used to enrich standard general equilibrium models with DE.

Under DE, agents form expectations about future outcomes giving more weight to those scenarios that became more likely with respect to agents' previous beliefs. Under the 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. The parameter θ controls the severity of the distortion, while the parameter J controls the lag of the reference beliefs.

As a first step, we study the implications of DE in a simple statistical example in which an agent forecasts a random variable that presents an exogenous and an endogenous component. In this stylized example, a variable known at time t affects outcomes of the variable of interest at time $t+1$. We use this statistical model to highlight three key properties of DE that will serve as guidelines to interpret the results of the paper. First, the endogenous response to the exogenous state variable affects the predictability of the variable that the agent is trying to forecast. This first feature of DE, that we label *endogenous predictability*, in turn, affects the extent of the over-reaction in expectations for a given level of DE distortion. Second, if the endogenous component completely removes uncertainty about future outcomes, then DE become irrelevant as the agent does not face any residual uncertainty about the variable she is interested in forecasting, even if she might still face uncertainty about the exogenous shocks. We refer to this property with the term *endogenous non-stochasticity.* This property is tightly linked to the heuristics of DE emphasized by Bordalo et al. (2018) and pointed out earlier by Gennaioli and Shleifer (2010): Limits to memory recall arise only when there is residual uncertainty about the variable that the agent is interested in predicting. The novelty here is that the existence of residual uncertainty can be affected by endogenous actions, like policymakers' interventions. Third, when the reference point for the DE distortion is not pinned down by the immediate past ($J > 1$), the law of iterated expectations 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. We refer to this property as the *failure of LOIE under distant memory.*

To understand the important ramifications of the first two properties, we analyze a two-period consumption-saving problem. To isolate the effects of the first property, *endogenous predictability*, we assume that the exogenous income process is an iid random variable. In this model, the endogenous component of the variable that the agent is trying to forecast, total future resources, is the result of the saving decision made by the agent herself. Following a current unusually high (low) income shock, and for a given response of current savings to this innovation, the agent correctly realizes that her future available resources and consumption are more likely to be higher (lower) than usual. With iid future income shocks, this conditional predictability of future resources comes entirely from the savings choice, which induces that *endogenous predictability* from the current income shock to the future available resources. 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 than usual, and importantly, more than under the true distribution, leading her make those saving decisions under an "as if" optimistic view of future resources.

Under quadratic utility, both DE and RE agents take optimal 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 that diagnostic information, she optimally consumes 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).

The two-period model also allows us to study the welfare implications of removing uncertainty under DE. In RE models, eliminating future uncertainty improves ex-ante utility by reducing risk and possibly by changing the intertemporal allocation of resources, as the precautionary savings motive disappears. Under DE a new channel arises, as a result of the second property, *endogenous non-stochasticity.* As explained above, DE are activated by the presence of residual uncertainty. Absent residual uncertainty, agents' expectations collapse to their RE counterpart. Thus, in line with the formulation of DE proposed by Bordalo et al. (2018) and adopted in this paper, allocations and welfare present a discontinuity: As long as residual uncertainty is present, DE are active; as soon as residual uncertainty is removed, DE distortions are silent. We show that policymakers can improve welfare by removing uncertainty about future outcomes and in this way silence the distortions due to DE.

We then move to showcase the consequences of the third property of DE, the *failure of LOIE under distant memory.* As argued above, DE generally violate the LOIE once the reference distribution is based on expectations formed more than one period earlier ($J > 1$). In a multi-period model the failure of the LOIE critically matters because it leads to time-inconsistent choices. To illustrate this result, we extend the two-period model to include a third period. The time 1 problem of the agent is now to choose actual savings in period 1 jointly with a contingent plan on how to choose savings in period 2, so to maximize current utility and the expected sum of future utilities. The key source of possible time-inconsistency is that when evaluating the optimal plan at time 1, the agent solves the trade-off between consumption in period 2 and period 3 in a way that reflects her distorted beliefs as of time 1. When $J > 1$, conditional on reaching period 2, she will evaluate the tradeoff between period 2 and 3 differently than she did in period 1 as her reference beliefs have changed. She will then choose a time 2 savings course of action, contingent on her inherited savings from period 1, that differs compared to what she initially planned in period 1. Thus, the behavior of an agent subject to the representative heuristic is generally time-inconsistent, except in the special case in which the comparison group is the immediate past ($J = 1$). In this case, the amount of savings chosen at time 1 will not constitute a surprise at time 2 because the saving choice made at time 1 also serves as reference point for the comparison group.

To address the issue of time inconsistency, and be able to 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.

When $J = 1$, we can show analytically three important and connected results: (a) choices are time-consistent because LOIE holds; (b) both the naïve and sophisticated problems lead to the same optimal savings functions in the first and second period, and (c) these DE savings policy functions feature a response to the endogenous economic state that equals that of the RE policy functions, but a muted response to the current income innovation. The intuitive reason behind these three results is that when $J = 1$ savings chosen in the first period also pins down reference expectations in the second period. Therefore, in that second period the new information to which agents over-react to only depends on the exogenous innovation and not the endogenous state represented by past savings. This over-sensitivity of beliefs to the period-two exogenous shock leads to a muted reaction to the exogenous shock, but not to the endogenous state chosen in the first period.

On the other hand, if $J > 1$ the selective memory recall is based on more distant information. We use an analytical illustration based on $J = 2$ to investigate the implications of time inconsistency with respect to the saving choice planned for the second period. In this case, sophistication and naïveté deliver different results. 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 anymore for the comparison group pinning down the selective memory recall process. Instead, the agent looks at expectations 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 hew 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.

We argue that the naïveté approach has three main advantages. First, it is portable, as it can be applied to any linearized model. Second, we argue that the solution under sophistication in the infinite horizon setting implies a level of ultra-rationality that is arguably in tension with the motivation of accounting for belief heuristics, usually viewed as a mental shortcut that allows agents to make judgments quickly and efficiently (eg. Tversky and Kahneman (1975)). Third, for the similar reasons of computational complexity, it is also more tractable and advantageous to characterize for us, as modelers.

Equipped with these results, we move to apply DE to a full-fledged infinite horizon RBC model that we solve based on the naïveté approach. We first explain how to solve any linear general equilibrium model in the presence of DE by using standard solution methods, such as Sims (2000). Intuitively, solving a model with DE requires building a shadow RE economy that keeps track of agents' reference points. We then study how the propagation of shocks changes under DE. We show that when $J = 1$, DE have modest effects, as expectations are quickly realigned with RE. On the other hand, when $J > 1$ DE can generate rich dynamics, bringing a very parsimonious model closer to the data. First, DE under distant memory leads to significant persistence and the possibility of abrupt changes in the propagation of the shocks as agents can suddenly come to realize that past misperceptions led to the wrong capital accumulation decisions. Thus, a parsimonious RBC model is able to generate episodes of boom and bust (Christiano et al. (2008)) and Minsky (1977) moments. Second, there are predictable differences (or 'wedges') between the typical realization of future return on capital and its expected value under DE beliefs. Thus, when $J > 1$, we obtain long-lived predictable wedges and reversal, as suggested for example by the evidence in López-Salido et al. (2017). Third, DE can also significantly increase the volatility of consumption and hours for standard parameterizations, typically a challenge for standard RBC models (King and Rebelo (1999)).

After having established that DE can generate rich propagation dynamics, we close our analysis by showing that policy makers' behavior can have profound effects on the properties of the macroeconomy by activating or silencing DE. This is because changes in policymakers' behavior can affect predictability and uncertainty about future outcomes. To illustrate this point, we consider a simple Fisherian model in which the real interest rate follows an iid exogenous process and the central bank moves the nominal interest rate in response to inflation. We study the consequences of switches in the conduct of monetary policy between rules that satisfy the Taylor principle and rules that do not. In the first case, the central bank reacts more than one-to-one to inflation, in the second case it does not. We find that DE are silent when the Taylor principle is satisfied, while they become relevant when policymakers violate the Taylor principle. Interestingly, under this second scenario, DE can be relevant even when inflation has *zero persistence* because exogenous shocks act as *news* about future inflation. In this context, sunspot shocks play a key role by generating uncertainty about future outcomes. Without such uncertainty, DE would be silent also when the Taylor principle is violated. This example showcases how DE arise in presence of two distinctive features: Predictability, from the real interest shocks, and residual uncertainty, from the sunspot shock.

Besides the work cited above, our paper is particularly related to some recent contributions that introduce DE into macro models. Bordalo et al. (2019a) study DE about an exogenous TFP process to account for credit cycles, L'Huillier et al. (2021) analyze the effects of introducing DE in RBC and NK models, while Maxted (2020) studies the consequences of introducing DE into a He and Krishnamurthy (2019) style macro-finance model. Our paper contributes to this literature in some important dimensions. Like Maxted (2020) and L'Huillier et al. (2021), we study general equilibrium models in which DE apply to both exogenous and endogenous variables. However, unlike these papers, we build on the three distinctive features of DE outlined above. First, in line with the paradigm developed in Bordalo et al. (2018), we leverage the notion that DE arise only when residual uncertainty exists. We show that this feature of DE introduces an interesting and important discontinuity linked to uncertainty and that different policy rules can activate or deactivate DE by introducing or removing uncertainty. In this respect, we take a different approach to L'Huillier et al. (2021) who assume that DE apply even absent any uncertainty. Second, we confront the problem of time inconsistency that arises in the empirically relevant case in which the reference beliefs are based on the more distant past ($J > 1$). We discuss the advantages and disadvantages of the naïve and sophisticated approaches to deal with time inconsistency and show that distant memory is key to obtain rich and novel propagation mechanisms in standard macro models.

The rest of the paper is organized as follows. Section 2 presents the three properties characterizing the DE paradigm using a simple statistical model. Sections 3 and 4 present the implications of these properties in a two-period and a three-period model, respectively. Section 5 studies the quantitative implications of DE in an RBC model. Section 6 presents a Fisherian model to highlight the interaction between policymakers' behavior and DE.

2 Conditional Distributions and Diagnostic Beliefs

We start off our analysis by considering a very stylized statistical description for how a random variable Z_{t+1} evolves and how the agent forms DE over it. In particular, conditional on time t information the true distribution for Z_{t+1} is

$$Z_{t+1} = K_t + \delta Y_{t+1}. \tag{1}$$

The coefficient δ here captures the exposure of the variable Z_{t+1}, which the agent is interested in forecasting, to the randomness in Y_{t+1}. The variable K_t is of key interest for us. In an economic model, we will emphasize its role as an endogenous action optimally taken as a response to the exogenous state Y_t. For this statistical description, what matters is that K_t is another random variable that affects the conditional mean of Z_{t+1} in equation (1). To give an economic context, in our running examples below, Z_{t+1} captures the agent's total financial resources at time $t+1$, Y_{t+1} is a random variable, capturing an exogenous income source which has a conditional normal distribution, and K_t is an endogenous choice of savings.

To fix ideas for this statistical model, consider a simple example where Y_t follows an AR(1) process and the random variable K_t is determined as a simple reaction to Y_t, as in

$$Y_{t+1} = \rho Y_t + \varepsilon_{t+1}; \;\; K_t = \alpha Y_t, \tag{2}$$

where ε_{t+1} are mean zero iid normal shocks with a variance $\sigma^2 > 0$. By substituting out this K_t and Y_{t+1} from equation (2), the conditional distribution of Z_{t+1} in equation (1) becomes

$$Z_{t+1} = \mu^Z_{t+1|t} + \delta\varepsilon_{t+1}. \tag{3}$$

Here $\mu^Z_{t+1|t}$ defines the time t conditional mean $\mathbb{E}_t(Z_{t+1})$, given in this case by

$$\mu^Z_{t+1|t} = (\alpha + \rho\delta)\, Y_t. \tag{4}$$

Therefore, in this case, the time t conditional mean of Z_{t+1} is proportional to the realization of current Y_t by a factor $(\alpha + \rho\delta)$, reflecting the exogenous persistence component (through ρ and exposure of δ) and the *response* of the variable K_t (through α). The conditional mean $\mu^Z_{t+1|t}$ therefore tracks Y_t and, using the law of motion of Y_t in equation (2), follows the process

$$\mu^Z_{t+1|t} = \rho\mu^Z_{t|t-1} + (\alpha + \rho\delta)\,\varepsilon_t. \tag{5}$$

Therefore, by using 'hats' when needed to emphasize the specific realization of any given random variables, equation (3) defines the true distribution associated to some realization $\widehat{Z}_{t+1}$, for a given current realization $\widehat{\mu}^Z_{t+1|t}$ of the conditional mean. For future reference we denote this true conditional normal distribution as $h(\widehat{Z}_{t+1}|\mu^Z_{t+1|t} = \widehat{\mu}^Z_{t+1|t})$.

Diagnostic Expectations and the Representativeness Heuristic. We build on the recent work of Gennaioli and Shleifer (2010), Bordalo et al. (2018) who formulate a behavioral model of *diagnostic expectations* (DE). The fundamental psychological first-principle basis for this 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 precise 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 more recently Bordalo et al. (2020a) and more broadly Bordalo et al. (2018)). The basic intuition brought forward by 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.

The representativeness heuristic and the DE belief is a model of imperfect probability judgements, or, in statistical terms, one of forecasting an otherwise uncertain event. For example, Bordalo et al. (2016) use this representativeness heuristic to build a model of stereotypes. Consider one of their illustrations. An agent is told that he will meet a person that is Irish. Hearing this information, the agent's imperfect memory retrieves more intensely the relative attribute of Irish people having a larger incidence of red hair, compared to other groups, even though this incidence is not that common in absolute terms. This over-reliance of the red hair 'representative' feature of the Irish, leads the agent to over-estimate the probability that the Irish person he will next meet has red hair.

2.1 Deterministic Processes

Gennaioli and Shleifer (2010) already describe the modeling of 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, the starting defining characteristic of this heuristic is that, naturally, it does not distort perceptions when there is no residual uncertainty. Indeed, since in that case the new data completely informs the agent about the realization of the variable that she is otherwise interested in forecasting, the conditional likelihood of observing any other scenario than the one she is now fully informed on has become degenerately equal to zero.[1] The agent does not need to appeal to the recollection of past data, possibly affected by its association with the new data, since she is now fully informed.

This characteristic serves as a useful preliminary step in our analysis given our interest in understanding the role that limited memory plays in biasing the informational content of a predetermined (or state) variable, like K_t in equation (1). The most immediate implication of this characteristic is 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

$$\mathbb{E}_t^{\theta}(K_t) = \mathbb{E}_t(K_t) = \widehat{K}_t, \tag{6}$$

where $\mathbb{E}_t^{\theta}(.)$ and $\mathbb{E}_t(.)$ denote the time t conditional belief of a variable based on limited memory and under the objective process, respectively.[2] Independent of the details of how limited memory affects the agent's probabilistic judgments, the representativeness heuristic does not influence behavior since memory recall is not activated if the new data completely eliminates uncertainty over the variable to be forecasted (as it does in this case for K_t).

The same observation immediately carries over if the agent is interested in forecasting Z_{t+1}, but that variable turns out to be fully pre-determined as of time t. In our simple statistical model, we see this characteristic by setting $\delta = 0$ in equation (3). Based on the

[1] In the context of the earlier illustration, suppose that the Irish person whom the agent has just met has black hair. The agent would report no judgment bias in evaluating the probability that this particular Irish person has red (or black) hair, since the probability of that event has just collapsed to zero (or one, respectively) conditional on her information. Similarly, if before meeting this Irish person the agent is told that the person has black hair, the agent would not show any bias as there is no residual uncertainty about the future event with respect to the hair color.

[2] In the language developed in Bordalo et al. (2018), which we will detail below, to compute the diagnostic expectation $\mathbb{E}_t^{\theta}(K_t)$, the realization $\widehat{K}_t$ constitutes its infinitely representative state (see appendix in Bordalo et al. (2018) on Corollary 1).

current time t data, all residual uncertainty about Z_{t+1} is eliminated, as it is known to equal

$$Z_{t+1} = K_t = \alpha Y_t. \tag{7}$$

Thus, in this case, as in equation (6), $\mathbb{E}_t^{\theta}(Z_{t+1}) = \mathbb{E}_t(Z_{t+1}) = \widehat{K}_t$.

Setting $\delta = 0$ also allows us to note that uncertainty is unconditionally present in this statistical model, since $\sigma^2 > 0$ and thus the agent observes fluctuations in Y_t. Nevertheless, in this case, such fluctuations do not imply uncertainty in the conditional distribution of Z_{t+1}. Here an agent that perfectly observes the conditional mean and knows that there is no residual uncertainty in Z_{t+1}, points to $\widehat{\mu}^Z_{t+1|t} = \alpha Y_t$ as the only possible value of Z_{t+1}, without having to resort to memory recall.

In the micro-founded models that we later study in the paper, we will showcase the important policy implications of eliminating conditional uncertainty (through what we intuitively will refer to as *endogenous non-stochasticity*), and thus, eliminating the possible distortions induced by imperfect memory.

2.2 Selective Memory Recall

Consider now the more general case of a positive amount of residual uncertainty about Z_{t+1}. We obtain this by simply allowing for the rest of this section to have $\delta \neq 0$ in our statistical model. The key point to note here is that even if the agent does not face uncertainty in now-casting K_t (as per equation (6)), when there is strictly positive conditional uncertainty about Z_{t+1}, the new information contained in the observation of K_t does activate memory recall in how to form beliefs in forecasting Z_{t+1}.

As anticipated, in order to model the particular analytical implementation of how that imperfect memory influences behavior, we follow the details of the formulation proposed in Gennaioli and Shleifer (2010) and Bordalo et al. (2018). In particular, in the context of our statistical model above, this work models the distortion in beliefs arising from the representativeness heuristic as the following density $h_t^{\theta}(\widehat{Z}_{t+1})$

$$h_t^{\theta}(\widehat{Z}_{t+1}) = h(\widehat{Z}_{t+1}|\mu^Z_{t+1|t} = \widehat{\mu}^Z_{t+1|t}) \left[\frac{h(\widehat{Z}_{t+1}|\mu^Z_{t+1|t} = \widehat{\mu}^Z_{t+1|t})}{h(\widehat{Z}_{t+1}|\mu^Z_{t+1|t} = \mathbb{E}_{t-J}\mu^Z_{t+1|t})} \right]^{\theta} \frac{1}{a} \tag{8}$$

where a is an integration constant, that ensures that $h_t^{\theta}(\widehat{Z}_{t+1})$ integrates to one.

There are three important elements in this distorted distribution. First, as introduced above, $h(\widehat{Z}_{t+1}|\mu^Z_{t+1|t} = \widehat{\mu}^Z_{t+1|t})$ is the true density. Second, $\mathbb{E}_{t-J}\mu^Z_{t+1|t}$ is the *comparison group* for the random variable $\mu^Z_{t+1|t}$, where $\mathbb{E}_{t-J}$ denotes the expectation operator for any arbitrary random variable conditional on $t-J$ information under the true density. 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, using the law of motion for $\mu^Z_{t+1|t}$ in equation (5), this comparison group takes the form $\mathbb{E}_{t-J}\mu^Z_{t+1|t} = \rho^J \widehat{\mu}^Z_{t+1-J|t-J}$.

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 back 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 particular formulation captures exactly the notion that the agent has the true distribution in the back of her mind, but selectively retrieves and overweighs realizations $\widehat{Z}_{t+1}$ that are representative (or diagnostic) of the group consisting of $\left\{\mu^Z_{t+1|t} = \widehat{\mu}^Z_{t+1|t}\right\}$ relative to the comparison group consisting of $\left\{\mu^Z_{t+1|t} = \mathbb{E}_{t-J}\mu^Z_{t+1|t}\right\}$. Because $h^\theta_t(\widehat{Z}_{t+1})$ overweighs the most diagnostic future outcomes, Bordalo et al. (2018) call these expectations *diagnostic*. Thus, the description of the true process for Z_{t+1} and for its conditional mean $\mu^Z_{t+1|t}$ together with the representativeness parameters $\theta \geq 0$ and $J \geq 1$ fully describe the DE beliefs in equation (8).

While in general it may be difficult to characterize analytically, Bordalo et al. (2018) show how the *normality* assumption on ε_{t+1} leads to a very tractable characterization of the conditional distribution of Z_{t+1} under $h^\theta_t(.)$. In particular, it remains normally distributed, with the same variance $\delta^2\sigma^2$, but a distorted mean

$$\mathbb{E}^\theta_t(Z_{t+1}) = \widehat{\mu}^Z_{t+1|t} + \theta\left(\widehat{\mu}^Z_{t+1|t} - \mathbb{E}_{t-J}\mu^Z_{t+1|t}\right), \tag{9}$$

where the extra term $\theta(\widehat{\mu}^Z_{t+1|t} - \mathbb{E}_{t-J}\mu^Z_{t+1|t})$ captures the *over-reaction of the conditional mean to the new information*.

The same tractable characterization of course applies more generally as long as the random variable to be forecasted is normally distributed. For example, in the context of our simple statistical model, DE satisfy the following additivity property (see the proof of Corollary 1 in Bordalo et al. (2018) for details), when there is residual uncertainty about $Z_{t+1} + Z_{t+2}$

$$\mathbb{E}^\theta_t(Z_{t+1} + Z_{t+2}) = \mathbb{E}^\theta_t(Z_{t+1}) + \mathbb{E}^\theta_t(Z_{t+2}), \tag{10}$$

where $\mathbb{E}^\theta_t(Z_{t+2})$ follows the same structure as in (9), by replacing the conditional mean $\widehat{\mu}^Z_{t+1|t}$ with the two-step ahead conditional mean $\widehat{\mu}^Z_{t+2|t}$ under the true density.

In the next sections we will highlight two important properties of the DE beliefs, driven by the imperfect memory recall that is at the heart of the representativeness heuristic. Briefly, these properties are: (1) DE matters for the sum of a random and a predetermined variable (a property that we intuitively refer to as *endogenous predictability*); and (2) the Law of

Iterated Expectations (LOIE) fails when the comparison group is such that $J > 1$ (i.e. *the failure of LOIE under distant memory*). The formalism behind these properties is not novel to our paper. As with the starting property that there is no DE distortion over purely deterministic variables, described in section 2.1, this formalism has been noticed and proposed as characterizing the DE operator in equation (8) by previous work, such as Bordalo et al. (2018). Our key contribution here is 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 macroeconomic models; (2) the role of policy in altering these distortions by possibly eliminating conditional uncertainty, and (3) how the role of past memory introduces additional informational state variables that can alter significantly the model's dynamics.

2.3 Endogenous Predictability

For the statistical model in equations (1) and (2), the DE conditional mean of equation (9) is

$$\mathbb{E}_t^\theta(Z_{t+1}) = \mathbb{E}_t^\theta(K_t + \delta Y_{t+1}) = (\alpha + \rho\delta)\left[Y_t + \theta\left(Y_t - \rho^J Y_{t-J}\right)\right]. \tag{11}$$

This simply reflects that the true conditional mean $\mu_{t+1|t}^Z$ tracks Y_t by the proportionality factor $(\alpha + \rho\delta)$, as in equation (4). Therefore, the new information in the realization of this conditional mean is also proportional to the new information $\left(Y_t - \rho^J Y_{t-J}\right)$ in the random variable Y_t, when compared to its $t - J$ expectation.

To analyze this first property of interest to us, i.e. the 'additivity' of a predetermined and random variable in equation (11), let us first consider a simple case where Z_{t+1} is just proportional to the exogenous variable Y_{t+1}. Through the lenses of equation (4) this means setting $\alpha = 0$, so that the conditional mean $\widehat{\mu}_{t+1|t}^Z$ in equation (4) becomes equal to $\rho\delta Y_t$. In this case, typically analyzed in the existing literature on DE (see, for example, Bordalo et al. (2019a)), the representativeness heuristic matters only through the effect of the exogenous persistence ρ. Indeed, by equation (11)

$$\mathbb{E}_t^\theta(Z_{t+1}) = \rho\delta\left[Y_t + \theta\left(Y_t - \rho^J Y_{t-J}\right)\right], \text{ when } \alpha = 0 \tag{12}$$

As in equation (11), the new information in Y_t is the same, but here the conditional mean reacts to that only by a factor of $\rho\delta$.

In this context, consider the stark case of $\rho = 0$. By equation (12) the representativeness parameter θ would not matter, as the conditional mean of Z_{t+1} in this case is always zero

$$\mathbb{E}_t^\theta(Z_{t+1}) = 0 = \mathbb{E}_t(Z_{t+1}), \text{ when } \alpha = 0 \text{ and } \rho = 0, \tag{13}$$

both under DE and under the true model. Intuitively, even if memory is imperfect in its nature, this does not play a role since the agent does not perceive any difference between the diagnostic group (equal to zero at all times) and the comparison group (the time $t-J$ expectation of observing that conditional mean of zero).

However, when $\alpha \neq 0$, the exogenous persistence in Y_t is not the only driver of movements in the conditional mean of Z_{t+1} at time t. To see this clearly, we turn off that exogenous persistence by continuing to set $\rho = 0$. Crucially, the conditional mean $\widehat{\mu}^Z_{t+1|t}$ responds one-to-one to K_t, which equals αY_t, so that the DE belief by equation (11) is

$$\mathbb{E}^\theta_t(Z_{t+1}) = \alpha \left[Y_t + \theta \left(Y_t - \rho^J Y_{t-J}\right)\right], \text{ when } \alpha \neq 0 \text{ and } \rho = 0. \tag{14}$$

There are four important remarks to make here. First, in equation (12) the conditional mean of Z_{t+1} moves with Y_t only because the exogenous Y_{t+1} is persistent, while in equation (14) it moves today only because its time t determined component K_t moves with Y_t. In either of these two extremes, DE beliefs over-react to the new information in how that conditional mean has changed today compared to its comparison group. We emphasize the movement of this conditional mean that appears through K_t as a form of *endogenous predictability*.

Second, note that even if Y_{t+1} is iid, and the conditional mean only responds to K_t, the sheer presence of those uncertain future shocks ε_{t+1} activates the need of memory recall, as we discussed in section 2.1. Putting this together with equation (7), we then have

$$\begin{aligned} \mathbb{E}^\theta_t(K_t + \delta Y_{t+1}) &\neq K_t + \delta \mathbb{E}^\theta_t\left(Y_{t+1}\right), \text{ if } \delta \neq 0 \\ &= K_t \qquad\qquad\qquad\;\;, \text{ if } \delta = 0 \end{aligned}$$

where also recall that by equation (6), $K_t = \mathbb{E}^\theta_t\left(K_t\right)$.

The inequality above is the behavioral manifestation of the DE beliefs over a sum of random variable when one of those is predetermined. While this formal result has appeared already in Bordalo et al. (2018), we bring 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_t here.

Third, note that when $\alpha \neq 0$ and $\rho = 0$, the forecast in equation (10), $\mathbb{E}^\theta_t(Z_{t+1} + Z_{t+2}) = \mathbb{E}^\theta_t(Z_{t+1})$, because the two-step ahead forecast $\mathbb{E}^\theta_t(Z_{t+2}) = \mathbb{E}_t(Z_{t+2}) = 0$. Indeed, in this statistical model the persistence of K_{t+1} is intimately linked to the persistence of Y_{t+1}. This showcases how the persistence in the state variable matters for the informational role of the current state in forecasting multi-steps ahead. In economic models, we expect that the state variables themselves have such persistence and therefore this multi-steps ahead forecasts are distorted by the current observation of K_t.

Fourth, when $\rho \neq 0$, but $\alpha = -\rho\delta$, the K_t component exactly offsets the persistence coming from Y_{t+1}, so that the conditional mean is zero under both the true process and the DE beliefs: $\mathbb{E}_t^\theta(Z_{t+1}) = \mathbb{E}_t(Z_{t+1}) = 0$. Similar to equation (13), this is another case where even if memory is imperfect, the selective retrieval affect does not end up mattering because the conditional mean is always zero, both in the representative and the diagnostic group.

In the micro-founded models that we study later this formalism implies that we need to characterize *jointly* how a given response α affects DE beliefs and, critically, how these beliefs in turn affect the optimal response α. Before we do so, we find it useful to analyze an additional property of DE that will be important for the informational role of state variables, namely the Law of Iterated Expectations (LOIE). The reason that LOIE is important in dynamic models such as those we are interested is that its failure generally leads to time-inconsistency.

2.4 Failure of LOIE under Distant Memory

To analyze this issue, consider some arbitrary periods $t > J$, integers $m \geq 1, n \geq 1$, and some comparison group $t - J$ in equation (8), where $J \geq 1$. The same formalism that lead to equation (9) for $J = 1$, can be extended to (as discussed in Bordalo et al. (2018), Corollary 1)

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

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

$$\mathbb{E}_t^\theta\left[\mathbb{E}_{t+m}^\theta Z_{t+m+n}\right] = (1+\theta)\left[\mathbb{E}_t Z_{t+m+n} + \theta\left(\mathbb{E}_t Z_{t+m+n} - \mathbb{E}_t\mathbb{E}_{t+m-J} Z_{t+m+n}\right)\right] - \theta\mathbb{E}_{t-J} Z_{t+m+n}. \tag{15}$$

We are then interested in establishing whether the LOIE holds, i.e. whether

$$\mathbb{E}_t^\theta\left[\mathbb{E}_{t+m}^\theta Z_{t+m+n}\right] = \mathbb{E}_t^\theta\left[Z_{t+m+n}\right]. \tag{16}$$

As we discuss below, the key term in the conditional belief of equation (15) that will matter for LOIE is the perceived surprise

$$\mathbb{E}_t Z_{t+m+n} - \mathbb{E}_t\mathbb{E}_{t+m-J} Z_{t+m+n}. \tag{17}$$

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

To prove this, 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 $\mathbb{E}_t\mathbb{E}_{t+m-J} Z_{t+m+n} = \mathbb{E}_t Z_{t+m+n}$ for $J \leq m$. It follows that the surprise in equation (17) is zero and the LOIE holds under the DE operator:

$$\mathbb{E}_t^\theta\left[\mathbb{E}_{t+m}^\theta Z_{t+m+n}\right] = \mathbb{E}_t Z_{t+m+n} + \theta\left(\mathbb{E}_t Z_{t+m+n} - \mathbb{E}_{t-J} Z_{t+m+n}\right) = \mathbb{E}_t^\theta\left[Z_{l+m+n}\right]. \tag{18}$$

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 (17) is not zero and constitutes an additional source of variation for $\mathbb{E}_t^\theta\left[\mathbb{E}_{t+m}^\theta Z_{t+m+n}\right]$ in equation (15), which now becomes

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

Critically, by comparing this result to that of equation in (18), for the generic case of $\mathbb{E}_t Z_{t+m+n} \neq \mathbb{E}_{t+m-J} Z_{t+m+n}$, which occurs with probability one, the LOIE in equation (16) 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 Z_{t+m+n} introduces an *additional* lagged forecast (here $\mathbb{E}_{t+m-J} Z_{t+m+n}$) which would not be otherwise included in the time t DE forecast of Z_{t+m+n} itself. This case of $J > m$ is not just a theoretical curiosity. We note for example that 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 the best 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 LOIE holds under DE only when $J = 1$. Indeed, in that case, the term in equation (17) necessarily becomes zero, since $\mathbb{E}_t\mathbb{E}_{t+m-1} Z_{t+m+n} = \mathbb{E}_t Z_{t+m+n}$ 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 $\mathbb{E}_t^\theta Z_{t+m+n}$.

We build on the three main DE properties analyzed in this simple statistical model (*endogenous predictability*, *endogenous non-stochasticity* and the *failure of LOIE under distant memory*) to study, in the next two sections, simple consumption-savings problems that extract intuition on how endogenous states matter in jointly distorting beliefs and actions.

3 A Two-Period Consumption-Smoothing Model

In the previous statistical model, the response α of the action K_t to the state Y_t was taken as given to showcase that the action affects the conditional mean μ_t of future resources Z_{t+1}. We now describe in a two-period consumption-savings model how DE beliefs affect the agent's optimal decision over α.

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

$$C_1 + K_1 = Y_1 + (1+r)K_0;\ C_2 = Y_2 + (1+r)K_1, \tag{19}$$

where C_1 and C_2 are her consumption choices, yet to be determined. 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 assume $r = 0$, a discount factor $\beta = 1$, and a per-period utility $u(C) = bC - .5C^2$, with $b > 0$ and $C < b$ so that utility is increasing in consumption in that region.

Moreover, we work throughout these consumption-savings models assuming that the exogenous income Y has zero persistence

$$Y_{t+1} = \overline{Y} + \varepsilon_{t+1},$$

where ε_{t+1} are mean zero iid normal shocks with variance $\sigma^2 > 0$. This way we isolate the *endogenous predictability* mechanism, in which the action K_1 induces persistence in resources.

Under perfect memory ($\theta = 0$) the agent maximizes:

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

where (C_1, C_2) are determined by the choice K_1 and the budget constraints in equation (19). The Euler equation for K_1 solves the tradeoff implied by equating the current and future expected marginal utility, i.e. $u'\left(Y_1 - K_1\right) = \mathbb{E}_1 u'\left(Y_2 + K_2\right)$. Using the quadratic utility assumption and $\rho = 0$, this immediately results in

$$K_1^{RE} = \alpha^{RE}\varepsilon_1;\ \alpha^{RE} = 0.5, \tag{20}$$

where the superscript RE denotes choices when $\theta = 0$ (or 'rational expectations'). This action leads to a desired perfectly flat expected consumption profile ($C_1^{RE} = \mathbb{E}_1 C_2^{RE}$), where

$$C_1^{RE} = \overline{Y} + 0.5\varepsilon_1;\ C_2^{RE} = \overline{Y} + \varepsilon_2 + 0.5\varepsilon_1. \tag{21}$$

3.1 A Diagnostic Euler Equation

Under DE ($\theta > 0$) the agent maximizes

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

where $\mathbb{E}_1^\theta$ is formed under a distorted conditional density h_1^θ, to be specified below. Under this density, the first-order condition that characterizes the optimal choice in (22) is given by

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

The random C_2 is known by the time-2 budget constraint to equal $\overline{Y} + \varepsilon_2 + K_1^\theta$. Similarly, C_1 can also be substituted from the time-1 budget constraint, and thus we obtain the tradeoff

$$u'\left(\overline{Y} + \varepsilon_1 - K_1^\theta\right) = \mathbb{E}_1^\theta \left(u'\left(\overline{Y} + \varepsilon_2 + K_1^\theta\right)\right), \tag{23}$$

which gives the 'diagnostic' Euler equation, where the superscript θ denotes choices when $\theta > 0$. We conjecture a response of the optimal K_1^θ to the current state $Y_1 = \overline{Y} + \varepsilon_1$, given by

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

and substitute this in (23) to obtain

$$u'\left(\overline{Y} + \varepsilon_1 - K_1^\theta\right) = \mathbb{E}_1^\theta \left(u'\left(\overline{Y} + \varepsilon_2 + \alpha^\theta \varepsilon_1\right)\right). \tag{24}$$

Equation (24) gives the marginal tradeoff that guides agent's choices. Our approach is to characterize the effect of the representativeness heuristic and diagnostic expectations on these optimal choices. To do so, we rely on the analytical tractability of the conditional normality introduced in section 2.2. Because here future marginal utility is linear in future consumption, the tradeoff simply involves predicting C_2, which in turn follows a conditionally normal distribution under the conjectured K_1^θ.

We reiterate here the formalism of equation (8) and note that the conditional mean, given time 1 information, is $\mu_{2|1}^C \equiv \mathbb{E}_1 C_2 = \overline{Y} + \alpha^\theta \varepsilon_1$. As in equation (8), the representativeness heuristic then amounts to specifying

$$h_1^\theta(\widehat{C}_2) = h_1(\widehat{C}_2 | \mu_{2|1}^C = \overline{Y} + \alpha^\theta \widehat{\varepsilon}_1) \left[\frac{h_1(\widehat{C}_2 | \mu_{2|1}^C = \overline{Y} + \alpha^\theta \widehat{\varepsilon}_1)}{h_1(\widehat{C}_2 | \mu_{2|1}^C = \mathbb{E}_{1-J}\left(\overline{Y} + \alpha^\theta \varepsilon_1\right))} \right]^\theta \frac{1}{a}, \tag{25}$$

where $\mathbb{E}_{1-J}$ is the belief under the true density and the information set J periods before the generic time 1. Because of normality, this distorted density has the very appealing property of simplifying to the conditional expectation formulation described in (9). Thus, under this distorted h_1^θ density, we have

$$\mathbb{E}_1^\theta\left(C_2\right) = \mathbb{E}_1\left[\overline{Y} + \varepsilon_2 + \alpha^\theta \varepsilon_1\right] + \theta\left[\mathbb{E}_1\left(\varepsilon_2 + \alpha^\theta \varepsilon_1\right) - \mathbb{E}_{1-J}\left(\varepsilon_2 + \alpha^\theta \varepsilon_1\right)\right]. \tag{26}$$

Before we detail the implications of DE for this model we make a more general comment on our approach here, based on the observation that the distorted density h_t^θ formulated in equation (8), 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 the representativeness heuristic in affecting beliefs, and (2) this formulation is particularly convenient to employ when the processes over which it applies are conditionally normal.

As illustrated by the tractable characterization of equation (26), we leverage these joint properties 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 (24) will not be conditionally normal. Indeed, in the class of models we analyze in section 5, 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 future return on capital, which linearly load on Gaussian shocks. We exploit the convenient formulation of the representativeness heuristic based on the density h_t^θ in equation (8) by applying it on those relevant Gaussian objects that enter into the log-linearized perceived tradeoffs, leading to distorted expectations which qualitatively resemble that in (26).

Moreover, 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 (22) is influenced by the representativeness heuristic by being evaluated under the distorted conditional density of C_2, i.e. $h_1^\theta(\widehat{C}_2)$ defined in (25). Technically, this evaluation will generally differ from applying the formulation of h_1^θ in equation (8) to $u(C_2)$ itself, since the quadratic utility and normality of C_2 imply that $u(\widehat{C}_2)$ follows a χ^2-distribution, for which a similar tractable description as in (26) 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 (8). Our primitive approach, common with that proposed in Bordalo et al. (2018), is thus to emphasize 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 find the direct modeling of perceptions of linearized marginal tradeoffs as distorted by the density h_t^θ

of equation (8) appealing because: (a) in linearized models these perceptions actually guide decisions, and (b) in standard Gaussian environments these tradeoffs can also be tractably characterized, a property that we leverage throughout the paper.

3.2 Consumption smoothing under DE

Turning to the economics behind the resulting consumption smoothing under DE, we first note that, as with the case of $\theta = 0$, given the current income Y_1, the agent chooses savings K_1^θ to achieve perfect smoothing of C_1 (the LHS of (24)) with the expected C_2 (the RHS), but this time that expectation is formed under DE.

Proposition 1. *The optimal marginal propensity α^θ to save out of a transitory income shock ε_1 is lower than for the RE policy and decreases with θ*

$$\alpha^\theta = \frac{1}{2+\theta} < \alpha^{RE} = \frac{1}{2}. \tag{27}$$

The proof proceeds by evaluating the conditional expected $\mathbb{E}_1^\theta \left(Y_2 + K_1^\theta\right)$, for which we use the analytics of section 2, and in particular the DE conditional mean obtained in equation (9). The tradeoff in equation (23) becomes

$$\varepsilon_1(1-\alpha^\theta) = \mathbb{E}_1\left[\varepsilon_2 + \alpha^\theta \varepsilon_1\right] + \theta\left[\mathbb{E}_1\left(\varepsilon_2 + \alpha^\theta\varepsilon_1\right) - \mathbb{E}_{1-J}\left(\varepsilon_2 + \alpha^\theta\varepsilon_1\right)\right], \tag{28}$$

where $\mathbb{E}_{1-J}$ is the belief under the true density and the information set J periods before the generic time 1. Since income innovations are unpredictable, the tradeoff further simplifies to $\varepsilon_1(1-\alpha^\theta) = \alpha^\theta(1+\theta)\varepsilon_1$, which obtains the result in Proposition 1.

Importantly, when $\theta > 0$ the marginal propensity to save α^θ is lower than the corresponding optimal action $\alpha^{RE} = 1/2$. Thus, 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 the choice K_1^θ and evaluated under DE, is perfectly flat, since $C_1^\theta = Y_1 - K_1^\theta$, and thus by Proposition 1

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta = \overline{Y} + \frac{1+\theta}{2+\theta}\varepsilon_1, \tag{29}$$

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 (29).

Interestingly, this apparent puzzling behavior of a lack of consumption smoothing has been documented by a large empirical literature. This literature, using various identifying approaches (see Jappelli and Pistaferri (2010) for a survey) finds that the marginal propensity to consume (MPC), or the additional consumption brought upon by an unexpected income increase, is puzzlingly large, even for agents that are not financially constrained.[3]

The intuition for this result strongly connects to the *endogenous predictability* mechanism analyzed in detail for the statistical model in subsection 2.3. Here $Y_2 + K_1^\theta = C_2^\theta$ is the random variable that gives the future total financial resources available to the agent, and in equilibrium future consumption. The agent can perfectly observe the income realization Y_1 and thus K_1^θ. As detailed in subsection 2.1, to forecast K_1^θ, conditional on observing it, the agent does not need to resort to her memory and thus the representativeness heuristic does not distort her belief about K_1^θ (i.e. formally $K_1^\theta = \mathbb{E}_1^\theta \left(K_1^\theta \right)$, as per equation (6)).

However, as detailed in subsection 2.3, memory recall matters in how she forms beliefs about her future resources. 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 itself is iid, as analyzed here, this conditional predictability of future resources (given by $\overline{Y} + K_1^\theta$) comes just from K_1^θ, which through the response α^θ induces *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 $\overline{Y} + K_1^\theta$. In diagnostic terminology, the new information is the difference between the diagnostic, i.e. the current realization of expected resources, and comparison group, i.e. the realization she has previously expected to observe, which is just $\overline{Y}$ in this iid case. 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.

Therefore, given high (low) assets today, the agent is more optimistic (pessimistic) about future available resources than usual, and importantly, more than under the true distribution. Thus, 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, as summarized by Proposition 1, resulting in a

[3]Liquidity frictions (as in Kaplan and Violante (2014)) are typically used to account for the high MPC of rich but liquidity-constrained agents. However, Kueng (2018) and Lewis et al. (2020) are some examples of empirical evidence that documents that even rich people with high liquid wealth have significantly higher MPCs than implied by standard models.

puzzling high MPC and a lack of consumption smoothing for an outside observer.

3.3 Policy and Endogenous Non-Stochasticity

In a standard RE consumption-savings framework, eliminating future income uncertainty improves ex-ante utility through the standard reduction in risk, as well as possibly changing the intertemporal allocation of resources. The latter effect occurs because agents may not longer engage in precautionary savings behavior. In this context, our two-period model has been purposefully kept simple — it features quadratic utility, so this usual precautionary motive is absent in the first place.

In the DE model, eliminating uncertainty also improves welfare through a reduction in risk and a change in allocation. However, our analysis emphasizes a novel effect, on top of a standard precautionary savings one, through an *endogenous non-stochasticity*, where the elimination of residual uncertainty arising from policy actions matters. In particular, as detailed in subsection 2.1, the intertemporal allocation now changes also because DE agents no longer needs to recall memory to forecast future resources, and hence their policy function becomes identical to the RE one. Importantly, this change in allocation and its welfare effect is discontinuous when the residual uncertainty about future resources collapses to zero.

To understand this in the context of our two-period model, suppose the government announces at $t = 0$ a lump sum tax that exactly offsets the innovation in income at $t = 2$: $T_2 = -\varepsilon_2$. Then the budget constraint at $t = 2$ is

$$C_2 = Y_2 + T_2 + K_1. \tag{30}$$

Denote the unconditional variance of the income innovations as $\mathbb{E}[\varepsilon_1] = \sigma_1^2$ and $\mathbb{E}[\varepsilon_2] = \sigma_2^2$.[4] We further assume $2\overline{Y} - b > 0$, which ensures that the utility function is not "too linear".

In Proposition 2 we show that the ex-ante utility U evaluated at $t = 0$ under the true density is higher with a tax policy ($U^{\theta,policy}$) than without (U^{θ}), with a discontinuity at $\sigma_2 = 0$. For that, let us denote by $\overline{U}$ the ex-ante utility achievable under a certain income stream ($Y_1 = Y_2 = \overline{Y}$).

[4] While in the previous section, $\sigma_1^2 = \sigma_2^2 = \sigma^2$, the reason for its formal distinction here is to separate the role of time 1 and 2 uncertainty in evaluating time 0 ex-ante welfare.

Proposition 2. *The ex-ante utility U^θ without the policy under the limit $\sigma_2 \to 0$ is*

$$\lim_{\sigma_2 \to 0} U^\theta - \overline{U} = -\frac{2\overline{Y} - b}{2\overline{Y}} \left[\underbrace{\left(\frac{1+\theta}{2+\theta}\right)^2 \sigma_1^2}_{\textit{utility loss at } t=1} + \underbrace{\left(\frac{1}{2+\theta}\right)^2 \sigma_1^2}_{\textit{utility loss at } t=2} \right],$$

which is decreasing in θ. The ex-ante utility with the tax policy is given by

$$U^{\theta,policy} - \overline{U} = -\frac{2\overline{Y} - b}{2\overline{Y}} \left[\underbrace{\left(\frac{1}{2}\right)^2 \sigma_1^2}_{\textit{utility loss at } t=1} + \underbrace{\left(\frac{1}{2}\right)^2 \sigma_1^2}_{\textit{utility loss at } t=2} \right],$$

The utility gain from adopting the tax policy, $U^{\theta,policy} - U^\theta$, is then increasing in the representativeness parameter θ.

Proof. See the Appendix. □

The limit $\sigma_2 \to 0$ drives uncertainty to be 'small', but still present in the random variable Y_2, which helps us isolate the distortions coming from the DE beliefs on top of the reduction in risk. The key role of the policy is to undo with transfers that income uncertainty and thus entirely eliminate the residual uncertainty in the variable of interest to be forecasted by the agent, C_2. Indeed, by the budget constraint in equation (30), in this case $C_2 = \overline{Y} + K_1$. As detailed in subsection 2.1, the absence of residual uncertainty about C_2 renders the imperfect nature of memory recall a moot friction. Thus, $\mathbb{E}_1^\theta C_2^\theta$ becomes equal to $\overline{Y} + K_1^\theta$, without being affected by θ, and the perceived consumption smoothing tradeoff $C_1^\theta = \mathbb{E}_1^\theta C_2^\theta$ is now identical to the one under RE. Thus, under the tax policy, this leads the agent to take the same optimal savings rate as under RE

$$C_1^{\theta,policy} = C_1^{RE} = \overline{Y} + \frac{1}{2}\varepsilon_1; \; C_2^{\theta,policy} = C_2^{RE} = \overline{Y} + \frac{1}{2}\varepsilon_1. \tag{31}$$

The RE policy efficiently spreads out, in ex-ante utility terms, the income shock ε_1 across $t = 1$ and $t = 2$. In contrast, in the absence of the policy, even a small amount of uncertainty σ_2 activates the need of imperfect memory recall, leading to the actions (as in (29))

$$C_1^\theta = \overline{Y} + \frac{1+\theta}{2+\theta}\varepsilon_1; \; C_2^\theta = \overline{Y} + \frac{1}{2+\theta}\varepsilon_1.$$

Thus, under DE and in the absence of the policy, C_1^θ over-reacts to income shocks ε_1 and as a result C_2^θ ends up under–reacting to the same shock. Ex-ante, from the perspective of time

0, as shown in Proposition 2, these reactions result in an allocative inefficiency and utility loss, which are increasing in θ, relative to the RE policy.[5]

Finally, Proposition 2 brings forward the key role of policy communication. If agents are not aware of the tax policy when they make savings decisions at $t = 1$, then their actions will not be identical to the RE ones, and hence the welfare gain will not be realized.

4 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 detailed in subsection 2.4, a key challenge in taking the model to multi-periods is the *failure of the Law of Iterated Expectations (LOIE) under distant memory.*[6] While in the simple statistical model of section 2.4, that failure can be characterized in isolation, for a given path of the forecasted random variable, here we need to explore it *jointly* with the *endogenous predictability* created by optimal actions.

To analyze these interactions, we extend the two-period model to include a third period. The time 1 problem of the agent 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 $\beta = 1$)

$$\max_{K_1^\theta, K_2^{\theta,p}} \left\{ u(C_1^\theta) + \mathbb{E}_1^\theta \left[u(C_2^{\theta,p}) + u(C_3^{\theta,p}) \right] \right\}, \tag{32}$$

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

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

The key source of possible time-inconsistency is that in period 2, conditional on the inherited savings K_1^θ and the income realization ε_2, the agent can re-optimize over her initially planned $K_2^{\theta,p}$, by looking for a K_2^θ that solves her problem

$$\max_{K_2^\theta} \left[u(C_2^\theta) + \mathbb{E}_2^\theta u(C_3^\theta) \right]. \tag{33}$$

where $C_2^\theta = Y_2 + K_1^\theta - K_2^\theta(K_1^\theta, \varepsilon_2)$ and $C_3^\theta = Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2) - K_3$.[7]

[5] If $\sigma_2 = 0$, then the same argument of no residual uncertainty in C_2 renders the optimal solution under θ equivalent to (31), with or without tax policy. This is the formal source of discontinuity in Proposition 2.

[6] This property does not appear in the two-period model, since there only the one-step-ahead conditional expectation $\mathbb{E}_1^\theta C_2$ appears.

[7] Note that at the cost of extra notation the consumption path at time 2 and 3 under the actual choice

LOIE and perceived tradeoffs. As we have seen in section 2.4, the LOIE for the two-step-ahead expectation holds if and only if $J = 1$ leading, as we describe below, to possible time-inconsistency between planned and actual future choices. In particular, consider a given time 1 policy $K_1^\theta(K_0, \varepsilon_1)$. Conditional on reaching time 2, the optimal savings action $K_2^\theta(K_1^\theta, \varepsilon_2)$ in problem (33) implements the perceived tradeoff $C_2^\theta - \mathbb{E}_2^\theta C_3^\theta = 0$.

Proposition 3. *At time 1, the perceived consumption smoothing between C_2 and C_3, under $K_1^\theta(K_0, \varepsilon_1)$ and the optimal policy $K_2^\theta(K_1^\theta, \varepsilon_2)$, is $\mathbb{E}_1^\theta\left[C_2^\theta - C_3^\theta\right] = (1+\theta)\theta\left[\mathbb{E}_1 K_2^\theta - \mathbb{E}_{2-J} K_2^\theta\right]$, if $J > 1$, and equal to zero if $J = 1$.*

Proof. See the Appendix. □

Consider first the intuition for the $J = 1$ case. At time 2 the endogenous state K_1^θ is included in both the diagnostic and the comparison group and therefore does not constitute a surprise that gets selectively over-weighted by the selective memory. Instead, it is only the effect of the innovation ε_2 on $K_2^\theta(K_1^\theta, \varepsilon_2)$ that triggers memories that affects the perception of future available resources at time 3, which in equilibrium just equal $C_3^\theta = Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2)$. In the context of the two-period model, we have discussed how these over-reaction to ε_2 affects the optimal choice $K_2^\theta(K_1^\theta, \varepsilon_2)$ which implements $C_2^\theta = \mathbb{E}_2^\theta C_3^\theta$. As of time 1 the iid innovation ε_2 is expected to be zero on average under the true density, but importantly also under $\mathbb{E}_1^\theta$. Therefore the over-reaction of time-2 beliefs in $\mathbb{E}_2^\theta C_3^\theta$ (and thus action C_2^θ) is not expected under $\mathbb{E}_1$ to have a systematic direction, or bias, conditional on time-1 information. Thus, if $K_2^\theta(K_1^\theta, \varepsilon_2)$ implements a flat expected profile as of time 2, it is also expected to do so, on average, as of time 1 and $\mathbb{E}_1^\theta\left[C_2^\theta - C_3^\theta\right] = 0$.

In contrast, these conditional perceptions are altered intertemporally for an agent whose memory is based on more distant past, i.e. when $J > 1$. Intuitively, if K_1^θ increases with the current ε_1, at time 2 the agent observes an unusual high K_1^θ compared to the the more distant comparison group $\mathbb{E}_{2-J} K_1^\theta$. This leads her imperfect memory recall to forecast unusually high $\mathbb{E}_2^\theta C_3^\theta$ based on this representative state of high K_1^θ. She then chooses a consumption policy $K_2^\theta(K_1^\theta, \varepsilon_2)$ that as of time 2 implements a flat expected consumption profile, by keeping C_2^θ equal to this optimistic perceived $\mathbb{E}_2^\theta C_3^\theta$, from problem (33). The key difference from the $J = 1$ case is that now there is over-reaction to both ε_2 and K_1^θ - while the former source is not unpredictable as of time 1, the latter is since K_1^θ is in the time-1 information set. Thus, in evaluating her perceived future tradeoff $\mathbb{E}_1^\theta\left[C_2^\theta - C_3^\theta\right]$ the agent forecasts that, given her current information and memory recall process (based on the comparison group $\mathbb{E}_{1-J}$), she will save too little at time 2 under $K_2^\theta(K_1^\theta, \varepsilon_2)$ and thus end up implementing an expected downward sloping consumption path $C_2^\theta - C_3^\theta$, as formalized in Proposition 3.

taken at time 2 is indexed by θ, while the path under the plan is indexed by the superscript (θ, p).

Proposition 3 details how tradeoffs are perceived differently when imperfect memory recall is based on more distant past. This is crucial for the time-consistency of optimal actions, as we summarize in the following Corollary.

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

While for a given K_1^θ, the conditional optimal K_2^θ solves the tradeoff $C_2^\theta - \mathbb{E}_2^\theta C_3^\theta = 0$, the optimal time-1 plan $K_2^{\theta,p}$ in equation (32) is set such that $\mathbb{E}_1^\theta \left[C_2^{\theta,p} - C_3^{\theta,p}\right]$. However, immediately following from Proposition 3, for a given K_1^θ the conditional optimal K_2^θ implements exactly the time-1 desired consumption path under $K_2^{\theta,p}$ if and only if $J = 1$. Intuitively, as described above, for a given K_1^θ that increases with the current ε_1, when $J > 1$ the agent currently forecasts that she will save too little at time 2 under $K_2^\theta(K_1^\theta, \varepsilon_2)$ and would thus like instead to stick at time 2 with a plan $K_2^{\theta,p}(K_1^\theta, \varepsilon_2)$ that saves more than that. This misalignment of intertemporal perceived tradeoffs leads to time-inconsistency.

4.1 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 our analysis above. We now characterize the resulting optimal DE saving functions K_1^θ and K_2^θ. In particular, we first conjecture that the optimal policy functions under RE are

$$K_1^{RE} = \alpha_{K_0}^{RE} K_0 + \alpha_{\varepsilon_1}^{RE} \varepsilon_1;\ K_2^{RE} = \alpha_{K_1}^{RE} K_1^{RE} + \alpha_{\varepsilon_2}^{RE} \varepsilon_2. \tag{34}$$

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

$$K_1^\theta = \alpha_{K_0}^\theta K_0 + \alpha_{\varepsilon_1}^\theta \varepsilon_1;\ K_2^\theta = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta \varepsilon_2. \tag{35}$$

Proposition 4. *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.*

$$\alpha_{K_0}^\theta = \alpha_{K_0}^{RE} = \frac{2}{3}; \alpha_{K_1}^\theta = \alpha_{K_1}^{RE} = \frac{1}{2} \tag{36}$$

$$\alpha_{\varepsilon_1}^\theta = \frac{2}{3+\theta} < \alpha_{\varepsilon_1}^{RE} = \alpha_{K_0}^{RE};\ \alpha_{\varepsilon_2}^\theta = \frac{1}{2+\theta} < \alpha_{\varepsilon_2}^{RE} = \alpha_{K_1}^{RE}. \tag{37}$$

Proof. See the Appendix. □

To understand the intuition behind this Proposition, first consider the time 2 problem in (33). Conditional on reaching period 2, 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

$$\varepsilon_2 + K_1^\theta - K_2^\theta = \mathbb{E}_2(\varepsilon_3 + K_2^\theta) + \theta \left[\mathbb{E}_2(\varepsilon_3 + K_2^\theta) - \mathbb{E}_1\left(\varepsilon_3 + K_2^\theta\right)\right]. \tag{38}$$

Under the conjecture in equation (35), 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 (38), we obtain their values characterized in Proposition 4.

The key observation here is that, as observed in deriving Proposition 3, 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 ($\alpha_{\varepsilon_2}^\theta < \alpha_{\varepsilon_2}^{RE}$) for the same reason detailed in the two-period model (see equation (27) and the discussion around Proposition 1).

We now move back to the time 1 problem in (32), where the total derivative of the objective function with respect to K_1^θ is

$$C_1 = \mathbb{E}_1^\theta \left[C_2^{\theta,p} + \frac{\partial K_2^{\theta,p}}{\partial K_1^\theta} \left(C_3^{\theta,p} - C_2^{\theta,p} \right) \right]. \tag{39}$$

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, $K_2^\theta(K_1^\theta, \varepsilon_2)$. By the additivity property of the DE operator (see equation (10)), since C_2 and C_3 are conditionally normal and have residual uncertainty as of time 1, as arising here from the normally distributed income innovations and the conjectured K_1^θ and K_2^θ, the tradeoff in (39) can be broken in

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

By Proposition 3 and Corollary 1, when $J = 1$, the time-consistency in plans means that

K_2^{θ} implements the same path as the plan $K_2^{\theta,p}$ and thus the term $\mathbb{E}_1^{\theta}\left[C_3^{\theta,p} - C_2^{\theta,p}\right]$ equals zero, under both the plan and the anticipation of the future choice K_2^{θ}. Thus, the tradeoff in equation (39) 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.

4.2 Beliefs over Future Actions

Proposition 3 and Corollary 1 show for $J > 1$ the time-inconsistency between the planned K_2 and what she 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 of 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 here an agent that fails to take into account the fact that her future self's behavior is governed by the representativeness heuristic. Her time 1 problem is now

$$\max_{K_1^{\theta,n}} \left\{ u(C_1^{\theta}) + \mathbb{E}_1^{\theta}\left[u(C_2^{RE}) + u(C_3^{RE})\right]\right\} \tag{40}$$

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

$$\max_{K_2^{RE}} \left[u(C_2^{RE}) + \mathbb{E}_2 u(C_3^{RE})\right]. \tag{41}$$

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 (40) (and (41), respectively) and trivially $K_3^{RE}(.) = 0$.[8]

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 influenced by the representativeness heuristic, so her conditionally optimal action is

$$\max_{K_2^{\theta,n}} \left[u(C_2^\theta) + \mathbb{E}_2^\theta u(C_3^{RE})\right]. \tag{42}$$

where now

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

The behavioral interpretation of equations (40), (41) and (42) 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

$$\max_{K_2^{\theta,s}} \left[u(C_2^\theta) + \mathbb{E}_2^\theta u(C_3^\theta)\right] \tag{43}$$

where now

$$C_2^\theta = Y_2 + K_1^{\theta,s} - K_2^{\theta,s}(.); C_3^\theta = Y_3 + K_2^{\theta,s}(.) - K_3^\theta(.). \tag{44}$$

The agent's sophistication reflects the idea that, unlike the naïve agent, at time 1 she understands that her future action is dictated by equation (43). Thus, at time 1 the sophisticated agent anticipates the resulting policy function $K_2^{\theta,s}$ that solves (43), as well as $K_3^\theta(.) = 0$, and currently solves

$$\max_{K_1^{\theta,s}} \left\{u(C_1^\theta) + \mathbb{E}_1^\theta \left[u(C_2^\theta) + u(C_3^\theta)\right]\right\}, \tag{45}$$

where current $C_1^\theta = Y_1 + K_0 - K_1^{\theta,s}(.)$ while the forecasted consumption choices C_2^θ and C_3^θ are determined as in (44).

Comparison Groups for Memory Retrieval. The comparison group in the representativeness heuristic gives the state prevailing if there is no news, under the true density,

[8]There is no material distinction between $K_3^{RE}(.)$ and $K_3^\theta(.)$ since they are both equal to zero. However, we keep referring to these objects separately to highlight the conceptual difference between these policy functions being taken under different beliefs. Thus, in equation (43) below she would anticipate her future behavior to be governed by the $K_3^\theta(.)$ action rather than the $K_3^{RE}(.)$ that appears in (42).

compared to the immediate ($J = 1$), or more distant past ($J > 1$). In formulating the DE beliefs of the naïve and sophisticated agents we have to be precise about the construction of 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 $\mathbb{E}_2^\theta(Y_3 + K_2^{\theta,n})$. What is his comparison group that defines the new information that drives their over–reaction to the realization of $K_2^{\theta,n}$? Here we note that 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 agents that forms the forecast $\mathbb{E}_2^\theta(Y_3 + K_2^{\theta,s})$. In contrast to the naïve agent, note that at any past date before time 2, this sophisticated agent expects her future selves to act under the DE policy function. Therefore, for the same reason of 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 made J periods ago by the former self of this sophisticated agent of the DE savings choice at time 2, under the true density.[9]

Euler Equations. Per the objective function in (42) and (43), in both approaches, 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

$$\varepsilon_2 + K_1^{\theta,n} - K_2^{\theta,n} = \mathbb{E}_2^\theta[\varepsilon_3 + K_2^{\theta,n}];\ \varepsilon_2 + K_1^{\theta,s} - K_2^{\theta,s} = \mathbb{E}_2^\theta[\varepsilon_3 + K_2^{\theta,s}], \tag{46}$$

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

$$C_1^\theta = \mathbb{E}_1^\theta\left[C_2^{RE} + \frac{\partial K_2^{RE}}{\partial K_1^{\theta,n}}\left(C_3^{RE} - C_2^{RE}\right)\right], \tag{47}$$

while under sophistication it is

$$C_1^\theta = \mathbb{E}_1^\theta\left[C_2^\theta + \frac{\partial K_2^{\theta,s}}{\partial K_1^{\theta,s}}\left(C_3^\theta - C_2^\theta\right)\right]. \tag{48}$$

[9] As we will show soon below, when $J = 1$, these comparison groups collapse to the same object, because the laws of motion of $K_2^{\theta,s}$ and K_2^{RE} load in the same way on the endogenous states (see Proposition 4). However, when $J > 1$ the optimal laws of motion themselves behave differently and the formation of the two comparison groups matters, as we detail later.

As in equation (39), these tradeoff capture the direct effects of the current choice on tomorrow's consumption and the indirect effects through the capital choice at time 2 and the resulting consumption path. The latter effect appears as the elasticity of K_2^{RE} in (47) and $K_2^{\theta,s}$ in (48) with respect to the current action.[10]

Accounting for effects of current action on future actions. How important is the indirect effect of the agent internalizing that the current choice affects the future problem and decision? From Proposition 3, we know that for any given K_1^{θ}, the optimal $K_2^{\theta,s}$ induces a path for consumption smoothing $C_3^{\theta} - C_2^{\theta}$ that is perceived at time 1 as optimal, i.e. $\mathbb{E}_1^{\theta}\left[C_3^{\theta} - C_2^{\theta}\right] = 0$, if and only if $J = 1$, which has an immediate implication for the optimal $K_1^{\theta,s}$ that we state in Corollary 2.

Corollary 2. *The tradeoff in equation (48) solved by the time-1 optimal solution $K_1^{\theta,s}(K_0, \varepsilon_1)$ under sophistication simplifies to the one-step ahead consumption smoothing, $C_1^{\theta} = \mathbb{E}_1^{\theta} C_2^{\theta}$, 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 (48) that affects the current choice of the sophisticated agent when $J > 1$ and otherwise disappears when $J = 1$.

The problem for the current naïve self turns out to be simpler, which will also bring tractability for the solution method that we will later implement to solve infinite horizon models. In particular, we can derive a counterpart of Proposition 3 for the naïveté solution, as follows. For any given policy $K_1^{\theta,n}(K_0, \varepsilon_1)$, at time 2, the optimal linear savings action K_2^{RE}, conditional on the resulting state $K_1^{\theta,n}$ and the realized innovation ε_2, implements the perceived tradeoff $C_2^{RE} - \mathbb{E}_2 C_3^{RE} = 0$, where recall that C_3^{RE} is also determined by $K_3^{RE}(.) = 0$. Going backwards, at time 1, we can establish the following Lemma.

Lemma 2. *At time 1, 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 the policy K_2^{RE}, is $\mathbb{E}_1^{\theta}\left[C_2^{RE} - C_3^{RE}\right] = 0$ for $J \geq 1$. Therefore, the optimal naïve choice $K_1^{\theta,n}$ in equation (47) solves the one-step tradeoff*

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

[10]Note that in this paper we maintain the standard assumption of agents having free cognitive access to their optimal course of action, given beliefs about current and future circumstances, but compared to the standard model we enrich the latter aspect through the role of the representativeness heuristic in forecasting. See Ilut and Valchev (2020) for a recent approach that relaxes this standard assumption by modeling agents' costly reasoning about their unknown policy functions, given beliefs about the states.

Proof. See the Appendix. □

This result exploits the observation that the time 1 perceived savings behavior of the future self at time 2 under naïveté 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 (47), 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.[11] Critically, by anticipating future actions taken under perfect memory recall, this result holds for any $J \geq 1$. By the additivity property of DE beliefs (see equation (10)), the second, indirect, effect in equation (47) thus disappears, leading to the Euler equation in equation (49).

$\mathbf{J = 1}$ 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 4).

Proposition 5. *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.*

Proof. See the Appendix. □

First, note that naïve problem recovers the same solution as the time-consistent one because of the property, shown in Proposition 4, that when $J = 1$, K_1^{θ} and K_2^{θ} optimally 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.[12] Based on this property, on the one hand, 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 for the representativeness heuristic needed in the Euler equations of the naïve problem and the time-consistent problem are the same. Thus, given the same belief formation and optimality condition, the naïve solution for $K_2^{\theta,n}$ in (46) must recover the time-consistent policy at time 2 from (38). On the other hand, this time consistency with respect to the endogenous variables means that the consumption smoothing 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 4, where the agent was planning in a time-consistent way to follow the policy K_2^{θ}.

[11] In dynamic programming, this property reflects the envelope theorem, a characteristic that applies here for the naïve agent. When we extend this model to infinite horizon this property will allow us to tractably characterize recursively optimal actions taken under DE and naïveté for both $J = 1$ and $J > 1$. In contrast, for the sophisticated agent this 'envelope argument' applied only when $J = 1$, as we have discussed above.

[12] 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.

Second, when $J = 1$, by Proposition 3 and our earlier discussion, the time-1 tradeoff under sophistication in equation (48) is identical to the time-consistent one $C_1^\theta = \mathbb{E}_1^\theta C_2^\theta$ (and also to $C_1^\theta = \mathbb{E}_1^\theta C_2^{RE}$ per the naïveté argument above). The conjecture that $K_1^{\theta,s}$ and $K_2^{\theta,s}$ are equal to their time-consistent policy and naïveté counterparts is then immediately verified, by further noting that the construction of the comparison groups is identical across these models of belief formation when $J = 1$.

4.3 Dynamics with Memory Recall of More Distant Past

As indicated by our previous discussion, 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- and psychologically-coherent ways to speak to these issues. We analyze them below in the three period model when the comparison group is based on $J = 2$.

4.3.1 Role of Informational States for Comparison Groups

The most transparent way to see the effects of distant past in forming the comparison groups is to study the period 2 problem, 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.

We now characterize the optimal $K_2^{\theta,n}$ and $K_2^{\theta,s}$ that solve their respective time 2 problem in equation (50), given some inherited savings ($K_1^{\theta,n}$ in the naïveté and $K_1^{\theta,s}$ for the sophisticated agent economy, respectively), and some observed innovation ε_2, as follows:

Proposition 6. *When $J = 2$ the time-2 naïveté and sophisticated policy functions are*

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

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:

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

Proof. See the Appendix. □

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

$$\varepsilon_2 + K_1^{\theta,n} - K_2^{\theta,n} = K_2^{\theta,n} + \theta\left(K_2^{\theta,n} - \mathbb{E}_{2-J}K_2^{RE}\right). \tag{50}$$

The key difference between $J = 1$ and $J = 2$ is how the comparison group, $\mathbb{E}_{2-J}K_2^{RE}$, affects conditional beliefs about future consumption. We have characterized in Proposition 4 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

$$\mathbb{E}_0K_2^{RE} = \alpha_{K_1}^{RE}\mathbb{E}_0K_1^{RE}. \tag{51}$$

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 6, the DE beliefs over-react by a factor of θ to the new information comprised of

$$K_2^{\theta,n} - \mathbb{E}_0K_2^{RE} = \alpha_{K_1}^{\theta}K_1^{\theta,n} + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2 + \left(\alpha_{\mathbb{E}_0K_1}^{\theta} - \alpha_{K_1}^{RE}\right)\mathbb{E}_0K_1^{RE}.$$

By substituting this over-reaction in equation (50), we recover the optimal $K_2^{\theta,n}$ coefficients.[13]

The critical reason behind these novel state dynamics is that when $J = 2$ the economic state $K_1^{\theta,n}$ is *not* a sufficient statistic anymore (as it was when $J = 1$) for the comparison group characterizing the memory recall process. With $J = 2$, the conditional expectation $\mathbb{E}_0K_1^{RE}$ forms a separate *informational state* that affects time 2 choices. Indeed, the higher the $\mathbb{E}_0K_1^{RE}$, the higher is the comparison group for $K_2^{\theta,n}$ (since $\alpha_{K_1}^{RE} > 0$ in equation (51)) and thus the more the DE agent is typically 'disappointed' by the perceived innovation in the conditional mean of future consumption, given by $K_2^{\theta,n} - \mathbb{E}_0K_2^{RE}$. 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 $\alpha_{\mathbb{E}_0K_1}^{\theta}$ on the informational state $\mathbb{E}_0K_1^{RE}$ is positive for $\theta > 0$ in Proposition 6. Of course, this over-reaction caused by imperfect memory recall is absent in the RE case, where $\mathbb{E}_0K_1^{RE}$ does not matter for the choice K_2^{RE}.

[13] Note 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. First, Lian (2020) shows that (partial) sophistication is key for an agent to decide to save less today out of anticipations of future mistakes. Second, in Ilut and Valchev (2020) the feedback between endogenous reasoning over the optimal conditional consumption action (here taken as known) and wealth accumulation leads to learning traps where agents typically end up acting under steep estimated slopes of the unknown consumption functions.

The other manifestation of the separate role of the variable $K_1^{\theta,n}$ 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 4). This separate role can also be seen by rewriting the solution for $K_2^{\theta,n}$ in Proposition 6 as

$$K_2^{\theta,n} = \left(\alpha_{K_1}^{\theta} + \alpha_{\mathbb{E}_0 K_1}^{\theta}\right) K_1^{\theta,n} - \alpha_{\mathbb{E}_0 K_1}^{\theta} \left(K_1^{\theta,n} - \mathbb{E}_0 K_1^{RE}\right) + \alpha_{\varepsilon_2}^{\theta}\varepsilon_2. \tag{52}$$

The first part captures the role of $K_1^{\theta,n}$ as an economic state. Not surprisingly then, this economic state 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 4.3.2). 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 explains why the innovation $\left(K_1^{\theta,n} - \mathbb{E}_0 K_1^{RE}\right)$ enters with a negative sign in equation (52). 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. $\alpha_{K_1}^{\theta} < \alpha_{K_1}^{RE}$).[14]

The solution for the optimal choice $K_2^{\theta,s}$ follows the same logic as detailed above for $K_2^{\theta,n}$, leading to the result in Proposition 6 that the optimal coefficients are the same. The subtle difference between these the optimal actions is the comparison groups, which are consistent with two different views of the future action paths as of time 0. For sophistication this amounts to tracing $\mathbb{E}_0 K_2^{\theta,s}$, which under the verified conjecture in Proposition 6, as of time 0, necessitates the agent's understanding of $\mathbb{E}_0 K_1^{\theta,s}$. A similar re-formulation of $K_2^{\theta,s}$ can be written as in (52) where the innovation $\left(K_1^{\theta,s} - \mathbb{E}_0 K_1^{\theta,s}\right)$ enters with the same negative sign.

The two respective informational states $\mathbb{E}_0 K_1^{RE}$ and $\mathbb{E}_0 K_1^{\theta,s}$ capture the economics of the mechanism. In this context, we note that naïveté solution can be further simplified by taking advantage of the known law of motion for K_1^{RE}. Therefore, $\mathbb{E}_0 K_1^{RE}$ can be immediately plugged in the determination of $K_2^{\theta,n}$ as $\alpha_{K_0}^{RE} K_0$. 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

[14] Note that when $J=1$, this innovation does not enter as an additional relevant state for $K_2^{\theta,n}$ (see equation (35)) 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 $\alpha_{K_1}^{\theta} + \alpha_{\mathbb{E}_0 K_1}^{\theta} = \alpha_{K_1}^{RE}$.

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.

4.3.2 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. To capture the novel informational state due to the comparison group based on a more distant past ($J = 2$), we allow the $t = -1$ conditional expected value of K_0 to enter the solution, given that this expectation could be potentially different from the realized K_0.[15]

Consider first the naïveté case. We have shown earlier (see Lemma 2) that 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 solves the 'direct' effect of affecting perceived consumption smoothing between time 1 and 2, i.e. $C_1^\theta = \mathbb{E}_1^\theta C_2^{RE}$, even when $J > 1$. As this one-step 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 just one period back, the resulting qualitative properties of the optimal $K_1^{\theta,n}$ resembles those for $K_2^{\theta,n}$ in Proposition 6. We provide details in the Appendix, in the form of Proposition 7, on the optimal coefficients for the conjectured $K_1^{\theta,n} = \alpha_{\mathbb{E}_{-1}K_0}^{\theta,n}\mathbb{E}_{-1}K_0 + \alpha_{K_0}^{\theta,n}K_0 + \alpha_{\varepsilon_1}^{\theta,n}\varepsilon_1$. Similarly to $K_2^{\theta,n}$, we find that compared to $J = 1$, in this $J = 2$ case the solution has (i) a positive loading $\alpha_{\mathbb{E}_{-1}K_0}^{\theta,n}$ on the past informational state $\mathbb{E}_{-1}K_0$, (ii) a muted response $\alpha_{K_0}^{\theta,n}$ to the current economic state K_0, and (iii) an identical (while still muted), response $\alpha_{\varepsilon_1}^{\theta,n}$ to the current innovation. As with $K_2^{\theta,n}$, through the lenses of equation (52), the interpretation of this changed state dynamics is the separate role of K_0 as an economic and informational state, which can be gauged by observing that $\alpha_{\mathbb{E}_{-1}K_0}^{\theta,n} + \alpha_{K_0}^{\theta,n} = \alpha_{K_0}^{RE}$.

As detailed in our discussion of Proposition 3, 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, as seen in equation (48), her current optimal action $K_1^{\theta,s}$ aims to fix this misalignment by affecting the state of her

[15] For the special case of $\mathbb{E}_{-1}K_0 = K_0$, the time-1 comparison group becomes identical to that for the case of $J = 1$. Here we take $\mathbb{E}_{-1}K_0$ as exogenous, outside-the-model determined, but note that in a full infinite horizon model, like the one in section 5, these past expectations wil be part of the solution of the model, as captured by the model-implied informational states $\mathbb{E}_0 K_1^{RE}$ and $\mathbb{E}_0 K_1^{\theta,s}$ in Proposition 6.

anticipated policy function $K_2^{\theta,s}$, and solve

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta + \alpha_{K_1}^\theta \mathbb{E}_1^\theta \left(C_3^\theta - C_2^\theta \right), \tag{53}$$

where $\alpha_{K_1}^\theta = \partial K_2^{\theta,s} / \partial K_1^{\theta,s}$ is given in Proposition 6.

As with the naïveté case we conjecture and verify that the optimal solution takes the form $K_1^{\theta,s} = \alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} \mathbb{E}_{-1} K_0 + \alpha_{K_0}^{\theta,s} K_0 + \alpha_{\varepsilon_1}^{\theta,s} \varepsilon_1$. There are three conceptual forces that affect these coefficients compared to their naïveté case.

First, there is a direct effect through which 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. $\alpha_{K_1}^\theta < \alpha_{K_1}^{RE}$. This force alone, coming from the $\mathbb{E}_1^\theta C_2^\theta$ term in (53), leads the agent to consume more today out of ε_1 to achieve consumption smoothing between period 1 and 2. Second, as detailed in our discussion of Proposition 3, 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 (53), that leads to more saving in order to mitigate the relative under-consumption in $t = 3$ following a positive shock ε_1. The race between these two forces is dominated here by the former, direct effect, as $\alpha_{K_1}^\theta < 0.5$, and thus the agent ends up saving less out of ε_1 than under naïveté, i.e. $\alpha_{\varepsilon_1}^{\theta,s} < \alpha_{\varepsilon_1}^{\theta,n}$.

Third, there is a conceptual and technical difference in terms of the comparison group construction between naïveté and sophistication, as emphasized earlier. Now, there is a fixed point in how the informational state $\mathbb{E}_0 K_1^{\theta,s}$, which matters for the $K_2^{\theta,s}$ solution in Proposition 6, needs to be itself based on the policy function $K_1^{\theta,s}$, which in turn is affected by $\mathbb{E}_1^\theta K_2^{\theta,s}$ in equation (53). 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 $K_1^{\theta,s}$ to K_0 and $\mathbb{E}_{-1}K_0$ compared to the $K_1^{\theta,n}$ case. We provide details in the Appendix in Proposition 8.

4.3.3 Modeling imperfect memory recall in larger models

Building on the theoretical framework developed in this section in the specific context of a three-period consumption-savings model, we are interested in studying the role of beliefs formed under the representativeness heuristic in larger, standard business cycle models. These would usually involve multiple actions, a large state space and decisions taken over an infinite horizon. While in the three-period model we have fully characterized both the naïve and the sophistication solution, expanding those to more meaningful dynamic economies brings forward a set of conceptual and methodological issues that become particularly critical when memory retrieval is based on more distant past.

In this context, our analysis so far suggests that the proposed naïve approach offers a (1) model-coherent and (2) methodologically tractable method that (3) captures insightful behavioral response of current actions to the selective memory recall. First, as noted by a large theory literature it is a coherent micro-founded model of beliefs and behavior. In our case, it further implies that when memory recall is based on the immediate past, the same approach recovers the time-consistent solution. Second, our three-period model 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 (and specifically muted) response of the current optimal savings to the same given set of endogenous and exogenous states that would matter in a fully rational model (i.e. the economic states); 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 (53), 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.[16]

The fixed point between the perceived elasticity of the value function of the sophisticated agent with respect to its states (economic and informational) and the current optimal action taken under DE beliefs becomes computationally very demanding, especially as informational states proliferate, both for us as modelers, and presumably also for the economic agents (as also generally argued in Ilut and Valchev (2020)). In this sense, the solution is not only difficult

[16]Due to its simplicity, we did not encounter this extra indirect effect of current actions on the life-time utility in the three period model, but it would generally be present. For example, in a four period model extension, with $J = 2$, the agent at time 1 would internalize and compute how the choice $K_1^{\theta,s}$ affects the comparison group $\mathbb{E}_1 K_3^{\theta,s}$ that matters for the $K_3^{\theta,s}$ solution.

to characterize by us as outside observers, but this required hyper-rationality is arguably in tension with the motivation of modeling agents' beliefs about their future circumstances as influenced by a heuristic, usually viewed as a mental shortcut that allows agents to make judgments quickly and efficiently (Tversky and Kahneman (1975) and Kahneman (2011)).

Our theoretical framework lays the conceptual ground for the naïve approach, compared to the sophisticated one, as being more portable across richer settings, arguably more realistic, and a 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.

5 Applications to Real Business Cycle Models

In this section we leverage some of the previous insights in standard, infinite-horizon models. Methodologically, we formally rely on the naïve approach to model beliefs, as argued earlier. This allows us to tractably and recursively characterize equilibrium laws of motion when agents act under DE beliefs, while uncovering rich and novel state dynamics.

5.1 The Model

A representative household chooses capital K_t^θ, consumption C_t^θ and labor N_t^θ to solve

$$W(K_{t-1}^\theta, A_t, \{\mathbb{E}_{t-j}K_t^{RE}, \mathbb{E}_{t-j}A_{t+1}\}_{j=1}^J) = \max_{K_t^\theta, C_t^\theta, N_t^\theta} \left[\ln C_t^\theta - \frac{(N_t^\theta)^{1+\eta}}{1+\eta} + \beta \mathbb{E}_t^\theta V(K_t^\theta, A_{t+1})\right] \tag{54}$$

subject to the resource constraint

$$C_t^\theta + K_t^\theta - (1-\delta)K_{t-1}^\theta = A_t(K_{t-1}^\theta)^\alpha (N_t^\theta)^{1-\alpha},$$

where A_t is total factor productivity (TFP), following $\ln A_t = \rho \ln A_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2)$. The variables $\{\mathbb{E}_{t-j}K_t^{RE}, \mathbb{E}_{t-j}A_{t+1}\}_{j=1}^J$ are included as state variables for W because they are used to form the necessary comparison groups for the endogenous and exogenous states.[17]

[17]We include the whole lagged expectations from $t-1$ to $t-J$ to accommodate a general formulation of a comparison group that consists of a weighted average of past expectations, as we study below. If the comparison group consists only of $\mathbb{E}_{t-J}$, then the states are simply $\{\mathbb{E}_{t-J}K_t^{RE}, \mathbb{E}_{t-J}A_{t+1}\}$.

As introduced and explained in section 4.2, in this naïve approach, in evaluating the continuation value $V(.)$ in equation (54), the household assumes that her future conditional preferences and resulting conditionally optimal actions, given a value of capital of K_t^{θ} entering $t+1$, will be taken under perfect memory and RE. To construct that continuation value we thus we set up an alternative economy where a representative household chooses capital K_t^{RE}, consumption C_t^{RE} and labor N_t^{RE} under perfect memory:

$$V(K_{t-1}^{\theta}, A_t) = \max_{K_t^{RE}, C_t^{RE}, N_t^{RE}} \left[\ln C_t^{RE} - \frac{(N_t^{RE})^{1+\eta}}{1+\eta} + \beta \mathbb{E}_t V(K_t^{RE}, A_{t+1}) \right], \tag{55}$$

subject to the resource constraint

$$C_t^{RE} + K_t^{RE} - (1-\delta) K_{t-1}^{\theta} = A_t (K_{t-1}^{\theta})^{\alpha} (N_t^{RE})^{1-\alpha}.$$

In the Appendix, we provide the equilibrium conditions. In particular, we show that the household's problem (54) gives rise to a diagnostic Euler equation for the RBC model:

$$(C_t^{\theta})^{-1} = \beta \mathbb{E}_t^{\theta} \left[(C_{t+1}^{RE})^{-1} R_{t+1}^{K,RE} \right], \tag{56}$$

where $R_{t+1}^{K,RE}$ is the $t+1$ return on capital under RE policy function:

$$R_t^{K,RE} \equiv \alpha A_t (K_{t-1}^{\theta})^{\alpha-1} (N_t^{RE})^{1-\alpha} + 1 - \delta.$$

There are three key points to be made here that connect the solution method to our previous discussions. First, the expectation in the diagnostic Euler equation is taken under DE. In particular, as we discussed at more length in section 3.1, we leverage the tractability of the density h_t^{θ} in equation (8) by applying it on those relevant Gaussian objects that enter into the log-linearized perceived tradeoffs characterizing the Euler equations of the model. The log-linearized diagnostic Euler equation in (56) is

$$-\widehat{C}_t^{\theta} = \mathbb{E}_t^{\theta} (-\widehat{C}_{t+1}^{RE} + \widehat{R}_{t+1}^{K,RE}), \tag{57}$$

which provides the specific distorted beliefs that lead to the analyzed equilibrium outcomes. Second, expected endogenous variables, and in turn, the comparison group which is used to compute DE, are under the RE policy functions, a characteristic of this naïve approach. Third, technically, our previous discussion also indicates that, without residual uncertainty in the object we take expectations on, DE collapses to the RE. To check that all the relevant variables we take DE on indeed contains residual uncertainty, one can simply verify, for given parameter values, that the conditional variance under the RE law of motion (given specifically by equation (59) below) is strictly positive.[18]

Relative to the naïve approach, under sophistication the Euler equation features (i) the expectation and the comparison group evaluated under the DE equilibrium policy function, (ii) the term that reflects the fact that the equilibrium policy and future utility, evaluated under the current preference, are not aligned and (iii) the term that is associated with the fact that agents have control over future comparison group through current choice. We relegate additional discussions of the RBC model under sophistication to Appendix C, which point to the computational complexities of this problem, and thus also link to the qualitative arguments made in section 4.3.3.

5.2 Solution Method

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

1. The first step of the solution algorithm would be to obtain a RE law of motion that allows us to compute the evolution of endogenous variables which we then use to form DE on. We start from a linear RE system

$$\mathbf{\Gamma}_0 \mathbf{x}_t^{RE} = \mathbf{\Gamma}_1 \mathbf{x}_{t-1}^{RE} + \mathbf{\Psi}\varepsilon_t + \mathbf{\Pi}\eta_t^{RE}, \tag{58}$$

   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; linear equilibrium conditions where DE ($\mathbb{E}_t^\theta$) is simply replaced with RE ($\mathbb{E}_t$).

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

$$\mathbf{x}_t^{RE} = \mathbf{T}^{RE}\mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE}\varepsilon_t. \tag{59}$$

2. Consider a linear DE system

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t, \tag{60}$$

   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

[18] The check is under the RE law of motion because DE is taken under the RE policy function. For example, for the log-linearized diagnostic Euler equation in (56), the relevant variable to check its conditional variance is $(-\widehat{C}_{t+1}^{RE} + \widehat{R}_{t+1}^{K,RE})$. We provide more details in Appendix D.

RE system (58), which implicitly defines expectations in $\mathbf{x}_t^{RE}$ by using expectation errors η_t^{RE}, here the DE system (60) explicitly accommodates DE ($\mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE}$).

We then substitute in $\mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} = (1+\theta)\mathbb{E}_t \mathbf{y}_{t+1}^{RE} - \theta \sum_{j=1}^J \alpha_j \mathbb{E}_{t-j} \mathbf{y}_{t+1}^{RE}$, where $\{\alpha_j\}_{j=1}^J$ are weight parameters associated with lagged expectations (and hence $\sum_{j=1}^J \alpha_j = 1$):

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

Let $\mathbf{y}_t^{RE} = \mathbf{M}\mathbf{x}_t^{RE}$, where $\mathbf{M}$ is a selection matrix that selects variables from a vector $\mathbf{x}_t^{RE}$. Then (60) becomes

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \left[(1+\theta)\,\mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^\theta - \sum_{j=1}^J \theta \alpha_j \mathbf{M} \left(\mathbf{T}^{RE}\right)^{j+1} \mathbf{x}_{t-j}^\theta \right] + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t, \tag{61}$$

which makes clear that agents form DE based on state variables inherited from the DE economy but using RE policy going forward.

3. Rewrite the system (61) into a more compact form:

$$\mathbf{\Gamma}_0^{\mathbf{z}} \mathbf{z}_t^\theta = \mathbf{\Gamma}_1^{\mathbf{z}} \mathbf{z}_{t-1}^\theta + \mathbf{\Psi}^{\mathbf{z}} \varepsilon_t,$$

where we provide expressions for $\mathbf{\Gamma}_0^{\mathbf{z}}$, $\mathbf{\Gamma}_1^{\mathbf{z}}$ and $\mathbf{\Psi}^{\mathbf{z}}$ in Appendix D. $\mathbf{z}_t^\theta$ is a vector that includes not only $\mathbf{x}_t^\theta$ but also $\mathbb{E}_t^\theta \mathbf{y}_t^{RE}$ and lags $\mathbf{x}_{t-1}^\theta, \ldots, \mathbf{x}_{t-J+1}^\theta$. The DE decision rule can be obtained simply by inverting this system:

$$\mathbf{z}_t^\theta = \mathbf{T}^\theta \mathbf{z}_{t-1}^\theta + \mathbf{R}^\theta \varepsilon_t, \tag{62}$$

where

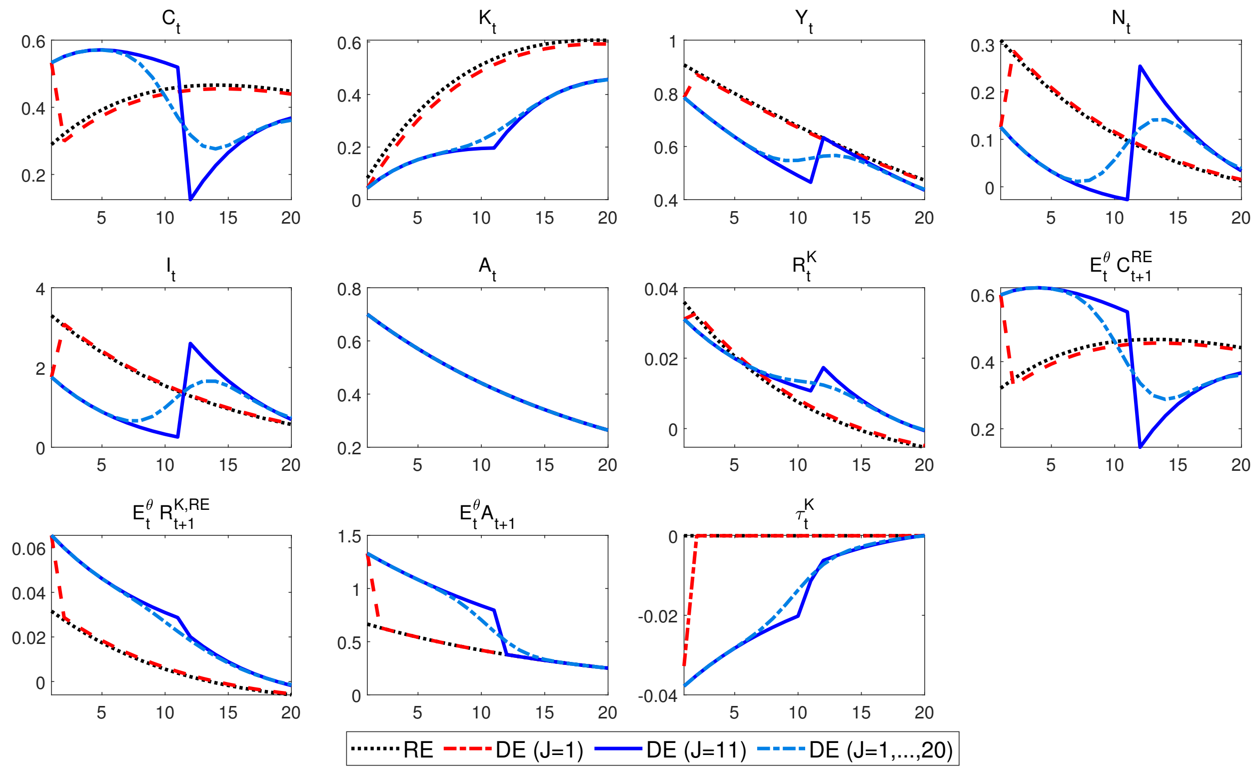
$$\mathbf{T}^\theta \equiv (\mathbf{\Gamma}_0^z)^{-1}\mathbf{\Gamma}_1^z; \quad \mathbf{R}^\theta \equiv (\mathbf{\Gamma}_1^z)^{-1}\mathbf{\Psi}^z.$$

The key advantage of our solution method is its *portability*: a researcher can transform any off-the-shelf linear dynamic equilibrium model (58) and compute the DE law of motion (62) with a few additional lines of code.

5.3 Results

We study our mechanism by simulating the impulse response function (IRF) under DE and compare it to a counterfactual RE economy where decisions at each point in time are taken

Figure 1: Impulse response to a positive TFP shock



Notes: We report percentage responses to a one-standard-deviation TFP shock in an RE model (dotted black line), DE model with $J = 1$ (dashed red line), DE model with $J = 11$ (solid blue line), and DE model with varying weights from $J = 1, \ldots, 20$ (dashed light blue line).

under RE ($\theta = 0$). We choose conventional values for many parameters. The inverse Frisch elasticity, η, is set to 1. We choose $\alpha = 0.33$ for the capital share and $\beta = 0.985$ for the discount factor. The depreciation rate, δ, is set to 0.025 and the TFP process is such that $\rho = 0.95$ and $\sigma = 0.007$. We set the representativeness parameter θ to 1. This value is in line with those used in Bordalo et al. (2019b) and Bordalo et al. (2019a). We consider here several cases for the agent's comparison group: $J = 1$ (immediate past) and a more realistic case of $J > 1$.[19] Figure 1 presents the IRFs under RE and under the various DE models to a one-standard-deviation increase in TFP.

Memory recall based on the immediate past. Let us first consider the $J = 1$ economy in Figure 1. As shown analytically in the context of the three-period model of section 4.1, and confirmed numerically in this RBC model, with $J = 1$ the response to the

[19] In Appendix E, we also consider $J = \infty$, i.e. the comparison group is fixed at the steady state.

state variable is the same as in the RE economy, while the reaction to the innovation is different due to DE beliefs. Indeed, we see here that on impact, at $t = 1$, the DE agent's consumption shoots up nearly as twice as high relative to the RE IRF. Given a surprise increase of the path of income, DE agents expect higher consumption tomorrow relative to what was expected as of yesterday. Faced with this new information, they overinflate the probability of high consumption tomorrow (high $\mathbb{E}_t^\theta C_{t+1}^{RE}$ in the diagnostic Euler equation (56)). This over-inflation of the current good news causes DE agents to raise consumption today through (perceived) consumption smoothing.[20] Hours, output and investment is lower on impact relative to the RE. This is due to the standard Barro and King (1984) logic: since the marginal utility of consumption is 'excessively' low compared to RE, this wealth effect lowers the incentive to work. Thus, compared to the RE case, we observe less labor, output and, through the resource constraint, also investment.

Because memory recall is based on the immediate past, starting at time $t = 2$, the high consumption is no longer surprising, since the comparison group is the expected consumption that takes into account the observed positive TFP shock at $t = 1$. As we argued earlier, this intuition manifests itself in the same DE response to the endogenous states as for the RE law of motion. Therefore, the only difference between the two economies starting from $t = 2$ is that the $t = 1$ muted impact on investment under DE (which falls by about 1.5% compared to RE) leads to less capital entering period $t = 2$ (which therefore falls only by 1.5δ%, compared to RE). Put differently, conditional on this difference in inherited capital, the transitional dynamics of the DE economy towards the steady state follow exactly the RE law of motion.

Importantly, since the difference in capital stock at $t = 2$ is small (as the depreciation rate is small, at $\delta = 0.025$), the resulting differences in the IRF paths of the two economies are also small starting from $t = 2$. In particular, while consumption at $t = 2$ drops significantly compared to its $t = 1$ over-optimistic expectation (see the $t = 1$ increase in $\mathbb{E}_t^\theta \widehat{C}_{t+1}$), it then returns near the counterfactual RE path starting from $t = 2$, as do the other variables.

Memory recall based on the distant past. As we saw above, while the $J = 1$ case is instructive, the effect of DE is transitory and negligible, as also found in L'Huillier et al. (2021). However, our proposed theoretical framework presented in section 4.3 has emphasized and allowed the possibility to study the arguably more plausible, and dynamically-richer case, where the comparison group is the state prevailing under an intermediate past or its combination. For example, it is plausible that, in the agents' mind, recent events take time

[20] Thus, the high MPC property we saw in the simplified model of section 3 also appears here. In this model the consumption smoothing tradeoff also incorporates the effect that DE agents over-react to the current news of future higher productivity, as can be seen in the response of $\mathbb{E}_t^\theta \widehat{A}_{t+1}$, because the TFP process is allowed here to be persistent.

to sink in, due to data lag or cognitive constraints, while memory of a distant past fade away. To explore these implications, in Figure 1, we report the case with $J = 11$. This value is from Bordalo et al. (2019b), who estimate the parameter using firm level earnings data and analysts' forecasts. We then also consider a case where the comparison group consists of a weighted average of lagged expectations from $J = 1$ to $J = 20$.[21]

Consider first $J = 11$ (solid blue lines). Initially, the impulse response is identical to that of $J = 1$ since the comparison group is the same as $J = 1$ case for $t = 1$. After $t = 1$, however, agents' beliefs continue to overinflate the high consumption state and hence consumption is persistently higher than RE, while hours, output and investment are lower. In particular, compared to the $J = 1$ case, while the capital entering period $t = 2$ is the same, the law of motion of this $J > 1$ economy is different as it continues to be characterized by overoptimism regarding future resources. At $t = 12$, the high consumption is no longer surprising as the comparison group is now the expected consumption after the TFP shock. Now DE agents realize that, after periods of overconsumption, the capital stock is too low. The strong resulting negative wealth effect leads to a sharp cut back on consumption, increase in hours, and investment, in order to build back the capital stock.

Finally, consider the combination case (dashed light blue lines). The general intuition behind these dynamics resemble that for the discrete version of $J = 11$. In both cases, we notice that setting $J > 1$ allows the model to generate long-lived effects of distortions caused by DE beliefs, as well as a sudden reversal phenomenon, or in other words, a boom-bust cycle. Naturally, these reversals and general dynamics are smoother under the combination case than the discrete one.[22]

Labor-in-advance extension. We also briefly mention here a modification of the economic structure of the standard model. As the literature on belief distortions has already noted (eg. Angeletos and La'O (2009), Ilut and Schneider (2014)) in a labor-in-advance version, where the labor input must be chosen before the realization of shocks, as opposed to being a static choice, those belief distortions matter for the equilibrium labor. To that end, we can easily include this feature in the standard model presented above so that the labor

[21] Specifically, we assign weights α_j attached to each $t - j$ expectations, $\sum_{j=1}^{J} \alpha_j \mathbb{E}_{t-j}(.)$, by fitting a Beta distribution with mean 0.5 and standard deviation 0.1. Thus, while $t - 10$ and $t - 11$ expectations receive the most weight ($\alpha_{10} = \alpha_{11} = 0.18$), surrounding lagged expectations also receive substantial weights (for example, $\alpha_9 = \alpha_{12} = 0.15$ and $\alpha_8 = \alpha_{13} = 0.10$).

[22] In Figure 2 in the Appendix, we also consider the case of $J = \infty$, which implies that the comparison set is the steady state instead of the previous period. Here the agent's belief continues to overinflate high consumption state after $t = 1$ and hence consumption is persistently higher than RE while hours, output and investment are lower. In this case, expectations are never re-aligned in the DE economy to equal the ones under the true density, and therefore there is a very large difference in the path taken under the DE economy and the counterfactual RE one.

optimality condition becomes:

$$(N_t^\theta)^\eta = \mathbb{E}_t^\theta \left[(C_{t+1}^{RE})^{-1}(1-\alpha)A_{t+1}(K_t^\theta)^\alpha (N_{t+1}^{RE})^{-\alpha} \right].$$

As we show in detail in Figure 2 in the Appendix, now in addition to consumption, labor on impact is higher than under RE. This is because the agent's perceived expected return on labor under DE, $\mathbb{E}_t^\theta R_{t+1}^{N,RE} \equiv \mathbb{E}_t^\theta \left[(1-\alpha)A_{t+1}(K_t^\theta)^\alpha (N_{t+1}^{RE})^{-\alpha}\right]$, is higher than the RE return on labor because of 'surprisingly' high TFP and capital.[23] Thus, the DE model can potentially address a major shortcoming of the standard RBC models: relative to data, consumption and labor volatility is too low (King and Rebelo (1999)).

Beliefs about returns. A characteristic property of the representativeness heuristic is that due to the selective memory recall, conditional on the same information observed by an outside econometrician, beliefs are distorted compared to the true density. In our model, this distortion manifests as a time-t predictable 'wedge' τ_t^K given by the difference between the typical realization of future return on capital $\widehat{R}_{t+1}^K$ and its expected value under DE beliefs, which dictates perceived intertemporal tradeoffs in diagnostic Euler equation (56). In particular, in response to a positive TFP shock, DE agents overinflate the possibility of high return on capital (high $\mathbb{E}_t^\theta \widehat{R}_{t+1}^{K,RE}$ in Figure 1). The econometrician, who measures ex-post the return on capital under the true density, does not find on average evidence of such a high return and rationalizes, in a reduced-form, the resulting difference as a negative τ_t^K (or an 'as if' subsidy that stimulates investment at time t).[24] As shown in Figure 1, when $J = 1$ this predictability of forecast errors on returns only occurs for one period following the shock, but is longer-lived when memory is based on the more distant past.

Put together, our proposed method to study the belief distortions implied by the representativeness heuristic when memory is based on more distant past is a promising way to generate rich macroeconomic dynamics. Qualitatively, our results point to the potential of this model to generate significant boom-bust cycles and long-lived belief distortions in general equilibrium models with endogenous states. In this respect, our model clearly shares with an existing literature on DE the ability to generate such reversals and predictable distortions in beliefs about returns. That literature typically obtains (i) such predictability and reversal on longer horizons in partial equilibrium models, which directly posit a law of motion for example for asset returns (eg. Bordalo et al. (2019b)), or (ii) predictability over one period ahead in equilibrium models where capital is endogenously chosen (eg. Bordalo et al. (2019a)

[23] When $J > 1$, we find that this amplification of consumption and labor relative to RE is persistent.

[24] A recent literature evaluates the quantitative role of such countercyclical wedges, taken either as independent 'confidence' shocks (eg. Ilut and Schneider (2014), Bianchi et al. (2018), Angeletos et al. (2018)), or propagated as a response to the aggregate state of the economy, through incomplete information (Angeletos and Lian (2020)) or ambiguity aversion (Ilut and Saijo (2020)).

and L'Huillier et al. (2021)). In our model, by studying memory recall based on more distant past, we can obtain longer-lived predictable wedges and reversal, as suggested for example by the evidence in López-Salido et al. (2017), in otherwise standard business cycle models.

As illustrated here through different specifications on the role of memory recall, the tractability underlying these results suggests that typical methods for quantitatively estimating linear DSGE models (e.g. full information likelihood-based, see An and Schorfheide (2007)), could be further employed to discipline and evaluate the quantitative potential of the DE beliefs to parsimoniously explain the data.

6 Diagnostic Expectations and Policy Rules

In the previous section, we established that DE can generate rich propagation dynamics as agents adjust their behavior in response to their perception of the data generating process. In this section, we show that another kind of endogeneity can have profound effects on the properties of the macroeconomy by activating or silencing DE: policymakers' behavior. Changes in policymakers' behavior can affect predictability and uncertainty about future outcomes. To illustrate this point, we consider a simple Fisherian model in which the real interest rate follows an iid exogenous process and the central bank moves the nominal interest rate in response to inflation. The response to inflation can be more than one-to-one, in which case the Taylor is satisfied, or not, in which case the Taylor principle is violated.

We find that DE are silent when the Taylor principle is satisfied, while they become relevant once the central bank violates the Taylor principle. When the Taylor principle is satisfied (and real interest rates are iid), inflation becomes unpredictable. Instead, when the Taylor principle is violated, *endogenous predictability* arises as a result of the policy rule in place. Interestingly, *endogenous predictability* holds even when inflation shows no persistence because exogenous real interest rate shocks have news effects. Under indeterminacy, sunspot shocks play an additional important role: By introducing uncertainty about inflation, they activate DE. Without sunspot shocks, *endogenous non-stochasticity* applies, as one-step-ahead inflation becomes fully predictable and DE are silenced even if the Taylor principle is violated. These results elucidate how changes in policymakers' behavior interact with the two conditions necessary for DE to manifest themselves: Predictability and residual uncertainty.

6.1 Rational Expectations

Before deriving the solution under DE, we briefly review the RE solution. Consider the Fisherian model under RE:

$$r_t = i_t - \mathbb{E}_t [\pi_{t+1}] ; \quad i_t = \psi_\pi \pi_t, \tag{63}$$

where $r_t \sim N(0, \sigma_r^2)$ is the exogenous real interest rate, i_t is the nominal interest rate chosen by the central bank based on the rule in (63), and π_t is net inflation. The two equations can be combined to obtain:

$$\psi_\pi \pi_t = r_t + \mathbb{E}_t [\pi_{t+1}] \tag{64}$$

Determinacy Suppose that the Taylor principle holds, i.e. $\psi_\pi > 1$. Then, determinacy holds and the unique solution is obtained by iterating (64) forward:

$$\pi_t = \psi_\pi^{-1} r_t \sim N\left(0, \psi_\pi^{-2} \sigma_r^2\right).$$

Indeterminacy If the Taylor principle does not hold, $\psi_\pi \leq 1$, the model admits multiple solutions (indeterminacy). If we replace $\mathbb{E}_t [\pi_{t+1}]$ with $\pi_{t+1}+ \eta_{t+1}^\pi$ with η_{t+1}^π a sunspot shock such that $\mathbb{E}_t \left[\eta_{t+1}^\pi\right] = 0$, the solutions to the model are given by:

$$\pi_{t+1} = \psi_\pi \pi_t - r_t + \eta_{t+1}^\pi$$

If we assume $\eta_{t+1}^\pi \sim N\left(0, \sigma_{\eta^\pi}^2\right)$, we have:

$$\pi_{t+1} \sim N\left(0, \left(1 - \psi_\pi^2\right)^{-1} \left(\sigma_{\eta^\pi}^2 + \sigma_r^2\right)\right).$$

Under determinacy, a shock to the real interest rate has an immediate impact on current inflation, as inflation expectations are anchored to zero. Inflation has no persistence as long as the real interest rate has no persistence and volatility is declining in ψ_π, the parameter controlling the strength with which the central bank responds to inflation. Under indeterminacy, inflation expectations are not anchored. A shock to the real interest induces a movement of expected inflation in the opposite direction instead of a change in current inflation. In equilibrium, this translates into a movement in *next period* inflation. Thus, absent sunspot shocks, inflation is predetermined with respect to the fundamental shocks to the real interest rate. In this sense, shocks to the real interest rate act as news, by moving expectations about future inflation. Sunspot shocks induce a further increase in volatility. Finally, once the Taylor principle is violated ($\psi_\pi \leq 1$), an increase in ψ_π increases inflation persistence and, consequently, its unconditional volatility.

6.2 Diagnostic Expectations

Now we introduce DE. To keep the analysis simple, we focus on the case $J = 1$. Note that given the assumption that the exogenous process for the real interest rate is i.i.d., DE would not have any role if applied only to the exogenous process. Under DE, the Fisherian model is:

$$r_t = i_t - \mathbb{E}_t^{\theta}\left[\pi_{t+1}\right]; \quad i_t = \psi_\pi \pi_t,$$

which is just the counterpart to the model in (63), but under DE beliefs over inflation. The two equations can be combined to obtain:

$$\psi_\pi \pi_t = r_t + \mathbb{E}_t^{\theta}\left[\pi_{t+1}\right]. \tag{65}$$

Determinacy Even under DE, if the Taylor principle holds ($\psi_\pi > 1$), the model is determinate. The unique solution is obtained iterating forward equation (65) and it us identical to the one obtained under RE:

$$\pi_t = \psi_\pi^{-1} r_t \sim N\left(0, \psi_\pi^{-2}\sigma_r^2\right).$$

Thus, if the Taylor principle holds, DE do not affect the solution of the model. Intuitively, this occurs because under determinacy inflation is a purely forward looking process that inherits the properties of the exogenous process r_t. Given that we have assumed that this process is i.i.d., there are no revisions in expectations about future outcomes and the conditional expectation of inflation is always zero independently of the time horizon.

Indeterminacy If the Taylor principle does not hold, $\psi_\pi \leq 1$, the model admits multiple solutions (indeterminacy). As a first step, we can replace $\mathbb{E}_t\left[\pi_{t+1}\right]$ with $\pi_{t+1} + \eta_{t+1}^{\pi}$:

$$r_t = \psi_\pi \pi_t - \left[(1+\theta)\,\mathbb{E}_t\left[\pi_{t+1}\right] - \theta \mathbb{E}_{t-1}\left[\pi_{t+1}\right]\right] = \psi_\pi \pi_t - (1+\theta)\left[\pi_{t+1} - \eta_{t+1}^{\pi}\right] + \theta \mathbb{E}_{t-1}\left[\pi_{t+1}\right]$$

where we assume that η_{t+1}^{π} is a normally distributed sunspot shock such that $\mathbb{E}_t\left[\eta_{t+1}^{\pi}\right] = 0$. As we explain in more detail below, the assumption of a random, normality distributed sunspot shock is important because it activates DE and allows us to write DE as a linear combination of present and lagged RE. Rearranging terms and substituting recursively the expression for π_{t+1} in $\mathbb{E}_{t-1}\left[\pi_{t+1}\right]$, we obtain:

$$\pi_{t+1} = \psi_\pi \pi_t - (1+\theta)^{-1} r_t + \eta_{t+1}^{\pi} - \theta\,(1+\theta)^{-1}\,\psi_\pi \eta_t^{\pi}.$$

When the Taylor principle is violated, the DE solution is different from the RE solution because the policy rule induces predictability of inflation. This occurs via two channels. First, inflation becomes persistent, despite the exogenous shocks not being persistent. Second, the real interest rate shock acts as a *news shock* that determines a revision in expectations about future inflation without a movement in current inflation. Both channels are absent when the Taylor principle holds and the solution is determinate.

Sunspot shocks play a key role under DE. Absent sunspot shocks, DE are not active even if the Taylor principle is violated. If the only shock occurring is the real interest rate shock, inflation is predetermined and there is no uncertainty about future inflation. Furthermore, in deriving the solution we have made use of $\mathbb{E}_t^\theta [\pi_{t+1}] = (1+\theta)\,\mathbb{E}_t [\pi_{t+1}] - \theta \mathbb{E}_{t-1} [\pi_{t+1}]$ that holds under the assumption of normality of the one-step-ahead distribution of inflation. This property holds in equilibrium if the sunspot shocks are normally distributed.

Sunspot shocks and fundamental shocks together activate DE even when inflation has no persistence ($\psi_\pi = 0$). On the one hand, the presence of the sunspot shock preserves uncertainty and normality about one-step-ahead inflation. On the other hand, the current real interest rate shock determines a revision in expectations even if inflation has no persistence. Absent sunspot shocks, inflation becomes predetermined. Absent the fundamental shock r_t, there are no revisions in expectations because inflation has zero persistence. Thus, indeterminacy introduces a *news effect* that activates the DE features as long as the one-step-ahead distribution of inflation is not degenerate.

The real interest shock r_t moves inflation expectations for time $t+1$ with respect to the previous period. Under DE, the effect of this new information on expectations about future inflation is enhanced by a factor $(1+\theta)$. Accordingly, in equilibrium, the actual response needs to be tempered down by a factor $(1+\theta)$. Thus, the over-reaction in expectations leads to an equilibrium that is *less* volatile with respect to the fundamental shock.

The impact of the sunspot shock at time t is the same under RE and under DE. But this is true by definition, as a sunspot shock can be interpreted as a shock to expectations that determines a one-to-one movement in actual inflation (a self-fulfilling prophecy). Under DE, however, the sunspot shock also induces misperception about future inflation. While the shock has evidently the same impact at time t, it has different implications for expected inflation and the future path of inflation. Thus, the effect of the shock is the same only in terms of its contemporaneous effect on inflation. Once the economy reaches time $t+1$, agents' DE are realigned with RE and inflation declines. The correction depends on the size of the initial movement, η_t^π, and the associated revision in expectations. This revision, in turn, depends on the persistence of inflation, ψ_π, and the severity of the DE distortion, θ.

If $\eta_t^\pi \sim N\left(0, \sigma_{\eta^\pi}^2\right)$, inflation follows the distribution:

$$\pi_t \sim N\left(0, \left(1-\psi_\pi^2\right)^{-1}\left[(1+\theta)^{-2}\sigma_r^2 + \left[1+\left(\theta\,(1+\theta)^{-1}\psi_\pi\right)^2\right]\sigma_{\eta^\pi}^2\right]\right).$$

Under DE, volatility is increasing in ψ_π for two reasons. First, when $\psi_\pi < 1$, a larger ψ_π implies a more persistent inflation process as it does under RE. Second, a larger ψ_π increases the impact of the lagged sunspot shock on future inflation for a given θ. When ψ_π increases, a sunspot shock leads to a larger revision in expectations about future inflation. Both current inflation and the distorted expectations about future inflation increase. In the next period, inflation experiences a correction that depends on ψ_π. An increase in the DE distortion, controlled by the parameter θ, determines an increase in the volatility of inflation due to the sunspot shock, but a decline in the volatility of inflation due to the fundamental shock r_t. Thus, the overall effect on the volatility of inflation is ambiguous.

Summarizing, this illustrative example shows how DE can be relevant under some policy regimes, but not others. Furthermore, the implications of DE following policy changes vary *depending on the source* of the disturbances. In this example, DE determine an increase in volatility due to sunspot shocks, but a decline in the volatility due to fundamental shocks. Sunspot shocks play a particularly important role because they preserve uncertainty about one-step-ahead inflation, activating the DE channel. These insights also apply when allowing for persistence in the exogenous process and in richer models with more shocks and more complex forms of policy changes. This is because policy changes typically induce chances in the propagation of the shocks trough the economy, affecting which shocks are relevant and how persistent their effects are. Thus, policy changes can potentially have important welfare implications as agents come to experience different volatilities and changes in the propagation of shocks, as a result of the interaction between DE and policymakers' behavior.

7 Conclusions

In this paper, we developed the micro-foundations for applying the DE paradigm to both exogenous and endogenous variables and characterized the general equilibrium formulations for DE in a large class of recursive macroeconomic models. Building on the paradigm developed by Bordalo et al. (2018), we established and studied the implications of three important properties of DE in the presence of endogenous variables. First, *endogenous predictability*: The predictability of future outcomes depends on the endogenous actions taken by agents and policymakers. Therefore, the optimal actions and distorted DE are determined jointly. Second, *endogenous non-stochasticity*: Actions taken by agents or policymakers can silence DE by removing residual uncertainty and making future outcomes fully predictable. This silencing constitutes a novel channel through which policy can affect allocations and welfare. Third, the *failure of LOIE under distant memory:* When current DE are affected by memories formed in the distant past, the LOIE generally fails. This leads to time inconsistency because optimal plans decided in the past become suboptimal as a result of the change in beliefs induced by imperfect memory. We proposed and studied two possible ways to address time-inconsistency: Naïveté or sophistication.

In the final part of the paper, we built on these results to show that DE can generate rich propagation dynamics in an otherwise standard RBC model and to study the pervasive effects of policymakers' behavior under DE. In future work, we aim to further develop this last set of results. First, we plan to formally establish the quantitative importance of DE in state-of-the-art business cycle models that can be confronted with the data. Second, while here DE are activated or de-activated discretely by uncertainty, we also plan to extend the DE paradigm to have smooth effects of uncertainty. This will also allow us to study environments in which the severity of DE distortions varies over time in response to changes in the volatility of the exogenous shocks or due to policymakers' behavior.

References

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

An, Sungbae and Frank Schorfheide, "Bayesian analysis of DSGE models," *Econometric Reviews*, 2007, *26* (2–4), 113–172.

Angeletos, George-Marios and Chen Lian, "Confidence and the Propagation of Demand Shocks," 2020. NBER WP 27702.

_ **and Jennifer La'O**, "Noisy Business Cycles," in "NBER Macroeconomics Annual 2009, Volume 24" 2009, pp. 319–378.

_ **, Fabrice Collard, and Harris Dellas**, "Quantifying Confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

Barro, Robert and Robert G. King, "Time-separable preferences and intertemporal-substitution models of business cycles," *Quarterly Journal of Economics*, 1984, *99* (4), 817–839.

Bianchi, Francesco, Cosmin L Ilut, and Martin Schneider, "Uncertainty shocks, asset supply and pricing over the business cycle," *The Review of Economic Studies*, 2018, *85* (2), 810–854.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *The Quarterly Journal of Economics*, 2016, *131* (4), 1753–1794.

_ **,** _ **,** _ **, 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, Cosmin Ilut, Roberto Motto, and Massimo Rostagno, "Monetary policy and stock market boom-bust cycles," 2008. ECB WP 955.

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 Hikaru Saijo, "Learning, confidence, and business cycles," *Journal of Monetary Economics*, 2020. forthcoming.

_ **and Martin Schneider**, "Ambiguous business cycles," *American Economic Review*, 2014, *104* (8), 2368–99.

_ **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.

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.

King, Robert G. and Sergio T. Rebelo, "Resuscitating real business cycles," in "Handbook of Macroeconomics" 1999, pp. 927–1007.

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.

Lewis, Daniel J, Davide Melcangi, and Laura Pilossoph, "Latent heterogeneity in the marginal propensity to consume," 2020. FRB of New York Staff Report No. 902.

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.

López-Salido, David, Jeremy C Stein, and Egon Zakrajšek, "Credit-market sentiment and the business cycle," *The Quarterly Journal of Economics*, 2017, *132* (3), 1373–1426.

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

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.

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

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

First, take a second-order Taylor approximation of period utility around the steady state,

$$u_t - \overline{u} \approx \overline{u}_c\overline{C}\left(\frac{C_t - \overline{C}}{\overline{C}}\right) + \frac{1}{2}\overline{u}_{cc}\overline{C}^2\left(\frac{C_t - \overline{C}}{\overline{C}}\right)^2$$

and we use

$$\frac{C_t - \overline{C}}{\overline{C}} = \widehat{C}_t + \frac{1}{2}\widehat{C}_t^2$$

where hat denotes log-deviations, and also note

$$\begin{aligned} u(C) &= bC - \frac{1}{2}C^2 \\ \overline{u}_c &= b - \overline{C} \\ \overline{u}_{cc} &= -1 \end{aligned}$$

so we have

$$\begin{aligned} u_t - \overline{u} &\approx \overline{u}_c\overline{C}\left(\widehat{C}_t + \frac{1}{2}\widehat{C}_t^2\right) + \frac{1}{2}\overline{u}_{cc}\overline{C}^2\widehat{C}_t^2 \\ &= (b - \overline{C})\overline{C}\left(\widehat{C}_t + \frac{1}{2}\widehat{C}_t^2\right) - \frac{1}{2}\overline{C}^2\widehat{C}_t^2 \\ &= \overline{C}\left[(b - \overline{C})\widehat{C}_t - \frac{1}{2}(2\overline{C} - b)\widehat{C}_t^2\right] \\ &= \overline{Y}\left[(b - \overline{Y})\widehat{C}_t - \frac{1}{2}(2\overline{Y} - b)\widehat{C}_t^2\right] \end{aligned}$$

where we use at the steady state $\overline{C}_1 = \overline{C}_2 = \overline{Y} = \overline{C}$.

The policy functions are

- Without tax policy

$$C_1^\theta = \overline{Y} + \frac{1+\theta}{2+\theta}\varepsilon_1; C_2^\theta = \overline{Y} + \varepsilon_2 + \frac{1}{2+\theta}\varepsilon_1.$$

- With tax policy

$$C_1^{\theta,policy} = \overline{Y} + \frac{1}{2}\varepsilon_1; C_2^{\theta,policy} = \overline{Y} + \frac{1}{2}\varepsilon_1.$$

Then, in log-deviations, the policy functions are

- Without tax policy

$$\overline{Y}\widehat{C}_1^{\theta} = \frac{1+\theta}{2+\theta}\varepsilon_1; \overline{Y}\widehat{C}_2^{\theta} = \varepsilon_2 + \frac{1}{2+\theta}\varepsilon_1.$$

- With tax policy

$$\overline{Y}\widehat{C}_1^{\theta,policy} = \frac{1}{2}\varepsilon_1; \overline{Y}\widehat{C}_2^{\theta,policy} = \frac{1}{2}\varepsilon_1.$$

So the ex-ante utility under the true DGP is

- Without tax policy

$$\begin{aligned}
U^{\theta} - \overline{U} &= E_0\left[(u_1 - \bar{u}) + (u_2 - \bar{u})\right] \\
&= \overline{Y}\left[(b-\overline{Y})E_0\widehat{C}_1^{\theta} - \frac{1}{2}(2\overline{Y}-b)E_0(\widehat{C}_1^{\theta})^2\right] + \overline{Y}\left[(b-\overline{Y})E_0\widehat{C}_2^{\theta} - \frac{1}{2}(2\overline{Y}-b)E_0(\widehat{C}_2^{\theta})^2\right] \\
&= -\frac{2\overline{Y}-b}{2\overline{Y}}E_0\left[\left(\frac{1+\theta}{2+\theta}\varepsilon_1\right)^2\right] - \frac{2\overline{Y}-b}{2\overline{Y}}E_0\left[\left(\varepsilon_2 + \frac{1}{2+\theta}\varepsilon_1\right)^2\right] \\
&= -\frac{2\overline{Y}-b}{2\overline{Y}}\left(\frac{1+\theta}{2+\theta}\right)^2\sigma_1^2 - \frac{2\overline{Y}-b}{2\overline{Y}}\left[\sigma_2^2 + \left(\frac{1}{2+\theta}\right)^2\sigma_1^2\right],
\end{aligned}$$

where

$$\lim_{\sigma_2\to 0} U^{\theta} - \overline{U} = -\frac{2\overline{Y}-b}{2\overline{Y}}\left[\left(\frac{1+\theta}{2+\theta}\right)^2 + \left(\frac{1}{2+\theta}\right)^2\right]\sigma_1^2.$$

- With tax policy

$$\begin{aligned}
U^{\theta,policy} - \overline{U} &= E_0\left[(u_1 - \bar{u}) + (u_2 - \bar{u})\right] \\
&= \overline{Y}\left[(b-\overline{Y})E_0\widehat{C}_1^{\theta,policy} - \frac{1}{2}(2\overline{Y}-b)E_0(\widehat{C}_1^{\theta,policy})^2\right] \\
&+ \overline{Y}\left[(b-\overline{Y})E_0\widehat{C}_2^{\theta,policy} - \frac{1}{2}(2\overline{Y}-b)E_0(\widehat{C}_2^{\theta,policy})^2\right] \\
&= -\frac{2\overline{Y}-b}{2\overline{Y}}E_0\left[\left(\frac{1}{2}\varepsilon_1\right)^2\right] - \frac{2\overline{Y}-b}{2\overline{Y}}E_0\left[\left(\frac{1}{2}\varepsilon_1\right)^2\right] \\
&= -\left(\frac{2\overline{Y}-b}{2\overline{Y}}\right)\frac{1}{2}\sigma_1^2.
\end{aligned}$$

Thus, taking the difference between the former and the latter, the utility gain from adopting the tax policy is given by

$$\left(\frac{2\overline{Y}-b}{2\overline{Y}}\right)\frac{\theta^2}{2(2+\theta)^2}\sigma_1^2,$$

which is increasing in the representativeness parameter θ.

A.2 Proof of Proposition 3

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 $J>1$

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

A.3 Proof of Proposition 4

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

$$\varepsilon_2+K_1-K_2^{RE}=\mathbb{E}_2\left(\varepsilon_3+K_2^{RE}\right), \tag{66}$$

which give us the RE coefficients for K_2^{RE} in Proposition 4.

For the DE agent, conditional on reaching period 2, the optimal K_2^θ solves the tradeoff

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

as implied by the problem in (33). When $J=1$, for a given state K_1 and exogenous innovation ε_2, by using the time 2 and 3 budget constraint, this tradeoff amounts to

$$\varepsilon_2+K_1-K_2^\theta=\mathbb{E}_2\left(\varepsilon_3+K_2^\theta\right)+\theta\left[\mathbb{E}_2\left(\varepsilon_3+K_2^\theta\right)-\mathbb{E}_1\left(\varepsilon_3+K_2^\theta\right)\right] \tag{67}$$

By substituting the conjectured coefficients $\alpha_{K_1}^\theta$ and $\alpha_{\varepsilon_2}^\theta$ into the tradeoff (67), we obtain their values characterized in Proposition 4.

Time 1 policy. Moving backward, let us characterize the time 1 problem. By Corollary 1 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

$$C_1^\theta=\mathbb{E}_1^\theta\left[C_2^\theta+\alpha_{K_1}^\theta\left(C_3^\theta-C_2^\theta\right)\right], \tag{68}$$

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 equation (10) and the proof of Corollary 1 in Bordalo et al. (2018) for details), so we can break the RHS of (68) into

$$\mathbb{E}_1^\theta(C_2^\theta) + \alpha_{K_1}^\theta \mathbb{E}_1^\theta \left(C_3^\theta - C_2^\theta\right). \tag{69}$$

which by Proposition 3 and Corollary 1, means

$$C_1^\theta = \mathbb{E}_1^\theta C_2^\theta. \tag{70}$$

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

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

Notice that the qualitative resemblance of this result to the one in the statistical model of equation (14). 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 4. In the case of $\theta = 0$, this also solves for the corresponding RE coefficients.

A.4 Proof of Proposition 5

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 tradeoff 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 tradeoff is given by

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

The RHS equals

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

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 tradeoff 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 2, the time 1 tradeoff 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.5 Proof of Proposition 6

Time 2 policy under naïveté. Consider the conjecture

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

The time 2 tradeoff is given by

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

The RHS equals

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

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

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

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

Time 2 policy under sophistication. Consider the conjecture

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

The time 2 tradeoff is given by

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

The RHS equals

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

Connecting this with the LHS, we have

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

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

A.6 Proof of Proposition 7

The Proposition below considers the time 1 savings policy under naïveté, referenced in section 4.3.2.

Proposition 7. *Compared to the $J = 1$ case, when $J = 2$ the naïveté policy function*

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

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.7 Proof of Proposition 8

The Proposition below considers the time 1 savings policy under sophistication, referenced in section 4.3.2.

Proposition 8. *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 7 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

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

The RHS equals

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

After some algebra, we find that this equals

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

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

$$
\begin{aligned}
\alpha_{\varepsilon_1}^{\theta,s} &= \frac{1}{1+(1+\theta)\left[(1-\alpha_{K_1}^{\theta})^2+(\alpha_{K_1}^{\theta})^2\right]} = \frac{(2+\theta)^2}{(2+\theta)^2+(1+\theta)\left[(1+\theta)^2+1\right]} \\
\alpha_{K_0}^{\theta,s} &= \frac{1}{1+(1+\theta)\left[(1-\alpha_{K_1}^{\theta})(1-\alpha_{\mathbb{E}_0K_1}^{\theta}-\alpha_{K_1}^{\theta})+\alpha_{K_1}^{\theta}(\alpha_{\mathbb{E}_0K_1}^{\theta}+\alpha_{K_1}^{\theta})\right]} \\
&= \frac{2(2+\theta)^2}{2(2+\theta)^2+(1+\theta)\left[(1+\theta)(1+2\theta)+3\right]} \\
\alpha_{\mathbb{E}_{-1}K_0}^{\theta,s} &= \frac{\theta\left[(1-\alpha_{K_1}^{\theta})(1-\alpha_{\mathbb{E}_0K_1}^{\theta}-\alpha_{K_1}^{\theta})+\alpha_{K_1}^{\theta}(\alpha_{\mathbb{E}_0K_1}^{\theta}+\alpha_{K_1}^{\theta})\right]}{1+(1-\alpha_{K_1}^{\theta})(1-\alpha_{\mathbb{E}_0K_1}^{\theta}-\alpha_{K_1}^{\theta})+\alpha_{K_1}^{\theta}(\alpha_{\mathbb{E}_0K_1}^{\theta}+\alpha_{K_1}^{\theta})}\alpha_{K_0}^{\theta,s} \\
&= \frac{\theta\left[(1+2\theta)(1+\theta)+3\right]}{2(2+\theta)^2+(1+2\theta)(1+\theta)+3}\alpha_{K_0}^{\theta,s}
\end{aligned}
$$

which give the specific coefficients in Proposition 8. When we compare this sophistication solution to the naïveté one, we find the patterns stated in Proposition 8. □

B Derivation of the Diagnostic Euler Equation and the Equilibrium Conditions in the RBC Model

First, we derive the diagnostic Euler equation. The first-order condition for capital in (54) implies

$$(C_t^{\theta})^{-1} = \beta\mathbb{E}_t^{\theta}V'(K_t^{\theta},A_{t+1}). \tag{71}$$

The envelope condition for (55) implies

$$V'(K_{t-1}^{\theta},A_t) = (C_t^{RE})^{-1}\left[A_t(K_{t-1}^{\theta})^{\alpha-1}(N_t^{RE})^{1-\alpha}+1-\delta\right].$$

Iterate one-period forward, we have

$$V'(K_t^{\theta},A_{t+1}) = (C_{t+1}^{RE})^{-1}\left[\alpha A_{t+1}(K_t^{\theta})^{\alpha-1}(N_{t+1}^{RE})^{1-\alpha}+1-\delta\right]. \tag{72}$$

Combining (71) and (72) yields the diagnostic Euler equation:

$$(C_t^{\theta})^{-1} = \beta\mathbb{E}_t^{\theta}\left[(C_{t+1}^{RE})^{-1}R_{t+1}^{K,RE}\right],$$

where $R_{t+1}^{K,RE}$ is the $t+1$ return on capital under RE policy function:

$$R_t^{K,RE} \equiv \alpha A_t(K_{t-1}^{\theta})^{\alpha-1}(N_t^{RE})^{1-\alpha}+1-\delta.$$

Since the derivations of other equilibrium conditions are standard, we simply list them below:

1. Labor supply: $(N_t^\theta)^\eta = (C_t^\theta)^{-1}(1-\alpha)Z_t(K_{t-1}^\theta)^\alpha(N_t^\theta)^{-\alpha}$
2. Diagnostic Euler equation: $(C_t^\theta)^{-1} = \beta\mathbb{E}_t^\theta\left[(C_{t+1}^{RE})^{-1}R_{t+1}^{K,RE}\right]$
3. Resource constraint: $C_t^\theta + I_t^\theta = Y_t^\theta$
4. Production function: $Y_t^\theta = Z_t(K_{t-1}^\theta)^\alpha(N_t^\theta)^{1-\alpha}$
5. Capital accumulation: $K_t^\theta = (1-\delta)K_{t-1}^\theta + I_t^\theta$
6. Return on capital: $R_t^{K,RE} \equiv \alpha Z_t(K_{t-1}^\theta)^{\alpha-1}(N_t^{RE})^{1-\alpha} + 1 - \delta$
7. Law of motion for the TFP shock: $\ln A_t = \rho\ln A_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,\sigma^2)$

C The RBC Model under Sophistication

Consider a representative household whose value W is given by

$$\begin{aligned} W(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}) = &\ln\widetilde{C}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}) - \frac{(\widetilde{N}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}))^{1+\eta}}{1+\eta} \\ &+ \beta\mathbb{E}_t^\theta V(\widetilde{K}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}), A_{t+1}, \Omega_t^{t+1-J}) \end{aligned} \tag{73}$$

subject to the resource constraint

$$\widetilde{C}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}) + \widetilde{K}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}) - (1-\delta)K_{t-1}^\theta = A_t(K_{t-1}^\theta)^\alpha(\widetilde{N}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}))^{1-\alpha},$$

where A_t is a TFP which follows $\ln A_t = \rho\ln A_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0,\sigma^2)$. Ω_{t-1}^{t-J} collects lagged expectations about capital and TFP: $\Omega_{t-1}^{t-J} \equiv \{\mathbb{E}_{t-j}K_t^\theta, \mathbb{E}_{t-j}A_{t+1}\}_{j=1}^J$ and $\Omega_t^{t+1-J} \equiv \{\mathbb{E}_{t+1-j}K_{t+1}^\theta, \mathbb{E}_{t+1-j}A_{t+2}\}_{j=1}^J$. Lagged expectations are included as states in W because they are used to construct the relevant comparison groups. In turn, they are included as states in V because it depends on the equilibrium policy that depends on lagged expectations. $\widetilde{K}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J})$, $\widetilde{C}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J})$ and $\widetilde{N}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J})$ are the optimal policies that solve the Bellman equation (73):

$$\begin{aligned} &\{\widetilde{K}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}), \widetilde{C}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}), \widetilde{N}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J})\} \\ &= \underset{K^\theta, C^\theta, N^\theta}{\operatorname{argmax}}\left[\ln C^\theta - \frac{(N^\theta)^{1+\eta}}{1+\eta} + \beta\mathbb{E}_t^\theta V(K^\theta, A_{t+1}, \Omega_t^{t+1-J})\right]. \end{aligned} \tag{74}$$

Under sophistication, the household evaluates the continuation value (75) under the policy rule that maximizes the Bellman equation (73). Thus we have

$$\begin{aligned} V(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}) = &\ln \widetilde{C}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}) - \frac{(\widetilde{N}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}))^{1+\eta}}{1+\eta} \\ &+ \beta \mathbb{E}_t V(\widetilde{K}^{\theta}(K_{t-1}^{\theta}, A_t), A_{t+1}, \Omega_t^{t-J+1}), \end{aligned} \tag{75}$$

subject to the resource constraint

$$\widetilde{C}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}) + \widetilde{K}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}) - (1-\delta)K_{t-1}^{\theta} = A_t(K_{t-1}^{\theta})^{\alpha}(\widetilde{N}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}))^{1-\alpha}.$$

To derive the Euler equation, first note that the first-order condition with respect to K_t^{θ} in (73) is given by

$$(\widetilde{C}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}))^{-1} = \beta \mathbb{E}_t^{\theta}\left[\frac{\partial V(K_t^{\theta}, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^{\theta}}\right]. \tag{76}$$

The envelope condition is

$$\frac{\partial W(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J})}{\partial K_{t-1}^{\theta}} = (\widetilde{C}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}))^{-1}\left[\alpha A_t(K_{t-1}^{\theta})^{\alpha-1}(\widetilde{N}^{\theta}(K_{t-1}^{\theta}, A_t, \Omega_{t-1}^{t-J}))^{1-\alpha} + 1 - \delta\right].$$

Comparing (73) and (75), we have

$$\begin{aligned} &V(K_t^{\theta}, A_{t+1}, \Omega_t^{t+1-J}) = W(K_t^{\theta}, A_{t+1}, \Omega_t^{t+1-J}) \\ &- \beta\left[\mathbb{E}_{t+1}^{\theta} V(\widetilde{K}^{\theta}(K_t^{\theta}, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J}) - \mathbb{E}_{t+1} V(\widetilde{K}^{\theta}(K_t^{\theta}, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})\right] \end{aligned}$$

Take a derivative with respect to K_t^θ and apply DE:

$$
\begin{aligned}
&\mathbb{E}_t^\theta\left[\frac{\partial V(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^\theta}\right] = \mathbb{E}_t^\theta\left[\frac{\partial W(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^\theta}\right] \\
&- \beta\mathbb{E}_t^\theta\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right]\Bigg\}\frac{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^\theta} \\
&- \beta\mathbb{E}_t^\theta\sum_{j=1}^{J}\Bigg(\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right]\Bigg\}\frac{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}{\partial K_t^\theta}\Bigg) \\
&= \mathbb{E}_t^\theta\left\{(\widetilde{C}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}))^{-1}\left[\alpha A_{t+1}(K_t^\theta)^{\alpha-1}(\widetilde{N}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}))^{1-\alpha} + 1 - \delta\right]\right\} \\
&- \beta\mathbb{E}_t^\theta\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right]\Bigg\}\frac{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^\theta} \\
&- \beta\mathbb{E}_t^\theta\sum_{j=1}^{J}\Bigg(\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right]\Bigg\}\frac{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}{\partial K_t^\theta}\Bigg),
\end{aligned}
$$

and combine it with (76):

$$
\begin{aligned}
(C_t^\theta)^{-1} &= \beta\mathbb{E}_t^\theta\left[(C_{t+1}^\theta)^{-1}R_{t+1}^{K,\theta}\right] \\
&- \beta\mathbb{E}_t^\theta\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}\right]\Bigg\}\frac{\partial \widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J})}{\partial K_t^\theta} \\
&- \beta\mathbb{E}_t^\theta\sum_{j=1}^{J}\Bigg(\Bigg\{\mathbb{E}_{t+1}^\theta\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right] \\
&- \mathbb{E}_{t+1}\left[\frac{\partial V(\widetilde{K}^\theta(K_t^\theta, A_{t+1}, \Omega_t^{t+1-J}), A_{t+2}, \Omega_{t+1}^{t+2-J})}{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}\right]\Bigg\}\frac{\partial \mathbb{E}_{t+2-j}K_{t+2}^\theta}{\partial K_t^\theta}\Bigg),
\end{aligned}
\tag{77}
$$

where $C_t^\theta = \widetilde{C}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J})$ and $R_{t+1}^{K,\theta}$ is the $t+1$ return on capital under the DE policy function:

$$R_t^{K,\theta} \equiv \alpha A_t (K_{t-1}^\theta)^{\alpha-1}(\widetilde{N}^\theta(K_{t-1}^\theta, A_t, \Omega_{t-1}^{t-J}))^{1-\alpha} + 1 - \delta.$$

The Euler equation (77) features (i) the expectation and the comparison group evaluated under the DE equilibrium policy function, (ii) the term that reflects the fact that the equilibrium policy and future utility, evaluated under the current preference, are not aligned (second and third lines) and (iii) the term that is associated with the fact that agents have control over future comparison group through current choice (fourth and fifth lines).

The Euler equation (77) highlights the computational demand for both us as modelers, as well as arguably for the agent, as we argued in section 4.3.3. One particular possibility to "simplify" this calculation is to obtain the Euler equation under "approximate" sophistication by focusing only on the direct, one-step ahead, term in (i):

$$(C_t^\theta)^{-1} = \beta \mathbb{E}_t^\theta \left[(C_{t+1}^\theta)^{-1} R_{t+1}^{K,\theta} \right]. \tag{78}$$

It is easy to solve for the recursive law of motion of the model featuring (78) by appropriately modifying existing solution methods for linear RE models. We find that, in an RBC model with "approximate" sophistication, the level of an increase in consumption on impact ($t = 1$) in response to a positive TFP shock is rising in J. That is, the $t = 1$ amplification of consumption is larger when the agent considers a more distant past as a comparison group. To see this, first consider $J = 2$. Consumption is higher at $t = 2$, relative to the RE path and the DE path when $J = 1$, because at $t = 2$ the agent overinflates the possibility of high future resources. The agent thus (correctly) expects higher consumption tomorrow relative to RE and $J = 1$ cases (higher $\mathbb{E}_1 C_2^\theta$), which makes expectations about higher consumption even more salient, and in turn causes her to further overinflate the possibility of high consumption tomorrow (higher $\mathbb{E}_1^\theta C_2^\theta$ in (78)). Consumption smoothing then implies the agent raises consumption today (higher C_1^θ) more than under RE and $J = 1$. As J increases, the agent (correctly) expects the high consumption path, relative to the RE path, to persist longer. Extending the above logic to $J > 2$, we deduce that the effect of high future consumption accumulates and front-loaded to $t = 1$, magnifying consumption on impact.

D 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:

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

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

$$\boldsymbol{\Gamma}_0^\theta \mathbf{x}_t^\theta = \boldsymbol{\Gamma}_2^\theta \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} + \boldsymbol{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \boldsymbol{\Psi}^\theta \varepsilon_t \tag{79}$$

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

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

Then:

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

Suppose that we do not need all elements in $\mathbf{x}_t^\theta$ to form expectations about the future.[25] In particular, we have

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

[25] 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.

but can be reduced to

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

Then (79) becomes

$$\mathbf{\Gamma}_0^{\theta}\mathbf{x}_t^{\theta} = \mathbf{\Gamma}_2^{\theta}\left[(1+\theta)\,\mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^{\theta} - \sum_{j=1}^{J}\theta\alpha_j\widetilde{\mathbf{M}}\left(\widetilde{\mathbf{T}}^{RE}\right)^{j+1}\widetilde{\mathbf{x}}_{t-j}^{\theta}\right] + \mathbf{\Gamma}_1^{\theta}\mathbf{x}_{t-1}^{\theta} + \mathbf{\Psi}^{\theta}\varepsilon_t. \tag{80}$$

This becomes:

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

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

We also expand the vector of endogenous variables to include $\mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE}$:

$$\underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ -(1+\theta)\mathbf{M}\mathbf{T}^{RE} & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \dots & \mathbf{I} \end{bmatrix}}_{\mathbf{\Gamma}_0^z} \underbrace{\begin{bmatrix} \mathbf{x}_t^\theta \\ \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} \\ \widetilde{\mathbf{x}}_{t-1}^\theta \\ \vdots \\ \widetilde{\mathbf{x}}_{t-J+1}^\theta \end{bmatrix}}_{\mathbf{z}_t^\theta}$$

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

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

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

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

The DE decision rule can be obtained simply by inverting this system:

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

where

$$\mathbf{T}^\theta \equiv (\mathbf{\Gamma}_0^z)^{-1}\mathbf{\Gamma}_1^z$$

$$\mathbf{R}^\theta \equiv (\mathbf{\Gamma}_1^z)^{-1}\mathbf{\Psi}^z.$$

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:

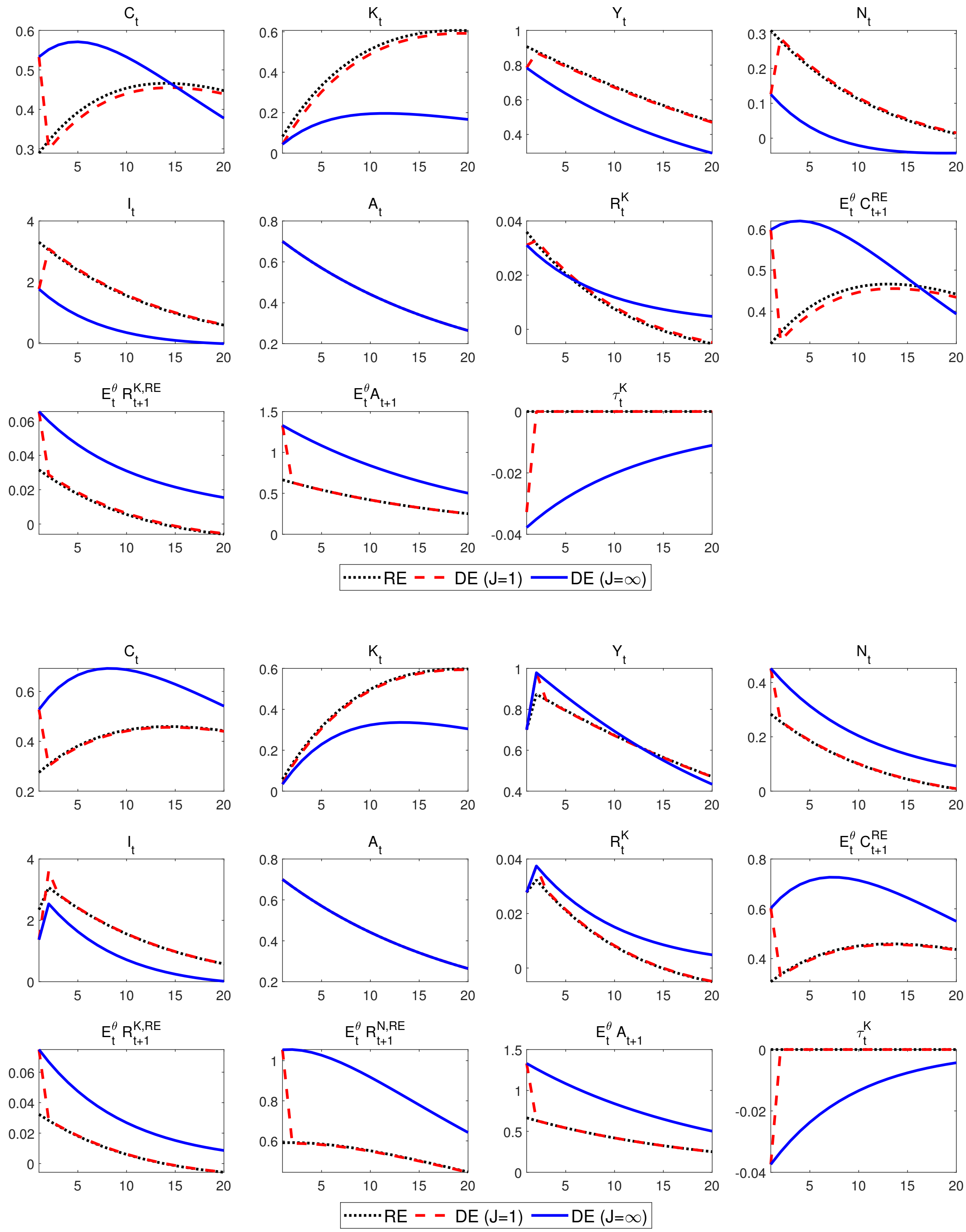
$$Var_t\left(w_{j,t+1}^{RE}\right) = (\mathbf{QR}^{RE}\mathbf{\Sigma}(\mathbf{QR}^{RE})')_{j,j} > 0,$$

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.

E Additional Figures

As referenced in section 5, in Figure 2 below we report the impulse response function (IRF) to a positive TFP shock for the standard model (top panel) and the labor-in-advance model (bottom panel). We do so for the model-implied path under DE with $J = 1$ (dashed red line), under DE with $J = \infty$ (solid blue line) and under the RE counterfactual model (dotted black line).

Figure 2: IRF for positive TFP shock: standard model (top panel) and labor-in-advance model (bottom panel)



Notes: We report percentage responses to a one-standard-deviation TFP shock in an RE model (dotted black line), DE model with $J = 1$ (dashed red line), and DE model with $J = \infty$ (solid blue line).