NBER WORKING PAPER SERIES

ECONOMIC AGENTS AS IMPERFECT PROBLEM SOLVERS

Cosmin L. Ilut
Rosen Valchev

Working Paper 27820
http://www.nber.org/papers/w27820

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020

We would like to thank our discussants George-Marios Angeletos, Paolo Bonomolo, Tarek Hassan and Luigi Iovino, as well as Ian Dew-Becker, Ryan Chahrour, Guido Lorenzoni, Filip Matejka, Kristoffer Nimark, Philipp Sadowski, Todd Sarver and Mirko Wiederholt, and seminar and conference participants at Bank of Finland, CERGE, CEU, EEA Congress, Green Line Macro Meetings, NBER EFBEM and EFCE Groups, Northwestern Macro Conference, Society for Economic Dynamics and Computing in Economics and Finance for helpful discussions and comments. This project is supported by NSF Award SES-1824367. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Economic Agents as Imperfect Problem Solvers
Cosmin L. Ilut and Rosen Valchev
NBER Working Paper No. 27820
September 2020
JEL No. C11,D83,D91,E21,E71

ABSTRACT

We develop a tractable model of limited cognitive perception of the optimal policy function, with agents using costly reasoning effort to update beliefs about this optimal mapping of economic states into actions. A key result is that agents reason less (more) when observing usual (unusual) states, producing state- and history-dependent behavior. Our application is a standard incomplete markets model with ex-ante identical agents that hold no a-priori behavioral biases. The resulting ergodic distribution of actions and beliefs is characterized by “learning traps”, where locally stable dynamics of wealth generate “familiar” regions of the state space within which behavior appears to follow past-experience-based heuristics. We show qualitatively and quantitatively how these traps have empirically desirable properties: the marginal propensity to consume is higher, hand-to-mouth status is more frequent and persistent, and there is more wealth inequality than in the standard model.

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

Rosen Valchev
Boston College
Department of Economics
Maloney Hall 396
140 Commonwealth Avenue
Chestnut Hill, MA 02467
rosen.valchev@bc.edu

1 Introduction

Standard models assume that agents freely and perfectly solve for their respective optimal policy functions. In this view, decision-makers face no limitations in their reasoning and are thus endowed with full knowledge of the *mapping* of beliefs about their economic environment (states, structure, etc.) into their best course of action. Our paper is motivated by a long-standing interest in relaxing this strong assumption of costless reasoning, an interest emerging in economics as early as Simon (1955) and also spanning behavioral and neuro-sciences.[1]

Departing from this extreme view of rationality raises two modeling questions, which constitute the focus of our paper. The first is how to model costly reasoning with a *portable* framework that is applicable across different economic environments, while (i) capturing broad empirical insights on how human reasoning works, and (ii) remaining tractable, flexible, and *parsimonious*. Our methodological contribution here is to build on a growing neuroscience, experimental and computational literature that allows us, as outside analysts, to cast agents' costly reasoning as a Bayesian non-parametric estimation of the unknown optimal policy function. The second question is to establish whether the resulting bounded-rationality mechanism matters for micro and macro behavior. Here, our applied contribution is to use a standard incomplete markets model as a laboratory to show that, even though reasoning errors are i.i.d. and agents have no a-priori behavioral biases, costly reasoning *endogenously* alters ergodic behavior in a systematic and quantitatively promising way.

Methodological contribution. The typical constrained-rational approach to modeling limitations in decision-making is to assume agents face uncertainty about the realization or the law of motion of the relevant state of the world. This uncertainty may reflect a theoretical interest in putting the agent inside the model on a similar footing as an outside econometrician.[2] Imperfect perceptions of the state may also arise from agents' limited attention capacity, as for example in a recent literature inspired by Sims (1998, 2003).[3]

A common feature of these approaches is that, *conditional* on beliefs about the state, the optimal policy *function* mapping those beliefs into the best perceived action is derived under no additional cognitive cost. Our complementary, but distinct, interest is to model the

[1]See Klaes et al. (2005) for a conceptual history of bounded rationality and Conlisk (1996) for an early review on the evidence and challenges in incorporating human cognition as a scarce resource in economics.

[2]For example, Sargent (1993) posits the difficulties faced by boundedly rational agents in computing general equilibrium effects as an estimation problem. Hansen (2007) argues for incorporating in the agents' decision problem the inherent doubts faced by econometricians when estimating probability distributions.

[3]Woodford (2003), Reis (2006), Maćkowiak and Wiederholt (2009), van Nieuwerburgh and Veldkamp (2010), Gabaix (2014), Matějka and McKay (2014) and Stevens (2020) are examples of modeling agents' choice of attention to their individual or aggregate state (see Sims (2010), Wiederholt (2010) and Gabaix (2019) for surveys). Angeletos and Lian (2018), Farhi and Werning (2019), and García-Schmidt and Woodford (2019) model agents' limited capacity to compute the law of motion of aggregate variables.

optimal policy function itself as an unknown object that the agents can estimate via costly reasoning effort. Intuitively, our agents are imperfect problem solvers as they do not have immediate access to the optimal solution to their decision-making problem. For example, to fix ideas and to preview our consumption-saving application, consider the relevant state, y, to be the available cash-on-hand and $c^*(y)$ the unknown optimal consumption function. Our agents therefore face uncertainty about the optimal consumption action in a given period t, $c_t \equiv c^*(y_t)$, even if the current level of the state, y_t, and its law of motion is perfectly known.[4]

We describe costly reasoning as a tractable and realistic framework by appealing to a standard problem of functional estimation in Bayesian statistics. In particular, we model the uncertainty about the unknown optimal policy function, eg. c^*, as a Gaussian Process (GP) distribution over which agents update beliefs.[5] This approach allows us to capture the accumulation of information about the optimal policy function over time in a parsimonious and flexible way, while being consistent with three broad motivating properties that emerge from a large literature on bounded rationality, as follows.

First, at its conceptual core, reasoning is "*fact-free*" learning, or internal reflection that helps the agent get closer to the optimal decision without what an outside observed would register as new objective information.[6] Consistent with this view, we model reasoning as the internal production of noisy and unbiased signals η_t about $c^*(y_t)$, that update prior beliefs over the unknown function c^*.[7] The resulting conditional expectation of $c^*(y_t)$ is the agent's actual time-t action, c_t, as it is the best guess of the optimal action given the state value y_t.

Second, the general idea that human cognition is a scarce resource subject to *cost-benefit tradeoffs* is central to a view, common across the theory, experimental and neuroscience literature, that economic agents are bounded, but "resource-rational".[8] To this end, we let the agents optimally choose the precision of their signals, trading off the reduction of uncertainty about the optimal current-period action $c^*(y_t)$ achieved by the new reasoning signal η_t, and a cognitive cost proportional to the amount of information about the optimal action carried in the reasoning signal, as measured by Shannon mutual information.

[4] Our interest is thus also substantially different from a behavioral literature that introduces a departure from the standard behavior (eg. present focus (Laibson (1997)) or lack of self-control (Gul and Pesendorfer (2004))) and then allows agents to costlessly optimize over their resulting optimal policy function.

[5] Intuitively, a GP distribution models a function as a vector of infinite length, where the vector has a joint Gaussian distribution. Gaussian Processes have been widely used in Bayesian statistics (see Liu et al. (2011)), and in machine learning over unknown functional relationships (see Rasmussen and Williams (2006)).

[6] See Aragones et al. (2005) and Alaoui and Penta (2016) as examples of this view in the theory literature.

[7] The stochastic nature of the reasoning signals implies that our agents exhibit stochastic choice, i.e. even conditioning on the same observed state their actions may differ. See Mosteller and Nogee (1951) and more recently Ballinger and Wilcox (1997) and Hey (2001) for experimental evidence of stochastic choice.

[8] Similar in spirit, Simon (1976) argues for modeling "procedurally rational" agents, who exhibit behavior that is the outcome of limited but appropriate deliberation. See Lieder and Griffiths (2020) for a recent survey on the general concept of cognition as "resource-rational", trading off accuracy and cost.

Third, the literature also often emphasizes that reasoning appears to be *local* in nature, in the sense that in solving their particular problem at hand, people tend to rely more heavily on past experiences and information derived from "similar" situations.[9] Our framework parsimoniously captures this local nature of reasoning by assuming that agents are not confident c^* belongs to a particular parametric family of functions, hence do not extrapolate infinitely away from available information. As a result, while a signal η_t updates beliefs about the function c^* across the whole state space, the signal's informativeness is decaying when updating beliefs about $c^*(y')$ at states y' further away from y_t (recall η_t is centered around $c^*(y_t)$). Thus, in coming up with the best perceived action today, our agents rely more heavily on past reasoning signals derived at state values y_{t-k} that are close to the current state y_t.

Overall, we interpret reasoning "as if" the agent engages in probabilistic inference, a view at the heart of the link between cognitive sciences and machine learning (see Barber (2012)). This interpretation means that we do not literally assume agents engage in statistical functional estimation in everyday life, but rather we use this abstraction as a modeling device, rich enough to capture realistic reasoning properties, but also agnostic enough about further specific deliberation details so to be tractable and portable for us as analysts.[10]

State- and history-dependent reasoning choice. The key qualitative implications of our framework are underpinned by the local nature of information, which leads to state and history dependent residual uncertainty as signals accumulate over time. In particular, for realizations of the state y_t where the beginning-of-period conditional variance over the best course of action $c^*(y_t)$ is relatively high (i.e. few past reasoning signals in that part of the state space), the agent optimally chooses to acquire a more precise current reasoning signal to reduce that uncertainty. In contrast, at state realizations where past deliberation has occurred more often and thus the agent has accumulated substantial information about the optimal action, much further deliberation is not optimal. Near such "familiar" state values the updating weight on the new reasoning signals is endogenously low and the resulting action is primarily driven by the beginning-of-period beliefs about $c^*(y_t)$, leading to behavior that appears to follow past-experience based heuristics in that "familiar" part of the state space. However, our agents do not follow mechanical rules, since they optimally choose precise new signals (which lead to significant beliefs revision on average) if the state y_t moves in an unfamiliar territory, where uncertainty over the optimal action is relatively high.[11]

[9] This similarity-based view of human reasoning is suggested by a recent experimental and neuroscience literature (eg. Bornstein et al. (2017) and Gershman and Daw (2017)), which finds that more familiar contexts make agents more likely to use previous experiences in guiding current decisions.

[10] Conditional on the class of reasoning properties incorporated in our framework, we do not take a further stand on the details of the reasoning mode (eg. search and satisficing (Simon (1955)), or a particular dual-processing structure (Kahneman (2011)).

[11] Nimark and Pitschner (2017) obtain state-dependency in a model where agents decide which information

Applied contribution. We inject our proposed framework of costly reasoning into a standard Aiyagari (1994) general equilibrium incomplete markets model. There is a continuum of ex-ante identical agents facing i.i.d. labor income shocks. Risk-free capital is the only asset available for saving, with a borrowing limit set to zero. The relevant economic state for the consumption decision of agent i is her perfectly observed cash-on-hand $y_{i,t}$ – i.e. the sum of the exogenous income shock and the endogenously accumulated asset. Agents face the same reasoning problem and share an identical time-0 prior over the unknown optimal consumption function c^*. To highlight the role of our costly-reasoning mechanism, we eliminate any ex-ante behavioral biases by centering the initial prior around the true policy function.

The feedback between the state- and history- dependent reasoning choices and the endogenous state (wealth) is crucial in shaping the ergodic properties of our model, since consumption choices affect wealth evolution, which affects the uncertainty facing the agent. In particular, as long as wealth drifts into new and uncertain parts of the state space, an agent's beliefs about the optimal policy function c^* are likely to continue to evolve, as she accumulates additional precise reasoning signals. On the contrary, if the reasoning signal realizations lead to a policy function estimate that establishes stable wealth dynamics, and thus a high likelihood for the agent's state variable to remain within a particular neighborhood of the state space, the evolution of beliefs slows down significantly. As wealth fluctuates in that familiar, low uncertainty region, the agent has little incentive to reason further. Thus, she continues to rely heavily on her past information and policy estimate, which perpetuates the stable dynamics in wealth. This "learning trap" underscores a powerful *selection effect* of *when* agents choose to reason intensely, which in turn endogenously determines which kind of errors in the policy function estimates are likely to be over-represented on average and thus characterize the ergodic steady state of the model.

Hand-to-Mouth status. One particular feature of the ergodic steady state is a frequent and persistent "Hand-to-Mouth" (HtM) status. To see the intuition behind this implication, consider a mass of agents that have initially received reasoning signals that over-estimate the consumption function c^*. Acting under these beliefs, agents draw on their individual assets $y_{i,t}$ and on average drift towards the borrowing constraint. The combination of the borrowing constraint, which acts as a barrier to the agents' downward drift in assets, and the high consumption policy estimate generates stable wealth dynamics. Indeed, for such agents, positive income shocks increase cash-on-hand away from the constraint, but in doing so the state moves back into a familiar region, where they just follow their "business-as-usual", over-consuming behavior. As a result, agents draw down on their assets again, leading wealth

provider to use in acquiring information over the relevant economic state. Relatedly, in Nimark (2014) the assumed information structure is such that signals are more likely to be available about more unusual events.

to fluctuate within this now-familiar region near the borrowing constraint.

High MPCs for rich, unconstrained agents. Consider now a mass of agents that have typically received consumption reasoning signals contaminated with negative noise, leading them to under-estimate c^*. As their assets drift up, these agents eventually enter into an unfamiliar territory of the state space, where they are likely to decide to reason more again. In turn, a fraction of these rich agents are bound to obtain new signals on c^* at their current (high) state $y_{i,t}$ that are contaminated with positive noise. Due to the local nature of information, the resulting beliefs revisions from these new signals change the *slope* of the estimated policy function. In particular, for these agents their internal reflection that the best course of action is "high-consumption" at the current, high wealth values, together with their previously accumulated reasoning signals pointing to "low-consumption" at the corresponding lower past $y_{i,t-k}$, leads to a posterior policy estimate with a *high* local slope. As a result, these rich agents now exhibit a high marginal propensity to consume (MPC).

The logic of the "learning traps" makes this high MPC behavior the norm, rather than an exception, because a consumption policy estimate with a high slope establishes locally stable wealth dynamics. As a result, it is precisely this type of high MPC consumption behavior that is endogenously and thus systematically selected in the ergodic distribution.

Wealth inequality. A related implication of this "business-as-usual" behavior is that even if agents are ex-ante identical, their specific history of reasoning errors eventually results in heterogeneous anchoring points for wealth. At the bottom of the wealth distribution, we have discussed how agents are likely to get stuck as HtM, with near-zero wealth. At the top of the distribution, we point out that wealth may stabilize at high levels for those agents that experience a sequence of reasoning errors that, more repeatedly or more strongly, point to saving a large fraction of their cash-on-hand as their best course of action.

Quantitative implications. In our numerical analysis we follow Aiyagari (1994) for the standard set of parameters. For the reasoning friction, we first aim to impose "model-consistent" priors and thus set the parameters governing the prior uncertainty over $c^*(y)$ equal to what an outside econometrician would estimate if he were to observe simulations from the model. Second, we calibrate the marginal cost of reasoning, as the only remaining degree of freedom, by targeting a particularly challenging fact for the standard model, namely that empirically the bottom 20% in the US wealth distribution have roughly zero net assets.[12]

The costly reasoning (CR) model significantly and parsimoniously improves along three key dimensions upon its full-information (FI) counterpart – the standard model where agents have perfect knowledge of their optimal policy function. First, the CR version produces more

[12] See Krueger et al. (2016) for a reference and a discussion on this challenge for the standard models.

wealth inequality, with the Gini coefficient rising by 50% relative to its FI counterpart. This higher inequality does not arise only from the (targeted) larger fraction of agents with zero assets but also from a significantly higher mass of rich agents. Second, the characteristic learning traps of the CR model imply that HtM status is not only prevalent, but also persistent, in line with the data. In contrast, because in the FI version agents save aggressively when they are close to the constraint, being HtM is both a low probability and transitory status.[13]

Third, the average MPC in the CR model is 0.29, in line with the data, while under FI it is just 0.05.[14] Crucially, while MPCs are naturally high for the HtM agents in both models, we find that in the CR model MPCs are high even for the rich - e.g. the average MPC across the richest 20% of agents is still very high, and equal to 0.15. In contrast, in the FI economy away from the constraint, consumption behavior is similar to the permanent-income-hypothesis, with a corresponding MPC for the rich agents of just 0.04.[15] Therefore, consistent with the empirical literature, the CR model delivers a high average MPC by generating *both* a large mass of constrained agents and high MPCs for the unconstrained agents.

Finally, we use two experiments to highlight the importance of modeling bounded, but "resource-rational" agents. We first discuss a counter-factual where agents have full-information about the optimal policy function c^* but every period make idiosyncratic mistakes ("trembles") in their actions. In this case, because errors are not systematically selected in the ergodic steady state, they tend to wash out and the model ends up resembling its standard FI version. In the second experiment we consider a one-time uncertainty shock that, partly motivated by the current COVID-19 crisis, renders useless the agents' accumulated information about c^*. We find that the average MPC falls substantially on impact, as agents abandon their "business-as-usual" consumption patterns. We then compute how an aggregate income increase has a significant impact on the economy in normal, high MPC times, but is considerably weaker in very uncertain times, showcasing the relevance of modeling behavioral "mistakes" that respond to changes in environment.

Section 2 describes the general framework of costly reasoning. Section 3 introduces the reasoning friction into the Aiyagari (1994) environment and analyzes its qualitative insights. A numerical analysis of the model is discussed in Section 4.

[13] This aggressive build up of wealth also implies that HtM status predicts unusually high future consumption growth in the FI model. However, Aguiar et al. (2020) find no evidence of this prediction in the PSID. The CR model is consistent with this evidence, due to the characteristic persistence of "learning traps" behavior.

[14] See Jappelli and Pistaferri (2010) for a survey on the measurement of MPCs. As detailed by Carroll et al. (2017) the results are varied, with credible estimates of annual MPCs appearing to range from 0.2 to 0.6.

[15] In Kaplan et al. (2014) rich agents may have high MPCs because they are HtM in terms of liquid wealth. Still, in the data, even agents with high-liquid wealth appear to have MPC levels significantly higher than implied by standard models (see Parker (2017), Olafsson and Pagel (2018) and Fagereng et al. (2019)).

2 General framework

We model deliberation as a problem of estimating an unknown optimal policy function c^* that maps the current state y_t to an action c_t. We represent uncertainty over the space of functions using a tractable, yet flexible Bayesian non-parametric approach. For simplicity, we restrict our exposition to the case where the state y is a scalar, hence $c^* : \mathbb{R} \to \mathbb{R}$ is a univariate function. The framework readily generalizes to multivariate settings.

2.1 The Gaussian Process distribution

A Gaussian Process (GP) distribution is the generalization of the Gaussian distribution to infinite-sized collections of real-valued random variables, and it is often used as a prior for Bayesian inference on functions (Liu et al. (2011)).[16] For our purposes, we assume that the agent's prior beliefs over the unknown function c^* are given by a GP distribution

$$c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0), \tag{1}$$

where $\widehat{c}_0 : \mathbb{R} \to \mathbb{R}$ and $\widehat{\sigma}_0 : \mathbb{R}^2 \to \mathbb{R}$. A Gaussian Process distribution has the defining feature that for any arbitrary pair of inputs y and y', the joint distribution of the resulting function values $c^*(y)$ and $c^*(y')$ is Gaussian:

$$\begin{bmatrix} c^*(y) \\ c^*(y') \end{bmatrix} \sim N\left(\begin{bmatrix} \widehat{c}_0(y) \\ \widehat{c}_0(y') \end{bmatrix}, \begin{bmatrix} \widehat{\sigma}_0(y,y) & \widehat{\sigma}_0(y,y') \\ \widehat{\sigma}_0(y,y') & \widehat{\sigma}_0(y',y') \end{bmatrix} \right),$$

where $\widehat{c}_0$ is known as the "mean function" and specifies the prior mean of $c^*(y)$ for any y,

$$\widehat{c}_0(y) = \mathbb{E}(c^*(y)),$$

and the "covariance function" $\widehat{\sigma}_0(y, y')$ specifies the unconditional covariance between the values of the function c^* at any pair of inputs y and y':

$$\widehat{\sigma}_0(y, y') = \mathbb{E}\left((c^*(y) - \widehat{c}_0(y))(c^*(y') - \widehat{c}_0(y'))\right).$$

[16] Ilut et al. (2020) use a Gaussian Process setup to model firms learning about the unknown demand function, leading to state-dependent uncertainty about their relevant state. However, there firms continue to use the mapping from that imperfect information about the state to the optimal pricing action derived under no cognition constraints. Similar recent approaches based on non-parametric learning about the state, but assumed knowledge of the mapping to the optimal actions, include Dew-Becker and Nathanson (2019) and Kozlowski et al. (2020).

A computational interpretation

Modeling the uncertainty over c^* as a GP distribution has a formal link to the common computational approach of expressing the policy function as a linear combination of a set of basis functions ϕ_k indexed by k. Assuming the set of $\{\phi_k\}$ forms a complete basis for the class of functions c^* belongs to, we can obtain an arbitrarily good approximation by including a large enough number of basis functions so that (with slight abuse of notation) we can write

$$c^*(y) = \sum_{k=1}^{N} \theta_k \phi_k(y). \tag{2}$$

Projecting on a set of basis transforms the problem of solving for the function c^* into finding the optimal weights θ_k – they are the "unknown" variables that one needs to solve for.

If the unknown θ_k's have a joint-Gaussian distribution, then the function c^* itself has a Gaussian Process distribution. Of particular interest to us is the case where the priors over the different θ_k are independent Normal distributions, which is formally treated in Lemma 1.

Lemma 1. *If $\theta_k \overset{iid}{\sim} N(\mu_k, \sigma_c^2)$, equation (2) implies that $c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$, with*

$$\widehat{c}_0(y) = \sum_{k=1}^{N} \mu_k \phi_k(y); \qquad \widehat{\sigma}_0(y, y') = \sigma_c^2 \sum_{k=1}^{N} \phi_k(y) \phi_k(y')$$

Proof. Details are in Appendix A. □

Thus, the mean and variance functions that characterize the GP prior in equation (1) can be interpreted as arising from specific choices for μ_k and the set of basis functions $\{\phi_k\}$.[17]

The formal link between the underlying source of uncertainty over θ_k and the implied uncertainty over c^* is conceptually useful when thinking about the fundamental source of uncertainty. However, we follow the Bayesian statistics literature (as detailed in Rasmussen and Williams (2006)) and treat the θ_k's as "nuisance" parameters, integrate them out as in Lemma 1 and work directly with the resulting GP distribution over the space of functions.

The reason is the same as in the standard problem faced by statisticians – ultimately the agent (and we as modelers) is interested in estimating the function c^* itself, and the coefficients θ_k are just an intermediate step. To this end it is both more tractable and transparent to work directly with the implied GP distribution of the function c^*. Therefore, throughout we focus on specifying priors and signal structures directly over c^* itself.

[17]The conceptual framework is thus quite general, and encompasses a number of different solution techniques, from Chebyshev polynomials to neural networks, as different choices of bases sets $\{\phi_k\}$, all of which can be represented as special cases of Gaussian Processes (for formal details see Rasmussen and Williams (2006)).

2.2 Reasoning as fact-free learning

The agent does not simply act on her prior beliefs about the optimal action $c^*(y_t)$, but can expend costly cognitive resources to obtain a better handle of the unknown optimal action. This is modeled by giving the agent access to unbiased signals about the actual optimal action at the current state y_t,

$$\eta_t = c^*(y_t) + \varepsilon_t,$$

where $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2_{\eta,t})$, and allowing the agent to choose the precision of those signals, i.e. $\sigma^2_{\eta,t}$.

The choice $\sigma^2_{\eta,t}$ reflects the agent's intensity of deliberation – the more time and effort spent on thinking about the optimal behavior, the more precise is the resulting signal, and thus the more accurate are the posterior beliefs. The subjective reasoning signals embody the idea of reasoning as "fact-free" learning, one of our motivating properties exposed in the introduction. Moreover, the stochastic nature of η_t implies that our agents exhibit stochastic choice, i.e. even conditioning on the same observed state their actions may differ.

The reasoning signals update the agent's beliefs about the unknown function c^*, with the conditional distribution of beliefs following a Kalman-filter like recursion.

Lemma 2. *Given the time-0 prior belief* $c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$, *the time-t conditional beliefs are* $c^* \big| \{\eta^t, y^t\} \sim \mathcal{GP}(\widehat{c}_t, \widehat{\sigma}_t)$ *with moments evolving according to the recursive expressions*

$$\widehat{c}_t(y) = \widehat{c}_{t-1}(y) + \frac{\widehat{\sigma}_{t-1}(y, y_t)}{\widehat{\sigma}^2_{t-1}(y_t) + \sigma^2_{\eta,t}}(\eta_t - \widehat{c}_{t-1}(y_t)), \tag{3}$$

$$\widehat{\sigma}_t(y, y') = \widehat{\sigma}_{t-1}(y, y') - \frac{\widehat{\sigma}_{t-1}(y, y_t)\widehat{\sigma}_{t-1}(y', y_t)}{\widehat{\sigma}^2_{t-1}(y_t) + \sigma^2_{\eta,t}} \tag{4}$$

where $\widehat{c}_t(y) \equiv E_t(c^*(y)|\eta^t)$ *and* $\widehat{\sigma}_t(y, y') \equiv \mathrm{Cov}(c^*(y), c^*(y')|\eta^t)$ *are the posterior mean and covariance functions. Lastly,* $\widehat{\sigma}^2_t(y) \equiv \widehat{\sigma}_t(y, y)$ *denotes the posterior variance at a given* y.

Proof. Details are in Appendix A. □

The framework represents an abstract description of deliberation about the policy function c^* as a Bayesian inference from "fact-free" reasoning signals by using the tractability and flexibility of a GP distribution. Given priors $\widehat{c}_0, \widehat{\sigma}_0$ and signal-noise variance $\sigma^2_{\eta,t}$, the recursively-determined mean and covariance functions in formulas (3) and (4) fully characterize the posterior distribution of beliefs about c^*. Next we discuss our approach to impose structure on $\widehat{c}_0, \widehat{\sigma}_0$ and $\sigma^2_{\eta,t}$ so as to capture the broad motivating properties of reasoning as (i) local in nature, and (ii) an outcome of a cost-benefit tradeoff, as motivated in the introduction.

We first note that any difference between the unknown, to the agent, optimal policy function c^* and the prior mean $\widehat{c}_0$ represents an ex-ante bias in beliefs. Our motivating properties of reasoning do not impose structure on this bias and so it represents a degree of freedom for us as analysts. In that sense, selecting the prior mean $\widehat{c}_0$ is the part of our costly reasoning model where "the wilderness of bounded rationality" may arise.

While our conceptual framework allows for any prior mean, we are generally interested in eliminating the role of this degree of freedom. To this end and in the spirit of Rational Expectations, throughout the rest of the analysis we assume that the prior mean coincides with the truth and set

$$\widehat{c}_0(y) = c^*(y)$$

for all y. Nevertheless, even though prior beliefs are centered at the actual optimal policy, the agent still faces uncertainty about $c^*(y)$, as encoded by the covariance function of the prior $\widehat{\sigma}_0$. Intuitively, the prior mean belief is correct, but the agent does not know that for certain.

We now turn to the objects that characterize the structure of this uncertainty, and hence are key to the dynamics of beliefs and their long-run behavior in our costly reasoning model. In particular, we describe a restriction on $\widehat{\sigma}_0$ designed to capture the local nature of reasoning and then we model a cost-benefit tradeoff in the choice of $\sigma^2_{\eta,t}$.

2.3 Local nature of reasoning

The prior covariance function $\widehat{\sigma}_0$ determines how new information about c^* is interpreted and combined with prior information to form posterior beliefs. In particular, it determines the extent to which the agent is willing to extrapolate the information contained in η_t to update beliefs about the optimal action $c^*(y)$ at state realizations different than y_t.

The local nature of reasoning emerges from imposing the restriction that $\widehat{\sigma}_0(y, y')$ is declining in the distance $||y - y'||$. Interestingly, while this restriction helps us as economists to generate the empirically appealing local nature of reasoning, it is also widely used in Bayesian statistics purely for estimation and computational reasons (see Rasmussen and Williams (2006)). Such "stationary" covariance functions are flexible enough to ensure asymptotic consistency for a wide variety of functional forms of the unknown c^*, and thus impose few ex-ante structural restrictions (see Rasmussen and Williams (2006) for details).

Specifically, drawing on this literature we assume the covariance function belongs to the popular squared exponential family, i.e.

$$\widehat{\sigma}_0(y, y') = \sigma^2_c \exp(-\psi(y - y')^2).$$

This is a commonly-used prior in Bayesian statistics because it offers an efficient trade-off between flexibility and degrees of freedom. It also connects to the computational interpretation of our framework, as using this covariance function is equivalent to assuming that the agent approximates c^* with Gaussian radial basis functions – i.e. assuming the functions ϕ_k in equation (2) are Gaussian kernels. In fact, this is a popular choice of basis functions in the machine learning literature, because this class of "local" basis function allows a modeler to use "solve-as-you-go" methods, where the solution is computed recursively over time. Those methods offer an efficient trade-off between time and accuracy, since the goal in each computational step is to only approximate the optimal policy function in a neighborhood of the current state (i.e. the most relevant region for the current objective), with knowledge about the global behavior of the function gradually building over time (see Bertsekas (2019)).

The squared exponential covariance function is characterized by just two parameters. First, σ_c^2 controls the prior uncertainty about the value of $c^*(y)$ at any given point y. We can think of this as the "quantity" of uncertainty agents face around $\widehat{c}_0(y)$. Second, ψ controls the extent to which information about the value of the function at a point y is informative about its value at a different point y'. Formally, ψ controls the prior probability that the unknown function c^* has a shape that deviates from the prior mean $\widehat{c}_0$, as per Lemma 3.

Lemma 3. *Let* $\alpha_t(y) \equiv \frac{\widehat{\sigma}_{t-1}(y,y_t)}{\widehat{\sigma}^2_{t-1}(y_t)+\sigma^2_{\eta,t}}$ *be the weight put on* η_t *in the time-t estimate* $\widehat{c}_t(y)$*.*

- *If* $\psi = 0$*, then* $\alpha_t(y)$ *is just a constant – i.e.* $\alpha_t(y) = \alpha_t$ *for all* $y \in \mathbb{R}$*, and thus*

$$\widehat{c}_t(y) = \widehat{c}_0(y) + \sum_{k=1}^{t} \alpha_k \prod_{j=k+1}^{t} (1-\alpha_j) u_k$$

where $u_k = \eta_k - \widehat{c}_0(y_k)$ *is the deviation of signal* η_k *from the time-0 prior mean belief.*

- *If* $\psi > 0$*, then the informativeness of the signal* η_t *is state-dependent –* $\frac{\partial \alpha_t(y)}{\partial y} \neq 0$ *– and hence the shape of the time-t estimate* $\widehat{c}_t$ *differs from the time-0 prior, i.e.:*

$$\widehat{c}_t(y) = \widehat{c}_0(y) + \sum_{k=1}^{t} \alpha_k(y) \prod_{j=k+1}^{t} (1-\alpha_j(y)) u_k$$

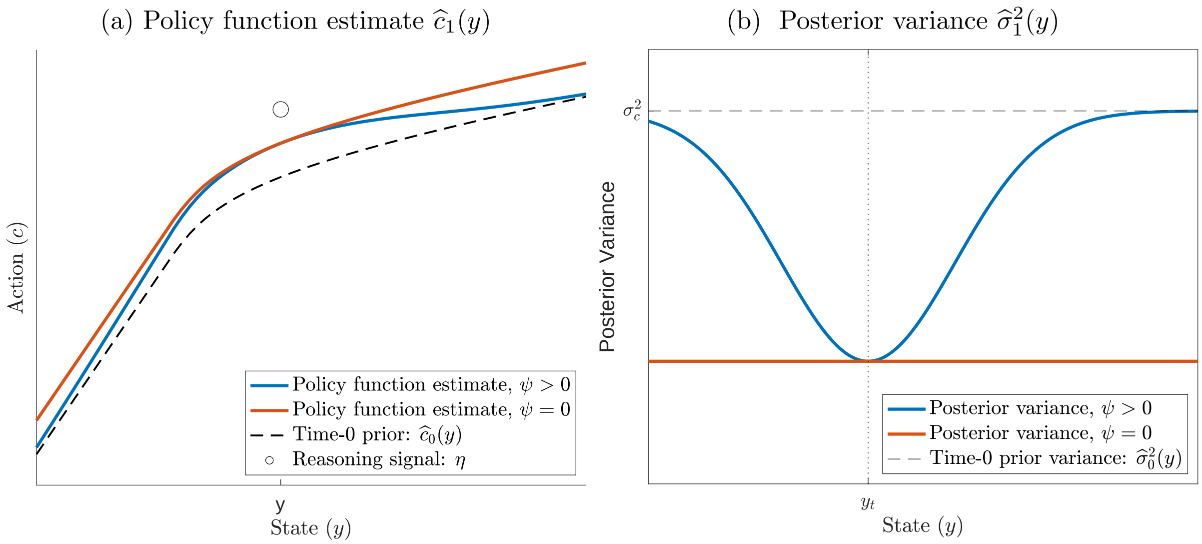
$$\frac{\partial\left(\widehat{c}_t(y) - \widehat{c}_0(y)\right)}{\partial y} \neq 0$$

The effect of the information in η_t *is also local to* y_t*, since* $\lim_{||y-y_t||\to\infty} \alpha_t(y) = 0$

Proof. Details are in Appendix A. □

Intuitively, in the limiting case of $\psi = 0$ the prior beliefs only entertain functions c^* that are generated by a constant shift of the prior mean $\widehat{c}_0$ – hence there is no uncertainty about the *shape* of $c^*(y)$, only its level. On the other hand, $\psi > 0$ allows for shapes that are

Figure 1: Conditional beliefs $\psi = 0$ vs $\psi > 0$



different from the prior mean $\widehat{c}_0$. In particular, as ψ increases, information about the optimal action at some value of the state y is decreasingly useful for inferring the optimal action at a different value y', and this is because the agent is less willing to extrapolate away from y based on the shape of the prior mean $\widehat{c}_0$. Overall, when $\psi > 0$ the agent's estimate of the optimal action at y, $c^*(y)$, puts more weight on reasoning signals obtained at states y' that are close to y itself, consistent with the similarity-based view of reasoning described in the introduction.

To illustrate the basic result, in Figure 1 we plot an example update of beliefs conditional on observing a single signal η centered at a value of the state y (e.g. we can interpret this as being period one). We show the resulting updated mean and variance functions $\widehat{c}_1$ and $\widehat{\sigma}_1^2$ for both $\psi = 0$ and $\psi > 0$ as a function of the state y. While in this example there is only one signal to update with, it changes beliefs about the policy function throughout the state space. In the case of $\psi = 0$, it is clear that the resulting conditional estimate $\widehat{c}_1$ is simply a level shift of $\widehat{c}_0$, and naturally the posterior variance of the unknown function c^* is reduced equally for all values of the state – there is no uncertainty about the *shape* of c^*, just its level.

On the other hand, in the case of $\psi > 0$ the signal η only exerts local influence on the estimate $\widehat{c}_1$, with the agent updating her beliefs that the function is higher and more curved in the neighborhood of the signal. The local effect of information is also reflected in the posterior variance $\widehat{\sigma}_1^2(y)$ – the signal η primarily reduces uncertainty about the optimal action at y itself, $c^*(y)$, and is less informative at other state values $y' \neq y$. This showcases how in our benchmark case of $\psi > 0$, uncertainty is state and history dependent.

2.4 Cost-benefit tradeoff of reasoning

In any given period t, the agent faces an intuitive cost-benefit tradeoff in choosing the reasoning noise variance $\sigma^2_{\eta,t}$. The benefit comes from a standard tracking problem where, given the state y_t, the agent wants to choose her action, c_t, to be as close as possible to the action implied by the unknown optimal policy function – i.e. $c^*(y_t)$. Thus, she minimizes the expected squared deviations

$$\min_{c_t,\sigma^2_{\eta,t}} \mathbb{E}_t(c_t - c^*(y_t))^2, \tag{5}$$

where $\mathbb{E}_t$ denotes the conditional expectation over the distribution of c^* with moments recursively determined by equations (3) and (4). The solution to the tracking problem is to act according to the conditional expectation of the unknown optimal action, i.e. to set $c_t = \widehat{c}_t(y_t)$. Substituting in this choice of c_t in equation (5), the agent's objective simplifies to choosing $\sigma^2_{\eta,t}$ to reduce the resulting posterior variance $\widehat{\sigma}^2_t(y_t)$.

A tradeoff emerges because the cognitive effort needed to produce precise reasoning signals is costly. We model this as a cost on the amount of information about the optimal action $c^*(y_t)$ carried in the signal η_t. We measure information flow with the reduction in entropy, i.e. Shannon mutual information, about $c^*(y_t)$, defined as

$$I(c^*(y_t); \eta_t|\eta^{t-1}) = H(c^*(y_t)|\eta^{t-1}) - H(c^*(y_t)|\eta_t, \eta^{t-1}), \tag{6}$$

where $H(X)$ denotes the entropy of a random variable X, and is the standard measure of uncertainty in information theory.[18] Thus, equation (6) measures the reduction in uncertainty about the unknown optimal action $c^*(y_t)$ from seeing the new signal η_t, given the history of past deliberation signals η^{t-1}. For analytical tractability, we further assume that the agent faces a simple reasoning cost linear in $I(c^*(y_t); \eta_t|\eta^{t-1})$.

Given prior information entering the current period, summarized by $\widehat{c}_{t-1}$ and $\widehat{\sigma}_{t-1}$, and the state y_t, by equation (4) a choice over $\sigma^2_{\eta,t}$ is equivalent to selecting a posterior variance $\widehat{\sigma}^2_t(y_t)$. Thus, the cost-benefit tradeoff of reasoning can be cast as the information problem

$$\min_{\widehat{\sigma}^2_t(y_t)} \widehat{\sigma}^2_t(y_t) + \kappa \ln\left(\frac{\widehat{\sigma}^2_{t-1}(y_t)}{\widehat{\sigma}^2_t(y_t)}\right). \tag{7}$$

s.t.

$$\widehat{\sigma}^2_t(y_t) \leq \widehat{\sigma}^2_{t-1}(y_t),$$

The first component is the benefit of reasoning, in the form of a lower dispersion of the

[18] Following Sims (2003) a large literature has studied the choice properties of information costs based on the Shannon mutual information. See for example Caplin et al. (2016) and Woodford (2014).

action $c_t = \mathbb{E}_t(c^*(y_t))$ around the unknown optimal action $c^*(y_t)$, which is equal to $\widehat{\sigma}_t^2(y_t)$. The second represents the cost of reasoning, where we use the analytical expression for the mutual information in a Gaussian framework as equal to one half of the log-ratio of prior and posterior variances. The parameter κ controls the marginal cost of a unit of information. For example, κ will be higher for individuals with a higher cost of deliberation – either because they have a higher opportunity cost of time or because their particular deliberation process takes longer to achieve a given improvement in precision. In addition, κ would also be higher if the economic environment facing the agent is more complex, and thus the optimal action is objectively harder to figure out – for example solving an objectively difficult math problem.

Lastly, the minimization in (7) is subject to the "no forgetting constraint" $\widehat{\sigma}_t^2(y_t) \leq \widehat{\sigma}_{t-1}^2(y_t)$ which ensures that the chosen variance of the noise in the signal, $\sigma_{\eta,t}^2$, is non-negative. Otherwise, the agent can gain utility by "forgetting" some of her prior information.

2.5 Optimal deliberation choice

The optimal deliberation choice that solves equation (7) is given by

$$\widehat{\sigma}_t^{*2}(y_t) = \min\left[\kappa, \widehat{\sigma}_{t-1}^2(y_t)\right],$$

meaning that the optimal target level for the posterior variance that equates the marginal benefit and cost of reasoning is simply κ. Intuitively, the desired precision in actions (and thus deliberation effort) is larger when the deliberation cost κ is lower. The *min* function enforces the no-forgetting constraint – if the agent's beginning-of-period conditional variance over the optimal action at y_t is lower than the optimal target κ, then she does not acquire any further information and the posterior variance at time t remains equal to $\widehat{\sigma}_{t-1}^2(y_t)$. This leads to the following optimal behavior.

Proposition 1. *The optimal signal noise variance is given by*

$$\sigma_{\eta,t}^{*2} = \begin{cases} \frac{\kappa \widehat{\sigma}_{t-1}^2(y_t)}{\widehat{\sigma}_{t-1}^2(y_t) - \kappa} & , \text{ if } \widehat{\sigma}_{t-1}^2(y_t) \geq \kappa \\ \infty & , \text{ if } \widehat{\sigma}_{t-1}^2(y_t) < \kappa \end{cases} \tag{8}$$

and this in turn implies the time-t action

$$c_t = \widehat{c}_t(y_t) = \widehat{c}_{t-1}(y_t) + \alpha_t^*(y_t)(c^*(y_t) + \varepsilon_t - \widehat{c}_{t-1}(y_t)), \tag{9}$$

where the optimal weight put on the new reasoning signal, $\alpha_t^*(y_t)$ depends on the current state

y_t and the history $\{y^{t-1}, \sigma_\eta^{t-1}\}$ of past signals' location and precision:

$$\alpha_t^*(y_t) \equiv \frac{\widehat{\sigma}_{t-1}^2(y_t)}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,t}^{*2}} = \max\left[1 - \frac{\kappa}{\widehat{\sigma}_{t-1}^2(y_t)}, 0\right]. \tag{10}$$

Proof. Details are in Appendix A. □

Thus, since posterior uncertainty $\widehat{\sigma}_t^2(y)$ is state and history dependent (when $\psi > 0$), both the optimal reasoning choice, in the form of signal-noise variance $\sigma_{\eta,t}^2$, and the effective action c_t are also state- and history-dependent.

Put together, the qualitative features of our general reasoning framework offer a narrative that connects to our basic motivation in the introduction. In particular, similar to us modelers, economic agents do not have costless cognitive access to the the optimal policy function $c^*(y)$. Agents invest costly cognitive effort into reasoning, as a form of fact-free learning, to reduce uncertainty about the best course of action. Reasoning leads to rich heterogeneity in beliefs, since even agents that are otherwise ex-ante identical end up with different views about the optimal conditional action.

The key qualitative implications of our framework are driven by the endogenous state and history dependence of actions and beliefs. First, for state realizations y_t where the precision of initial beliefs is relatively far from its target (high $\widehat{\sigma}_{t-1}^2(y_t)$), the agent chooses to acquire a more precise current signal and hence puts a bigger weight on it in the resulting action $\widehat{c}_t(y_t)$ (high $\alpha_t^*(y_t)$). In contrast, for state realizations close to the position of previous signals η_{t-k}, the precision of initial beliefs is high (low $\widehat{\sigma}_{t-1}^2(y_t)$) and the agent finds it optimal to not acquire much additional information. At such "familiar" state values the optimal $\alpha_t^*(y_t)$ is relatively small, and the resulting action will be primarily driven by the prior beliefs $\widehat{c}_{t-1}(y_t)$. This can give rise to behavior that appears to follow past-experience based heuristics in the "familiar" part of the state space, where prior deliberation has reduced uncertainty about the optimal action sufficiently so that $\alpha_t^*(y_t)$ is small. However, agents do not follow mechanical rules, since beliefs are likely to be revised if the state y_t moves into unfamiliar territory, where uncertainty over the optimal action is relatively high.

Second, while $\alpha_t^*(y_t)$ is a deterministic function of the state y_t and history $\{y^{t-1}, \sigma_\eta^{t-1}\}$, the reasoning signal η_t and the subsequent effective action is stochastic, even conditional on y^t. Therefore, even if agents are ex-ante identical, i.e. they face the same reasoning problem, states y^t and enter period t with the same prior beliefs, they take different actions $\widehat{c}_t(y_t)$.

3 A consumption-savings model with costly reasoning

The characteristic state and history dependence of the general framework suggests that it is particularly interesting to consider economic settings where the state y_t is endogenous. To this end, in this section we study the implications of our costly reasoning model of behavior in the context of an otherwise standard consumption-savings problem. This is a setting that features an endogenous state (wealth), and the feedback between states, reasoning and the resulting actions is crucial in understanding the ergodic properties of the model.

We consider a standard Aiyagari (1994) economy populated by a continuum of ex-ante identical agents that share the same preferences. Agents are indexed by i and derive non-satiable and concave felicity $u(c_{i,t})$ from the consumption of a numeraire good. Each agent inelastically supplies her stochastic endowment of labor $s_{i,t}$ at a constant wage w. The idiosyncratic income shocks $s_{i,t}$, are iid across time and agents and drawn from a time-invariant distribution $\mathcal{S}$ with a mean of one. The agents' ability to reduce the consumption exposure to this risk is limited because there is only one asset, in the form of a homogeneous physical capital that earns a constant rental rate $\widetilde{r}$. The asset holdings of the agent evolve as

$$a_{i,t} = (1-\delta)a_{i,t-1} + x_{i,t}$$

where $a_{i,t-1}$ is the amount of capital held by agent i at the end of period $t-1$, $\delta \in (0,1)$ is a depreciation rate and x_t is gross investment. The resulting budget constraint is

$$c_{i,t} + x_{i,t} = (1+\widetilde{r})a_{i,t-1} + ws_{i,t}.$$

The capital accumulation and the budget constraint can be combined to obtain

$$c_{i,t} + a_{i,t} = (1+r)a_{i,t-1} + ws_{i,t},$$

where $r \equiv \widetilde{r} - \delta$. As in standard models of incomplete markets, the optimal asset choice is subject to a borrowing constraint $\underline{a} \geq 0$ such that

$$a_{i,t} \geq -\underline{a}.$$

The aggregate production function is also standard - it takes as input the average capital $K = \int a_i di$ and employment $H = \int s_{i,t} di$, and produces $K^\alpha H^{1-\alpha}$, with $\alpha \in (0,1)$. The role of this side of the economy is to determine the rental rate and the wage from the static first-order conditions $\widetilde{r} = \alpha K^{\alpha-1}H^{1-\alpha}$ and $w = (1-\alpha)K^\alpha H^{-\alpha}$, respectively. Given the assumed inelastically supplied labor and i.i.d. labor supply shocks s_{it}, we have $H = 1$.

We organize the rest of this section as follows. First, we introduce the agents' decision problem and then define the stationary equilibrium Second, we illustrate the basic mechanics of the evolution of beliefs and the feedback between information and actions through a simple example. Third, in section 4, we describe the properties of the general equilibrium stationary distribution in an illustrative numerical simulation.

3.1 Decision problem

The consumption-savings problem is a specific example of the general recursive dynamic problem introduced in Section 2. Our assumption of i.i.d. exogenous income shocks $s_{i,t}$ means that the sufficient state for the agent's decision is the available "cash-on-hand" defined as

$$y_{i,t} \equiv (1+r)a_{i,t-1} + ws_{i,t}.$$

The key property here is that the future state $y_{i,t+1}$ is determined partly endogenously from the current choice of consumption $c_{i,t}$, as well as by the random realization of $s_{i,t+1}$.

Each agent is interested in solving the same problem

$$V(y_{i,t}) = \max_{c_{i,t},a_{i,t}} u(c_{i,t}) + \beta E_t V(y_{i,t+1}), \tag{11}$$

subject to the budget constraint $a_{i,t} + c_{i,t} = y_{i,t}$ and the borrowing limit $a_{i,t} \geq -\underline{a}$.

Reasoning

The consumption policy function that solves the problem in (11) can be written as

$$\widetilde{c}^*(y_{i,t}) = \min(y_{i,t} + \underline{a}, c^*(y_{i,t})), \tag{12}$$

where the kink at $y_{i,t} + \underline{a}$ arises from the borrowing limit. The policy $c^*(y_{i,t})$ gives the optimal action taking into account future borrowing constraints, but ignoring today's constraint.

The agents in our economy do not have free cognitive access to the full-information policy function c^*. However, they do perfectly observe $y_{i,t}$ and $\underline{a}$, hence the only uncertainty about optimal consumption is in the second component of equation (12) - the current-period borrowing constraint itself is a technological restriction that the agent can directly observe. Applying the framework laid out in Section 2, the agents are uncertain about c^* and obtain costly reasoning signals about the unknown optimal action $c^*(y_t)$

$$\eta_{i,t} = c^*(y_{i,t}) + \varepsilon_{i,t}, \;\; \varepsilon_{i,t} \sim N(0, \sigma_{\eta,i,t}). \tag{13}$$

Agents have the common time-0 prior that $c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$, which as discussed earlier is centered at the truth, i.e. $\widehat{c}_0 = c^*$, and has a squared exponential covariance function $\widehat{\sigma}_0$.

An agent i at time t makes two choices: reasoning intensity, by selecting the desired variance $\sigma^2_{\eta,i,t}$, and consumption choice $c_{i,t}$. In choosing the precision of the reasoning signal $\eta_{i,t}$, the agent optimally trades off its cost and benefit, as described in detail in Section 2, and optimally selects a signal precision that generates the posterior variance

$$\widehat{\sigma}^2_{i,t}(y_{i,t}) = \min\left[\kappa, \widehat{\sigma}^2_{i,t-1}(y_{i,t})\right]. \tag{14}$$

By Proposition 1, the optimal choice of $\sigma^2_{\eta,i,t}$, together with beginning-of-period beliefs $\widehat{c}_{i,t-1}(y_{i,t})$ and $\widehat{\sigma}_{i,t-1}(y_{i,t})$, lead to the following conditional expectation of $c^*(y_{i,t})$:

$$\widehat{c}_{i,t}(y_{i,t}) = \widehat{c}_{i,t-1}(y_{i,t-1}) + \alpha_{i,t}(y_{i,t})(c^*(y_{i,t}) + \varepsilon_{i,t} - \widehat{c}_{i,t-1}(y_{i,t-1})),$$

where the optimal weight put on the current reasoning signal $\eta_{i,t}$ can be expressed as

$$\alpha_{i,t}(y_{i,t}) = \max\left[1 - \kappa/\widehat{\sigma}^2_{i,t-1}(y_{i,t}), 0\right]. \tag{15}$$

Lastly, taking into account the borrowing limit, the agent sets the consumption level [19]

$$c_{i,t} = \min(y_{i,t} + \underline{a}, \widehat{c}_{i,t}(y_{i,t})).$$

Heterogeneity and equilibrium

Agents in our economy are heterogeneous for two reasons. The first is the idiosyncratic income shocks $s_{i,t}$, which are a common feature of standard models too. The second reason is that agents obtain stochastic histories of reasoning errors $\varepsilon_{i,t}$, leading to different information sets $\Omega_{i,t}$ and perceived optimal decision rules $\widehat{c}_{i,t}(y)$. Therefore, the distribution of agent types in the costly reasoning model is richer than in the standard full information Aiyagari (1994) model. In particular, an agent's type at time t is characterized by the following: (i) prior conditional variance function $\widehat{\sigma}^2_{i,t-1}$; (ii) prior conditional mean function $\widehat{c}_{i,t-1}$, (iii) observed cash on hand ($y_{i,t} = (1+r)a_{i,t-1} + ws_{i,t}$) and (iv) current period reasoning error $\varepsilon_{i,t}$. We denote the set of objects that determine an agent state at time t as $\tau_{i,t} \equiv (\widehat{\sigma}_{i,t-1}, \widehat{c}_{i,t-1}, y_{i,t}, \varepsilon_{i,t})$,

The type of heterogeneity in actions produced by our costly reasoning model is novel compared to a full-information model, where the state $\tau_{i,t}$ is simply $y_{i,t}$, since there agents act under the same policy function $c^*(y)$. Moreover, we will show that the beliefs-driven

[19] The Gaussian noise in signals suggests that extreme signals can lead to a negative estimate: $\widehat{c}_{i,t}(y) < 0$. In the numerical implementation we prevent this by imposing another constraint that $c_{i,t} > 0$. In practice, we find that this is not a problem at our calibration.

heterogeneity of our model matters in the aggregate, in the sense that it systematically changes properties of the average behavior and outcomes in the long-run as compared to the full-information model. To do so, we denote by $\lambda_t(\tau)$ the time-t probability distribution over the agent types $\tau_{i,t}$ and note that the constraint optimal behavior of the agents induces a law of motion for $\lambda_t(\tau)$. We are interested in characterizing the properties of this distribution at the stationary equilibrium, which we define below.

Definition 1. *A stationary equilibrium is a full-information policy function $c^*(y)$, a probability distribution $\lambda(\tau)$ and positive real numbers (K, r, w), such that*

(1) prices (r, w) satisfy

$$r = \alpha K^{\alpha-1} - \delta; \;\; w = (1-\alpha)K^{\alpha}$$

(2) the policy function $c^*(y)$ solves the full-information problem in equation (11)

(3) the reasoning choice $\sigma^2_{\eta,i,t}$ and the consumption choice $c_{i,t}$ satisfy Proposition 1, while conditional beliefs $\widehat{c}_{i,t}(y)$ and $\widehat{\sigma}_{i,t}(y, y')$ follow Lemma 2.

(4) given $y_{i,t}$ and a consumption choice $c_{i,t}$, cash on hand evolves as $y_{i,t+1} = (1 + r)(y_{i,t} - c_{i,t}) + ws_{i,t+1}$, where $s_{i,t+1}$ is an iid draw from the time-invariant distribution $\mathcal{S}$.

(5) the distribution $\lambda(\tau)$ is time-invariant, with a law of motion induced by (1)-(4).

(6) aggregate capital equals the average of the households' asset decisions

$$K = \int_{\tau} [y - c(\tau)]\, d\lambda(\tau)$$

3.2 Evolution of Beliefs

The key force behind the stationary equilibrium is the feedback between reasoning and wealth dynamics. Noisy consumption policy function estimates affect the endogenous evolution of wealth, while movements in wealth cause shifts in the uncertainty facing the agents, and in turn this generates different reasoning choices, both over time and across agents.

This feedback is crucial in shaping the ergodic properties of $\lambda(\tau)$ and is underpinned by a powerful selection force in regards to *when* agents choose to reason more or less intensely, which determines the kind of estimation errors that persist through time. In particular, learning and beliefs updating dramatically slows down once an agent accumulates a sequence of reasoning signals that imply wealth is locally stable in the neighborhood of those same signals' positions. In this case, the state $y_{i,t}$ cycles through familiar states at which uncertainty is already low (due to past learning), hence the agent perceives little need for further costly reasoning, and thus chooses to acquire little new information. We call this situation a "learning trap", and it characterizes the stochastic steady state at the individual level, since agents' beliefs keep trending (as described below), unless the agent settles into such a stable, familiar region. As we will see, a necessary condition for stable wealth dynamics is a relatively steep estimate $\widehat{c}_{i,t}$, hence a prominent feature of the stochastic steady state of the model is a high marginal propensity to consume (MPC) across the wealth distribution.

We illustrate the basic mechanism by following the evolution of beliefs of two agents that are perfectly identical except for the realization of their very first reasoning signal errors, $\varepsilon_{i,1}$. Otherwise, the agents have the same initial wealth, $a_{i,0} = a_0$, receive the same sequence of reasoning signal innovations after time-1 ($\varepsilon_{i,t} = \varepsilon_t$ for $t > 1$), and same income shocks, $s_{i,t} = s_t$, $\forall t \geq 1$. We use this example to showcase two results. First, how the interaction between the endogenous state and reasoning errors leads to a particular kind of stable and systematically biased beliefs. Second, how the differing initial reasoning errors can lead to different long-run economic outcomes, even though the agents are otherwise ex-ante identical.

Period 1: ex-post heterogeneous beliefs

At the beginning of time-1 agents are identical as they (i) face the same cash on hand $y_{i,1} = y_1 = (1+r)a_0 + ws_1$ and (ii) have the same prior beliefs centered around $c^*(y)$. As a result, both agents choose the same reasoning effort, i.e. the same $\sigma^2_{\eta,i,1}$ which by equation (14) is such that the posterior posterior variance $\hat{\sigma}^2_{i,1}(y_1) = \kappa$. Thus, the reasoning signals the agents draw have the identical distribution $\eta_{i,1} = c^*(y_1) + \varepsilon_{i,1}$, $\varepsilon_{i1} \sim N(0, \kappa/\alpha^*_1)$. Here $\alpha^*_1 \equiv \alpha_1(y_1) = 1 - \frac{\kappa}{\sigma^2_c}$ is the resulting signal-to-noise ratio, which is the same for both agents, and does not depend on the value of y_1 since $\widehat{\sigma}^2_0(y) = \sigma^2_c$ for all y.

Agents are ex-ante identical and choose the same reasoning strategy, but make different consumption choices due to the idiosyncratic reasoning errors, leading to the effective action

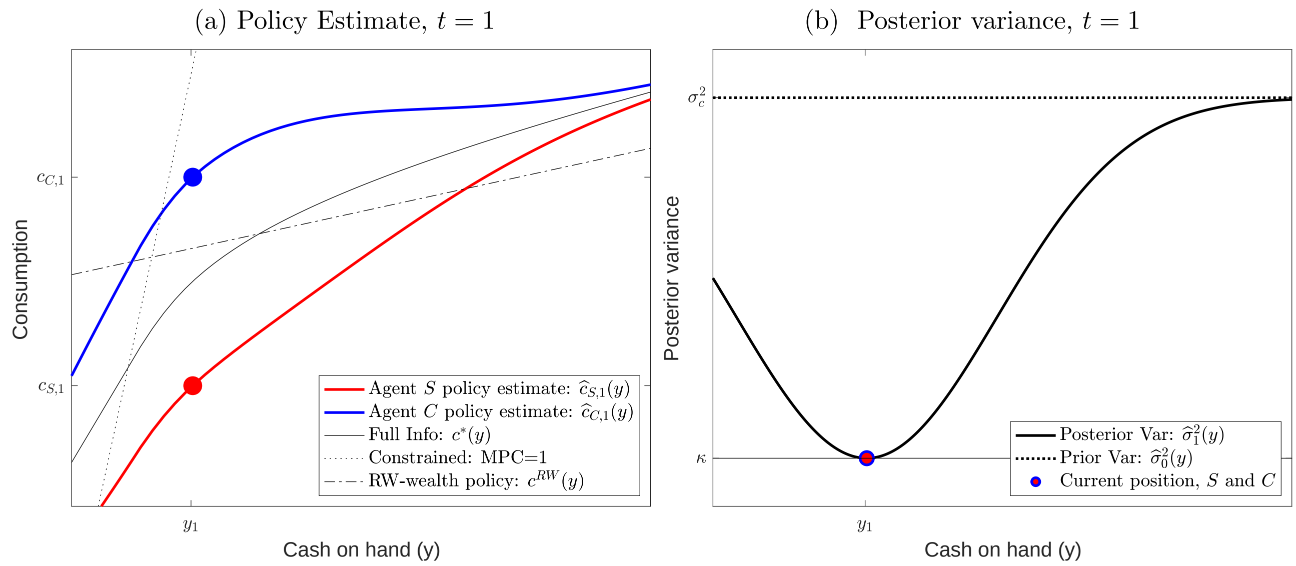
$$c_{i,1} = \min[y_1 + \underline{a}, c^*(y_1) + \alpha^*_1 \varepsilon_{i,1}].$$

At this stage, behavior is not systematically different across agents - it only differs due to the iid shock $\varepsilon_{i,1}$ which on average is zero across agents. However, the reasoning signals do not only generate stochastic choice in the current period, but also lead to systematic differences in the policy function estimates $\widehat{c}_{i,1}(y)$ across the state space. In particular, using equation (3) agent i's conditional estimate of the policy function for any value of y is

$$\widehat{c}_{i,1}(y) = c^*(y) + \exp(-\psi(y - y_1)^2)\alpha^*_1 \varepsilon_{i,1}.$$

For expositional purposes, we now assume that agent 1 receives a signal with a positive

Figure 2: Conditional beliefs at $t = 1$



error ($\varepsilon_{1,1} \equiv \varepsilon > 0$), while agent 2 receives a symmetrically biased signal with a negative error ($\varepsilon_{2,1} = -\varepsilon < 0$). To simplify exposition and intuition, we assume throughout this illustration that ε is small enough so that agent 1 does not hit the borrowing constraint. As the reasoning error ε makes the two agents over- or under-estimate, respectively, optimal consumption at *all* states y, we label agent 1 as a "consumer" or C, and similarly label agent 2 a "saver" or S.

To illustrate, in Figure 2 panel (a) we plot the updated estimates of the policy function $\widehat{c}_{i,1}$, for agent C (in blue) and agent S (in red), as induced by the reasoning errors $\varepsilon_{i,1}$. The blue and red circles represent the respective actions taken at time-1, $c_{i,1}$, and also mark the position of y_1 in this example. We also plot the full-information policy c^* (solid black line), which is also the mean prior belief, and two other objects that are helpful for gaining some geometric intuition. First, with the dot-dashed black line we plot the consumption policy

$$c^{RW}(y_{i,t}) = \frac{r}{1+r} y_{i,t} + \frac{1}{1+r} w,$$

which induces a random walk process for cash-on-hand $y_{i,t}$, i.e. if the agent was to follow this consumption rule, her wealth would be a random walk: $\mathbb{E}_t(y_{i,t+1}) = y_{i,t}$. Second, the dotted line plots the action implied by the borrowing constraint, i.e. $c(y_{i,t}) = y_{i,t}$.

Figure 2, panel (a) shows that agent C updates beliefs upward and hence ends up with an "over-consumption" bias relative to $c^*(y)$ throughout the state space, while the opposite is true for agent S. The plot also showcases that the shifts in the estimated policy functions $\widehat{c}_{C,1}(y)$ and $\widehat{c}_{S,1}(y)$, relative to the common prior belief of $c^*(y)$, are strongest at cash-at-hand values y close to y_1 – e.g. notice that both policy estimates converge to $c^*(y)$ for large y.

This reflects the local nature in the reduction of uncertainty – since we assume $\psi > 0$, agents face uncertainty in the shape of c^* and thus their signals $\eta_{i,1}$ primarily update beliefs locally.

To see the local reduction in uncertainty directly, consider the resulting posterior variance function $\widehat{\sigma}^2_{i,1}$, which is the same for both agents due to their identical choice of $\sigma^2_{\eta,1}$:

$$\widehat{\sigma}^2_{i,1}(y) = \widehat{\sigma}^2_1(y) = \sigma^2_c(1 - \alpha^*_1 \exp(-2\psi(y - y_1)^2)). \tag{16}$$

This function is plotted in Figure 2, panel (b) and has a characteristic U-shape that signifies the state and history dependence of posterior uncertainty. The state-dependence is embodied in the the fact that $\hat{\sigma}^2_1(y)$ is not a constant function, but varies with the value of the state y. The history dependence is due to the fact that $\widehat{\sigma}^2_1(y)$ is increasing in the distance between y and y_1, the state at which the time-1 signal is most informative.

Period 2: selection effects

The state and history dependence of uncertainty interacts with the endogenous dynamics of $y_{i,t}$, because it depends on past consumption choices – for example, entering time-2

$$y_{i,2} = (1 + r)\,(y_1 - c_{i,1}) + ws_2. \tag{17}$$

Even when income shocks are common across agents (i.e. $s_{i,2} = s_2$), the dispersion in the initial reasoning signals leads to wealth dispersion in period 2. We can infer the specific dynamics of cash-on-hand for each agent i by comparing their chosen level of time-1 consumption, $c_{i,1}$, to the consumption level implied by the RW-wealth policy function $c^{RW}(y)$, at the same level of cash-on-hand y_1. A consumption action below (above) the line $c^{RW}(y)$ implies an expected increase (decrease) in savings and thus an average drift up (down) in future assets.

The key property of this second period is a selection effect in how much each agent chooses to reason. As summarized in Proposition 2 below, the optimal time-2 reasoning intensity is increasing in the distance between the current state $y_{i,2}$ and y_1 – intuitively, if an agent finds herself facing a significantly different level of cash-on-hand than the level at which she reasoned previously, then uncertainty about the current optimal action is higher and will warrant further reasoning than otherwise.

Proposition 2. *The optimal reasoning intensity and the weight of the new signal in updating beliefs are both increasing in distance from location of the previous reasoning signal:*

$$\frac{\partial \sigma^2_{\eta,i,2}}{\partial ||y_{i,2} - y_1||} < 0 \textit{ and } \frac{\partial \alpha_{i,2}(y_{i,2})}{\partial ||y_{i,2} - y_1||} > 0.$$

Therefore, agent C reasons more than agent S, i.e. $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$, if and only if

$$s_2 < 1 + \frac{(1+r)}{w}(c^*(y_1) - c^{RW}(y_1)) \equiv \bar{s}$$

Proof. Details are in Appendix A. □

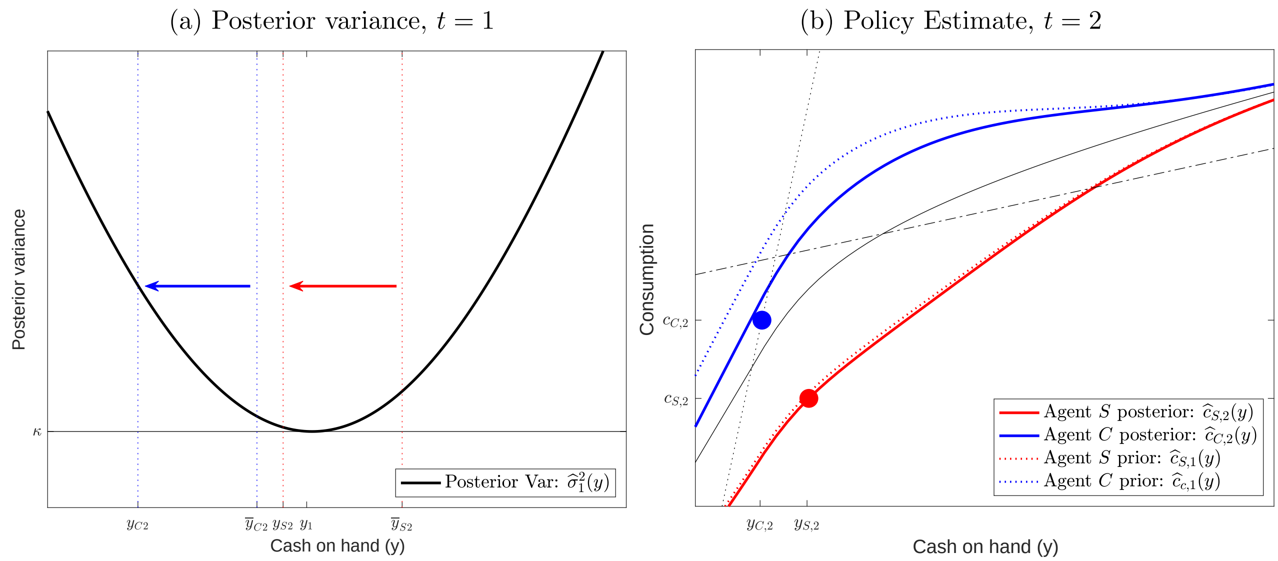
These results build on Proposition 1, where the optimal variance of the reasoning error was derived as $\sigma^2_{\eta,i,2} = \frac{\kappa\widehat{\sigma}^2_1(y_{i,2})}{\widehat{\sigma}^2_1(y_{i,2})-\kappa}$ and the resulting signal to noise ratio is $\alpha_{i,2}(y_{i,2}) = 1 - \frac{\kappa}{\widehat{\sigma}^2_1(y_{i,2})}$. Given that the variance function $\widehat{\sigma}^2_1(y)$ is lowest at y_1 and U-shaped around it, an agent i chooses to reason more the larger is the distance $||y_{i,2} - y_1||$. Taking into account the ex-post heterogeneity in $y_{i,2}$, we can show that $||y_{C,2} - y_1|| > ||y_{S,2} - y_1||$ if an only if $s_{i,2} = s_2 < \bar{s}$, and hence for such "low" realizations of income agent C perceives more uncertainty at his respective time-2 state, and thus a stronger incentive to reason about c^* than agent S.

We illustrate the implications of Proposition 2 in Figure 3, panel (a). There we plot again the posterior variance function entering time-2, $\widehat{\sigma}^2_1(y)$, but now zoom in on the region of the state space around y_1, and mark on the x-axis two possible time-2 cash-on-hand positions for each agent. First, we indicate the expected values $\bar{y}_{i,2} \equiv \mathbb{E}_1(y_{i,2})$, obtained by setting $s_{i,2} = 1$ in equation (17). Naturally, the "consumer" type who over-estimates optimal consumption, has lower time-2 wealth on average: $\bar{y}_{C,2} < \bar{y}_{S,2}$. But even though the reasoning signals $\eta_{i,1}$ are symmetric around $c^*(y_1)$, time-2 assets $\bar{y}_{i,2}$ are not symmetric around y_1.

The reason is that the full-information policy c^* is below c^{RW} in that region of the state space, reflecting a precautionary savings motive. Thus, the reasoning errors disperse actions around a positive drift in assets. In this illustration, the initial under-estimation of optimal consumption by agent S therefore further magnifies that positive drift, making it on average more likely for her to see a larger change in the time-2 state. Thus, $\widehat{\sigma}^2_1(\bar{y}_{C,2}) < \widehat{\sigma}^2_1(\bar{y}_{S,2})$ implying agent S faces more uncertainty about the optimal action and thus choose to reason more at the average future wealth. Through the lenses of Proposition 2, it follows that in this case the threshold $\bar{s}$ is lower than the mean of the labor income shocks (i.e. one). Absent the systematic difference between $c^*(y)$ and $c^{RW}(y)$, as for example in a standard PIH model, $\bar{s} = 1$ and agents would reason the same at their respective average states $\bar{y}_{i,2}$.

The key message of Proposition 2 is the state-dependence of the agents' optimal reasoning choices, which depend on the specific income shock realization. To illustrate, in the same panel (a) we also indicate the resulting cash-on-hand positions $y_{i,2}$ for the case of a below-average income shock realization $s_2 < \bar{s}$. Naturally, this results in lower than average cash-on-hand levels, as indicated by the two equal leftward arrows. However, the crucial observation is that this shock affects asymmetrically the distance $||y_{i,2} - y_1||$, and through it, perceived uncertainty and thus the optimal reasoning choices of the two agents.

Figure 3: Conditional beliefs: evolution from $t = 1$ to $t = 2$



In particular, note that this low income realization s_2 essentially compounds the "over-consumption" bias of agent C, and thus pushes his cash-on-hand even further away from y_1 than otherwise. On the other hand, the same shock acts to counter-balance the "under-consumption" bias of agent S, leaving her cash-on-hand relatively unchanged compared to y_1. Thus, in this situation, agent C faces significantly higher uncertainty about his time-2 optimal action than agent S, i.e. $\widehat{\sigma}_1^2(y_{C,2}) > \widehat{\sigma}_1^2(y_{S,2})$, and hence agent C chooses to reason with a higher intensity. As a result, agent C updates his beliefs by more, as seen in panel (b), where we also plot for comparison the previous period estimates $\widehat{c}_{i,1}(y)$ (dashed lines). As predicted by Proposition 2, because in this case $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$, the change in beliefs of agent C is substantially bigger than the revision in agent S's beliefs, which hardly change (red dashed line vs solid red line) as the latter chooses to acquire little new information. Naturally, by Proposition 2, the reverse type of behavior would emerge when $s_2 > \bar{s}$. In that case, agent S faces a more unfamiliar part of the state space than agent C, for whom the shock effectively constitutes a return towards the familiar state y_1.

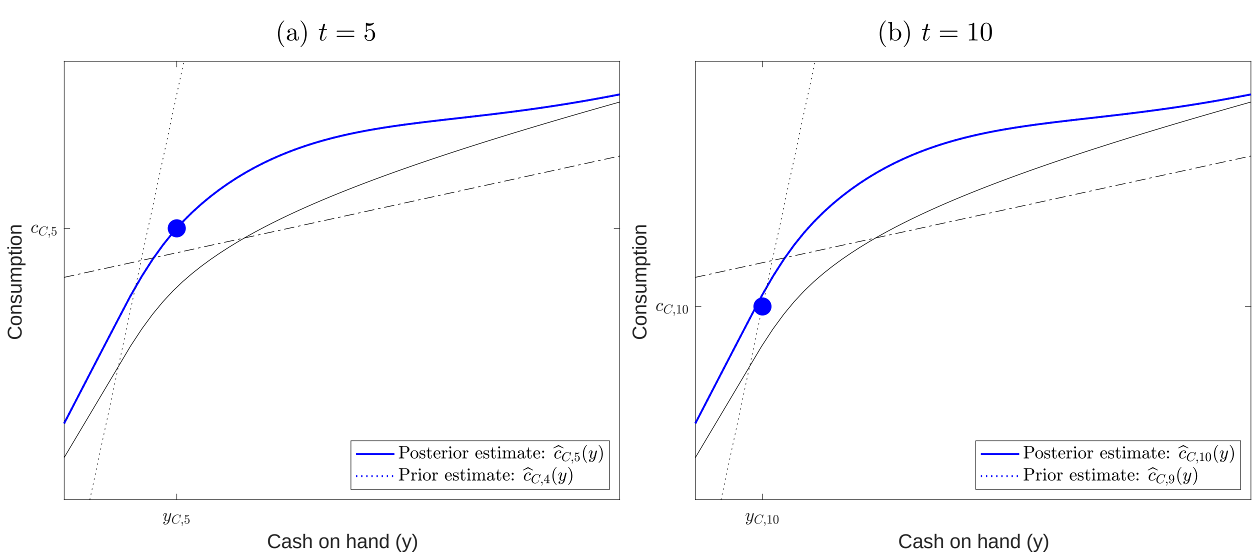
This analysis highlights a fundamental selection effect in *which* agents' beliefs get significantly updated and when. While a "saver" type is more likely to revise beliefs when experiencing a high income shock, a "consumer" type behaves in the opposite way. This selection reinforces the effects of the time-1 errors on agents' wealth, and helps generate persistent economic differences, which we illustrate next by following the agents' beliefs dynamics through time.

Learning traps: remaining hand-to-mouth

Notice that in our illustration the combination of a high consumption policy estimate and a negative income shock leads agent C to hit the borrowing constraint in period 2, seen from the fact that the blue-circle (current action) is on the 45-degree line and below the blue solid line in panel (b) of Figure 3. In this situation, negative income shocks will necessarily have a muted impact on cash-on-hand, as there is no borrowing. On the other hand, positive income shocks will move the agent in the part of the state space where his policy function estimate $\widehat{c}_{i,t}(y)$ lies above $c^{RW}(y)$. Importantly, this over-estimate of optimal consumption means that even if positive income shocks increase cash-on-hand, wealth is likely to remain high only temporarily as agent C draws down his assets $a_{i,t}$ and drifts back towards the constraint.

Therefore, the combination of the borrowing constraint, which acts as a barrier to the agents' downward drift in assets, and the high consumption estimate in the adjacent part of the state space generates stable wealth dynamics for agent C. As he is hit by income shocks in the future, his cash-on-hand is likely to fluctuate in the same familiar region where uncertainty is already low thanks to previous reasoning done in the neighborhood of y_1. Hence new signals are not likely to change the policy function estimate by much locally, because he will not choose very informative new signals $\eta_{C,t}$ unless there is a dramatic change in his cash-on-hand $y_{C,t}$, at which point the new signal will be primarily informative for a different part of the state space. As a result, on average his wealth fluctuates in a familiar region near the borrowing constraint, and the agent follows his established "over-consumption" behavior.

Figure 4: Policy estimate evolution, agent C



This type of dynamic is illustrated by Figure 4, where we plot two snapshots of beliefs and actions further into the future, for period-5 on the left and period-10 on the right. In both pictures, the dashed line, representing the agent's beliefs entering the period, lies virtually under the solid blue line of the end-of-period beliefs, signifying a small update in the policy function estimate. Similarly, there is little change in the policy estimate between periods five and ten, seen by comparing across panels (a) and (b). In fact, in the latter period agent C hits the borrowing constraint again, similar to his experience in period two, per panel (b) in Figure 3, illustrating how this agent settle into a low-wealth, "hand-to-mouth" status.

We label this situation a "learning trap". Its defining feature is a sequence of reasoning errors that establishes stable wealth dynamics, and thus a high likelihood for the agent's state variable to remain within a particular neighborhood of the state space. As wealth fluctuates in that familiar, low uncertainty region, the agent has little incentive to reason further.

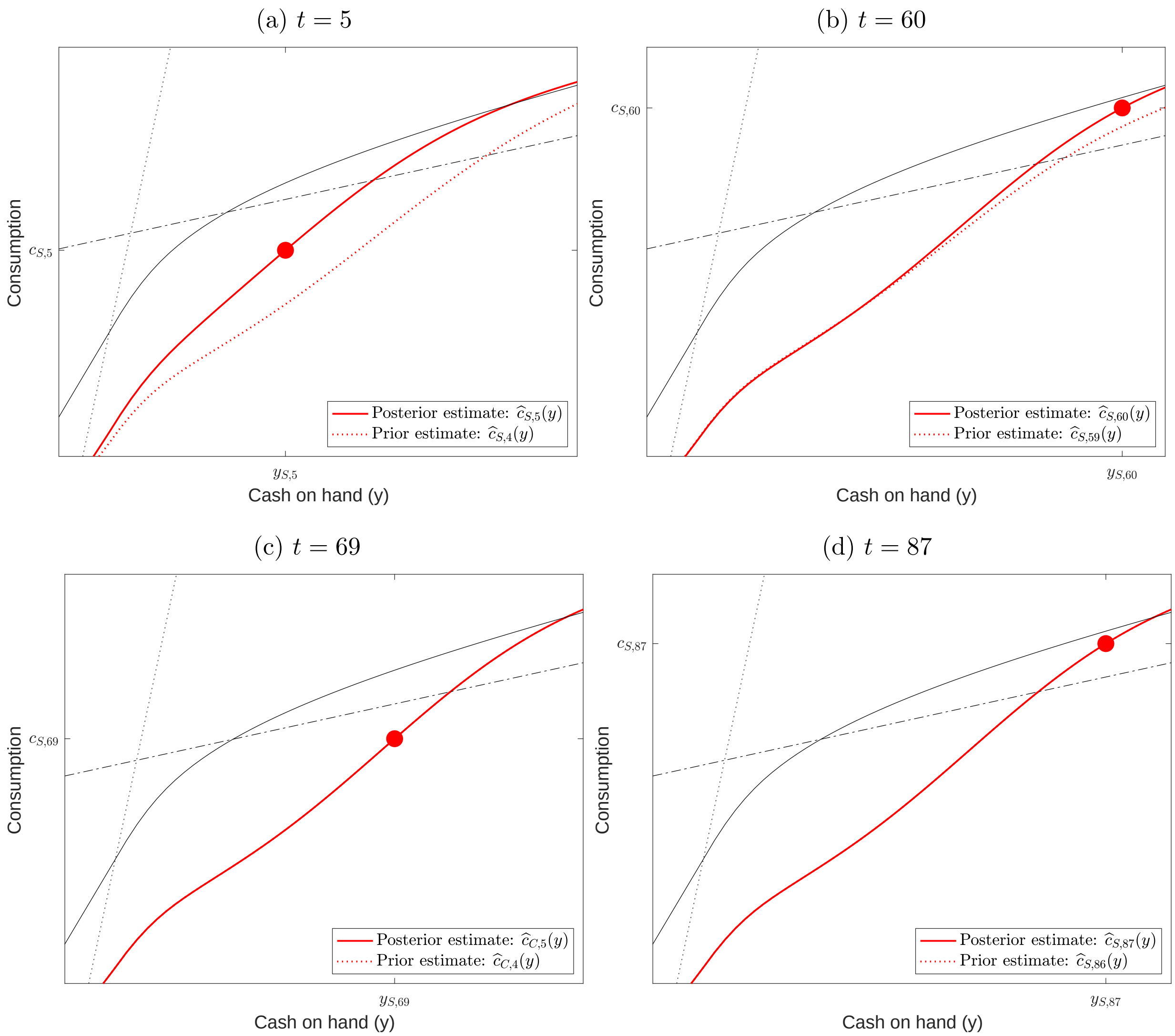
Learning traps: becoming and remaining rich

The "saver" agent is also subject to similar learning traps. However, her eventual steady state wealth level is different, owing to her different early life experiences and choices. First, the initial negative reasoning error generates an upward drift in assets. Second, this upward drift is compounded by the asymmetric effects of income shocks discussed earlier, where this agent is likely to draw new informative signals, and thus change her "under-consumption" bias, only at higher levels of wealth. Third, as her wealth drifts up over time, she accumulates new signals that revise her policy function estimate $\widehat{c}_{S,t}(y)$ largely locally, due to the local nature of information. Thus, rather than a uniform shift of beliefs up, the revisions will instead change the estimated slope of the policy function at high y.

This dynamic is illustrated in Figure 5 which plots several snapshots of this agent's policy estimate and respective action at four periods further into this example simulation: for $t = 5, 60, 69, 87$. Panel (a) shows that as time has progressed the agent has collected new information which has left to a revised policy estimate, mainly through an increase in the slope of $\widehat{c}_{S,5}(y)$. In panel (b), we show the eventual estimate and action taken at $t = 60$, the last period (in this simulation run) before the agent settles into a stable slow-learning region.

By that time, the agent has acquired numerous signals at lower levels of cash-on-hand, hence if wealth drifts down she will find little need for further reasoning. Meanwhile, her relatively steep estimate of the policy function ensures that it crosses the $c^{RW}(y)$ policy from below. This *upward crossing* establishes stable wealth dynamics because: (1) any action above the $c^{RW}(y)$ line leads to dis-saving and a downward drift in assets, and (2) in turn, dis-saving moves cash-on-hand back to a familiar region, where the agent's best estimate is to consume less than implied by the $c^{RW}(y)$ line and to build assets back up. Therefore, the resulting back-and-forth leads to a steady-state level of wealth given by the intersection of

Figure 5: Policy estimate evolution, agent S



$c^{RW}(y)$ and $\widehat{c}_{S,t}(y)$. We illustrate this stable, "business-as-usual" behavior in panels (c) and (d) as two more snapshots of the agent's beliefs and actions, showing that they have started fluctuating around the upward crossing point, with little further changes. We also note that this upward crossing point is specific to the history of reasoning signals experienced by agent S and is therefore different with probability one from the intersection of $c^{RW}(y)$ with $c^*(y)$ (which defines the steady-state level of wealth under full-information).

Taking stock of qualitative implications

As long as wealth drifts into previously uncertain parts of the state space, an agent's beliefs about the optimal policy function c^* are likely to continue to evolve, as she accumulates new reasoning signals. Eventually, reasoning errors lead to a policy function estimate $\widehat{c}_{i,t}(y)$ that establishes stable wealth dynamics in the neighborhood of past reasoning signals, and in such "learning traps" cash-on-hand starts cycling in a familiar, low uncertainty region. Such a situation can happen in two ways – either an agent gets stuck against the borrowing constraint while carrying an overly high estimate of the consumption function, or if away from the constraint, her estimate $\widehat{c}_{i,t}(y)$ forms an upward crossing with the $c^{RW}(y)$ policy.

These belief dynamics have important implications about the stochastic steady state of the model, which we analyze numerically in Section 4. First, agents' beliefs are on average trending towards policy function estimates that imply high MPCs across all agents, both rich and poor. The reason for the high MPCs for the poor agents are standard – they are close to the borrowing constraint, and hence beliefs there are likely to inherit the steepness of the full-information $c^*(y)$. The reason for the high MPCs for rich agents is unique to our model – those arise from the fact that beliefs endogenously stabilize around an upward crossing between the policy estimate $\widehat{c}_{i,t}(y)$ and the $c^{RW}(y)$ policy.

A second characteristic feature is illustrated by agent C's near-zero steady-state wealth. This outcome is very unlikely under full-information, where agents save aggressively when low on assets. In our model, however, being close to the constraint can be stable behavior. Intuitively, Bayesian agents have no reason to question their conditional policy estimates and some are bound to end up with estimates that push them close to the constraint.

A third fundamental implication of the model is the significantly larger wealth heterogeneity relative to the full-information case. This arises from agents effectively establishing individual stable levels of wealth that are dispersed across the whole state space, in contrast to the single steady-state level under $c^*(y)$. Hence, under full information, wealth heterogeneity is entirely due to income shocks, as wealth is trending to the same steady-state level for all agents. In our model, instead, the specific history of reasoning errors of each agent results in heterogeneous anchoring points for wealth. This is starkly illustrated by the experiences of agents S and C – two agents that are identical except for their reasoning errors in the first period, yet have very different long-run average wealth levels.

In closing, we make two further literature comparisons. First, changing the focus from imperfect perception of the state $y_{i,t}$ to imperfect perception of the optimal action conditional on the state allows the model to easily and robustly handle budget (and borrowing) constraints. Our agents both observe $y_{i,t}$ and understand how the budget constraint imposes a deterministic restriction between any given consumption policy function $c(y)$ and the implied savings rule $a(y) = y - c(y)$. As a result, given primitive informational parameters (eg. σ_c^2 and ψ), our model produces the same ergodic behavior whether reasoning occurs over $c^*(y)$ or over the optimal savings rule $a^*(y)$. The latter case is simply a deterministic translation of the Gaussian Process distribution over $c^*(y)$ to one over $a^*(y) = y - c^*(y)$, and hence there is no change in the structure of uncertainty facing the agent. Intuitively, the ergodic behavior will again be characterized by a situation in which assets have stable dynamics, implying a high MPC, or equivalently, a relatively flat effective savings policy function $\widehat{a}_t(y)$.

In contrast, if the state $y_{i,t}$ would be imprecisely perceived, as it is often assumed in the canonical literature on imperfect information, such equivalence may not hold. In that case, given signals about $y_{i,t}$, behavior may be substantially different whether consumption $c_{i,t}$ or savings $a_{i,t}$ is assumed to be the "residual" action that is imposed to ex-post clear the otherwise imperfectly perceived budget constrained.[20]

Second, in an environment with ex-ante identical agents and in the absence of any a-priori biases, costly reasoning about the optimal action endogenously generates two types of results that are typically challenging to produce, especially jointly. In particular, the higher MPCs and larger wealth heterogeneity than in the standard model, with both more poor and rich agents, indicates that our economy is different both in terms of the local *slope* of the consumption function as well as the dispersion of asset *levels* across the wealth distribution.[21]

4 Numerical analysis

Next we turn to a numerical implementation of the Aiyagari (1994) economy augmented with our reasoning friction as described above. We evaluate the key properties of the stationary equilibrium and contrast with the standard full-information version of the model.

4.1 Parametrization

Whenever possible, we follow the standard parametrization considered in Aiyagari (1994). Households have log-utility with a discount factor of $\beta = 0.96$. The i.i.d. labor income shock is drawn from a log-normal distribution $\ln(s_{i,t}) \sim N(-\frac{\sigma_s^2}{2}, \sigma_s^2)$, with $\sigma_s = 0.2$, and there is no borrowing allowed, i.e. $\underline{a} = 0$. On the production side, the capital share and the annual

[20] The typical assumption in the consumption-savings literature on imperfect perception of wealth is to let savings be the residual action (see for example Sims (2003), Maćkowiak and Wiederholt (2015) and Luo (2008)). In a model of inattention, Reis (2006) obtains different behavior if the residual action is consumption or savings. Gabaix (2014) discusses assumptions for how to model multiple actions within the sparse-max operator in the presence of budget constraints.

[21] The typical departures from the standard model are largely aimed to specifically explain one of these challenging properties. On the one hand, heterogeneity in preferences or rates of return are typical extensions to generate more wealth dispersion (see De Nardi and Fella (2017) for a survey). On the other hand, high MPCs for rich agents that hold low liquid wealth emerge from Kaplan et al. (2014), while a range of behavioral models can lead to high MPCs even for highly-liquid agents (see Lian (2020) for a general analysis).

depreciation are set to standard values of $\alpha = 0.36$ and $\delta = 0.08$, respectively. We now turn to the parameters governing the reasoning friction.

First, as suggested by the discussion in the previous section, the beliefs of individual agents can evolve very slowly for long periods of time. To help with the computation of the stationary steady state, we introduce a form of discounting of past reasoning signals. We opt for the tractable modeling assumption that agents face i.i.d. Poisson information shocks, where with probability θ an agent's history of accumulated reasoning signals becomes obsolete and that agent's beginning-of-period beliefs reset to the time-0 prior.

One interpretation of this discounting scheme is based on viewing agents as finitely lived, where conditional on death, the agent transfers his assets and the resulting continuation utility off to the offspring. However, the transfer of reasoning information about the optimal policy is imperfect across generations, which for simplicity we assume leads to full discounting of past information. In our parametrization we set $\theta = 0.02$, so that the economy is continuously repopulated with agents that on average re-start their learning problem every 50 years.

Second, as also discussed in Section 2, we aim to reduce the degrees of freedom intrinsic in specifying the prior beliefs, through restrictions that resemble the rational-expectations idea of utilizing "model-consistent" priors. These restrictions constrain the inherent "wilderness of bounded rationality" as follows. First, the common prior mean function $\widehat{c}_0(y)$ is set equal to the full-information policy function $c^*(y)$. This leaves us with the parameters σ_c^2 and ψ of the covariance function $\widehat{\sigma}_0(y, y')$, which govern the uncertainty around $\widehat{c}_0(y)$. We aim to impose a model-consistent value for those, by setting them equal to what an econometrician would estimate if he were to observe simulations from the model. In particular, we look for a fixed point such that given the values of $\{\sigma_c^2, \psi\}$, if an econometrician uses the resulting ergodic distribution of reasoning signals $\eta_{i,t}$ as data, he would recover the same values of $\{\sigma_c^2, \psi\}$ as used to simulate $\eta_{i,t}$. This essentially restricts the assumption on prior uncertainty to be model consistent, and also has a connection with the practice of estimating hyper-parameters in Bayesian statistics. More details on this fixed-point procedure are given in the Appendix.

Employing this model-consistent priors strategy, we are left with one remaining degree of freedom, namely the marginal cost of reasoning κ. We calibrate this last parameter by exploiting the tendency for agents in our model to settle in "learning traps". Among other things, these traps link our mechanism to one particularly challenging fact - while empirically the bottom 20% in the US wealth distribution have roughly zero net assets, standard model predicts that very few agents should be in that position due to strong precautionary saving motives.[22] This insight motivates us to set κ to target this moment, given the fixed-point

[22]See Krueger et al. (2016) as a reference for this moment and for a detailed discussion on how the standard model has a difficult time in predicting it.

restriction over $\{\sigma_c^2, \psi\}$ and the rest of parameters described above.

Putting everything together, the resulting calibration for the reasoning parameters is $\{\sigma_c^2, \psi, \kappa\} = \{0.77, 0.05, 0.48\}$. Those values suggest that agents indeed face non-trivial amount of uncertainty in the optimal policy ($\sigma_c^2 > 0$) and its shape ($\psi > 0$). The particular numerical values do not have a direct interpretation, however, and are best understood in the context of the resulting moments we discuss below.

4.2 Stationary equilibrium

We argue that the stationary distribution of our costly reasoning model significantly improves upon its full-information counterpart along three key dimensions: (1) larger wealth inequality; (2) more frequent and persistent hand-to-mouth (HtM) status, and (3) higher marginal propensities to consume (MPC), especially for rich, unconstrained agents. In our discussion, we connect to the existing empirical literature and show that our model tilts an otherwise standard incomplete markets model a-la Aiyagari (1994) towards a closer empirical fit.

To compute the stationary distribution we iteratively simulate an economy with 10,000 periods and 5,000 agents, and search for the value of the interest rate r which satisfies the definition of stationary equilibrium in Section 3. We compute the reported moments over the last 5,000 periods of the simulation. We compare the stationary distribution of our costly reasoning (CR) model to a standard full-information (FI) model driven by the same sequence of income shocks, but where there is no uncertainty about c^* (i.e. $\sigma_c^2 = 0$). Naturally, the two models achieve equilibrium at different levels of the interest rate r, as discussed below.

Wealth distribution

As foreshadowed by the discussion in Section 3, the stationary distribution of assets in the CR model is influenced by the underlying distribution of individual agent wealth "steady-states". Those vary across agents and depend on the location of the upward crossing pattern between individual policy estimates $\widehat{c}_{i,t}(y)$ and the $c^{RW}(y)$ policy. Around such crossings wealth dynamics are stable and agents do not change behavior much, as their respective state remains in a familiar, low uncertainty region. In contrast, in the FI model the wealth steady state is common to all agents, since they follow the same $c^*(y)$ policy. As a result, the CR model features both more wealth heterogeneity, and a greater persistence in assets (i.e. lower social mobility).

Figure 6 plots the stationary distribution of assets a_i in the benchmark CR economy (red line) and the counter-factual FI economy (blue line). A striking visual difference between the two distributions is the large mass of low wealth agents in the costly reasoning model.

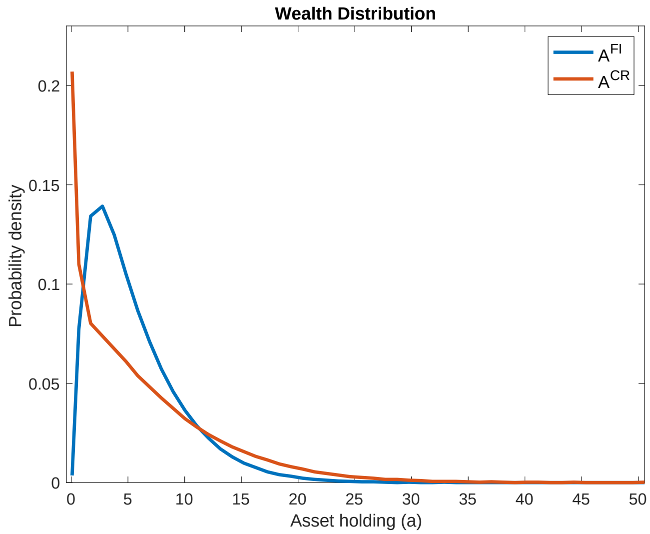


Figure 6: Wealth distribution in stationary equilibrium. Solid red line (A^{CR}) is for the benchmark costly reasoning economy, while the solid blue line (A^{FI}) is for the full information version.

The FI policy $c^*(y)$ is characterized by a strong precautionary savings motive which directs agents away from the borrowing constraint. In contrast, as discussed in the example of Section 3, our model has the potential to explain why agents acting under their best estimate of the optimal course of action could consistently maintain low levels of wealth.

The second qualitative difference between the two economies is the larger density of rich agents in the CR model as compared to the FI economy, exemplified by the slower decay of the right tail in the asset distribution – i.e. red line is consistently above the blue one in the right tail. The reason is again the presence of learning traps, which creates a mass of persistently rich agents that remain in the right tail, while the very rich agents in the FI model tend to spend down their wealth (since $c^*(y) > c^{RW}(y)$ for large y).

The net effect of the larger left and right tails can be deduced by the equilibrium interest rate, which is significantly lower in the CR economy – 3.87% versus 4.15% in the FI model. This indicates that the larger accumulation of assets by rich agents more than compensates the larger mass of poor agents, as compared to the FI model, resulting in higher mean level of assets for a given interest rate. Put together, our reasoning mechanism leads to an economy with more assets on average, where those are also more unequally distributed due to an increase in both the left and right tails of wealth.

As a summary statistic of the wealth inequality, in Panel (A) of Table 1 we show that our benchmark model (column (2)) produces a significantly higher Gini coefficient, bringing the model half-way from its full-information counter-part (column (3)) to the US data (column (1)). Consistent with the intuition discussed above, this increased inequality comes from more mass both in the left tail and in the right tail of the distribution. To see this, we split the wealth distribution by its quintiles and compute the amount of assets held by agents as a share out of the total. The asset share of the first quintile (Q1) is targeted by our calibration, and Table 1 shows that the model successfully achieves zero net assets for those agents.[23] Given the rich cross-equation restrictions on the joint stationary distribution of beliefs and actions, we view this match as success into itself. In contrast, as also illustrated by Figure 6, the standard FI version of the model is at odds with this moment - it predicts that agents in Q1 have largely managed to save away from the borrowing constraint and have accumulated assets that amount to 5% of the total.

In terms of the right tail of the wealth distribution, we note that while in the FI model the top quintile of agents holds 44% of all assets in the economy, in our benchmark model agents in the top quintile hold 57% of all assets. This share is still lower than the extreme concentration of wealth at the top that we see in the data, but our costly reasoning mechanism still brings the standard model closer to reality without the use of fat-tailed shocks.[24]

Moments	Data	Benchmark	Full info	Trembles
	(1)	(2)	(3)	(4)
(A) Wealth Distribution				
Gini coefficient	0.77	0.58	0.39	0.39
Bottom 20%: share of total assets	-0.01	0	0.05	0.04
Top 20%: share of total assets	0.83	0.57	0.44	0.46
(B) Hand-to-Mouth (HtM)				
Fraction of Hand-to-Mouth	0.23	0.23	0.01	0.02
HtM persistence: $\frac{P(HtM_{t+4}\|HtM_t)}{P(HtM_{t+2}\|HtM_t)}$	0.9	0.99	0.66	0.66
Future Δ in consumption $\| HtM_t$	0	0.001	0.015	0.028
(C) MPC				
Unconditional Mean	0.2-0.6	0.29	0.05	0.06
Mean | top 20% of assets	0.2-0.6	0.15	0.04	0.04

Table 1: We report moments from data in column (1) and the stationary distribution in our benchmark costly reasoning model in column (2). In our counterfactuals we keep parameters at their benchmark values but set $\sigma_c^2 = 0$ in column (3), or assume agents make iid mistakes around their known optimal policy function, per equation (18), in column (4), respectively. The data moments in Panel (A) are reported by Krueger et al. (2016) for PSID (in 2006), and in Panel (B) by Aguiar et al. (2020) who use the PSID panel structure. In Panel (C) we report a range summarized in Carroll et al. (2017) as credible empirical estimates of MPCs.

[23] In the data people in Q1 actually have slightly negative wealth: -0.9% as a share of total assets. Since in our model, as in Aiyagari (1994), the borrowing limit is set to zero, we have simply targeted a zero asset share for Q1. Introducing a negative borrowing limit would allow the model to obtain negative net wealth in Q1, but the data is so close to zero it makes little difference.

[24] See Benhabib and Bisin (2018) for a review on model extensions to generate more wealth concentration.

Hand-to-Mouth Status

In section 3 we discussed an example simulation of a "high consumption" agent (type C) whose reasoning experience lead to a low steady-state wealth, and the intuition behind that example underpins the model's ability to generate a large mass of low-wealth agents as observed in the data. Here we zoom-in on the ergodic behavior of such agents, and connect our model to the widely-used concept of "Hand-to-Mouth" (HtM) by using empirical moments reported in Aguiar et al. (2020). Implementing their measure of HtM based on net-worth (which follows Zeldes (1989)) we define an agent i at time t as being in HtM status if $a_{i,t-1}$ is less than two months of labor earnings.[25]

In the stationary distribution of the costly reasoning model 23% of agents are HtM, in line with the data. In contrast, under full-information agents save aggressively when they find themselves close to the constraint so the ergodic mass of HtM agents is essentially zero, a typical result in standard incomplete markets models. In such models, HtM status is the outcome of an unlikely long sequence of low income shocks, which is both a low probability and transitory status, as agents quickly build up their wealth once income shocks mean-revert.

However, in the data HtM status is not only prevalent, but also very persistent at the individual level, as is also true in our model. In particular, the probability of an agent being in a HtM status at $t+2$ conditional on currently being HtM is 90% in the CR model, while that probability is only 36% under full information. Furthermore, these probabilities become 89% and 24%, respectively, when we predict the future HtM status at $t+4$ instead. Aguiar et al. (2020) compute the same moments in their biannual PSID panel data and find high persistence of HtM status, which they also note is a challenge to the standard model. Their empirical estimates of that persistence are lower than in our model (65% for $t+2$ and 58% for $t+4$, respectively), but as they note, survey reporting errors for income and wealth are likely to create spurious transitions. Given this measurement error concern, we therefore emphasize that the HtM status in the data is remarkably persistent in relative terms – the probability of staying HtM over the next four years is 90% of the probability of staying HtM over the next two years, showing a very low rate of decay. The second row of Panel (B) of Table 1 shows that this high persistence is close to the costly reasoning model's implication, but significantly larger than the 66% delivered by the full information version.

Aguiar et al. (2020) also look at the dynamic properties of HtM status from another angle, by computing the expected future consumption growth of HtM agents, $\mathrm{E}(\Delta c_{t+2}|HtM_t)$, which we also report in Panel (B) of Table 1. Empirically, they find that this growth rate is essentially zero. As they note, this runs counter to the standard model's implication that HtM

[25] With i.i.d. risk, in our annual model this threshold is $w/6$, where w is the stationary equilibrium wage.

agents are on average expected to grow out of the constraint quickly, and thus experience high future consumption growth. Indeed, consistent with this intuition, we find that in the FI version of the model consumption growth conditional on current HtM status is 1.5% higher than otherwise. In contrast, in the CR model, the HtM agents are also characterized by a low wealth steady state, and their savings and assets are likely to bounce back and forth around that familiar low-wealth region. Therefore, consistent with the data, these agents appear to an outsider as having a low "target wealth". While this behavior could be interpreted as ex-ante differences in preferences, as in Aguiar et al. (2020), in our model it reflects the reality of costly reasoning for otherwise ex-ante identical agents.

Marginal Propensity to Consume

The empirical measurement of MPCs has been the object of a large literature, which uses various identifying approaches (see Jappelli and Pistaferri (2010) for a survey). As detailed by Carroll et al. (2017) the results are varied, with credible estimates of annual MPCs appearing to range from 0.2 to 0.6. Panel (C) of Table 1 reports the average MPCs computed from the considered models. The CR model is well in line with the data – with a mean MPC of 0.29 – while under full information the mean MPC is counter-factually low and just 0.05.

In our costly reasoning model there are two forces behind the high average MPC. The first one is a compositional effect. As in standard models, agents close to the constraint (e.g. Hand-to-Mouth agents) are less able to smooth out income shocks and are thus inherently characterized by high MPCs. Having a large mass of such agents in the ergodic distribution is an important contributor to the large average MPC the model generates. The HtM agents in the full-information model have similarly high MPCs, but unlike in the CR model, they are a negligible fraction of the population as a whole.

In addition to the composition effect, the CR model also features high individual MPCs for the rich, unconstrained agents. This is because the typical agent at the stationary equilibrium is likely in the midst of a "learning trap". As discussed qualitatively in Section 3, this occurs when an agent's wealth dynamics are sufficiently stable, which requires an upward crossing of the policy estimate $\widehat{c}_{i,t}(y)$ and the $c^{RW}(y)$ policy line. In practice, this means that the typical agent's policy function estimate is relatively steep, and this holds true for agents across the wealth distribution. Quantitatively, we find that the median MPC is 0.17. Moreover, we report in Table 1 that even for agents in the top quintile of the wealth distribution the average MPC is still very high, and equal to 0.15. In contrast, in the FI economy we find the standard result that away from the constraint, the consumption response to income shocks is muted and similar to the permanent-income-hypothesis, with an average MPC for the richest 20% of just 0.04 .

The empirical literature provides increasing support in favor of a model that delivers both a large mass of constrained agents *and* high MPCs for the unconstrained agents. Being "constrained" may be interpreted in terms of the overall net worth, as in the one asset economy that we study, or more specifically in terms of liquid wealth in a model that differentiates assets by liquidity, like in Kaplan et al. (2014). Still, Parker (2017), Olafsson and Pagel (2018) and Fagereng et al. (2019) point out that even for agents with high liquid wealth the MPC level is significantly higher than implied by standard models. Therefore, the costly reasoning friction brings an otherwise standard incomplete markets model closer to the data.

4.3 Endogenous patterns of mistakes

Motivated by the broad literature reviewed in the introduction, our model generates behavioral "mistakes" that are not mechanical, but the outcome of a constrained maximization. As a result, our "resource rational" agents make errors but not in an invariant, exogenous way. Instead, their reasoning intensity and patterns of mistakes depends on changes in the environment. We highlight these properties in this section through two additional experiments.

Costly reasoning vs simple mistakes

We first discuss a counter-factual model where agents have full-information about the optimal policy function but make idiosyncratic mistakes in their actions.[26] Namely, we consider that agents are suffering from a simple "trembling-hand" kind of control problem, where they set an approximately accurate action that is contaminated with i.i.d. noise:

$$c_{i,t}^{trmb} = c^*(y_{i,t}) + \sigma_\tau^2 \varepsilon_{i,t}. \tag{18}$$

When simulating this model, we use the exact same sequence of noise shocks $\varepsilon_{i,t}$ that affect the reasoning signals in the costly-reasoning model. Hence, the stochastic choice of agents in this alternative model is driven by the same source of exogenous disturbances that generates contemporaneous dispersion in our benchmark model. We calibrate the standard deviation of these shocks, $\sigma_\tau = 0.18$, so as to match the dispersion of actions around the full-information action in the benchmark CR model – i.e. to match $Var(c_{i,t} - c^*(y_{i,t}))$.

In this counterfactual "trembles" model, with moments reported in column (4) of Table 1, agents similarly display stochastic choice and make mistakes in their actions, as in the CR

[26] A related 'near-rational' approach (eg. Akerlof and Yellen (1985) and Hassan and Mertens (2017)) assumes a constant cost of implementing the otherwise known optimal action and studies the general equilibrium effects of the resulting individual errors. While those models typically assume some correlation in the errors, in this counter-factual we impose that they are iid across time and agents.

model. This creates some additional wealth heterogeneity compared to the FI model, but that is quantitatively negligible – the share of assets held by the top 20% of agents increases from 44% to 46%, and the fraction of HtM agents increases from 1% to 2%. Overall, the Gini coefficient remains the same at 0.39.

Similarly, while the "trembles" generate consumption volatility, on average an agent behaves as under c^*, hence the MPCs for the unconstrained agents are also low. There is a slight increase in the average MPC (0.06 vs 0.05), but that is completely due to the compositional effect of more HtM agents – the MPC of the unconstrained agents are equivalent in the "trembles" and the full-information models. This showcases that the endogeneity in the reasoning intensity choice and the implied selection effect we have discussed in Section 3.2 are crucial for obtaining the systematically higher MPCs in our benchmark model.

The key to these results is that the errors in this counter-factual are not systematic, hence they tend to wash out over the long-run. For example, an agent might under-consume a few times, but he is not likely to become a "saver" type that persistently under-consumes, like it is possible in the CR model due to the endogenous, state-dependent choice of reasoning intensity. This is exemplified by the expected consumption growth of HtM agents in the "trembles" version of the model, which is even higher than under FI (2.8% vs 1.5%), underscoring the strong mean-reversion in actions in this counter-factual model.

The fact that the ergodic implications of this "trembles" counter-factual are similar to the full-information model illustrates a standard justification in the literature that the latter model may still be a good approximation for an underlying model where agents do end up making errors. In contrast, our mechanism shows that when behavioral "mistakes" are modeled in a "resource rational" way, the joint-distribution of beliefs and actions differ qualitatively and quantitatively from its full-information version.

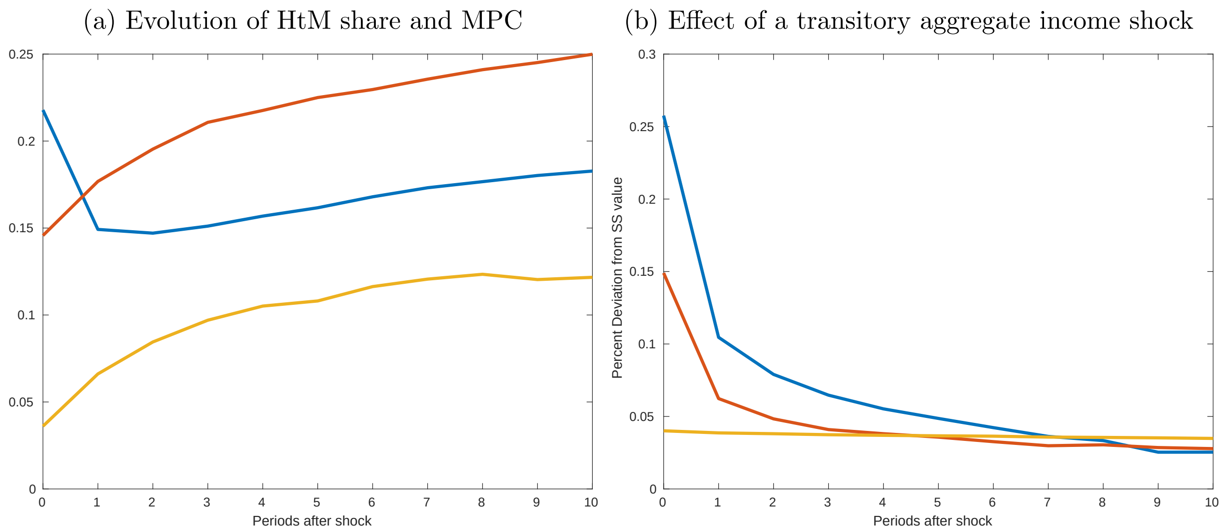
Uncertainty shock and policy effects

In the second experiment we compute the effect of an aggregate uncertainty shock on the endogenous reasoning choices and their importance for a policy maker. In particular, partly motivated by the current COVID-19 crisis, we consider a shock that, starting at the ergodic steady state of the CR economy, resets all agents' beliefs back to the time-0 prior.[27] There is no fundamental change in the economic structure, hence over time the economy converges back to the same stationary distribution.

On impact, all agents perceive a large increase in uncertainty about c^*, leading to temporary aggregate dynamics that showcase the endogenous nature of the behavioral

[27]We chose a simple, but stark, example of an uncertainty shock for illustration purposes. Naturally, a shock where only some, but not all, information is rendered useless has qualitatively similar implications.

Figure 7: Conditional beliefs following an uncertainty shock



“mistakes” of our agents. To visualize these effects, in Figure 7, panel (a) we plot the impulse responses of three key moments – (i) share of HtM agents (blue line), (ii) average MPC (red line) and (iii) average MPC among richest 20% of agents (yellow line).

There are two key forces that shape the response of the economy. On the one hand, all agents perceive significantly increased uncertainty and hence choose to reason intensely. On the other hand, since both the time-0 priors and the new reasoning signals are centered at the full-information policy c^*, on impact aggregate consumption behaves as under full-information.

This is underscored by the dramatic fall in the MPC across agents. We can see that both in the mean MPC in the economy, which is essentially halved to 0.15, and in the mean MPC of the richest agents, which falls to 0.04. The economy-wide average MPC does not fall all the way to its level under FI, because of the wealth distribution, which is inherited from the stationary distribution of the CR economy. In particular, even though upon the shock all agents start with a clean slate of beliefs, there is still a large fraction of agents with virtually no assets, and for those “hand-to-mouth” agents, MPCs are high even under the full-information policy. Over time, the economy transitions back to the stochastic steady state described earlier, with the average MPC rising back to 0.29.

The sharp fall in MPCs on the impact of the uncertainty shock is of particular interest to policy makers. High MPCs are often seen as important in the propagation of policy efforts to affect the economy, and thus our model suggests that policy becomes less effective during uncertain times, as our agents abandon their “business-as-usual”, high MPC consumption patterns. To make this point concrete, in the panel (b) of Figure 7 we plot the impulse response of a one-time, transitory aggregate income shock equal to 1% of income, both with and without a concurrent uncertainty shock. The income shock can be interpreted as a simple implementation of expansionary fiscal or monetary policy (under sticky prices).

Without a concurrent uncertainty shock, a 1% increase in average income leads to a significant response in our benchmark costly-reasoning economy, increasing aggregate consumption by 0.26% on impact. Meanwhile, as it is well known, the impact in the full-information economy is muted, with consumption increasing by just 0.04%. However, if the income shock happens at the same time as an uncertainty shock, the effect on aggregate consumption in the CR model is roughly halved, with aggregate consumption increasing by only 0.15% in this case. Thus, while a policy maker can expect to have a significant impact on the economy in normal times, the same policy will be considerably weaker in times of high uncertainty, showcasing the relevance of modeling behavioral "mistakes" that respond to changes in environment.

5 Conclusion

This paper is motivated by a long-standing interest in the economics literature of relaxing the typically convenient, but otherwise extreme, assumption of decision-makers having free cognitive access to their optimal policy function. In this context, the first contribution of the paper is methodological in nature, as we propose a framework to model costly reasoning that is (i) tractable and portable across specific economic models, and (ii) well grounded in a broad neuroscience, experimental and computational literature. The second contribution is applied. We show that the proposed costly reasoning framework is a parsimonious and novel mechanism that generates rich and intuitive joint dynamics of beliefs and actions in a standard incomplete markets model. These dynamics can help bring the model closer to the data, and also hold important lessons for policy makers.

References

Aguiar, M., M. Bils, and C. Boar (2020): "Who are the Hand-to-Mouth?" NBER Working Paper No. 26643.

Aiyagari, S. R. (1994): "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 109, 659–684.

Akerlof, G. A. and J. L. Yellen (1985): "Can small deviations from rationality make significant differences to economic equilibria?" *The American Economic Review*, 75, 708–720.

Alaoui, L. and A. Penta (2016): "Cost-Benefit Analysis in Reasoning," Working Paper.

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

ARAGONES, E., I. GILBOA, A. POSTLEWAITE, AND D. SCHMEIDLER (2005): "Fact-Free Learning," *The American Economic Review*, 95, pp–1355.

BALLINGER, T. P. AND N. T. WILCOX (1997): "Decisions, error and heterogeneity," *The Economic Journal*, 107, 1090–1105.

BARBER, D. (2012): *Bayesian reasoning and machine learning*, Cambridge University Press.

BENHABIB, J. AND A. BISIN (2018): "Skewed wealth distributions: Theory and empirics," *Journal of Economic Literature*, 56, 1261–91.

BERTSEKAS, D. P. (2019): *Reinforcement learning and optimal control*, Athena Scientific Belmont, MA.

BORNSTEIN, A. M., M. W. KHAW, D. SHOHAMY, AND N. D. DAW (2017): "Reminders of past choices bias decisions for reward in humans," *Nature Communications*, 8, 1–9.

CAPLIN, A., M. DEAN, AND J. LEAHY (2016): "Rational Inattention, Optimal Consideration Sets and Stochastic choice," Working paper.

CARROLL, C., J. SLACALEK, K. TOKUOKA, AND M. N. WHITE (2017): "The distribution of wealth and the marginal propensity to consume," *Quantitative Economics*, 8, 977–1020.

CONLISK, J. (1996): "Why bounded rationality?" *Journal of Economic Literature*, 34, 669–700.

DE NARDI, M. AND G. FELLA (2017): "Saving and wealth inequality," *Review of Economic Dynamics*, 26, 280–300.

DEW-BECKER, I. AND C. G. NATHANSON (2019): "Directed attention and nonparametric learning," *Journal of Economic Theory*, 181, 461–496.

FAGERENG, A., M. B. HOLM, AND G. J. J. NATVIK (2019): "MPC heterogeneity and household balance sheets," *Available at SSRN 3399027*.

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

GABAIX, X. (2014): "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 129, 1661–1710.

——— (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

GARCÍA-SCHMIDT, M. AND M. WOODFORD (2019): "Are low interest rates deflationary? A paradox of perfect-foresight analysis," *American Economic Review*, 109, 86–120.

Gershman, S. J. and N. D. Daw (2017): "Reinforcement learning and episodic memory in humans and animals: an integrative framework," *Annual review of psychology*, 68, 101–128.

Gul, F. and W. Pesendorfer (2004): "Self-control and the theory of consumption," *Econometrica*, 72, 119–158.

Hansen, L. P. (2007): "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," *American Economic Review*, 97, 1–30.

Hassan, T. A. and T. M. Mertens (2017): "The social cost of near-rational investment," *The American Economic Review*, 107, 1059–1103.

Hey, J. D. (2001): "Does repetition improve consistency?" *Experimental Economics*, 4, 5–54.

Ilut, C., R. Valchev, and N. Vincent (2020): "Paralyzed by Fear: Rigid and Discrete Pricing under Demand Uncertainty," *Econometrica*, forthcoming.

Jappelli, T. and L. Pistaferri (2010): "The Consumption Response to Income Changes," *Annual Review of Economics*, 2, 479–506.

Kahneman, D. (2011): *Thinking, fast and slow*, Macmillan.

Kaplan, G., G. L. Violante, and J. Weidner (2014): "The wealthy hand-to-mouth," *Brookings papers on economic activity*, 77–153.

Klaes, M., E.-M. Sent, et al. (2005): "A conceptual history of the emergence of bounded rationality," *History of political economy*, 37, 27–59.

Kozlowski, J., L. Veldkamp, and V. Venkateswaran (2020): "The tail that wags the economy: Beliefs and persistent stagnation," *Journal of Political Economy*, 128.

Krueger, D., K. Mitman, and F. Perri (2016): "Macroeconomics and household heterogeneity," in *Handbook of Macroeconomics*, Elsevier, vol. 2, 843–921.

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

Lian, C. (2020): "Mistakes in future consumption, high MPCs now," MIT, Working Paper.

Lieder, F. and T. L. Griffiths (2020): "Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources," *Behavioral and Brain Sciences*, 43.

Liu, W., J. C. Principe, and S. Haykin (2011): *Kernel adaptive filtering: a comprehensive introduction*, vol. 57, John Wiley & Sons.

Luo, Y. (2008): "Consumption dynamics under information processing constraints," *Review of Economic Dynamics*, 11, 366–385.

Maćkowiak, B. and M. Wiederholt (2009): "Optimal sticky prices under rational inattention," *The American Economic Review*, 99, 769–803.

——— (2015): "Business cycle dynamics under rational inattention," *The Review of Economic Studies*, 82, 1502–1532.

Matějka, F. and A. McKay (2014): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *The American Economic Review*, 105, 272–298.

Mosteller, F. and P. Nogee (1951): "An experimental measurement of utility," *Journal of Political Economy*, 59, 371–404.

Nimark, K. (2014): "Man-bites-dog business cycles," *The American Economic Review*, 104, 2320–2367.

Nimark, K. P. and S. Pitschner (2017): "News Media and Delegated Information Choice," Working Paper.

Olafsson, A. and M. Pagel (2018): "The liquid hand-to-mouth: Evidence from personal finance management software," *The Review of Financial Studies*, 31, 4398–4446.

Parker, J. A. (2017): "Why Don't Households Smooth Consumption? Evidence from a $25 Million Experiment," *American Economic Journal: Macroeconomics*, 9, 153–83.

Rasmussen, C. E. and C. K. Williams (2006): *Gaussian processes for machine learning*, vol. 1, MIT press Cambridge.

Reis, R. (2006): "Inattentive consumers," *Journal of Monetary Economics*, 53, 1761–1800.

Sargent, T. J. (1993): *Bounded Rationality in Macroeconomics*, Oxford University Press.

Simon, H. A. (1955): "A behavioral model of rational choice," *The Quarterly Journal of Economics*, 69, 99–118.

——— (1976): "From substantive to procedural rationality," in *25 years of economic theory*, Springer, 65–86.

Sims, C. A. (1998): "Stickiness," in *Carnegie-Rochester Conference Series on Public Policy*, vol. 49, 317–356.

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

——— (2010): "Rational inattention and monetary economics," in *Handbook of Monetary Economics*, ed. by B. M. Friedman and M. Woodford, Elsevier, vol. 3, 155–181.

Stevens, L. (2020): "Coarse pricing policies," *The Review of Economic Studies*, 87, 420–453.

van Nieuwerburgh, S. and L. Veldkamp (2010): "Information acquisition and under-diversification," *The Review of Economic Studies*, 77, 779–805.

Wiederholt, M. (2010): "Rational Inattention," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Palgrave Macmillan, vol. 4.

Woodford, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, 25.

——— (2014): "Stochastic choice: An optimizing neuroeconomic model," *The American Economic Review*, 104, 495–500.

Zeldes, S. P. (1989): "Consumption and liquidity constraints: an empirical investigation," *Journal of Political Economy*, 97, 305–346.

Appendix

A Proofs

Lemma 1. *If* $\theta_k \overset{iid}{\sim} N(\mu_k, \sigma_c^2)$*, equation (2) implies that* $c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$*, with*

$$\widehat{c}_0(y) = \sum_{k=1}^{N} \mu_k \phi_k(y); \qquad \widehat{\sigma}_0(y, y') = \sigma_c^2 \sum_{k=1}^{N} \phi_k(y)\phi_k(y')$$

Proof. Given that $\theta_k \overset{iid}{\sim} N(\mu_k, \sigma_c^2)$, it immediately follows that for any pair of real scalars, $y, y' \in \mathbb{R}$, the vector $\left[\sum_{k=1}^{N} \theta_k \phi_k(y), \sum_{k=1}^{N} \theta_k \phi_k(y')\right]'$ has the following joint Gaussian distribution:

$$\begin{bmatrix} \sum_{k=1}^{N} \theta_k \phi_k(y) \\ \sum_{k=1}^{N} \theta_k \phi_k(y') \end{bmatrix} \sim N\left(\begin{bmatrix} \sum_{k=1}^{N} \mu_k \phi_k(y) \\ \sum_{k=1}^{N} \mu_k \phi_k(y) \end{bmatrix}, \begin{bmatrix} \sigma_c^2 \sum_{k=1}^{N} (\phi_k(y))^2 & \sigma_c^2 \sum_{k=1}^{N} \phi_k(y)\phi_k(y') \\ \sigma_c^2 \sum_{k=1}^{N} \phi_k(y)\phi_k(y') & \sigma_c^2 \sum_{k=1}^{N} (\phi_k(y'))^2 \end{bmatrix} \right),$$

Assuming a complete set of basis functions $\{\phi_k\}_{k=1}^{N}$ that is big enough to achieve an arbitrarily good approximation of c^* so that

$$c^*(y) \approx \sum_{k=1}^{N} \theta_k \phi_k(y) \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$$

where

$$\widehat{c}_0(y) = \sum_{k=1}^{N} \mu_k \phi_k(y)$$

$$\widehat{\sigma}_0(y, y') = \sigma_c^2 \sum_{k=1}^{N} \phi_k(y)\phi_k(y')$$

□

Lemma 2. *Given the time-0 prior belief* $c^* \sim \mathcal{GP}(\widehat{c}_0, \widehat{\sigma}_0)$*, conditional beliefs are given by* $c^* \big| \{\eta^t, y^t\} \sim \mathcal{GP}(\widehat{c}_t, \widehat{\sigma}_t)$ *with moments evolving according to the recursive expressions*

$$\widehat{c}_t(y) = \widehat{c}_{t-1}(y) + \frac{\widehat{\sigma}_{t-1}(y, y_t)}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,t}^2}(\eta_t - \widehat{c}_{t-1}(y_t)), \tag{19}$$

$$\widehat{\sigma}_t(y, y') = \widehat{\sigma}_{t-1}(y, y') - \frac{\widehat{\sigma}_{t-1}(y, y_t)\widehat{\sigma}_{t-1}(y', y_t)}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,t}^2} \tag{20}$$

where $\widehat{c}_t(y) \equiv E_t(c^*(y)|\eta^t)$ *and* $\widehat{\sigma}_t(y, y') \equiv \mathrm{Cov}(c^*(y), c^*(y')|\eta^t)$ *are the posterior mean and covariance functions. Lastly,* $\widehat{\sigma}_t^2(y) \equiv \widehat{\sigma}_t(y, y)$ *denotes the posterior variance at a given* y*.*

Proof. We prove this by way of induction. Consider the first update of beliefs at $t = 1$. By the definition of the Gaussian Process distribution, and the fact that $\eta_1 = c^*(y_1) + \varepsilon_1$ where ε_1 is Gaussian scalar and independent of the Gaussian Process c^*, it follows that for any $y, y' \in \mathbb{R}$

$$\begin{bmatrix} c^*(y) \\ c^*(y') \\ \eta_1 \end{bmatrix} \sim N\left(\begin{bmatrix} \widehat{c}_0(y) \\ \widehat{c}_0(y') \\ \widehat{c}_0(y_1) \end{bmatrix}, \begin{bmatrix} \widehat{\sigma}_0^2(y) & \widehat{\sigma}_0(y,y') & \widehat{\sigma}_0(y,y_1) \\ \widehat{\sigma}_0(y',y) & \widehat{\sigma}_0^2(y') & \widehat{\sigma}_0(y',y_1) \\ \widehat{\sigma}_0(y_1,y) & \widehat{\sigma}_0(y_1,y') & \widehat{\sigma}_0^2(y_1) + \sigma_{\eta,1}^2 \end{bmatrix} \right),$$

where we have used the short-hand notation $\sigma_0^2(y) \equiv \sigma_0(y, y)$.

By the standard property of multivariate Gaussian distributions (and given that y_1 is a known, deterministic scalar), the conditional distribution $\begin{bmatrix} c^*(y) \\ c^*(y') \end{bmatrix} \Big| \eta_1$ is also Gaussian:

$$\begin{bmatrix} c^*(y) \\ c^*(y') \end{bmatrix} \Bigg| \eta_1 \sim N\left(\begin{bmatrix} \widehat{c}_1(y) \\ \widehat{c}_1(y') \end{bmatrix}, \begin{bmatrix} \widehat{\sigma}_1^2(y) & \widehat{\sigma}_1(y,y') \\ \widehat{\sigma}_1(y',y) & \widehat{\sigma}_1^2(y') \end{bmatrix} \right),$$

where by standard Bayesian updating formulas

$$\begin{bmatrix} \widehat{c}_1(y) \\ \widehat{c}_1(y') \end{bmatrix} = \begin{bmatrix} \widehat{c}_0(y) \\ \widehat{c}_0(y') \end{bmatrix} + \begin{bmatrix} \widehat{\sigma}_0(y,y_1) \\ \widehat{\sigma}_0(y',y_1) \end{bmatrix} \frac{\eta_1 - \widehat{c}_0(y_1)}{\widehat{\sigma}_0^2(y_1) + \sigma_{\eta,1}^2}$$

$$\begin{bmatrix} \widehat{\sigma}_1^2(y) & \widehat{\sigma}_1(y,y') \\ \widehat{\sigma}_1(y',y) & \widehat{\sigma}_1^2(y') \end{bmatrix} = \begin{bmatrix} \widehat{\sigma}_0^2(y) & \widehat{\sigma}_0(y,y') \\ \widehat{\sigma}_0(y',y) & \widehat{\sigma}_0^2(y') \end{bmatrix} - \begin{bmatrix} \widehat{\sigma}_0(y,y_1) \\ \widehat{\sigma}_0(y',y_1) \end{bmatrix} \frac{1}{\widehat{\sigma}_0^2(y_1) + \sigma_{\eta,1}^2} \begin{bmatrix} \widehat{\sigma}_0(y,y_1) \\ \widehat{\sigma}_0(y',y_1) \end{bmatrix}'$$

which is simply equations (19) and (20) in matrix form, evaluated at $t = 1$. Thus, $c^* | \eta_1 \sim \mathcal{GP}(\widehat{c}_1, \widehat{\sigma}_1)$, where the functions $\widehat{c}_1$ and $\widehat{\sigma}_1$ are defined above. This confirms the result for $t = 1$.

For the induction step, assume that equations (19) and (20) hold for $t - 1$ and that $c^* | \eta^{t-1} \sim \mathcal{GP}(\widehat{c}_{t-1}, \widehat{\sigma}_{t-1})$. Now consider the update at time t, again it follows that for any $y, y' \in \mathbb{R}$, the joint distribution $[c^*(y), c^*(y'), \eta_t] | \eta^{t-1}$ is Gaussian, with means given by the $\widehat{c}_{t-1}$ function and a variance-covariance matrix fully characterized by the $\widehat{\sigma}_{t-1}$ function. Then, by steps similar to those above

$$\begin{bmatrix} c^*(y) \\ c^*(y') \end{bmatrix} \Bigg| \eta^t \sim N\left(\begin{bmatrix} \widehat{c}_t(y) \\ \widehat{c}_t(y') \end{bmatrix}, \begin{bmatrix} \widehat{\sigma}_t^2(y) & \widehat{\sigma}_t(y,y') \\ \widehat{\sigma}_t(y',y) & \widehat{\sigma}_t^2(y') \end{bmatrix} \right),$$

where by the same standard Bayesian updating formulas for any $y, y' \in \mathbb{R}$

$$\begin{bmatrix} \widehat{c}_t(y) \\ \widehat{c}_t(y') \end{bmatrix} = \begin{bmatrix} \widehat{c}_{t-1}(y) \\ \widehat{c}_{t-1}(y') \end{bmatrix} + \begin{bmatrix} \widehat{\sigma}_{t-1}(y,y_t) \\ \widehat{\sigma}_{t-1}(y',y_t) \end{bmatrix} \frac{\eta_t - \widehat{c}_{t-1}(y_t)}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,t}^2}$$

$$\begin{bmatrix} \widehat{\sigma}_t^2(y) & \widehat{\sigma}_t(y,y') \\ \widehat{\sigma}_t(y',y) & \widehat{\sigma}_t^2(y') \end{bmatrix} = \begin{bmatrix} \widehat{\sigma}_{t-1}^2(y) & \widehat{\sigma}_{t-1}(y,y') \\ \widehat{\sigma}_{t-1}(y',y) & \widehat{\sigma}_{t-1}^2(y') \end{bmatrix} - \begin{bmatrix} \widehat{\sigma}_{t-1}(y,y_t) \\ \widehat{\sigma}_{t-1}(y',y_t) \end{bmatrix} \frac{1}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,t}^2} \begin{bmatrix} \widehat{\sigma}_{t-1}(y,y_t) \\ \widehat{\sigma}_{t-1}(y',y_t) \end{bmatrix}'$$

$\square$

Lemma 3. *Let* $\alpha_t(y) \equiv \frac{\widehat{\sigma}_{t-1}(y,y_t)}{\widehat{\sigma}^2_{t-1}(y_t)+\sigma^2_{\eta,t}}$ *be the weight put on* η_t *in the time-t estimate* $\widehat{c}_t(y)$*.*

- *If* $\psi = 0$*, then* $\alpha_t(y)$ *is just a constant – i.e.* $\alpha_t(y) = \alpha_t$ *for all* $y \in \mathbb{R}$*, and thus*

$$\widehat{c}_t(y) = \widehat{c}_0(y) + \sum_{k=1}^{t} \alpha_k \prod_{j=k+1}^{t} (1-\alpha_j) u_k$$

where $u_k = \eta_k - \widehat{c}_0(y_k)$ *is the deviation of signal* η_k *from the time-0 prior mean belief.*

- *If* $\psi > 0$*, then the informativeness of the signal* η_t *is state-dependent –* $\frac{\partial \alpha_t(y)}{\partial y} \neq 0$ *– and hence the shape of the time-t estimate* $\widehat{c}_t$ *differs from the time-0 prior, i.e.:*

$$\frac{\partial \left(\widehat{c}_t(y) - \widehat{c}_0(y)\right)}{\partial y} \neq 0$$

The effect of the information in η_t *is also local to* y_t*, since* $\lim_{||y-y_t||\to\infty} \alpha_t(y) = 0$

Proof. Consider the updated conditional estimate $\widehat{c}_t(y)$, where by Lemma 2 for any $y \in \mathbb{R}$:

$$\widehat{c}_t(y) = \widehat{c}_{t-1}(y) + \alpha_t(y)(\eta_t - \widehat{c}_{t-1}(y_t))$$

Iterating backwards, it follows that

$$\widehat{c}_t(y) = \widehat{c}_0(y) + \sum_{k=1}^{t} \alpha_k(y) \prod_{j=k+1}^{t} (1-\alpha_j(y)) u_k \tag{21}$$

where $u_k = \eta_k - \widehat{c}_0(y_k)$ is the deviation of signal η_k from the prior mean belief.

To prove the two parts of the Lemma, we will proceed by induction. First, if $\psi = 0$, for $t = 1$ we have

$$\alpha_1(y) = \frac{\widehat{\sigma}_0(y, y_1)}{\widehat{\sigma}^2_0(y_1) + \sigma^2_{\eta,1}} = \frac{\sigma^2_c}{\sigma^2_c + \sigma^2_{\eta,1}} \text{ for all } y$$

In this case, $\alpha_1(y) = \alpha_1$ is just a constant. Moreover, the updated covariance function is

$$\widehat{\sigma}_1(y, y') = \widehat{\sigma}_0(y, y') - \alpha_1 \widehat{\sigma}_0(y, y') = \widehat{\sigma}^2_1 = \sigma^2_c(1-\alpha_1)$$

which is again a constant independent of y and y'. Now consider the induction step; assuming that

$$\alpha_k(y) = \alpha_k \text{ for all } y \text{ and } k < t,$$

$$\widehat{\sigma}_{t-1}(y, y') = \widehat{\sigma}^2_{t-1} \equiv \sigma^2_c \prod_{k=1}^{t-1} (1-\alpha_k),$$

it follows that the effective signal-to-noise ratio for the time t signal is again a constant invariant to y:

$$\alpha_t(y) = \alpha_t \equiv \frac{\widehat{\sigma}_{t-1}(y, y_t)}{\widehat{\sigma}^2_{t-1}(y_t) + \sigma^2_{\eta,t}} = \frac{\widehat{\sigma}^2_{t-1}}{\widehat{\sigma}^2_{t-1} + \sigma^2_{\eta,t}}$$

Similarly, the resulting posterior variance at t is also invariant to y:

$$\widehat{\sigma}_t(y,y') = \widehat{\sigma}_{t-1} - \alpha_t \widehat{\sigma}_{t-1} = \widehat{\sigma}_t^2 = \widehat{\sigma}_{t-1}^2(1-\alpha_t)$$

Hence, $\alpha_{t+1}(y)$ is also a constant invariant to y and so on. Thus, for any time t the effective signal-to-noise ratio is invariant to y, hence the conditional estimate is simply a constant shift away from the time-0 prior, with the value of that shift given by a weighted average of signal surprises:

$$\widehat{c}_t(y) = \widehat{c}_0(y) + \sum_{k=1}^{t} \alpha_k \prod_{j=k+1}^{t} (1-\alpha_j) u_k$$

To prove the second part, when $\psi > 0$, we need to show that $\frac{\partial \alpha_t(y)}{\partial y} \neq 0$ almost everywhere and that $\lim_{||y-y'||\to\infty} \alpha_t(y) = 0$. We will do both by induction, and the key is the evolution of the conditional covariance $\widehat{\sigma}_{t-1}(y,y')$. Starting with the case of $t=1$, for any pair $y, y' \in \mathbb{R}$

$$\widehat{\sigma}_0(y,y') = \sigma_c^2 \exp(-\psi(y-y')^2)$$

which is decreasing in the distance $||y-y'||$, and $\frac{\partial \widehat{\sigma}_0(y,y')}{\partial y} \neq 0$ except for $y = y'$. The updated covariance function is

$$\widehat{\sigma}_1(y,y') = \sigma_c^2 \exp(-\psi(y-y')^2) - \frac{\sigma_c^4 \exp(-\psi\left((y-y_1)^2 + (y'-y_1)^2\right))}{\sigma_c^2 + \sigma_{\eta,1}^2}$$

hence similarly $\frac{\partial \widehat{\sigma}_1(y,y')}{\partial y} \neq 0$ outside of a measure 0 set and $\lim_{||y-y'||\to\infty} \widehat{\sigma}_1(y,y') = 0$.

On the other hand, given that

$$\alpha_1(y) = \frac{\sigma_c^2 \exp(-\psi(y-y_1)^2)}{\sigma_c^2 + \sigma_{\eta,1}^2}$$

we clearly have $\frac{\partial \alpha_1(y)}{\partial y} \neq 0$ except for when $y = y_1$, which is measure 0. Also, $\lim_{||y-y_1||\to\infty} \alpha_1(y) = 0$.

For the induction step, assume that $\frac{\partial \widehat{\sigma}_{t-1}(y,y')}{\partial y} \neq 0$ except for possibly on a set of measure 0, and that $\lim_{||y-y'||\to\infty} \widehat{\sigma}_{t-1}(y,y') = 0$. The updated covariance function is

$$\widehat{\sigma}_t(y,y') = \widehat{\sigma}_{t-1}(y,y') - \frac{\widehat{\sigma}_{t-1}(y,y_t)\widehat{\sigma}_{t-1}(y',y_t)}{\widehat{\sigma}_{t-1}^2(y_t) + \sigma_{\eta,1}^2}$$

thus, $\frac{\partial \widehat{\sigma}_t(y,y')}{\partial y} \neq 0$ except for possibly on a set of measure 0, and $\lim_{||y-y'||\to\infty} \widehat{\sigma}_t(y,y') = 0$.

Thus, for any arbitrary t, since $\alpha_t(y) = \frac{\sigma_{t-1}(y,y_t)}{\sigma_{t-1}^2(y_t) + \sigma_{\eta,t}^2}$,

$$\frac{\partial \alpha_t(y)}{\partial y} \neq 0$$

$$\lim_{||y-y_t||\to\infty} \widehat{\alpha}_t(y) = 0$$

Lastly, by equation (21),

$$\widehat{c}_t(y) - \widehat{c}_0(y) = \sum_{k=1}^{t} \alpha_k(y) \prod_{j=k+1}^{t} (1 - \alpha_j(y)) u_k$$

and since $\frac{\partial \alpha_t(y)}{\partial y} \neq 0$, it follows that

$$\frac{\partial \left(\widehat{c}_t(y) - \widehat{c}_0(y)\right)}{\partial y} \neq 0$$

□

Proposition 1. *The optimal signal noise variance is given by*

$$\sigma^{*2}_{\eta,t} = \begin{cases} \frac{\kappa \widehat{\sigma}^2_{t-1}(y_t)}{\widehat{\sigma}^2_{t-1}(y_t) - \kappa} & , \text{ if } \widehat{\sigma}^2_{t-1}(y_t) \geq \kappa \\ \infty & , \text{ if } \widehat{\sigma}^2_{t-1}(y_t) < \kappa \end{cases}$$

and this in turn implies the time-t action

$$c_t = \widehat{c}_t(y_t) = \widehat{c}_{t-1}(y_t) + \alpha^*_t(y_t)(c^*(y_t) + \varepsilon_t - \widehat{c}_{t-1}(y_t)),$$

where the optimal weight put on the new reasoning signal, $\alpha^*_t(y_t)$ depends on the current state y_t and the history $\{y^{t-1}, \sigma^{t-1}_\eta\}$ of past signals' location and precision:

$$\alpha^*_t(y_t) \equiv \frac{\widehat{\sigma}^2_{t-1}(y_t)}{\widehat{\sigma}^2_{t-1}(y_t) + \sigma^{*2}_{\eta,t}} = \max \left[1 - \frac{\kappa}{\widehat{\sigma}^2_{t-1}(y_t)}, 0\right].$$

Proof. The agent's reasoning decision is governed by

$$\min_{\widehat{\sigma}^2_t(y_t)} \widehat{\sigma}^2_t(y_t) + \kappa \ln \left(\frac{\widehat{\sigma}^2_{t-1}(y_t)}{\widehat{\sigma}^2_t(y_t)} \right). \tag{22}$$

s.t.

$$\widehat{\sigma}^2_t(y_t) \leq \widehat{\sigma}^2_{t-1}(y_t),$$

The first-order condition implies that the optimal posterior variance choice is

$$\widehat{\sigma}^2_t(y_t) = \min \left[\kappa, \widehat{\sigma}^2_{t-1}(y_t)\right] \tag{23}$$

By Lemma 2,

$$\widehat{\sigma}^2_t(y_t) = \frac{\widehat{\sigma}^2_{t-1}(y_t)\sigma^2_{\eta,t}}{\widehat{\sigma}^2_{t-1}(y_t) + \sigma^2_{\eta,t}}$$

Using this expression and equation (23), we obtain the expression for the optimal reasoning noise variance $\sigma^2_{\eta,t}$.

Similarly, using the solution for $\sigma^2_{\eta,t}$, Lemma 2 and the fact that $\eta_t = c^*(y_t) + \varepsilon_t$, it follows directly that

$$c_t = \widehat{c}_t(y_t) = \widehat{c}_{t-1}(y_t) + \max\left[1 - \frac{\kappa}{\widehat{\sigma}^2_{t-1}(y_t)}, 0\right](c^*(y_t) + \varepsilon_t - \widehat{c}_{t-1}(y_t)),$$

□

Proposition 2. *The optimal reasoning intensity and the weight of the new signal in updating beliefs are both increasing in distance from location of the previous reasoning signal:*

$$\frac{\partial \sigma^2_{\eta,i,2}}{\partial ||y_{i,2} - y_1||} < 0 \text{ and } \frac{\partial \alpha_{i,2}(y_{i,2})}{\partial ||y_{i,2} - y_1||} > 0.$$

Therefore, agent C reasons more than agent S, i.e. $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$, if and only if

$$s_2 < 1 + \frac{(1+r)}{w}(c^*(y_1) - c^{RW}(y_1)) \equiv \bar{s}$$

Proof. Consider the time-1 posterior variance function (expressed as a function of the distance $||y - y_1||$):

$$\widehat{\sigma}^2_1(y) = \sigma^2_c(1 - \frac{\sigma^2_c}{\sigma^2_c + \sigma^2_{\eta,1}} \exp\left(-2\psi||y - y_1||^2\right)$$

Note that this function is the same for both agents $i \in \{S, C\}$. Moreover,

$$\frac{\partial \widehat{\sigma}^2_1(y)}{\partial ||y - y_1||} = \frac{\sigma^4_c}{\sigma^2_c + \sigma^2_{\eta,1}} \exp\left(-2\psi||y - y_1||^2\right) 4\psi||y - y_1||^2 > 0$$

except for the knife-edge case $||y - y_1|| = 0$ when this derivative is zero. Using the expression for the optimal signal-noise variance $\sigma^2_{\eta,2}$ from Proposition 1:

$$\sigma^2_{\eta,i,2} = \frac{\kappa \widehat{\sigma}^2_1(y_{i,2})}{\widehat{\sigma}^2_1(y_{i,2}) - \kappa}$$

outside of the measure-zero case $y_{i,2} = y_1$. Then, it follows directly that

$$\frac{\partial \sigma^2_{\eta,i,2}}{\partial ||y_{i,2} - y_1||} = \frac{\kappa}{\widehat{\sigma}^2_1(y_{i,2}) - \kappa} \frac{\partial \widehat{\sigma}^2_1(y_{i,2})}{\partial ||y - y_1||} \underbrace{\left(1 - \frac{\widehat{\sigma}^2_1(y_{i,2})}{\widehat{\sigma}^2_1(y_{i,2}) - \kappa}\right)}_{<0} < 0$$

Similarly, outside of the measure-zero case $y_{i,2} = y_1$:

$$\alpha_{i,2}(y_{i,2}) = 1 - \frac{\kappa}{\widehat{\sigma}^2_{t-1}(y_{i,2})}$$

and thus

$$\frac{\partial \alpha_{i,2}(y_{i,2})}{\partial ||y - y_1||} = \frac{\kappa}{(\widehat{\sigma}^2_{t-1}(y_t))^2} \frac{\partial \widehat{\sigma}^2_1(y)}{\partial ||y - y_1||} > 0$$

For the last part of the proposition, note that

$$\alpha_{C,2}(y_{C,2}) - \alpha_{S,2}(y_{S,2}) = \frac{\kappa(\sigma_c^2 - \kappa)\left[\exp(-2\psi(y_{C,2} - y_1)^2) - \exp(-2\psi(y_{S,2} - y_1)^2)))\right]}{(\sigma_c^2 - (\sigma_c^2 - \kappa)\exp(-2\psi(y_1 - y_{C,2})^2))(\sigma_c^2 - (\sigma_c^2 - \kappa)\exp(-2\psi(y_1 - y_{S,2})^2))}$$

Hence, $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$ if and only if

$$\exp(-2\psi(y_{C,2} - y_1)^2) > \exp(-2\psi(y_{S,2} - y_1^2))$$
$$\iff$$
$$(y_{C,2} - y_1)^2 < (y_{C,2} - y_1)^2$$

substituting in the law of motion for cash-on-hand:

$$y_{i,2} = (1+r)(y_1 - \widehat{c}_{i,1}(y_1)) + ws_2 = (1+r)(y_1 - c^*(y_1) - \alpha_1\varepsilon_{i,1}) + ws_2$$

it follows that $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$ if and only if

$$s_2 < \frac{(1+r)c^*(y_1) - ry_1}{w} + \frac{(1+r)\alpha_1}{2w}\frac{\varepsilon_{C,1}^2 - \varepsilon_{S,1}^2}{\varepsilon_{C,1} + |\varepsilon_{S,1}|}$$

and using the definition of $c^{RW}(y)$ and our assumption that $\varepsilon_{C,1} = -\varepsilon_{S,1}$, it follows that $\alpha_{C,2}(y_{C,2}) > \alpha_{S,2}(y_{S,2})$ if and only if

$$s_2 < 1 + \frac{(1+r)}{w}(c^*(y_1) - c^{RW}(y_1))$$

□

B Calibration Strategy

To calibrate the learning-related parameters $\{\sigma_c^2, \psi, \kappa\}$ (given values for the rest of the parameters, which are discussed in Section 4.1) we look for a fixed point, such that (i) an econometrician trying to estimate the distribution of the reasoning signals $\eta_{i,t}$ would indeed recover the actual values of $\{\sigma_c^2, \psi\}$ used for the simulation and (ii) the model implies zero net-assets for the bottom 20% of the wealth distribution (matching the PSID data).

We are motivated to look for a fixed point in $\{\sigma_c^2, \psi\}$ in order to ensure that agents have model consistent priors, in the sense that the prior beliefs properly capture the features of the true $c^*(y)$ policy function. Intuitively, we would like to parameterize the prior in an "efficient" way so that it does not imply any ex-ante biases that impede the estimation process.

To do so, we consider the problem of an agnostic econometrician who attempts to estimate the function $c^*(y)$ out of a data set that constitutes the reasoning signals $\eta_{i,t}$ from the simulation of our model. The econometrician uses the same Bayesian methods as the agents in our model, and has a Gaussian Process prior over the unknown function c^*. The econometrician is "agnostic", however, in the sense that instead of treating the prior distribution of c^* as a primitive, he treats the prior as a "hyper-parameter" that is optimized over during the estimation procedure. In this way, the econometrician looks for the "optimal" or most efficient prior for the data at hand (which is the simulated reasoning signals from the ergodic distribution of the model).

In our calibration strategy, we want to ensure that the agents in our model indeed hold those same "optimal" priors. To do so, we look for a fixed point between the assumed parameterization of the agents' priors, i.e. $\{\sigma_c^2, \psi\}$, and the values for the prior hyper-parameters the agnostic econometrician recovers. – which we label $\{\tilde{\sigma}_c^{*,2}, \tilde{\psi}^*\}$.

In particular, this econometrician

1. is given the history of reasoning signals η_i^t for all agents in the simulated economy as a data set which we denote $\boldsymbol{\eta}$

2. is aware of the structure of the signals, i.e. that

$$\eta_{i,t} = c^*(y_{i,t}) + \varepsilon_{i,t} \ , \ \varepsilon_{i,t} \sim N(0, \sigma_{\eta,i,t}^2).$$

3. Observes $y_{i,t}$ and $\sigma_{\eta,i,t}^2$, which we collect in the observed vector $\mathbf{y}$, but is uncertain about the function $c^*(y)$ and holds the following Gaussian Process prior over it:

$$c^* \sim \mathcal{GP}(0, \tilde{\sigma}_0)$$

As is standard practice in Bayesian statistics, the econometrician assumes that the prior mean is the constant zero function – in this sense, he is "agnostic" about the particular functional form of c^*. The econometrician's prior is also parameterized by a squared exponential function, just as the agents in our model:

$$\tilde{\sigma}_0(y, y') = \tilde{\sigma}_c^2 \exp(-\tilde{\psi}(y - y')^2)$$

To differentiate with the agents' covariance function, we label the parameters of the econometrician's covariance function with tildes.

4. Given the collection of reasoning signals $\boldsymbol{\eta}$ the econometrician forms the posterior distribution $c^*|\boldsymbol{\eta}$, and finds the optimal hyper-parameters $\{\tilde{\sigma}_c^{*,2}, \tilde{\psi}^*\}$ by maximizing the resulting marginal likelihood of the data (as a function of $\{\tilde{\sigma}_c^2, \tilde{\psi}\}$):

$$\max_{\tilde{\sigma}_c^2, \tilde{\psi}} p(\boldsymbol{\eta}|\mathbf{y}, \tilde{\sigma}_c^2, \tilde{\psi}) = -\frac{1}{2}\mathbf{y}' K_\eta^1 \mathbf{y} - \frac{1}{2}\ln(K_\eta) - \frac{n}{2}\ln(2\pi)$$

where K_η is the covariance matrix of the econometrician's data vector $\boldsymbol{\eta}$, with (i, j) element

$$K_\eta(i, j) = \tilde{\sigma}_c^2 \exp\left(-\tilde{\psi}(\boldsymbol{y}(i) - \boldsymbol{y}(j))^2\right) + \mathbb{1}(i = j)\boldsymbol{\sigma}_\eta^2(i)$$

Note: since c^* has a Gaussian Process distribution with a squared exponential covariance function, the covariance between two data points $\boldsymbol{\eta}(i)$ and $\boldsymbol{\eta}(j)$ depends on the position of the y state values at which the two respective η signals are observed. The diagonal entries of K_η are also affected by the variance of the idiosyncratic reasoning noise $\sigma_{\eta,i,t}^2$, which the econometrician observes and takes into account.

5. This results in maximized values of the prior hyper-parameters $\tilde{\sigma}_c^2, \tilde{\psi}$:

$$\{\tilde{\sigma}_c^{*,2}, \tilde{\psi}^*\} = \arg\max_{\tilde{\sigma}_c^2, \tilde{\psi}} p(\boldsymbol{\eta}|\mathbf{y}, \tilde{\sigma}_c^2, \tilde{\psi})$$

Thus, for any given simulation of our model we can obtain the agnostic econometrician's inferred values of σ_c^2 and ψ by following steps 1-5 above. In addition, we can then vary the parameter κ, which controls the extent or magnitude of the reasoning friction, in order to hit the additional target of zero net-wealth for the poorest 20% of agents at the ergodic steady state of the model.

We use the following numerical strategy to find the necessary fixed-point in $\{\sigma_c^2, \psi, \kappa\}$:

(a) Given an initial guess for $\{\sigma_c^2, \psi, \kappa\}$ (taking rest of the parameters as given), we simulate the model using the benchmark simulation size of $T = 10,000, N = 5,000$.

(b) We discard the first half of the simulated time-series and are left with a 5000x5000 panel data set $\boldsymbol{\eta}$.

(c) This is a very large dataset, so to speed up the hyper-parameter estimation outlined above (and thus make our fixed point search feasible), we select a random samples of length $\frac{1}{\theta}$ (i.e. an average life-cycle of information) out of the full dataset $\boldsymbol{\eta}$ and perform steps 1-5 above on each of those samples.

We repeat, with replacement, 500 times and then take the average of the resulting 500 pairs of estimated hyper-parameters which we call $\{\bar{\sigma}_c^{*,2}, \bar{\psi}^*\}$.

(d) We check whether we have achieved a fixed point defined as satisfying both

(i) $||\{\sigma_c^2, \psi\} - \{\bar{\sigma}_c^{*,2}, \bar{\psi}^*\}|| < 1e - 5$

(ii) Share of total assets of bottom 20% of agents $< 1e - 5$

(e) if the two conditions above are satisfied we stop and use those coefficients $\{\sigma_c^2, \psi, \kappa\}$. If not, we update the guess of the parameter values as needed and go back to step (a).