Economic agents as imperfect problem solvers^{*}

Cosmin Ilut Duke University & NBER Rosen Valchev Boston College

September 2017

Abstract

In this paper we develop a tractable model where agents choose how much cognitively costly reasoning effort to allocate in computing the optimal policy function, even if objective state variables are perfectly observed. Information about the unknown policy function accumulates through time and it is most useful locally, around the state where reasoning occurs. A key property is that agents choose to reason more intensely when observing more unusual state, leading to state- and history-dependent priors over the optimal conditional action and as a result to a set of empirically plausible individual and aggregate behavior. In particular, the typical decision rule exhibits non-linearity, in the form of both inertial and 'salience' effects, stochastic choice and endogenous persistence in the action. In the cross-section, the different agents' prior experiences lead to predictable patterns of heterogeneity in biases, to endogenous persistence in the aggregate action, and to correlated clusters of volatility in the time-series and cross-sectional dispersion of actions. When extended to two actions, the model leads to state-dependent comovement and accuracy in different actions taken by the same agent.

^{*}Email addresses: Ilut cosmin.ilut@duke.edu, Valchev valchev@bc.edu. We would like to thank Ryan Chahrour, Philipp Sadowski, Todd Sarver and Mirko Wiederholt, as well as conference participants at the Society of Economic Dynamics and Computing in Economics and Finance for helpful discussions and comments.

1 Introduction

Contrary to standard models where rational agents act optimally, in both real world and in experimental settings, economic agents often choose what appear to be sub-optimal actions, especially when faced with complex situations. To account for that, the literature has become increasingly interested in modeling cognitive limitations that constrain the agents' ability to reach the full information rational decision. This interest has come from different fields including decision theory, behavioral economics, macroeconomics, finance and neuroscience and has led to a diverse set of approaches. The common principle of these approaches is that agents have limited cognitive resources to process payoff relevant information, and thus face a trade-off in the accuracy of their eventual decision and the cognitive cost of reaching it.

A key modeling choice is the nature of the costly payoff-relevant information that agents can choose to acquire. In general, a decision can be represented as a mapping from the information about the objective states of the world to the set of considered actions. Therefore, cognitive limitations in reaching the optimal decision may be relevant to two layers of imperfect perception. First, compared to an analyst that perfectly observes the objective states, such as income or interest rates, agent's cognitive limitations may imply limited perception about these state variables. Second, compared to an analyst that perfectly computes the agent's optimal action, such as consumption or labor, conditional on possibly noisy information about the states, cognitive limitations may imply limited perception about the states, cognitive limitations may imply limited perception about the mapping that achieves those conditional optimal actions.

The standard approach in the macroeconomics literature is to focus on the first layer of uncertainty and assume that agents perceive the objective states of the world with noise, but use a mapping of information about the states to actions derived under full rationality. Therefore, an analysts recovers errors in the resulting actions as arising from the imperfect perception of the objective state variables. This standard approach is exemplified by a literature inspired by the Rational Inattention model of Sims (1998, 2003).¹A similar idea of imperfect attention to state variables is also present in the sparse optimization framework of Gabaix (2014, 2016), and more generally, the idea of allowing the agents to choose their information about the unknown state variables is present in other parts of the literature.²

In this paper, we develop a tractable framework that focuses on the second layer of

¹Applications in macroeconomics include consumption dynamics (Luo (2008), Tutino (2013)), price setting (Maćkowiak and Wiederholt (2009), Stevens (2014), Matějka (2015)), monetary policy (Woodford (2009), Paciello and Wiederholt (2013)), business cycle dynamics (Melosi (2014), Maćkowiak and Wiederholt (2015a)) and portfolio choice (van Nieuwerburgh and Veldkamp (2009, 2010), Kacperczyk et al. (2016)). See Wiederholt (2010) and Sims (2010) for recent surveys on the rational inattention in macroeconomics literature.

²See for example Reis (2006a,b) and Woodford (2003). See Veldkamp (2011) for a review on imperfect information in macroeconomics and finance.

imperfect perception. In our model agents observe all relevant objective state variables perfectly, in the same way as an econometrician does. However, agents have limited cognitive resources that prevent them from computing their optimal policy *function* and coming up with the fully specified, optimal state-contingent plan of action.³ Agents can expend costly reasoning effort that helps reduce the uncertainty over the unknown best course of action, and do so optimally.⁴ Given their chosen reasoning effort, agents receive signals on the best course of action at the current state of the world, and this knowledge accumulates over time. Thus, while being imperfect problem solvers, agents are 'procedurally rational' in the sense of Simon (1976) and exhibit behavior that is the outcome of appropriate deliberation.

The defining feature of our framework is the agent's accumulation of knowledge about the best response to the observed state. Tractability is partly attained by our focus on the linear-quadratic (LQ) Gaussian framework that has become the canonical setup in the broader Rational Inattention literature.⁵ Formally, we model the unknown optimal policy function as a draw from a Gaussian Process distribution over which the agent updates beliefs.⁶ A more intense reasoning effort is beneficial because it lowers the variance of the noise in the incoming signals. In turn, a more intense reasoning is costly, which we model as a cost on the total amount of information about the optimal action carried in the new signal. Using a standard approach in the Rational Inattention literature, we measure information flow with the reduction in entropy, i.e. Shannon mutual information, about the unknown function.⁷

We focus on a simple environment that highlights the specific endogenous features generated by the cognition friction. For this purpose, the marginal benefit and marginal cost of reasoning are set to be time-invariant, the true optimal decision rule is a linear function, and for most of the analysis the state itself is iid. Finally, the time zero prior belief is also

³Such a friction is consistent with a large field and laboratory experimental literature on how the quality of decision making is negatively affected by the complexity of the decision problem. See for example Caplin et al. (2011), Kalayci and Serra-Garcia (2016) and Carvalho and Silverman (2017)) for experimental evidence. Sethi-Iyengar et al. (2004), Abaluck and Gruber (2011) and Schram and Sonnemans (2011) document significantly poor choices for savings and healthcare plans. Carlin et al. (2013) finds that complexity in asset valuation leads to significant decreases in the trade efficiency. More generally, see Deck and Jahedi (2015) for a recent survey of experiments on complexity and decision making.

⁴Generally, reasoning processes are characterized in the literature as 'fact-free' learning (see Aragones et al. (2005)), i.e. because of cognitive limitations additional deliberation helps the agent get closer to the optimal decision even without additional objective information as observed by an econometrician.

⁵This also helps highlight novel predictions for behavior and contrast with the previous literature.

⁶Intuitively, a Gaussian Process distribution models a function as a vector of infinite length, where the vector has a joint Gaussian distribution. In addition to its wide-spread use in Bayesian statistics, Gaussian Processes have also been applied in machine learning over unknown functional relationships – both in terms of supervised (Rasmussen and Williams (2006)) and non-supervised (Bishop (2006)) learning.

⁷Following Sims (2003) a large literature has studied the choice properties of attention costs based on the Shannon mutual information between prior and posterior beliefs. See for example Matějka and McKay (2014) Caplin et al. (2016), Woodford (2014) and Matějka et al. (2017).

parsimonious. We set the initial mean function of the Gaussian process as a constant so that the average prior belief equals the average optimal action.⁸ The prior covariance function is controlled by a key smoothing parameter that encodes the correlation between beliefs over the actions at different values of the state. Importantly, since there is learning about a function, the information acquired is most useful *locally* to the state realization where reasoning occurs.

The emerging key property of the decision to reason is intuitive: agents find it optimal to reason more intensely when observing more unusual state realizations. These are states at which the agents, given their history of reasoning entering that period, have a higher prior uncertainty over what is the best course of action, conditional on the observed state. Thus, deliberation choice and actions are both state and history dependent. Because of this, we focus on analyzing implications about the agent's behavior at the ergodic distribution, where the agent has seen a long history of signals that were all the outcome of optimal deliberation at the time. In other words, we look for the typical behavior of the agent.

At an individual level the typical action has some key properties. First, it is non-linear and it exhibits inertia and salience effects. At the states that have been deliberated at most often in the past, which are concentrated around the unconditional mean of the state, the agent has accumulated enough information that he chooses not to reason further and instead relies on the prior experiences. In turn, these prior signals point towards an effective action that is relatively unresponsive to changes in the state.⁹ Intuitively, the agent has a good understanding of the average level of the optimal action around the mean value of the state, hence there is not much reason to think hard about the way the policy function changes with small movements in the state. Thus, his effective action is relatively unresponsive in the neighborhood of the average state. In contrast, the action exhibits salience effects at more unusual states, which tend to stand out given his ergodic beliefs. There the agent reasons more intensely, leading to more informative signals, which point towards a more responsive conditional action than around the usual states.¹⁰

Second, there is endogenous persistence in the average action. This mechanism naturally arises as long as the information acquired about the optimal action at some particular state realization is perceived to be informative about the optimal action at a different state

⁸The flatness of the prior mean captures the no prior "free" information about the shape of the optimal policy function. Any such understanding comes at the cost of deliberation effort, and the agent is free to engage in that, solving a trade-off that we characterize.

⁹Such an inertial behavior is consistent with a widely documented 'status-quo' bias (Samuelson and Zeckhauser (1988)). The reliance on prior experiences and lack of additional reasoning characterizing inertia in our model relates to a broader literature on cognitive limitations according to which status quo alternatives require less mental effort to maintain (see Eidelman and Crandall (2009) for a survey).

¹⁰The stronger state-dependent response is consistent with the so-called salience bias, or availability heuristic (Tversky and Kahneman (1975)), that makes behavior more sensitive to vivid, salient events.

realization. Moreover, the interplay between the flat responses around the usual states and the salience effects at the more unusual states generate local convexities in the ergodic policy function. The average movement in the next period action may be dominated by salience type of reasoning so that, even if the exogenous state is mean reverting, the agent takes an even more reactive action than before, leading to hump-shape dynamics.

Third, the endogenous persistence also manifests in the volatility of actions. In times of more volatile state realizations the agent not only takes more volatile actions but also chooses to reason more intensely. The more frequent unusual state realizations lead to more frequent careful deliberation, which results in a posterior belief that the optimal action is more responsive. As a consequence, in the periods following such volatile times, the actions appear to respond more to states, leading to clusters of volatility in actions.

Stochastic choice is an important behavior characteristic observed in experiments.¹¹ In our model, an individual action is stochastic even conditioning on the perfectly observed state, as it is driven by the random reasoning signals. At the same time, agents tend to observe different histories of reasoning signals leading to different policy functions in the cross-section. Therefore, an analyst observes variation in behavior not only as a current stochastic element of choice for a given individual, but also in the systematic part of decision rules, through the cross-sectional heterogeneity in the priors entering that period.

We study cross-sectional effects by introducing a continuum of ex-ante identical agents who solve the same reasoning problem by observing the same state. However, these agents differ in the specific history of reasoning signals that they have obtained about the otherwise identical optimal policy function. At the more unusual state realizations, where there has been less relevant information accumulated through time, the agents' priors are more anchored by their common initial prior so that the dispersion of beliefs entering the period tends to be smaller. At the same time, at these states agents decide to rely more on their newly obtained idiosyncratic reasoning signals. We find that typically the latter effect dominates so that the cross-sectional dispersion of actions is larger at the more unusual states.¹²

Overall, we find that costly cognition can act as a parsimonious friction that generates several important characteristics of many aggregate time-series: endogenous persistence, non-linearity and volatility clustering, as well as possible hump-shape dynamics. Therefore, our findings connect three important directions in macroeconomics. One is a large literature that proposes a set of frictions, typically taking the form of adjustment costs, to explain sluggish behavior (eg. investment adjustment cost, habit formation in consumption, rigidities

¹¹See Mosteller and Nogee (1951) and more recently Hey (2001) and Ballinger and Wilcox (1997).

¹²The implied correlated volatility at the 'micro' (cross-sectional dispersion) and 'macro' (volatility clustering of the aggregate action) level is consistent with recent evidence surveyed in Bloom (2014).

in changing prices or wages). A second direction is to introduce frictions aimed at obtaining non-linear dynamics (eg. search and matching labor models, financial constraints).¹³ Third, there is a significant literature that documents and models exogenous variation in volatility and structural parameters to better fit macroeconomic data.¹⁴

Compared to the literature on imperfect attention to actions, the key property of our model is in the dynamics of the prior distribution over actions. In comparison to the standard Rational Inattention literature in macroeconomics, which analyzes imperfect perception of objective states, the prior and the deliberation choice in our model is conditional on such a state.¹⁵ Our emphasis on reasoning about optimal policy functions builds on a literature, mostly in decision theory, that analyzes the costly perception of unknown subjective states.¹⁶ We contribute to both of these strands of literature by developing a tractable model of accumulation of information about the optimal action in the form of an unknown function of the objective state. We characterize how this form of learning leads to endogenous timevariation in the mean and variance of the state-dependent prior distributions. In particular, these state-dependent priors are instrumental to producing non-linear action responses. Indeed, when we shut down the accumulation of information about the optimal decision rule, we recover linear actions and uniform under-reaction to the state, a result that is typical in the standard LQ Gaussian Rational Inattention analysis.¹⁷

The proposed cognition friction also relates to two types of literatures on bounded rationality in macroeconomics. One is a 'near-rational' approach that assumes a constant cost of implementing the otherwise known optimal action. There the interest is in studying the general equilibrium (GE) effects of the resulting individual errors, taken as state-independent stochastic forces.¹⁸ A second recent literature addresses the potential complexities of comput-

 16 As in the model of costly contemplation over tastes in Ergin and Sarver (2010) or the rationally inattentive axiomatization in Oliveira et al. (2017). Alaoui and Penta (2016) analyze a reasoning model where acquiring information about subjective mental states indexing payoffs is costly, but leads to higher accuracy.

 $^{^{13}\}mathrm{See}$ Fernández-Villaver de et al. (2016) for a recent survey of non-linear methods.

¹⁴For example Stock and Watson (2002), Cogley and Sargent (2005) and Justiniano and Primiceri (2008).

¹⁵In that work the choice of attention to a state realization is typically made ex-ante, conditional on a prior distribution over those realizations. Some recent models of ex-post attention choices, i.e conditional on the current realization, include: the sparse operator of Gabaix (2014, 2016) that can be applied ex-ante or ex-post to observing the state, information processing about the optimal action after a rare regime realizes (Maćkowiak and Wiederholt (2015b)) and a decision over which information provider to use (Nimark and Pitschner (2017)). The latter model can generate stronger agents' responses to more extreme events because these events are more likely to be widely reported and be closer to common knowledge. Nimark (2014) obtains these stronger responses as a result of an assumed information structure where signals are more likely to be available about more unusual events.

¹⁷While there is interest in departures from that result, most of the focus in the literature has been on obtaining optimal information structures that are non-Gaussian, taking the form of a discrete support for signals (see for example Sims (2006), Stevens (2014) and Matějka (2015)). Instead, we maintain the tractability of the LQ Gaussian setup but obtain non-linear dynamics.

¹⁸See for example Akerlof and Yellen (1985a,b), Dupor (2005) and Hassan and Mertens (2017).

ing GE effects faced by an agent with different forms of bounded rationality.¹⁹ There the individual decision rule is the same as the fully rational one but the GE effects are typically dampened. Compared to these two approaches, we share an interest in cognitive costs, and, while we abstract from GE effects for our aggregate-level implications, we focus on how imperfect reasoning generates an endogenous structure of errors in actions.

Finally, we extend our analysis to two actions that differ in their marginal cost of making errors. We show how the optimal deliberation model implies state-dependent comovement between actions because the information flow friction is not over the single state variable but it is specific to the cognitive effort it takes to make decisions about the two actions. The time-varying comovement stands in contrast with the standard approach in the literature of assuming an attention friction on the state, but endowing the agent with full knowledge of the two individual optimal policy functions. In that case, the mistakes in the two actions tend to be perfectly correlated, being driven by imperfect observation of the same state variable.

The paper is organized as follows. In Section 2, we develop our model of cognitively costly reasoning. In Section 3 we describe key implications at the individual and aggregate level. In Section 4 we present an extension of the model to two actions.

2 Model of Cognitively Costly Reasoning

In this section we develop our costly decision-making framework. To ease comparisons with the existing literature, we focus on what has become the canonical setup in the Rational Inattention literature and use a quadratic tracking framework with Gaussian uncertainty. Section 2.1 describes the basic setup and Section 2.2 presents the optimal deliberation choice.

2.1 Quadratic-Gaussian Framework

Our focus is on limiting the information flow not about an objective state but instead about the policy function. We model the tracking problem of an agent that knows the current value of the objective state, y_t , but does not know the optimal policy function $c^*(y_t)$ and chooses his actual action $\hat{c}(y_t)$ to minimize expected quadratic deviations from the optimum:

$$U = \min_{\hat{c}_t} W_{cc} E_t (\hat{c}(y_t) - c^*(y_t))^2.$$

¹⁹These include reflective equilibrium (García-Schmidt and Woodford (2015)), level k-thinking (Farhi and Werning (2017)), or lack of common knowledge as a result of cognitive limits (Angeletos and Lian (2017)).

The parameter $W_{cc} < 0$ measures the utility cost of suboptimal actions. We focus on this quadratic setup in order to present the mechanism in the most transparent way. In addition, it also facilitates comparisons with the existing Rational Inattention literature, since this setup has become the canonical framework in that literature following Sims (2003).²⁰

Given the tracking problem, the agent optimally chooses to act according to his conditional expectation of the true, unknown optimal action:

$$\hat{c}(y_t) = E_t(c^*(y_t))$$

2.1.1 Prior beliefs about the optimal action

The agent's prior beliefs over the unknown function $c^*(y_t)$ are given by a Gaussian Process distribution with a mean function $\mu(y)$ and a covariance function $\Sigma(y, y')$:

$$c^*(y) \sim \mathcal{GP}(\mu(y), \Sigma(y, y')),$$

where $\mu(y)$ specifies the unconditional mean of the function for any input value y,

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

and the covariance function specifies the unconditional covariance between the values of the function at any pair of inputs y and y':

$$\Sigma(y, y') = E\left((c^*(y) - \bar{c}(y))(c^*(y') - \bar{c}(y'))\right).$$

A Gaussian Process distribution is the generalization of the Gaussian distribution to infinite-sized collections of real-valued random variables, and is often used as a prior for Bayesian inference on functions (Liu et al. (2011)). Intuitively, a Gaussian Process distribution models a function as a vector of infinite length, where the whole vector has a joint Gaussian distribution. Often, (especially in high frequency econometrics) Gaussian Processes are defined as a function of time – e.g. Brownian motion. In this paper, however, we use it as a convenient and tractable way of modeling the uncertainty of the agent over his unknown policy function $c^*(y)$ and thus the indexing set is the real line (and in more general applications with multiple state variables also \mathbb{R}^N , see Section 4 below).

²⁰The framework could be viewed as a quadratic approximation to the value function of the agent, and W_{cc} as the second derivative of the value function with respect to the action. While extensions are possible, this relatively simple setup is already rich in economic insights. At a more general level, while we do not formalize the complexities in the type of tracking problems that the agent may choose to solve, we, as model builders, share a similar interest in tractability as the agent inside the model.

A Gaussian process is completely characterized by its mean and covariance functions, so that for any finite collection of points in the domain of $c^*(.)$, $\mathbf{y} = [y_1, \ldots, y_N]$, the resulting distribution of the vector of function values $c(\mathbf{y})$ is a joint normal distribution given by:

$$c^*(\mathbf{y}) \sim N\left(\begin{bmatrix} \mu(y_1) \\ \vdots \\ \mu(y_N) \end{bmatrix}, \begin{bmatrix} \Sigma(y_1, y_1) & \dots & \Sigma(y_1, y_N) \\ \vdots & \ddots & \vdots \\ \Sigma(y_N, y_1) & \dots & \Sigma(y_N, y_N) \end{bmatrix} \right).$$

Using this feature, we can draw samples from the distribution of functions evaluated at any arbitrary finite set of points \mathbf{y} , and hence this fully describes the prior uncertainty of the agent about the underlying policy function $c^*(y)$.

We assume that the agent's prior beliefs are centered around the steady state optimal action which we call \bar{c} . Note that this does not mean that the agent *knows* the optimal action at the steady state value of the state variable \bar{y} – he still faces uncertainty about that as the prior variance is non-zero. We are simply making the minimal assumption that the *average* prior belief is equal to the steady state optimal action \bar{c} .²¹ Formally, the prior mean function $\mu(y)$ is equal to the constant function \bar{c} :

$$\mu(y) = \bar{c} , \ \forall y.$$

Thus, the agent's prior beliefs are appropriately centered on average. Moreover, they also do not respond to movements in y. If the prior belief about $c^*(y)$ was itself a function of y, then that would be akin to giving the agent information about his optimal action "for free", since the knowledge of y will imply a change in the belief about c^* . We want to avoid any such correlation between the action and the state in the agent's prior beliefs, since the basic idea of our framework is to make any information about $c^*(y)$ subject to costly deliberation. Thus, without investing any effort in thinking about what to do, the agent is unsure how to adjust his action as y varies, but his prior belief is unbiased and correct on average.²²

Importantly, the prior mean function $\mu(y)$ is only the prior belief of an agent that has spent no time thinking about the economic environment and optimal decision making. In our framework, however, the agent optimally deliberates and accumulates information over time, and thus the typical information set he enters a period with is very different from $\mu(y)$. Characterizing this typical information set and resulting behavior at the ergodic steady state distribution of beliefs is the main focus of the paper.

²¹This parallels the typical Rational Expectations assumption that expectations are correct on average – since \bar{c} is the true optimal action on average, the agent's prior beliefs are indeed correct on average.

²²Building biases in the average prior belief is nevertheless possible in the analysis.

The other important component of the agent's prior beliefs is the covariance function. It determines how new information itself will be interpreted and combined with the prior to form the posterior beliefs of the agent. We assume that the covariance function is of the widely used squared exponential class (see Rasmussen and Williams (2006)):

$$\Sigma(y, y') = \sigma_c^2 \exp(-\psi(y - y'^2))$$

The function has two parameters $-\sigma_c^2$ controls the prior variance or uncertainty about the value of $c^*(y)$ at any given point y, and ψ controls the smoothness of the function and the extent to which information about the value of the function at point y is informative about its value at a different point y'. A Gaussian Process with a higher ψ has a higher rate of change and its value is more likely to experience a big change for the same change in y. In particular, it can be shown that the mean number of zero-crossings over a unit interval is given by $\frac{\psi}{\sqrt{2\pi}}$ (see Rasmussen and Williams (2006) for details). Hence the larger is ψ , the more "wiggly" is the average function drawn from that Gaussian Process distribution, and thus the smaller is the correlation between the function values at any pair of distinct points. Intuitively, a higher ψ parameterizes a prior belief that the underlying function $\hat{c}^*(y)$ could have a higher derivative and thus information about the optimal action at one value of the state y is less useful for inferring the optimal action at another value y'. This parameter can have profound effects on the deliberation choice of the agent, as discussed in detail below.²³

2.1.2 Costly Deliberation

A key feature of the model is the costly deliberation choice. The agent does not simply act on his prior beliefs about the unknown optimal action $c^*(y)$, but can expend costly cognitive resources to think about his underlying economic problem, and obtain a better handle of the unknown optimal policy $c^*(y_t)$. This is formalized by giving him access to unbiased signals about his actual optimal action (given state of the world y_t) of the form,

$$\eta(y_t) = c^*(y_t) + \varepsilon_t^{\eta},$$

where $\varepsilon_t^{\eta} \sim N(0, \sigma_{\eta,t}^2)$, and allowing him to choose the precision of those signals $\sigma_{\eta,t}^2$. The chosen precision of the signal models the amount of deliberation the agent does – the more time and effort he spends on thinking about what is the optimal action, the more precise is

²³We focus on the squared exponential covariance function because it presents a good trade-off between flexibility and the number of free parameters. However, our main results would hold under other covariance functions as well, as long as the correlation between $c^*(y)$ and $c^*(y')$ is decreasing in the distance between yand y' – which is true for most stationary covariance functions.

his resulting signal, and thus the more accurate are his resulting posterior beliefs and the effective action $\hat{c}(y_t) = E_t(c^*(y_t))$ that he takes.

However, the cognitive effort required to acquire a more accurate understanding of the optimal action is costly, which we model as a cost on the total amount of information about the optimal action carried in the signal η_t . We measure information flow with the reduction in entropy, i.e. Shannon mutual information, about $c^*(y_t)$. Mutual information is defined as

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

where H(X) denotes the entropy of a random variable X, and is a measure of uncertainty often used in Information Theory. Thus, we see that equation (1) measures the reduction in uncertainty about the unknown $c^*(y_t)$, given the history of past deliberation and resulting signals η^{t-1} , that is brought about by seeing the new signal η_t . This reduction in uncertainty, over and above what was already known is the new additional information contained in the current signal η_t and represents the effect of the additional deliberation performed in the current period. We model the cognitive deliberation cost facing the agent as an increasing function of $I(c^*(y_t); \eta(y_t)|\eta^{t-1})$, i.e. the informativeness of the chosen signal η_t .

Two remarks are in order. First, we do not explicitly model the mental deliberation process and the specific mental processes involved in it. Instead, we model the accuracy-cognitive effort trade-off inherent in any reasonable such process. Thus we do not make strong assumptions about the particular way people reason, but simply assume that obtaining a more accurate plan of action, in the sense of getting closer to the true optimal action, takes more mental effort, and is thus costlier.²⁴ Second, we have chosen to formalize the notion of information flow with the Shannon mutual information measure in part because it leads to particularly transparent results, but the qualitative features of the model are unchanged as long as the agent finds it costly to increase the precision of his posterior beliefs about $c^*(y_t)$.

2.1.3 Updating

In this section we detail how the beliefs of an agent evolve as he acquires information about his optimal policy function $c^*(y)$. For illustration purposes, first we consider the case where the agent receives signals with exogenously fixed precision and later turn to the question of optimal deliberation choice, which endogenizes the precision of η .

The key feature of the beliefs updating process is that since the agent is learning about the function $c^*(y)$, the information he acquires through his deliberation process is most useful

 $^{^{24}}$ Costly cognition of the optimal action may reflect for example costly contemplation over tastes (Ergin and Sarver (2010)), or the costly search and satisficing heuristic of Simon (1955), and more generally acquiring costly information over mental states indexing subjectively perceived payoffs (Alaoui and Penta (2016)).

locally to the actual value of the state y_t at which the deliberation occurred. In other words, if the agent spends some time thinking about what to do when the state of nature is $y_t = y$, this will clearly be informative not only about the optimal action at y itself, but also about the optimal action at other values of the state $y' \neq y$ close to y. However, the original deliberation is going to be increasingly less informative for values of y' further away from y, and the rate at which the informativeness decays with the distance between y and y' will depend on the covariance function $\Sigma(y, y')$.

For illustration purposes we consider a couple of specific example. We start with the simple case of updating the prior with a single signal $\eta_0(y_0)$ at some value of the state y_0 . One can think of this as the beginning of time, when the agent has done no prior deliberation and only has his prior beliefs characterized by his ex-ante prior functions $\mu(.)$ and $\Sigma(.,.)$. For an arbitrary y the joint distribution of $c^*(y)$ and the signal $\eta_0(y_0)$ is jointly Gaussian:

$$\begin{bmatrix} c^*(y) \\ \eta_0(y_0) \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{c} \\ \bar{c} \end{bmatrix}, \begin{bmatrix} \sigma_c^2 & \sigma_c^2 \exp(-\psi(y-y_0)^2) \\ \sigma_c^2 \exp(-\psi(y-y_0)^2) & \sigma_c^2 + \sigma_{\eta,1}^2 \end{bmatrix}\right)$$

From this, it further follows that the conditional mean of $c^*(y)$ given η_0 is

$$\hat{c}_0(y) = \mathbb{E}(c^*(y)|\eta_0) = \bar{c} + \underbrace{\frac{\sigma_c^2 e^{-\psi(y-y_0)^2}}{\sigma_c^2 + \sigma_{\eta,0}^2}}_{=\alpha_0(y,y_0)} (\eta_0 - \bar{c})$$

Notice that when updating the belief about the optimal action at y_0 , the specific value of the state at which the agent has deliberated, the updating formula reduces to the familiar Bayesian one based on the signal-to-noise ratio $\alpha_0 = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\eta,0}^2}$:

$$\hat{c}_0(y_0) = \mathbb{E}(c^*(y_0)|\eta_0) = \left(1 - \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\eta,0}^2}\right)\bar{c} + \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\eta,0}^2}\eta_0$$
(2)

However, when updating beliefs of $c^*(.)$ at an arbitrary value y, the effective signal-tonoise ratio of the signal is a decreasing function of the distance between y and y_0 ,

$$\alpha_0(y, y_0) = e^{-\psi(y-y_0)^2} \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\eta,0}^2}.$$

This shows how the informativeness of η_0 is dissipating as we move away from y_0 . Naturally, this effect also shows up in the posterior variance:

$$\hat{\sigma}_0^2(y) = \operatorname{Var}(c^*(y)|\eta_0) = \sigma_c^2 \left(1 - \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{\eta,0}^2} e^{-2\psi(y-y_0)^2}\right)$$



Figure 1: Posterior Mean and Variance. The figure illustrates an example with parameters $\sigma_c^2 = 1, \psi = 1$, and $\sigma_{\eta,0}^2 \in \{\sigma_c^2, \frac{\sigma_c^2}{4}\}$

Figure 1 illustrates both of these effects. The left panel shows the posterior mean, $\mu_0(y)$, while the right panel plots the posterior variance $\hat{\sigma}_0^2(y)$. The figure is drawn for an example where the true optimal action $c^*(y) = y$, and the realization of the signal η_0 equals the truth, so $\eta_0 = y_0$. Both figures draw the posterior moments for two values of the signal error variance, one high and one low, to showcase how increased precision affects them.

In the left panel, the solid lines plot the resulting posterior means, the red one corresponding to the more precise signal, and the yellow dashed line plots the true $c^*(y)$, while the purple dash-dot line plots the prior mean function $\mu(y) = \bar{c}$. We can see two important things. First, the signal clearly has a stronger effect on the posterior mean at values of ycloser to y_0 – it pulls the conditional expectation further away from its prior mean more strongly the closer is the underlying y to y_0 . And naturally, the red line lies closer to the value of the signal, since in that case there is a higher signal precision, and hence the agent puts more weight on it. Second, observing just one signal is only weakly informative about the *shape* of the function $c^*(y)$, although it is useful in determining the *level* of the optimal action in the neighborhood of the signal. The positive realization of η_0 (relative to the prior \bar{c}) tends to increase the posterior mean belief for all values of y, and hence it displays only a relatively gentle slope upwards. In other words, the positive signal leads the agent to update beliefs that in general the optimal action is higher than previously thought, but by itself does not yield too much information about the shape of the policy function. As we will see below, for the agent to acquire a good grasp of the shape of the unknown function $c^*(y)$, he needs several precise signals at different values of y.

In the right panel, we can directly see the effect of the local reduction in uncertainty. The posterior variance $\hat{\sigma}_0^2(y)$ is lowest at y_0 , and increases as we move away from y_0 . Hence, while the signal is helpful in learning about the value of the policy function locally, it does not provide too much information about the optimal response at values of y in the opposite extreme of the spectrum. Naturally, we also see that the more precise signal tends to have a stronger effect on the variance, and perhaps even stronger local effects.

In the next example, we add a second signal, η_1 , at some other value of the state y_1 . In this case, we use the fact that the 3 element vector $[c^*(y), \eta_0, \eta_1]'$ is jointly Gaussian, and apply the standard Bayesian updating formulas to obtain the posterior mean and variance:

$$\hat{c}_{1}(y) = \mathbb{E}(c^{*}(y)|\eta_{0},\eta_{1}) = \bar{c} + \underbrace{\frac{\sigma_{c}^{2}(e^{-\psi(y-y_{0})^{2}} + \frac{\sigma_{c}^{2}}{\sigma_{c}^{2} + \sigma_{\eta,1}^{2}}e^{-\psi((y-y_{1})^{2} + (y_{1}-y_{0})^{2})}}_{=\alpha_{10}(y;y_{0},y_{1})}}(\eta_{0} - \bar{c}) + \underbrace{\frac{\sigma_{c}^{2}(e^{-\psi(y-y_{1})^{2}} + \frac{\sigma_{c}^{2}}{\sigma_{c}^{2} + \sigma_{\eta,0}^{2}}e^{-\psi((y-y_{0})^{2} + (y_{1}-y_{0})^{2})}}_{=\alpha_{11}(y;y_{0},y_{1})}}(\eta_{1} - \bar{c})}_{=\alpha_{11}(y;y_{0},y_{1})}$$

$$\hat{\sigma}_1^2(y) = \operatorname{Var}(c^*(y)|\eta_0, \eta_1) = \sigma_c^2 \left(1 - (\alpha_{10}(y; y_0, y_1)e^{-\psi(y-y_0)^2} + \alpha_{11}(y; y_0, y_1)e^{-\psi(y-y_1)^2} \right).$$

where we define the notation $\alpha_{10}(y; y_0, y_1)$ and $\alpha_{11}(y; y_0, y_1)$ as the effective signal-to-noise ratios of the two signals, η_0 and η_1 respectively, when updating beliefs about $c^*(y)$. Even though the formulas are now more complicated, the updating equations have much of the same features as before. Signals are more informative locally, and in particular, the agent puts relatively more weight on the signal η_0 , as opposed to η_1 , if y_0 is relatively closer to ythan is y_1 . So for y between y_0 and y_1 , the agent may lean on both signals about equally, but for y's close to y_1 he relies mostly on η_1 . If we further take the limit of either y_1 or y_0 going to infinity, its weight in the updating equation falls to zero, while the weight on the other signal converges to the weight in the single signal case. Similarly, the posterior variance is affected more strongly by one or the other signal, depending on whether y is closer to y_0 or y_1 .

We illustrate this example in Figure 2. We plot the posterior mean in the left panel and the posterior variance in the right panel, and for this example we again use signal realizations that are equal to the truth, $\eta_0 = y_0$ and $\eta_1 = y_1$. Lastly, both η_0 and η_1 carry the same precision. We keep the location of η_0 as before, and consider a y_1 that is its mirror image around the mean, i.e. lays below the mean at the same distance that y_0 is above the mean \bar{y} .



Figure 2: Posterior Mean and Variance after seeing two signals at distinct $y_0 \neq y_1$. The figure illustrates an example with parameters $\sigma_c^2 = 1, \psi = 1$, and $\sigma_{\eta,0}^2 \in \{\sigma_c^2, \frac{\sigma_c^2}{4}\}$

Notice that the posterior mean now displays a steeper slope, and captures the overall shape of the true $c^*(y)$ better than in the single signal case. This is because now the agent has learned that his optimal action is relatively high for high realizations of y, but relatively low for low realizations. Thus, he can deduce that the optimal policy is most likely upward sloping between his two signals. This is different from the one signal case, where the single signal observation was not sufficient to provide good information about the shape of the unknown $c^*(y)$. It mainly served to update beliefs about the overall level of the policy function. Here, however, the posterior mean does not display a bias in its overall level, since given the two distinct observations, the agent is able to infer that the function is not higher than expected under the prior, but rather has a slope. Accumulating signals at distinct points is effective in allowing the agent to learn about the shape of the underlying optimal policy function.

We can also see that the posterior mean belief is closest to the truth in the interval between y_0 and y_1 . This is because for $y \in [y_0, y_1]$, the agent is able to lean strongly on his knowledge of η_0 and η_1 . However, at points outside of this interval his information becomes relatively worse, since it is mostly driven by just one of the two signals. This effect can be seen more clearly in the right panel, where we plot the posterior variance. Interestingly, the lowest posterior variance is not achieved right at either y_0 or y_1 , but in between the two, because there the agent has two relatively informative signals to use. This plot also displays one of the defining features of our setup, which is that the agent's beliefs are more precise within the region of the state space that he has previously deliberated about, and become increasingly imprecise at points far away from the types of situations considered in the past.

More generally, as the agent accumulates information, beliefs follow the recursion

$$\hat{c}_t(y) = \hat{c}_{t-1}(y) + \frac{\sum_{t-1}(y, y_t)}{\sum_{t-1}(y_t) + \sigma_{\eta, t}^2} (\eta_t - \mu_{t-1}(y))$$
(3)

$$\Sigma_t(y, y') = \Sigma_{t-1}(y, y') - \frac{\Sigma_{t-1}(y, y_t)\Sigma_{t-1}(y', y_t)}{\Sigma_{t-1}(y_t) + \sigma_{\eta, t}^2}$$
(4)

where $\mu_t(y) = E_t(c^*(y)|\eta^t)$ and $\Sigma_t(y, y') = \text{Cov}(c^*(y), c^*(y'^t))$ are the posterior mean and covariance functions after conditioning on all available deliberation signals up to and including time t. These two objects fully characterize the posterior beliefs of the agent about the unknown $c^*(y)$, which is itself also a Gaussian Process, just as his priors. Note that we use $\hat{\sigma}_t^2(y)$ to denote the posterior variance at a given value of y, i.e. $\Sigma_t(y, y)$. Hence we can write,

$$c^*(y)|\eta^t \sim \mathcal{GP}(\hat{c}_t(y), \Sigma_t(y, y')).$$

2.2 Optimal Deliberation

In this section we study the optimal deliberation choice of the agent and the resulting optimal signal precision and actions. Here we assume that the agent faces a simple linear cost in the total information contained in η_t . Thus, the information problem of the agent is,

$$U = \max_{\substack{\sigma_{\eta,t}^2 \\ \hat{\sigma}_t^2}} -W_{cc} E_t (\hat{c}_t(y_t) - c_t^*(y_t))^2 - \kappa I(c^*(y_t); \eta_t(y_t) | \eta^{t-1})$$

$$= \max_{\hat{\sigma}_t^2} -W_{cc} \hat{\sigma}_t^2(y_t) - \kappa \frac{1}{2} \ln \left(\frac{\hat{\sigma}_{t-1}^2(y_t)}{\hat{\sigma}_t^2(y_t)} \right)$$

The second line in the above equality follows from the fact that the mutual information in a Gaussian framework is simply one half of the log-ratio of prior and posterior variances, and κ is a parameter that measures the marginal cost of a unit of information. For example, κ will be higher for individuals with a higher opportunity 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 precision in decisions. In addition, κ would also be higher if the particular economic environment of the agent is more complex, and thus the optimal action is objectively harder to figure out (e.g. they are given a harder math problem).

Lastly, the maximization is subject to the "no forgetting constraint"

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

which ensures that the chosen value of the noise in the signal, σ_{η}^2 is non-negative. Essentially, the constraint is there in order to ensure that the agent cannot gain utility by "losing" some of his prior information and thus making his current posterior beliefs, after observing $\eta_t(y_t)$, more uncertain than the beliefs he enters the period with.

Taking first order conditions, we see that the optimal deliberation choice satisfies

$$\hat{\sigma}_t^{*2}(y_t) = \frac{\kappa}{2W_{cc}}.$$
(5)

Hence, with a cognition cost that is linear in mutual information the agent has an optimal target level for the posterior variance of the unknown optimal action that is a simple, and intuitive function of deep parameters. He would like to obtain a higher precision in his actions (and thus deliberate more) when the cost of making mistakes (W_{cc}) is high and when the cost of deliberation (κ) is low. Moreover, this optimal posterior precision of his beliefs is not a function of the actual realization of the current state y_t .

Imposing the no forgetting constraint, optimal reasoning leads to the posterior variance:

$$\hat{\sigma}_t^{*2}(y_t) = \min\{\frac{\kappa}{2W_{cc}}, \hat{\sigma}_{t-1}(y_t)\}.$$

So unless beliefs are already more precise than desired, the agent engages in just enough deliberation, and hence obtain just the right $\sigma_{\eta t}^2$, so that the time t posterior variance equals the target level $\frac{\kappa}{2W_{ex}}$. Hence, the intensity of deliberation is both state and history dependent.

To give an example, suppose that the agent enters the period having seen only one signal $\eta_0(y_0)$ in the past. Moreover, let us suppose that the resulting posterior variance, after seeing η_0 , is represented by the blue line in the right panel of Figure 1 and that $\frac{\kappa}{2W_{cc}} = 0.6$. Then we see that the agent will not do any additional deliberation in the current period, and hence choose $\sigma_{\eta,t}^2 = \infty$ for values of y that are roughly between 1 and 1.5. In that region, the initial beliefs about $c^*(y_t)$, which are based on the knowledge of η_0 , are already sufficiently precise and it is not beneficial to engage in any further deliberation.

On the other hand, the agent will engage in active deliberation if the current value of the state is below 1, since then the initial beliefs are not quite precise enough. Crucially, the intensity of deliberation is state dependent. The lower is y, and hence the further away it is from the region where the agent is fairly confident in his estimate, the more deliberation effort the agent elects to perform. Since the initial posterior variance is higher for values of yfarther away from y_0 , it will take more effort, and thus a higher precision of $\eta_t(y_t)$, in order to bring the precision of the posterior beliefs up to the optimal level $\frac{\kappa}{2W_{cc}}$. In particular, we can show that the optimal signal noise variance is:

$$\sigma_{\eta,t}^{2} = \begin{cases} \frac{\frac{\kappa}{2W_{cc}}\hat{\sigma}_{t-1}^{2}(y_{t})}{\hat{\sigma}_{t-1}^{2}(y_{t}) - \frac{\kappa}{2W_{cc}}} & , \text{ if } \hat{\sigma}_{t-1}^{2}(y_{t}) \ge \frac{\kappa}{2W_{cc}}\\ \infty & , \text{ if } \hat{\sigma}_{t-1}^{2}(y_{t}) < \frac{\kappa}{2W_{cc}} \end{cases}$$
(6)

The agent opts for higher precision, and hence higher deliberation cost, when the posterior variance of his initial beliefs, at y_t , are relatively far from the target level $\frac{\kappa}{2W_{cc}}$. In turn, the resulting optimal weight $\alpha^*(y_t; \eta_t, \eta^{t-1})$ of the current period signal η_t is:

$$\alpha^*(y_t; \eta_t, \eta^{t-1}) = \max\{1 - \frac{\kappa/(2W_{cc})}{\hat{\sigma}_{t-1}^2(y_t)}, 0\}$$
(7)

Thus the deliberation, and especially its effect on the resulting action $\hat{c}_t(y_t)$,

$$\hat{c}_t(y_t) = \hat{c}_{t-1}(y_t) + \alpha^*(y_t; \eta_t, \eta^{t-1})(\eta_t - \hat{c}_{t-1}(y_t))$$
(8)

is both state and history dependent. When the precision of initial beliefs is relatively far from its target, then the agent acquires a more precise current signal and puts a bigger weight on it (a higher $\alpha^*(y_t; \eta_t, \eta^{t-1})$). Thus, for "usual" values of y_t , i.e. ones for which the agents has deliberated at often in the past and thus has lower posterior variance at, the agent is unlikely to deliberate much more again. As a result, $\alpha^*(y_t; \eta_t, \eta^{t-1})$ will be relatively small, and his resulting action will be primarily driven by his beginning of period beliefs, $\hat{c}_{t-1}(y_t)$. On the other hand, unusual values of y_t , where the initial beliefs are more imprecise, are more likely to trigger more intensive deliberation. Such a deliberation choice will result in a higher $\alpha^*(y_t; \eta_t, \eta^{t-1})$ and an effective action \hat{c}_t that deviates more from the initial beliefs.

2.2.1 Comparison with rational inattention to the objective state

The typical approach in the Rational Inattention (RI) literature is to study attention allocation over the state variable y_t , but not over the policy function $c^*(y_t)$. Hence the standard RI approach would be to assume that the agent knows that the optimal action is

$$c^*(y_t) = c_y y_t \tag{9}$$

with a known c_y , and assume that the agent does not observe y_t , but has the correct prior

$$y_t \sim \operatorname{iid} N(\bar{y}, \sigma_y^2).$$

Learning is then about the unknown state variable y_t , but proceeds in a similar way,

with the agent observing unbiased signals,

$$\tilde{\eta}_t = y_t + \tilde{\varepsilon}_t^{\eta}.$$

Uncertainty about the state can be mapped directly into uncertainty about the optimal action c_t through the *known* optimal policy function $c^*(y_t)$, given here by equation (9):

$$\hat{c}_t = c_y E_t(y_t) = c_y \left(\bar{y} + \underbrace{\frac{\sigma_y^2}{\sigma_y^2 + \sigma_\eta^2}}_{=\tilde{\alpha}_y} (\tilde{\eta}_t - \bar{y}) \right)$$
(10)

In terms of the objective function, the agent is similarly facing a quadratic loss function and a linear cost in the Shannon information of his signal η_t :

$$U = \max_{\substack{\sigma_{\eta,t}^2 \\ \hat{\sigma}_t^2}} -W_{cc} E_t (c_y \hat{y}_t - c_y y_t)^2 - \kappa^{RI} I(y_t; \eta_t | \eta^{t-1})$$
$$= \max_{\hat{\sigma}_t^2} -W_{cc} c_y^2 \hat{\sigma}_{y,t}^2 - \kappa^{RI} \frac{1}{2} \ln \left(\frac{\hat{\sigma}_{y,t-1}^2}{\hat{\sigma}_{y,t}^2}\right)$$

where we use the notation $\hat{\sigma}_{y,t}^2 = \operatorname{Var}(y_t|\eta^t)$ to denote the posterior variance of y_t . The cost of information in this framework is also linear in mutual information, but the marginal cost κ^{RI} could be different than κ , since this captures information flow of a different activity – paying attention to an unknown state.

Similarly to before, the optimal attention allocation implies an optimal target level for the posterior variance, this time given by

$$\hat{\sigma}_{y,t}^2 = \frac{\kappa^{RI}}{2W_{cc}c_y^2} , \ \forall t \tag{11}$$

Importantly, in this case the optimal attention allocation and hence the resulting choice for the precision of the signals is not state or history dependent and $\tilde{\sigma}_{\eta,t}^2 = \tilde{\sigma}_{\eta}^2$, $\forall t$. The assumption that y_t is iid over time makes the optimal attention solution particularly simple, but the fact that it is not state or history dependent is a general result. If y_t was persistent, then we would simply substitute the steady state Kalman filter second moments for σ_y^2 in the equations below. With iid states it follows that the optimal signal-to-noise ratio $\tilde{\alpha}_y$ that appears in the updating equation (10), given the "no forgetting constraint", is

$$\tilde{\alpha}_{y}^{*} = \max\{1 - \frac{\kappa^{RI}/(2W_{cc}c_{y}^{2})}{\sigma_{y}^{2}}, 0\}.$$
(12)

The key difference between the optimal $\tilde{\alpha}_y^*$ in equation (12) and that of our benchmark framework, given by (7), is that in our model the prior uncertainty $\hat{\sigma}_{t-1}^2(y_t)$ is state and history dependent. In turn, this makes the optimal signal-to-noise ratio also state and history dependent, since the agent optimally targets a constant level of variation in the errors of his action. The state and history dependence of the deliberation choice is a fundamental feature of our setup, and it comes about as a result of the fact that the agent is not learning about the value of the state y_t , but about his optimal action as a *function* of the state. Hence deliberation at different past values of the state carry different amount of information about the optimal action at today's state y_t .

It is also useful to observe that the deliberation choice can be rewritten 'as if' it is a control cost problem where the agent has a 'default' distribution of actions, conditional on the state, and chooses a current distribution, subject to a cost that is increasing in the distance between the two distributions, as measured by relative entropy.²⁵ In turn, a tighter chosen distribution is beneficial because errors are smaller. In our model, the default distribution is the prior over the unknown conditional optimal action and the posterior is the updated distribution controlled by the chosen signal-to-noise ratio. While many models of imperfect attention to actions fit in such an 'as if' control cost interpretation, the key property of our model is in the endogenous dynamics of the prior distribution.²⁶

One special case of our reasoning setup delivers the same qualitative feature as the standard approach. In particular, right in the first period of reasoning, the prior uncertainty $\hat{\sigma}_{t-1}^2(y_t)$ is not state- or history- dependent. Hence, the optimal signal-to-noise ratio for the initial period, in equation (7), becomes qualitatively similar to that in equation (12):

$$\alpha^*(y_0;\eta_0) = \max\{1 - \frac{\kappa/(2W_{cc})}{\sigma_c^2}, 0\}.$$
(13)

Moreover, when the attention costs are the same and the optimal response $c_y = 1$, then the formulas produce equivalent predictions for the behavior of actions in the first period.

More generally, our framework shares with the standard RI framework the feature that agents find it costly to pay attention and therefore will make mistakes in their actions. However, in our framework the mistakes are driven by a lack of knowledge of the functional form of the optimal response $c^*(.)$, and not a lack of knowledge of the unknown state y_t . As a result, the patterns of errors made by agents in our framework are history and state dependent, while in the standard RI benchmark they are typically not.

 $^{^{25}}$ The control cost approach appears for example in game theory (see Van Damme (1987)) and the entropy based cost function is studied in probabilistic choice models such as Mattsson and Weibull (2002).

²⁶Matějka and McKay (2014) and Matějka et al. (2017) show the mapping between an entropy-based attention cost and a control problem leading to logit choices in static and dynamic environments, respectively.

3 Implications About Behavior

In this section we describe key implications of the model of costly reasoning described above.

3.1 Ergodic Steady State: Beliefs and Actions

Since the deliberation choice of the agent is history dependent, we focus on analyzing the ergodic steady state behavior of our agent – i.e. the optimal deliberation and resulting action after having seen and deliberated about a long history of y_t realizations. The goal is to analyze the typical policy function and behavior.

Discounting information

As specified, however, the model is non-stationary since the accumulation of information never stops, and signals even very far into the past remain equally informative as current ones. In other words, there is no steady state or ergodic distribution of beliefs. In order to achieve one, we can follows two routes. One is to introduce shocks to the true unknown policy function $c^*(y_t)$, for example model it as an AR(1) Gaussian Process:

$$c_t^*(y_t) = \bar{c}(1 - \rho_c) + \rho_c c_{t-1}^*(y) + \varepsilon_t(y)$$

In this case the object the agent is trying to learn is changing over time and hence eventually the information content of past signals will decay to zero. Or alternatively, we could exogenously assume that the information content of past signals is decaying over time, either because of time-variation in the policy function or because of costly or imperfect memory recall. For example, we could assume that at the beginning of each period, the precision of past signals is discounted at a constant rate $\delta \leq 1$ so that

$$\frac{1}{\sigma_{\eta,t,t-k}^2} = \frac{\delta}{\sigma_{\eta,t-1,t-k}^2}$$

where $\frac{1}{\sigma_{\eta,t,t-k}^2}$ is the effective precision of the t-k signal at time t. In the benchmark results presented in the main text we follow the second approach, as it appears a bit more general. But we note that the time varying $c_t^*(y_t)$ framework leads to essentially identical results.

Ergodic distribution

To obtain the ergodic distribution, we simulate the economy for 100 times, each of a length of 1000 periods, where in each period we draw a new value of y_t and the agent makes optimal deliberation choices given the history of idiosyncratic signals η_t . Then we analyze the average moments over the resulting distribution, and consider the optimal deliberation



Figure 3: Ergodic Uncertainty – Posterior Variance and the resulting optimal signal-to-noise ratio after 1000 periods of optimal deliberation. The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1, \text{ and } \delta = 0.9$

choice and the resulting optimal action of the agent at time t, given the long and typical history that came before it.

Figure 3 plots the posterior variance of beliefs, conditional on a typical long history of signals $(\hat{\sigma}_{t-1}^2(y_t))$, and the associated optimal signal-to-noise ratio for the new signal $\alpha^*(y_t; \eta_t, \eta^{t-1})$. There is an interval of y_t values right around the mean, \bar{y} , where the posterior variance is below the target level $\frac{\kappa}{2W_{cc}}$, and then outside of that interval, the posterior variance grows symmetrically towards the extreme values of y_t . As a result, the optimal signal to noise ratio of the current signal η_t is zero in the interval close to the mean, and then grows for realizations of y_t further away from the mean.

The reason that the posterior variance dips below the optimal level is because the majority of the signals that the agent has seen in the past center around the mean value of y_t , as this is the most likely realization of the state and is the region of the state space that the agent has deliberated the most about in the past. And while no single signal in the history of signals η^{t-1} is precise enough to lower the posterior variance below the optimal level $\frac{\kappa}{2W_{cc}}$, the combination of a large concentration of them indeed does so. To better illustrate this, the left panel of Figure 4 plots the distribution of the values of y_t where the agent has deliberated in the past. Note that the agent does not necessarily deliberate (and thus obtain an informative signal) every single period and thus for every single value of y_t that has occurred in the past. Costly deliberation is only triggered if the posterior uncertainty is relatively large, as evidenced by the right panel of Figure 3. Hence, the distribution of y_t values at which the



Figure 4: Long-Run distribution of signals and associated precision, after 1000 periods of optimal deliberation. The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$

agent has deliberated at is not necessarily identical to the unconditional distribution of y_t , but as we can see, it shares a lot of its key features, as it is symmetric and centered at \bar{y} .

Lastly, while this graph shows the distribution for the incidence of deliberation, it does not paint a full picture because the optimal signal precision (i.e. the intensity of deliberation) is state dependent, and is likely to be different for different values of y_t . For this purpose the right panel of Figure 4 plots the average signal precision chosen at each value of y_t . Unsurprisingly, we see that individual signals are most precise for values of y_t further away from the mean, and the least precise right around \bar{y} . This conforms with the main characteristic of the optimal deliberation choice discussed above – the agent has the least incentives to deliberate around the mean realization of y_t , because that is where the history of signals provides the most information.

In other words, the agent tends to have deliberated most often right around the mean y. His typical deliberation close to \bar{y} is relatively imprecise, since the sum total of past deliberation has already informed him relatively well about the optimal action in that region, hence he does not feel the need to invest much additional effort. At the same time, the agent tends to see and deliberate at unusual values of y_t more rarely, but conditionally on doing so, he invests a significant amount of effort, and thus obtains more precise signals. As a result, the typical history of deliberations delivers a large concentration of relatively imprecise signals around the \bar{y} , and fewer but much more precise signals at values away from \bar{y} . Overall, the region around \bar{y} is the one where the agent's beliefs are most certain, because

there the agent can equally well draw on information encoded in signals at both high and low y's, and has also seen the highest concentration of past signals.

3.2 Optimal action – non-linearity, inertia and salience effects

Figure 5 plots the ergodic action, $\hat{c}_t(.) = E(c^*(.)|\eta_t, \eta^{t-1})$. Because of the state and history dependent nature of the optimal deliberation choice, the effective policy function $\hat{c}(y_t)$ is non-linear, even though the underlying optimal action that the agent is learning about $(c^*(y_t))$ is linear. As a result, the behavior of the agent displays a number of interesting features, including exhibiting both inertia and salience.

The inertia occurs for realizations of y_t close to its mean, and is the result of two forces. On the one hand, the ergodic belief about $c^*(.)$,

$$\overline{\mu}(.) = E(c^*(.)|\eta^{t-1}),$$

the belief that the agent enters the typical period with, is relatively flat (see yellow dotted line in Figure 5).²⁷ On the other hand, these initial beliefs are rather precise in the region around \bar{y} (see Figure 3), and hence the agent optimally chooses to not do much additional deliberation in that region, resulting in a small $\alpha^*(y_t; \eta_t, \eta^{t-1})$. In other words, in that part of the state space, the agent feels confident in his initial understanding of $c^*(y_t)$ and does not seek to do any further costly deliberation in the current period. The agent thus follows his prior understanding of the optimal action, which happens to be rather flat. This generates the inertia in the middle part of the effective policy function $\hat{c}(.)$, as seen from the fact that the resulting action is only mildly responsive to changes in y_t .

The key to understanding this result is the fact that the ergodic beliefs of the agent around \bar{y} are both rather flat and also precise. This occurs from the state and history dependent nature of the optimal deliberation choice, which leads the agent to a precise estimate of the *average* level of the optimal action, but not the *shape* of the optimal policy in that part of the state space. As we saw in Figure 4, the agent tends to accumulate many, but individually imprecise signals around the mean value of the state. The large number of signals helps pin down the average level of the optimal action, and indeed the agent's beliefs about $c^*(.)$ are most accurate on average around \bar{y} . But because the agent does not reason intensely about $c^*(.)$ at distinct values of y_t in that neighborhood, he does not get a good estimate of the shape of the true optimal policy. The mistakes incurred for small deviations of y_t from \bar{y} are small, so the agent optimally does not invest much effort in thinking through

²⁷We use the notation $\overline{\mu}(.)$ to highlight that the typical long run prior integrates over various possible paths of histories η^{t-1} .



Figure 5: Ergodic Action. The figure plots the ergodic average action, computed from a cross-section of 100 agents and 1000 time periods, with parameters $W_{cc} = 1$, $\kappa = 0.5$, $\sigma_c^2 = 1$, $\psi = 1$, and $\delta = 0.9$

the details of the local shape of $c^*(.)$.

However, the optimal deliberation choice of the agent changes for more unusual realizations of y_t . As y_t moves further away from its mean \bar{y} , the agent's prior learning becomes less and less useful. As a result, he chooses to invest an increasing amount of effort into the current period deliberation, and hence the current signal η_t becomes increasingly precise. In turn, the agent's action becomes more heavily influenced by his current deliberation, which (on average) is accurately informing about increasing the action for higher values of y_t . Thus one observes non-linearity in the effective action. The optimal reasoning therefore generates a salience effect in the agent's behavior in the form of deliberation only for things that tend to stand out, given his prior beliefs.

Lastly, we make a remark on our initial choice of the ex-ante prior $\mu(y) = \bar{c}$. On the one hand, this initial prior does not necessarily matter for the results presented here because the agent has seen a long history of signals, and hence has updated the ergodic beliefs $\mu_{t-1}(.)$ accordingly. On the other hand, as we saw above, around the mean \bar{y} the ergodic beliefs are accurate on average but tend to not capture well the shape of the underlying $c^*(.)$. Hence, the *level* of the ex-ante prior $\mu(y)$ does not necessarily matter, but the steady state beliefs $\bar{\mu}(.)$ do tend to be influenced by the shape of $\mu(.)$ in that region. As we argued before, a conceptually coherent choice of $\mu(.)$ is \bar{c} , since in this case the ex-ante prior is not a function of y, hence the agent has no prior "free" information about the relationship between $c^*(.)$ and y. Any such understanding comes at the cost of deliberation effort, and the agent is free



Figure 6: Comparative statics. Benchmark values – $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$. Changing values: $\kappa = 0.15, \psi = 3$, and $\delta = 0.99$.

to engage in that. It is a result that his optimal deliberation is such that his understanding of the shape of $c^*(.)$ is not good around \bar{y} , but better at more extreme values of y_t .

Nevertheless, extending the model to a case where the prior $\mu(.) \neq \bar{c}$ is straightforward. We would still obtain a non-linear ergodic policy function, where in the interval of y_t realizations around the mean \bar{y} , the resulting action $\hat{c}(.)$ will be accurate on average, but its shape will be biased towards the local shape of $\mu(.)$. At realizations of y_t further away from its steady state mean, the shape of $\mu(.)$ will continue to matter little, as this is where the agent will be doing the most active deliberation, and follow the true $c^*(.)$ more closely.

3.2.1 Comparative statics

The comparative statics of the model are quite intuitive. In this univariate framework (for multivariate extensions see following sections) the effect of increasing the cost of making errors in the action (W_{cc}) is observationally equivalent to the effect of lowering the deliberation cost κ . A lower κ means that deliberation is cheaper, and as a result the agent will deliberate more often and more intensively. Hence, the resulting effective action $\hat{c}(.)$ tracks closer the true action $c^*(.)$. Still, as we can see in the example illustrated in the left panel of Figure 6 there is no fundamental change in the basic properties of the effective action $\hat{c}(.)$ – it is non-linear, being relatively more flat in the middle, and more upward sloping towards the ends. However, with a lower κ the overall level of the non-linearity is smaller, and in fact the action will converge to the true underlying linear $c^*(.)$ as $\kappa \to 0$.

Increasing ψ , the parameter that controls the correlation between $c^*(y)$ and $c^*(y')$ for two distinct values of the state $y' \neq y$, has a similar effect. A higher ψ makes the informativeness of any given signal more localized. Hence, past deliberations have generally lower effects on today's deliberation choice (unless today's y_t happens to be right in the neighborhood of a past η_{t-j}), and this weakens the history dependence in the deliberation choice. Since past deliberation is less useful, the agent finds it optimal to invest more effort in deliberation in any given period. Thus, his effective action $\hat{c}(.)$ becomes closer to the true underlying $c^*(.)$, and becomes more upward sloping everywhere. In particular, the resulting ergodic distribution of signals will feature a more individually precise signals at distinct values of y_t , which transmits more information about the shape of $c^*(.)$.

Lastly, we also consider the effect of increasing δ , which controls discounting of past information. This parameter has slight effects on the results presented here, as exemplified by the right panel of Figure 6. Keeping the sequence of past signals η^{t-1} fixed, increasing δ increases their overall informativeness, and leads to a lower optimal level of current deliberation. Since past signals are more informative the agent finds it less necessary to invest much effort in reasoning at any given period. As a result, at the new ergodic distribution with a higher δ the agent tends to have an effectively longer history η^{t-1} , but each of the signals in the history has a lower precision, since at any given period the agent deliberates less extensively. Thus, the effective residual uncertainty and the resulting policy function are similar to the case of lower δ , even if we choose a very high value, such as $\delta = 0.99$.

3.3 Persistent actions

Reasoning about the optimal function $c^*(y)$ endogenously leads to persistence in actions even if the state y is iid. This mechanism naturally arises as long as the information acquired about the optimal action at some particular realization of the state $y_t = y$ is perceived to be informative about the optimal action at a different value y'.

Equation (3) describes the recursive formulation of $\mu_t(y)$, the posterior belief about the average value of $c^*(y)$. Consider the ergodic version of that equation, where the agent enters period t with a prior mean and variance function that have settled on some long-run values, $\overline{\mu}(y)$ and $\overline{\Sigma}(y, y')$, respectively. Following a particular realization of the state y_t , equations (6) and (7) give the optimal noise variance and resulting signal-to-noise ratio at y_t , respectively, so that the posterior mean function updates at the entire vector y as follows:

$$\mu_t(y) = \overline{\mu}(y) + \alpha^*(y_t) \frac{\overline{\Sigma}(y, y_t)}{\overline{\Sigma}(y_t)} \left[\eta_t - \overline{\mu}(y)\right]$$

where the signal $\eta_t = c^*(y_t) + \sigma_{\eta,t}\varepsilon_t$. There $\sigma_{\eta,t}$ is given by the optimal intensity of the reasoning process, implying a signal to noise ratio $\alpha^*(y_t) = \max\{1 - 0.5\kappa/W_{cc}\overline{\Sigma}(y_t), 0\}$, and ε_t is the particular noise realization in the internal signal about $c^*(y_t)$.

3.3.1 Impulse response

Consider a shock to the state that moves it from its ergodic value of \overline{y} to some particular realization y_t . To understand the persistent effect of this change on the observed action, it is useful to separate the intuition into the impact and the propagation effects of that change.

The impact is given by the ergodic policy function $\mu_t(y)$ explained in section 3.2. If the agent does not find it optimal to spend additional cognitive effort at the state y_t , so that $\alpha^*(y_t) = 0$, then $\mu_t(y)$ is entirely driven by the ergodic prior $\overline{\mu}(y)$. As no new information about the optimal action is acquired, there are no updating effects on the agent's beliefs about $c^*(y)$ and so there are no persistent effects of the change in the state on the actions.

In contrast, if the agent does choose to invest cognitive effort, then the new signal η_t shifts beliefs about optimal action at y_t , as well as at other values in the state space. The strength of the shift is given by two forces. First, the impact effect at y_t , given by $\alpha^*(y_t)$. The action responds more to the signal η_t if the state is more unusual, in the sense that the ergodic prior uncertainty $\overline{\Sigma}(y_t)$ is larger, so that the agent decides to invest significant effort in reasoning about the optimal action. Second, the shift affects beliefs through the covariance function $\overline{\Sigma}(y, y_t)$. The higher is this prior covariance, the more informative is the signal at y_t about optimal actions at other state values y. The covariance is larger if the state y is closer to y_t and if the parameter ψ is smaller.

Let us turn now to the propagation of that change in the observed state. The rate at which the prior covariance decays with the distance between y and y_t is a key determinant of persistence. For a large enough value of ψ , the agent's prior is that the signals acquired about $c^*(y_t)$ are essentially useless about $c^*(y)$ at some $y \neq y_t$, as then the prior covariance function decays very quickly to zero as y differs from y_t . In that case, the only significant effect is an update of the optimal action at y_t . Since the probability of the future realizations to be very close to that specific value of y_t is close to zero, there are no significant persistent effects.

In contrast, for a smaller ψ , signals obtained at different values of y_t are used to shift beliefs about the entire unknown optimal action function. As new states will realize following period t, this shift in beliefs persists until it converges back to the ergodic $\overline{\mu}(y)$. Therefore, a signal about the optimal action at y_t is likely to have persistent effects on observed actions, even if the objective states are iid.

Consider a graphical representation of the impulse response function. To isolate the average effects, we focus on the case where the signal realization equals its average value, so

we directly set $\varepsilon_t = 0$ in this analysis.²⁸ The blue solid line in the left panel of Figure 7 plots the impulse response starting from t for a value of the state $y_t = -2\sigma_y$ for the benchmark parametrization. On impact, the action incorporates about 25% of the change in the state, dropping by 0.05 compared to the reduction of 0.2 in y_t . While such an underreaction appears more generally also in models of partial attention to the states, an important additional property of our model is the resulting propagation. Indeed, as described above, the negative shock to y_t leads to a persistently lower average action following the shock.



Figure 7: Impulse response function. The figure plots the typical path of the action following an innovation of $y_t = -2\sigma_y$. The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$. In the left panel we change parameters to $\kappa = 0.1, \psi = 5$ and $\delta = 1$, respectively. In the right panel we change persistence to $\rho_y = .9$, or we feed in future values of y_{t+j} that always equal to \overline{y} , or set the current shock to $y_t = -3\sigma_y$.

The left panel also illustrates the role of the parameters in driving the propagation. The response on impact is stronger and there is less persistence when the cognition cost is lower ('low κ ', dotted yellow line), or when the signal at y_t is less informative at other states ('high ψ ', dashed orange line). When there is no discounting of past information (' $\delta = 1$ ', starred purple line), the agent has accumulated so much information that it does not further deliberate and instead only uses the ergodic prior mean at y_t for the best guess on $c^*(y_t)$. The lack of further updating leads to no persistent effects on the average action.

The right panel of Figure 7 illustrates the strong non-linearity of the model by plotting three alternative assumptions on the shock realizations. First, in the dotted orange line, we make the state realization even more unusual, at $y_t = -3\sigma_y$. On impact, the action now incorporates about 33% of the change in the state, dropping by 0.11 compared to the

²⁸Given the non-linear nature of our model we compute a generalized impulse response function to a particular realization of the state y_t as $IRF_j(y_t) = E(c_{t+j}|y_t) - E(c_{t+j}|y_t = \overline{y})$.

reduction of 0.3 in y_t . The larger proportional reduction compared to the baseline case indicates the role of the state-dependent reasoning. Since the state is now more unusual, the prior uncertainty is large and the agent finds it optimal to reason intensely. The higher intensity leads to a proportionally stronger response in the average action.

Second, in the starred purple line, we calculate a different object for the impulse response, using the typical approach of computing the impulse response in a linear model. In particular, here we ignore the distribution of future states and instead feed in realizations for y_{t+j} that always equal to \bar{y} . The resulting path is more persistent since in this case the agent continues to see values of \bar{y} and does not reason further so the updating done at y_t is long lived. However, in our model the state-dependency of the deliberation choice means that when there are shocks to future y_{t+j} , it is more likely that some of these realizations trigger reasoning compared to the case of feeding always the average \bar{y} . The larger chance of future reasoning tends to decrease the weight that will be put in the future on the information obtained at y_t , which generates less persistence.

3.3.2 Hump-shape dynamics

The model is able to predict hump-shape dynamics in the action even when the exogenous state is mean reverting. For that we note that an important characteristic of the non-linearity of the ergodic policy function plotted in Figure 5 is a convex shape that appears for a significant part of the state space, both to the left and the right of the mean \overline{y} . Suppose now that y_t moves into those areas, for example being close to a value of 1.15, and record the implied action $\widehat{c}(y_t)$. From mean reversion, the distribution of future states has a mean that is closer to \overline{y} , i.e. $E(y_{t+1}) < y_t$. If the response function $\widehat{c}(y)$ would be linear we would also observe on average mean reversion in the action, i.e. $E(\widehat{c}(y_{t+1})) < \widehat{c}(y_t)$. However, given the local convexity of the function, the Jensen inequality implies that

$$E(\widehat{c}(y_{t+1})) > \widehat{c}(E(y_{t+1}).) \tag{14}$$

The inequality in equation (14) allows for the possibility that $E(\hat{c}(y_{t+1})) > \hat{c}(y_t)$, so that even if the future state on average is closer to its mean, the average action is not and one observes hump-shape dynamics. This convex effect can dominate the mean reversion effect when the response function is convex enough and when the persistence in the state is high enough.

Intuitively, the convexity of the policy function captures the interplay between the inertial effects around the usual states and the salience effects at the more unusual state. If the state moves closer to its mean, the agent uses more heavily past experiences which, at those usual state realizations, tend to be more informative about the average level of

the action rather than the slope so that the resulting action tends to be flatter. If the state moves even further away from the ergodic value, it enters into parts of the state space that are relatively more unusual, where the agent finds it optimal to reason more intensely. These salience type of effects are more informative about the local slope and lead to stronger responses. Therefore, the average movement in the next period action may be dominated by salience type of reasoning so that the agent takes an even more reactive action than before.

In the right panel of Figure 7 we plot as the dashed line the non-linear impulse response function for the case of $\rho_y = 0.9$. On impact, the effect is smaller because the size of the innovation has been reduced to match the same unconditional variance.²⁹ Importantly, for this parameterization the Jensen inequality effect of equation (14) is strong enough so the propagation is characterized by hump-shape dynamics.

3.4 Persistence of volatility of actions

Reasoning about the optimal function endogenously leads to changes in the time-series volatility of actions.

3.4.1 Volatility clustering

We illustrate how the model generates clusters of volatility through the following experiment. Once the economy reaches its long run distribution, at some point T we feed in a randomly drawn sequence of states of length s. Denote that vector of draws as y^s . We directly affect the variability of that vector of innovations by multiplying it by a scalar V. Denote the resulting vector of actions that the agent chooses in this first stage, when observing the vector Vy^s , as a_1 . At time T + s, we feed in another random sequence of states of length s, also multiplied by V. We collect the vector of actions that the agent chooses in this second stage as a_2 . We simulate repeatedly over y^s , we vary V and report $\sigma(a_2)$, the standard deviation of actions in the vector a_2 , compared to the action vector a_1 , $\sigma(a_1)$. Notice that, on average, the standard deviation of the states generating the vectors a_1 and a_2 is the same, given by $V\sigma_y$.

The main finding here is that the ratio $\sigma(a_2)/\sigma(a_1)$ is increasing in V, or, put differently, it is increasing in $\sigma(a_1)$, since the latter is increasing in V. Therefore, we find that there is an endogenous persistence of volatility in actions. An econometrician analyzing data produced from the model finds that, conditioning on a sample where the variability of innovations is larger than usual, the variation in actions following that sample is also larger than usual.

The intuition of this result is driven by the logic behind the ergodic reasoning behavior. In that ergodic distribution, when observing an unusual sequence of states, there are two

²⁹We adjust $\sigma_y = .1\sqrt{1-\rho_y^2}$ to keep the same unconditional variation in the state as in the baseline model.

main implications. One is a direct effect through which the agent's actions are also more unusual, since they respond to the observed states. From this direct effect, the measured volatility of actions is larger. Absent any further propagation mechanisms, the model would attribute a larger variability of actions to a larger variability of states.

However, there is a second, propagating effect, through the implications for optimal reasoning. Given more unusual states, the agent chooses to reason more intensely, since the prior uncertainty at those states is relatively large. By reasoning, the agent obtains more informative signals about the optimal policy function, which, on average, leads to an updated prior belief at T + s, $\mu_{T+s-1}(y)$, that is characterized by stronger action responses to the state y. This updated belief generates on average more variable actions starting at period T + s, as the belief converges back to the ergodic posterior mean $\overline{\mu}(y)$. As a consequence, the indirect effect of more intense reasoning, incentivized by more variable realizations of the state, leads to an endogenous persistent increase in the variability of actions.

The left panel of Figure 8 illustrates the effect graphically. We use the baseline parametrization and simulate repeatedly over a vector y^s of length s = 20. We set V = 0.5 and V = 1.5, corresponding to a 'low' and 'high' variance sample, respectively. Figure 8 plots the resulting policy functions. As described above, the posterior belief assigns stronger responses following more volatile innovations, leading to subsequent more variable actions.



(a) Conditional on recently low or high variance of (b) Conditional on recently unusual states states

Figure 8: Optimal action and variability of states. In the left panel the policy function conditions on samples of shocks with lower or higher than typical variability. In the right panel the policy function conditions on the previous state realization ranging from $-\sigma$ to -3σ . The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$.

3.4.2 More variable actions after a large shock

We use a second experiment to illustrate how after a 'big' shock, the measured response of actions endogenously changes. In particular, consider again the ergodic distribution where at some point T we feed in a state $y_T = \alpha \sigma_y$. Starting at T + 1, we repeatedly simulate a vector of length s of draws for the state and report how the resulting standard deviation of actions based on that vector changes as we vary α .

The main finding is that the variability of actions following the shock y_T increases with $|\alpha|$. After more unusual states, the agent finds it optimal to reason more intensely. On average, the signals received by the agent lead to an updated belief $\mu_T(y)$ that is more responsive around the realized state y_T . The agent thus enters the subsequent periods with a prior belief of stronger reaction, which generates a larger variability in actions.

The right panel of Figure 8 illustrates the point graphically. There we plot the policy functions for α taking three values, -1, -2 and -3, respectively, where we also note that the effects discussed here are symmetric for $\alpha = 1, 2$ and 3, respectively. There are two clear important effects. One is that the average level of the posterior belief is shifted down, as the agent updates over the whole function. This shift illustrates a previous point, namely that there is endogenous persistence in the mean action. The size of the effect is also non-linear: the update is larger for more unusual state realizations, since there the prior uncertainty is larger and thus there the agent chooses to reason more intensely.

The second effect is on the shape of the posterior belief. The downward shift implied by the average signal obtained at y_T is stronger locally, around the realized state. Importantly, the more unusual the state it, the steeper is the posterior belief. The reason is that the average signal obtained at y_T indicates to the agent that the optimal policy function should be generally more responsive, not only at the specific realized y_T , but also at the previously often visited states. The local informativeness of signals thus transmits the acquired information at the unusual states into subsequent stronger action responses.

In Appendices A and B we document in a more systematic way how an econometrician analyzing time-series data on actions from our model recovers evidence of persistence in the mean and variance of actions. For example, we find that even the benchmark case of an iid state generates significant volatility clustering, as well as a mean action whose persistence is given by an autoregressive coefficient of 0.6. Both types of persistence decrease when either the cognition cost, κ , or the local informativeness of signals, ψ , is lower.

3.5 Cross-sectional distributions

In our analysis so far we have described the average individual behavior by focusing on an average realization of the reasoning error. In this section we expand the analysis to include the effects of such errors. We do so at the individual level, where we emphasize the stochastic nature of choices, and then analyze implications for the cross-sectional distribution.

3.5.1 Cross-sectional distribution of reasoning errors

Here we introduce a continuum of agents indexed by *i* that are ex-ante identical in the sense that they have the same initial prior $\mu_{i,0}(y_t)$, they solve the same problem and face the same parameters, including their cognitive cost. Each agent's objective is to choose the current action that minimizes expected quadratic deviations from the otherwise identical optimal action function $c^*(y_t)$. There are no other strategic considerations between agents that affect this problem. The only source of heterogeneity in this economy is the specific history of reasoning signals that each agent has received about $c^*(y_t)$.

The history of observed states is common to all agents, given by the history of aggregate states y^t . Therefore, the choice of the reasoning intensity is also common across i at each t, since this choice only depends on the observed states and the rest of the structural parameters, which we have assumed to be the same across agents. It follows that agents share the same covariance function $\Sigma_{t-1}(y_t)$ at each time t. There is still potentially heterogeneity in the prior mean $\mu_{i,t-1}(y_t)$ due to the particular history of realizations of the reasoning errors.

Therefore, the signal structure is

$$\eta_{i,t} = c^*(y_t) + \sigma_{\eta,t}\varepsilon_{i,t},$$

where $\sigma_{\eta,t}$ is the optimally chosen variance of noise, common to all agents, and $\varepsilon_{i,t}$ is the resulting idiosyncratic reasoning error made by agent *i*.

The posterior beliefs about the optimal action function follow

$$\mu_{i,t}(y) = \mu_{i,t-1}(y) + \alpha_t^*(y_t) \frac{\sum_{t-1}(y, y_t)}{\sum_{t-1}(y_t)} \left[\eta_{i,t} - \mu_{i,t-1}(y)\right]$$
(15)

where, entering period t, $\mu_{i,t-1}$ and Σ_{t-1} denote as before the prior mean and the covariance functions, respectively. The previous analysis on the optimal signal-to-noise ratio applies, and we obtain that when $\Sigma_{t-1}(y_t) < \frac{\kappa}{2W_{cc}}$, then agents choose not to reason and set $\sigma_{\eta,t} = \infty$. Otherwise, the optimal signal-to-noise ratio is given by equation (7) as:

$$\alpha_t^*(y_t) = 1 - \frac{\kappa}{2W_{cc}\Sigma_{t-1}(y_t)} \tag{16}$$

3.5.2 Individual stochastic choice

Let us now focus on the observed behavior for an individual agent i. In this context, experimental studies have widely documented a stochastic element in observed choices. More precisely, a given subject does not make always the same choice even when the same set of alternatives is presented. Our model of optimal deliberation generates such stochastic choices as a result of the reasoning errors drawn from a distribution whose dispersion is controlled by the procedurally rational agent.

The average observed action at time t is given by equation (15) where $\eta_{i,t} = c^*(y_t)$, while the standard deviation of the action around that mean equals $\sqrt{.5\alpha_t^*(y_t)\kappa/W_{cc}}$. Figure 9 plots the resulting 95% confidence interval for the observed action. The solid blue line and the dashed orange lines are the same policy functions as in Figure 5 indicating the agent's average action and the true optimal function, respectively. The solid yellow line and the dotted purple lines now plot the lower and the higher bound of the 95% confidence interval, respectively. We note that when the state is more unusual the agent chooses to exert more effort into reasoning, while for more typical states the agent relies more heavily on previous experiences in the form of the prior $\mu_{i,t-1}(y_t)$. Consistent with a large experimental literature in our model the observed action is more volatile for the currently more unusual states.



Figure 9: 95% confidence interval around the ergodic policy function. The figure plots the ergodic average action together with lower and upper bounds on the 95% confidence interval around it. The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$.

3.5.3 Cross-sectional distribution of policy functions

While the stochastic nature of choices for the same individual is an intrinsic part of the model, the fact that the agent learns about a function implies systematic predictions from the history of signals to the observed behavior, as indicated by equation (15). Since different agents will tend to observe specific histories of reasoning signals leading to different policy functions, an analyst observes variation in behavior not only as individual stochastic elements of choice, through the current realized $\varepsilon_{i,t}$, but also in its systematic part of it, through the heterogeneity in the prior $\mu_{i,t-1}(y_t)$.



Figure 10: Policy functions for different history of reasoning signals. The figure illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$.

In contrast to the typical policy function that we showed in Figure 5, based on feeding to one agent a history of realizations of $\varepsilon_t = 0$, $\forall t$, in Figure 10 we illustrate two different histories of reasoning signals that lead to two very different policy functions. In the left panel one agent has experienced a string of reasoning signals that point towards a low average action, indicated by the low average prior $\mu_{i,t-1}(y)$ in the dotted line. Interestingly, in this case, the prior itself is non-monotonic in the state. We then feed in the average current signal $\eta_{i,t} = c^*(y)$, given by the dashed line, to focus on how different histories of signals shape the current decision. On average the new signal tends to indicate a significant rise in the best course of action for positive state innovations, leading to a strong upward change in the posterior. Around the mean innovation however, this agent has accumulated enough information that no further reasoning occurs so the posterior equals the prior. The resulting posterior belief about the best course of action has therefore interesting properties: it is (i) on average lower than the true optimal action; (ii) non-monotonic and changing curvature with the state; (iii) steeper for positive state innovations than the true optimal action. The right panel is an illustration of a history of signals for an agent that generates the opposite behavior. Here the prior is on average high and when we feed in the average current signal, we observe properties of the action that are the reverse of those in the left panel.

Our analysis therefore points to a possible interpretation of observed stochastic and heterogeneous behavior that is based on optimal deliberation. An analyst using data from our model recovers stochastic choice for each agent i, as well as heterogeneity in their typical behavior. For example, the agent described by the left (right) panel of Figure 10 looks like having a negative (positive) average bias in the action, with local changes in behavior that are non-monotonic as well as particularly large for positive (negative) innovations.

The recovered differences are not the manifestation of typical behavioral biases in the form of ex-ante individual differences that do not adjust as the environment structurally changes. Instead, our agents are ex-ante identical: they start with the same prior and learn about the same optimal action, but are ex-post different because they experience different internal signal realizations on the best course of action. In turn, these agents do not mechanically follow rules but are procedurally rational so that the degree of their reasoning errors respond to the state and to the characteristics of the environment.

3.5.4 First and second moments in the cross-section

Let \hat{c}_t denote the time t cross-sectional average action, i.e. $\hat{c}_t \equiv \int \mu_{i,t} di$, where $\mu_{i,t}$ is the agent i's action and $\overline{\sigma}_t \equiv \int (\mu_{i,t} - \hat{c}_t)^2 di$ denote the cross-sectional dispersion of actions.

The law of large numbers implies that the reasoning errors $\varepsilon_{i,t}$ average out, so that the average action has a recursive structure

$$\hat{c}_t = (1 - \alpha_t^*(y_t))\hat{c}_{t-1} + \alpha_t^*(y_t)y_t,$$
(17)

which is the same as the typical evolution of the posterior mean μ_t for an individual agent in equation (8). Therefore, the same properties for \hat{c}_t apply here, including non-linear responses to the state, persistence, possible hump-shape dynamics and volatility clustering.

To characterize the dispersion of actions, we use the optimal choice of $\alpha_t^*(y_t)$ given by equation (16), substitute the average action from equation (17) and note the orthogonality property of the reasoning errors to obtain the second moment $\hat{\sigma}_t$ as

$$\hat{\sigma}_t = (1 - \alpha_t^*(y_t))^2 \int \left[\mu_{i,t-1}(y_t) - \int \mu_{i,t-1}(y_t) di \right]^2 di + \frac{\alpha_t^*(y_t)\kappa}{2W_{cc}},\tag{18}$$

where the last term follows from $\int \varepsilon_{i,t}^2 di = 1$.

Consider now the cyclical movements in $\hat{\sigma}_t$. There are two types of state-dependency present in equation (18). First, the dispersion of priors is state-dependent, since agents have different priors about the optimal action at y_t . The source of this dispersion is the variability in the past updates about the optimal action obtained by different agents. It is now useful to relate back to Figure 4 which plots the probability distribution of the values of y_t where a typical agent has deliberated in the past. A defining characteristic of the ergodic behavior is that the region around \bar{y} is the one where the agents' beliefs are most certain. This larger certainty is obtained by accumulating and responding to the reasoning signals that each agent has obtained. Stronger responses to these idiosyncratic signals create larger dispersion of priors $\mu_{i,t-1}(y)$. In contrast, further away from \bar{y} , there is less relevant information accumulated through time and as a consequence the agents' priors $\mu_{i,t-1}(y)$ are less dispersed. Therefore, we expect that the dispersion of priors decreases with $|y_t|$. The dashed line in Figure 11 illustrates this property at the ergodic distribution.



Figure 11: Cross-sectional dispersion from reasoning errors. The figure plots then ergodic dispersion of actions, as well as in priors, together with the ergodic signal-to-noise ratio. illustrates an example with parameters $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$.

Second, the signal-to-noise ratio is state-dependent: as $|y_t|$ increases, the prior uncertainty is larger, which in turn increases the optimal $\alpha_t^*(y_t)$ and agents deliberate about the optimal action. The effect of a larger $\alpha_t^*(y_t)$ works in two opposite directions. On the one hand, the larger reliance on new idiosyncratic signals reduces the dispersion through the lower weight attached to differences in priors, as given by the first term of equation (18). This force amplifies the fact that the dispersion of priors tends to be lower for larger $|y_t|$. On the other hand, the new signals have errors with an optimally chosen standard deviation of $\sigma_{\eta,t}$ and update the posterior by $\alpha_t^*(y_t)\sigma_{\eta,t}$. These idiosyncratic signals increase the dispersion in posterior beliefs by a factor proportional to $\alpha_t^*(y_t)$, given by the second term of equation (18). The dotted line in Figure 11 illustrates this effect.

Overall, the effect of the absolute size of the state $|y_t|$ on the measured $\hat{\sigma}_t$ depends on the relative strength of these state-dependencies. The increase in dispersion of actions when the state is unusual, following the newly acquired reasoning signals, tends to dominate when the dispersion of priors is relatively flat with respect to the aggregate state, as it is in the numerical examples we have shown so far. The solid line in Figure 11 illustrates the relationship between $|y_t|$ and $\hat{\sigma}_t$ at the ergodic distribution for the baseline parametrization.

4 Two Actions

In this section we consider the extension of our model to a case where the agent takes multiple actions, which we call c and l (e.g. consumption and labor). The agent seeks to minimize the sum of squared deviations of both actions from the unknown policy functions $c^*(y)$ and $l^*(y)$:

$$W_{cc}E_t(\hat{c}(y_t) - c^*(y_t))^2 + W_{ll}E_t(\hat{l}(y_t) - l^*(y_t))^2,$$

where W_{cc} and W_{ll} parameterize the cost of mistakes in terms of the first and the second action respectively. We model the uncertainty over the vector of policy functions $[c^*(y), l^*(y)]$ as a vector Gaussian Process distribution,

$$\begin{bmatrix} c^*(y) \\ l^*(y) \end{bmatrix} \sim \mathcal{GP}\left(\begin{bmatrix} \mu^c(y) \\ \mu^l(y) \end{bmatrix}, \begin{bmatrix} \Sigma^c(y,y') & \Sigma^{cl}(y,y') \\ \Sigma^{cl}(y,y') & \Sigma^l(y,y') \end{bmatrix}\right),$$

where now $\mu^{c}(y)$ and $\mu^{l}(y)$ represent the prior mean functions over the $c^{*}(.)$ and $l^{*}(.)$ policy functions respectively, and $\Sigma^{c}(y, y')$ and $\Sigma^{l}(y, y')$ are the covariance functions within the two functions $c^{*}(.)$ and $l^{*}(.)$ respectively, and $\Sigma^{cl}(y, y')$ is the covariance function across the two policy functions $c^{*}(.)$ and $l^{*}(.)$. All covariance functions are of the squared exponential family

$$\operatorname{Cov}(c^*(y), c^*(y')) = \sigma_c^2 e^{-\psi_c(y-y'^2)}; \quad \operatorname{Cov}(l^*(y), l^*(y')) = \sigma_l^2 e^{-\psi_l(y-y'^2)};$$

including the cross-term $\operatorname{Cov}(c^*(y), l^*(y')) = \sigma_{cl}^2 e^{-\psi_{cl}(y-y'^2)}$, where the respective parameters play the same role as before. We focus on the case where the decay rate of information in the distance between y and y' is the same so that $\psi_c = \psi_l = \psi_{cl} = \psi$.

The deliberation process is modeled as a choice over the precision of unbiased signals

$$\eta^c(y_t) = c^*(y_t) + \varepsilon_t^{\eta_c}; \ \eta^l(y_t) = l^*(y_t) + \varepsilon_t^{\eta_t}$$

about the underlying policy functions $c^*(.)$ and $l^*(.)$. This amounts to choosing the respective variances of the idiosyncratic noise terms in the above signals. Lastly, the cognition cost is again a linear function of the total information the agent acquires about his (vector of) optimal actions, as measured by Shannon mutual information:

$$I(\begin{bmatrix} c^*(y)\\ l^*(y) \end{bmatrix}; \boldsymbol{\eta}(y_t), \boldsymbol{\eta}^{t-1}) = H(\begin{bmatrix} c^*(y)\\ l^*(y) \end{bmatrix} | \boldsymbol{\eta}^{t-1}) - H(\begin{bmatrix} c^*(y)\\ l^*(y) \end{bmatrix} | \boldsymbol{\eta}(y_t), \boldsymbol{\eta}^{t-1})$$

where we define the bold symbol $\eta(y_t)$ as the vector of signals η^c and η^l :

$$oldsymbol{\eta}(y_t) = \left[egin{array}{c} \eta^c(y_t) \ \eta^l(y_t) \end{array}
ight]$$

Hence the costly deliberation framework is a straight extension of the univariate case. The agent faces the following reasoning problem:

$$U = \max_{\hat{\sigma}_{ct}^2, \hat{\sigma}_{lt}^2} -W_{cc}\hat{\sigma}_{ct}^2(y_t) - W_{ll}\hat{\sigma}_{lt}^2(y_t) - \kappa I\left(\begin{bmatrix} c^*(y) \\ l^*(y) \end{bmatrix}; \boldsymbol{\eta}(y_t), \boldsymbol{\eta}^{t-1}\right)$$

s.t. $\hat{\sigma}_{ct}^2(y_t) \le \hat{\sigma}_{c,t-1}^2(y_t); \ \hat{\sigma}_{lt}^2(y_t) \le \hat{\sigma}_{l,t-1}^2(y_t)$

where for convenience we define the notation

$$\hat{\sigma}_{ct}^2(y_t) = \operatorname{Var}_t(c^*(y_t)); \ \hat{\sigma}_{lt}^2(y_t) = \operatorname{Var}_t(l^*(y_t))$$

Given an optimal choice of the signal error variance, there are corresponding effective signal to noise ratios $\alpha_c^*(y_t; \boldsymbol{\eta}_t, \boldsymbol{\eta}^{t-1})$ and $\alpha_l^*(y_t; \boldsymbol{\eta}_t; \boldsymbol{\eta}^{t-1})$ respectively, and the resulting effective actions taken by the agent are given by:

$$\hat{c}_t(y_t) = E_t(c^*(y_t)) = \hat{c}_{t-1}(y_t) + \alpha_c^*(y_t; \boldsymbol{\eta}_t, \boldsymbol{\eta}^{t-1})(\eta_t^c - \hat{c}_{t-1}(y_t))$$
(19)

$$\hat{l}_t(y_t) = E_t(l^*(y_t)) = \hat{l}_{t-1}(y_t) + \alpha_l^*(y_t; \boldsymbol{\eta}_t, \boldsymbol{\eta}^{t-1})(\eta_t^l - \hat{l}_{t-1}(y_t))$$
(20)

4.1 **Optimal Deliberation**

We focus on the case where the agent has a prior belief that there is no correlation between the values of his two actions and hence $\sigma_{cl}^2 = 0.30$ The optimal deliberation choice is represented

³⁰This is a particularly straightforward case to analyze because the optimal deliberation about each action is independent of the choice of deliberation about the other. Moreover, it is still an interesting special case because some of the fundamental features of the more general problem are quite transparent here.

by, this time potentially action specific, optimal target levels for the posterior uncertainty:

$$\hat{\sigma}_{ct}^{*2}(y_t) = \min\{\frac{\kappa}{2W_{cc}}, \hat{\sigma}_{c,t-1}(y_t)\}; \ \hat{\sigma}_{lt}^{*2}(y_t) = \min\{\frac{\kappa}{2W_{ll}}, \hat{\sigma}_{l,t-1}(y_t)\}$$

The overall features of the ergodic distribution of beliefs is thus similar to the univariate case, and hence we focus on the new features of the multidimensional problem presented here. The most interesting such features arise in the case when there is an asymmetry in the costliness of mistakes, $W_{cc} \neq W_{ll}$. Without loss of generality we assume that

$$W_{cc} > W_l$$

and illustrate the ergodic effective policy functions of such an example in Figure 12.



Figure 12: Two Actions Optimal Policy. Benchmark values $-W_{cc} = 1, W_{ll} = 0.5, \kappa = 0.5, \sigma_c^2 = 1, \sigma_l^2 = 1, \sigma_{cl}^2 = 0, \psi = 1, \text{ and } \delta = 0.9.$ Changing values: $\kappa = 0.15, \psi = 3, \text{ and } \delta = 0.99.$

As we can see, both policy functions display the characteristic non-linearity that we saw in the univariate case, with a region of relative inertia around the mean values of y_t , and salience-like effects for more distant values. But while the basic implications of the previous section readily extend to the multidimensional case, there are interesting differences in the extent to which the two separate effective policy functions, \hat{c} and \hat{l} , display those basic non-linear features. Since the agent finds it relatively less costly to make mistakes in terms of his second action, \hat{l} , he optimally ends up deliberating less about it. At the ergodic distribution, this results in an effective policy function \hat{l} that is much flatter than the effective policy \hat{c} , and that also displays less reaction even in the tails of y_t realizations.

4.2 Comovement and inference on the cognitive friction

The optimal deliberation model implies that the *comovement* between the two actions is time-varying. For realizations of y_t closer to the mean \bar{y} , the agent adjusts \hat{c} much more than he does \hat{l} , resulting in a relatively low comovement between the two. This is due to two forces. First, the ergodic distribution of beliefs differs across the two actions. Since the agent cares more about \hat{c} , he tends to accumulate more precise signals about it over time, which results in a more precise ergodic belief, $\bar{\mu}^c(y)$, that more accurately tracks the true $c^*(y)$. Second, for any given current y_t , the agent chooses to deliberate more about \hat{c} than \hat{l} , but this difference declines for more unusual realizations of y_t . Essentially, while in a typical period the agent spends less time reasoning about the less important action, when presented with an unusual state the agent finds it useful to deliberate about both. This results in time-varying comovement between the two actions, and to an econometrician who does not take into account the non-linear implications of our costly deliberation framework, it would look like a change in preferences, or another type of a regime shift.

The standard Rational Inattention (RI) framework cannot deliver such time-varying comovement between actions. Since there the friction is in terms of observing the objective state, any information about the unknown y_t informs both actions equally. The agent has full knowledge of the two individual optimal policy functions, and hence applies the information about y_t equally to both. As a result, the co-movement between the two effective actions $\hat{c}(.)$ and $\hat{l}(.)$ is constant for all values of y_t . Moreover, the overall choice of how much information to process about y_t will depend on the weighted average of the costliness of mistakes in terms of both \hat{c} and \hat{l} together. Thus, increasing W_{cc} , for example, will not only make the optimal $\hat{c}(.)$ more accurate, but will in fact increase the accuracy of both actions at the same time.

Our setup affects an econometrician's inference about the size of the underlying cognitive friction. To visualize these inference effects, the dashed lines in Figure 12 illustrate the resulting RI action estimated on data about \hat{c} only. While the RI model cannot deliver the non-linear features of the optimal \hat{c} , it will obtain the best linear fit, and hence imply behavior that is on average correct with the underlying \hat{c} . However, if we are to then extrapolate the estimated model to predict the agent's behavior in terms of the other action, \hat{l} , the prediction will have a significant bias to it, as illustrated in the right panel of Figure 12. The resulting optimal RI action in terms of \hat{l} is the same as that for \hat{c} , since they are equally loading on the same underlying belief about y_t . When the RI model sees data from \hat{c} , where the agent tracks his actual optimal $c^*(.)$ better, it infers that the agent pays a lot of attention to y_t and hence tracks it well. But given such precise beliefs about y_t the RI model would imply that the agent is freely using that information in setting his other action, resulting in a quite responsive \hat{l} . To the contrary, however, in our model the agent generally does not respond much to \hat{l} , and in particular not nearly as much as he does in terms of \hat{c} . Similar issues arise if the econometrician instead only uses data from \hat{l} and makes inference about behavior in terms of the other action, or tries to fit a single model to data on both actions.

5 Conclusion

In this paper we have developed a tractable model to study costly cognition. We have assumed that agents perfectly observe state variables, but have limited cognitive resources that prevent them from computing their optimal policy *function*. We have shown that the resulting actions are characterized by several empirically plausible features: (i) endogenous persistence: even if the observed states are iid, there is persistence in the agent's beliefs about the conditional optimal course of action; (ii) non-linearity: typical behavior exhibits inertia and 'salience' effects; (iii) stochasticity: the action is stochastic, even conditioning on the perfectly observed objective state, as it is driven by random reasoning signals; (iv) cross-sectional heterogeneity in policy functions: agents' experiences may lead to average biases and local changes that are non-monotonic; (v) endogenous persistence in the volatility of actions: following unusual times both the time-series and the cross-sectional variance increases; (vi) state-dependent comovement and accuracy in different actions.

The model has potentially important policy implications. First, it offers a cohesive framework to understand an array of features for the individual behavior of a procedurally rational agent. The errors made by this agent do not arise from mechanical decision rules, but instead respond to the state and the characteristics of the environment, including policy changes, in the spirit of Lucas (1976). Second, the friction may help understand macroeconomic phenomena such as non-linearity, persistence, volatility clustering, and in the process may change inference on the underlying sources of economic mechanisms and shocks. Indeed, an econometrician that fits a standard fully rational agent model to the equilibrium outcomes generated by our model may conclude that there are policy-invariant sources of non-linearity and time-variation in parameters. Instead, through the lenses of our model, these non-linearities and apparent time-varying parameters are a manifestation of a single cognitive friction that leads economic agents inside the model to act according to state-dependent reasoning rules.

Finally, our aggregate-level implications have abstracted from general equilibrium type of interactions. Incorporating these interactions involves studying the propagation effects of the state-dependent individual errors in actions, as well as modeling the agents' reasoning choice of computing potentially complex general equilibrium effects, issues that we find promising for future research.

References

- ABALUCK, J. AND J. GRUBER (2011): "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program," *The American Economic Review*, 101, 1180–1210.
- AKERLOF, G. A. AND J. L. YELLEN (1985a): "Can small deviations from rationality make significant differences to economic equilibria?" *The American Economic Review*, 75, 708–720.
- (1985b): "A near-rational model of the business cycle, with wage and price inertia," *The Quarterly Journal of Economics*, 100, 823–838.
- ALAOUI, L. AND A. PENTA (2016): "Cost-Benefit Analysis in Reasoning," Working Paper.
- ANGELETOS, G.-M. AND C. LIAN (2017): "Dampening General Equilibrium: From Micro to Macro," NBER Working Paper No. 23379.
- 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.
- BISHOP, C. M. (2006): Pattern recognition and machine learning, Springer.
- BLOOM, N. (2014): "Fluctuations in Uncertainty," Journal of Economic Perspectives, 28, 153–176.
- CAPLIN, A., M. DEAN, AND J. LEAHY (2016): "Rational Inattention, Optimal Consideration Sets and Stochastic choice," Working paper.
- CAPLIN, A., M. DEAN, AND D. MARTIN (2011): "Search and satisficing," *The American Economic Review*, 101, 2899–2922.
- CARLIN, B. I., S. KOGAN, AND R. LOWERY (2013): "Trading complex assets," *The Journal* of Finance, 68, 1937–1960.
- CARVALHO, L. AND D. SILVERMAN (2017): "Complexity and Sophistication," Working Paper.
- COGLEY, T. AND T. J. SARGENT (2005): "Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of Economic Dynamics*, 8, 262–302.
- DECK, C. AND S. JAHEDI (2015): "The effect of cognitive load on economic decision making: A survey and new experiments," *European Economic Review*, 78, 97–119.
- DUPOR, B. (2005): "Stabilizing non-fundamental asset price movements under discretion and limited information," *Journal of Monetary Economics*, 52, 727–747.

- EIDELMAN, S. AND C. S. CRANDALL (2009): "A psychological advantage for the status quo," Social and psychological bases of ideology and system justification, 85–106.
- ERGIN, H. AND T. SARVER (2010): "A unique costly contemplation representation," *Econo*metrica, 78, 1285–1339.
- FARHI, E. AND I. WERNING (2017): "Monetary Policy, Bounded Rationality, and Incomplete Markets," NBER Working Paper No. 23281.
- FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, AND F. SCHORFHEIDE (2016): "Solution and estimation methods for DSGE models," *Handbook of Macroeconomics*, 2, 527–724.
- GABAIX, X. (2014): "A sparsity-based model of bounded rationality," *The Quarterly Journal* of *Economics*, 129, 1661–1710.
- ——— (2016): "Behavioral macroeconomics via sparse dynamic programming," NBER Working Paper 21848.
- GARCÍA-SCHMIDT, M. AND M. WOODFORD (2015): "Are low interest rates deflationary? A paradox of perfect-foresight analysis," NBER Working Paper No. 21614.
- 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.
- JUSTINIANO, A. AND G. E. PRIMICERI (2008): "The time-varying volatility of macroeconomic fluctuations," *The American Economic Review*, 98, 604–641.
- KACPERCZYK, M., S. VAN NIEUWERBURGH, AND L. VELDKAMP (2016): "A rational theory of mutual funds' attention allocation," *Econometrica*, 84, 571–626.
- KALAYCI, K. AND M. SERRA-GARCIA (2016): "Complexity and biases," *Experimental Economics*, 19, 31–50.
- LIU, W., J. C. PRINCIPE, AND S. HAYKIN (2011): Kernel adaptive filtering: a comprehensive introduction, vol. 57, John Wiley & Sons.
- LUCAS, R. E. (1976): "Econometric policy evaluation: A critique," in *Carnegie-Rochester* conference series on public policy, vol. 1, 19–46.
- 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.

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

- (2015b): "Inattention to rare events," Working Paper.
- MATĚJKA, F. (2015): "Rationally inattentive seller: Sales and discrete pricing," *The Review* of *Economic Studies*, 83, 1125–1155.
- 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.
- MATĚJKA, F., J. STEINER, AND C. STEWART (2017): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85, 521–553.
- MATTSSON, L.-G. AND J. W. WEIBULL (2002): "Probabilistic choice and procedurally bounded rationality," *Games and Economic Behavior*, 41, 61–78.
- MELOSI, L. (2014): "Estimating models with dispersed information," American Economic Journal: Macroeconomics, 6, 1–31.
- 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.
- OLIVEIRA, H., T. DENTI, M. MIHM, AND K. OZBEK (2017): "Rationally inattentive preferences and hidden information costs," *Theoretical Economics*, 12, 621–654.
- PACIELLO, L. AND M. WIEDERHOLT (2013): "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *Review of Economic Studies*, 81, 356–388.
- RASMUSSEN, C. E. AND C. K. WILLIAMS (2006): *Gaussian processes for machine learning*, vol. 1, MIT press Cambridge.
- REIS, R. (2006a): "Inattentive consumers," Journal of Monetary Economics, 53, 1761–1800.
- (2006b): "Inattentive producers," The Review of Economic Studies, 73, 793–821.
- SAMUELSON, W. AND R. ZECKHAUSER (1988): "Status quo bias in decision making," Journal of risk and uncertainty, 1, 7–59.
- SCHRAM, A. AND J. SONNEMANS (2011): "How individuals choose health insurance: An experimental analysis," *European Economic Review*, 55, 799–819.
- SETHI-IYENGAR, S., G. HUBERMAN, AND W. JIANG (2004): "How much choice is too much? Contributions to 401 (k) retirement plans," *Pension design and structure: New lessons from behavioral finance*, 83, 84–87.

- 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.
- (2006): "Rational inattention: Beyond the linear-quadratic case," *The American* economic review, 96, 158–163.
- (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. (2014): "Coarse Pricing Policies," Manuscript, Univ. of Maryland.
- STOCK, J. H. AND M. W. WATSON (2002): "Has the US business cycle changed and why?" NBER Macroeconomics Annual, 17, 159–218.
- TUTINO, A. (2013): "Rationally inattentive consumption choices," *Review of Economic Dynamics*, 16, 421–439.
- TVERSKY, A. AND D. KAHNEMAN (1975): "Judgment under uncertainty: Heuristics and Biases," in *Utility, probability, and human decision making*, 141–162.
- VAN DAMME, E. (1987): Stability and perfection of Nash equilibria, Springer Verlag.
- VAN NIEUWERBURGH, S. AND L. VELDKAMP (2009): "Information immobility and the home bias puzzle," *The Journal of Finance*, 64, 1187–1215.
- (2010): "Information acquisition and under-diversification," *The Review of Economic Studies*, 77, 779–805.
- VELDKAMP, L. L. (2011): Information choice in macroeconomics and finance, Princeton University Press.
- 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.
 - (2009): "Information-constrained state-dependent pricing," Journal of Monetary Economics, 56, S100–S124.
 - (2014): "Stochastic choice: An optimizing neuroeconomic model," *The American Economic Review*, 104, 495–500.

A Statistical fit of persistence in the time-series

In this appendix section we document in a more systematic way how an econometrician analyzing time-series data on actions from our model recovers evidence of persistence as well. For this purpose, we consider an Autoregressive process of order p(AR(p)) for the action:

$$\hat{c}_t = \bar{c} + \sum_{i=1}^p \rho_i \hat{c}_{t-i} + \sigma \epsilon_t \tag{21}$$

Table 1: Time-series properties of the action

Moment \Model	Baseline	Lower κ	Higher ψ
Persistence $\hat{\rho}_1$.61	.18	.19
	[.59, .63]	[.17, .21]	[.17, .21]

Note: This table reports results for estimating ρ_1 in equation (21). The baseline parametrization is $W_{cc} = 1, \kappa = 0.5, \sigma_c^2 = 1, \psi = 1$, and $\delta = 0.9$. In the third column $\kappa = .1$, while in the fourth column $\psi = 5$. In squared brackets we report the 95% confidence interval.

Table 1 summarizes results for the baseline model and some alternative specifications.³¹ The model is characterized by significant evidence against the null of no persistence in actions. We focus there on an AR(1) process and find that the estimated ρ_1 parameter is significantly positive indicating a strong departure from the iid assumption on the evolution of the state y_t . The intuition for the presence of endogenous persistence follows the logic of the impulse response function presented in section 3.3. The key mechanism is that the information acquired about the optimal action at some particular realization of the state y_t is perceived to be informative about the optimal action at a different value y'. The discussion in sections 3.2 and 3.3 highlights the two essential structural forces for these time-series results. In Table 1 both for the alternative parameterizations of 'Lower κ ' and 'Higher ψ ', the effects of reasoning about the optimal function are weaker, in the form of smaller estimated ρ_1 .

In Table 2, we also estimate an AR(2) process and the best fitting AR(p) model. For the baseline parametrization the AR(2) improves the fit over the AR(1) with significant coefficients at both lags. The best fit, according to the Akaike information criterion, has 6 lags with a sum of coefficients equal to 0.71. While in the baseline model the state is iid, we explore the effects of increasing that persistence to 0.5 and 0.9, respectively. In the former case, the cognition friction leads to a higher persistence in actions, equal to 0.83. In the latter case, the model also generates significant hump-shape dynamics. In particular, when we fit an AR(2) process, the impulse response function transmits a raise in c_t into an increase in c_{t+j} that mean reverts only at j = 6. This confirms the hump-shape dynamics in Figure 7.

³¹We simulate 10000 periods and drop the first 500 so to compute long run moments.

Model/Moment	AR(1)	AR(2)		AR(p)	
Persistence in actions	ρ	$ ho_1$	ρ_2	p^*	$\sum_{i=1}^{p^*} \rho_i$
Baseline	.61	.48	.2	6	.71
Shock persistence $\rho_y = .5$.83	.85	02	5	.83
Shock persistence $\rho_y = .9$.97	1.15	18	8	.95

Table 2: AR(p) processes for action

Note: This table reports results for estimating AR(p) processes for the average action, as in equation (21).

B Statistical fit of volatility clustering in the time-series

To measure volatility clustering, we first use the AR(p) regression in equation (21) to compute squared residuals $\hat{\epsilon}_t^2$. We then regress them on the previous absolute value of actions $\tilde{c}_{t-1} \equiv c_{t-1}I(|c_{t-1}| > \overline{\sigma})$, where the latter indicator equals one if the absolute value of the action c_{t-1} is larger than a threshold proportional to the measured unconditional standard deviation of actions, denoted by $\overline{\sigma}$. The regression states:

$$\epsilon_t^2 = \alpha + \beta |\tilde{c}_{t-1}| + \tilde{\epsilon}_t \tag{22}$$

Table 3: Time-series properties of the action

Moment \Model	Baseline	Lower κ	Higher ψ
Volatility cluster $\widehat{\beta}$.6	.48	.5
	[.3,1]	[01,1]	[0,1]

Note: This table reports results for estimating β in equation (22) for the baseline and alternative parametrizations of Table 1. In squared brackets we report the 95% confidence interval.

Table 3 reports evidence for volatility clustering, in the form of a significantly positive β in equation (22).³² The intuition for this finding is presented in section 3.4.1. After more unusual states, which in turn generate more unusual actions compared to the ergodic action, in the form of a higher $|c_{t-1}|$, the agent finds it optimal to reason more intensely. This increased deliberation leads on average to an updated belief that is more responsive, which on average generates a larger variability in actions in the next period, resulting in larger residuals ε_t^2 . The effects are smaller for the 'Lower κ ' and 'Higher ψ ' cases.

³²If instead of the threshold $\overline{\sigma}$ in the indicator function in (22), we use a value of 0, then the point estimate for β is still positive but not significant at a 95% confidence level.