DISCUSSION PAPER SERIES

No. 10720

RATIONAL INATTENTION DYNAMICS: INERTIA AND DELAY IN DECISION-MAKING

Jakub Steiner, Colin Stewart and Filip Matejka

INDUSTRIAL ORGANIZATION and MACROECONOMICS AND GROWTH

Centre for Economic Policy Research

ISSN 0265-8003

RATIONAL INATTENTION DYNAMICS: INERTIA AND DELAY IN DECISION-MAKING

Jakub Steiner, Colin Stewart and Filip Matejka

Discussion Paper No. 10720
July 2015
Submitted 11 July 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

RATIONAL INATTENTION DYNAMICS: INERTIA AND DELAY IN DECISION-MAKING[†]

Abstract

We solve a general class of dynamic rational-inattention problems in which an agent repeatedly acquires costly information about an evolving state and selects actions. The solution resembles the choice rule in a dynamic logit model, but it is biased towards an optimal default rule that depends only on the history of actions, not on the realized state. We apply the general solution to the study of (i) the status quo bias; (ii) inertia in actions leading to lagged adjustments to shocks; and (iii) the tradeoff between accuracy and delay in decision-making.

JEL Classification: D81, D83 and D90
Keywords: adjustment delay, dynamic logit, information acquisition and rational inattention

Jakub Steiner jakub.steiner@cerge-ei.cz
CERGE-EI, University of Edinburgh and CEPR

Colin Stewart colinbstewart@gmail.com
University of Toronto

Filip Matejka filip.matejka@cerge-ei.cz
CERGE-EI

[†] We thank Victor Aguirregabiria, Mehmet Ekmekci, Štěpán Jurajda, Jakub Kastl, Philipp Kircher, Alessandro Pavan, Marcin Pęski, Balázs Szentes, and participants in seminars and conferences at Barcelona GSE, Boston College, CERGE-EI, Dauphine, Edinburgh, IESE, Northwestern, Pittsburgh, SED-Warsaw, Surrey, and Washington University for their comments. Maxim Goryunov provided excellent research assistance. Stewart is grateful to SSHRC for a grant supporting this research.

Rational Inattention Dynamics: Inertia and Delay in Decision-Making*

Jakub Steiner†
CERGE-EI and University of Edinburgh

Colin Stewart‡
University of Toronto

Filip Matějka§
CERGE-EI

July 10, 2015

Abstract

We solve a general class of dynamic rational-inattention problems in which an agent repeatedly acquires costly information about an evolving state and selects actions. The solution resembles the choice rule in a dynamic logit model, but it is biased towards an optimal default rule that depends only on the history of actions, not on the realized state. We apply the general solution to the study of (i) the status quo bias; (ii) inertia in actions leading to lagged adjustments to shocks; and (iii) the tradeoff between accuracy and delay in decision-making.

1 Introduction

Timing of information plays an important role in a variety of economic settings. Delays in learning contribute to lags in adjustment of macroeconomic variables, in adoption of new technologies, and in prices in financial markets. The speed of information processing is a

*We thank Victor Aguirregabiria, Mehmet Ekmekci, Štěpán Jurajda, Jakub Kastl, Philipp Kircher, Alessandro Pavan, Marcin Pęski, Balázs Szentes, and participants in seminars and conferences at Barcelona GSE, Boston College, CERGE-EI, Dauphine, Edinburgh, IESE, Northwestern, Pittsburgh, SED-Warsaw, Surrey, and Washington University for their comments. Maxim Goryunov provided excellent research assistance. Stewart is grateful to SSHRC for a grant supporting this research.

†email: jakub.steiner@cerge-ei.cz
‡email: colinbstewart@gmail.com
§email: filip.matejka@cerge-ei.cz

crucial determinant of response times in psychological experiments. In each of these cases, the timing is shaped in large part by individuals' efforts to acquire information.

We study a general dynamic decision problem in which an agent chooses *what* and *how much* information to acquire, as well as *when* to acquire it. In each period, the agent can choose an arbitrary signal about a payoff-relevant state of the world before taking an action. The state follows an arbitrary stochastic process, and the agent's flow payoff is a function of the histories of actions and states. Following Sims (2003), the agent pays a cost to acquire information that is proportional to the reduction in her uncertainty as measured by the entropy of her beliefs. We characterize the stochastic behavior that maximizes the sum of the agent's expected discounted utilities less the cost of the information she acquires.

We find that the optimal choice rule coincides with dynamic logit behavior (Rust, 1987) with respect to payoffs that differ from the agent's true payoffs by an endogenous additive term.[1] This additional term, which we refer to as a "predisposition", depends on the history of actions but does not depend directly on the history of states. Relative to dynamic logit behavior with the agent's true payoffs, the predisposition increases the relative payoffs associated with actions that are chosen with high probability on average across all states given the history of actions up to that point.

If states are positively serially correlated, the influence of predispositions can resemble switching costs; because learning whether the state has changed is costly, the agent relies in part on her past behavior to inform her current decision, and is therefore predisposed toward repeating her previous action.

Our results provide a new foundation for the use of dynamic logit in empirical research with an important caveat: the presence of predispositions affects extrapolation of behavior based on identification of utility parameters. An econometrician applying standard dynamic logit techniques to the agent in our model would correctly predict her behavior in repetitions of the same dynamic decision problem. However, problems involving different payoffs or distributions of states typically lead to different predispositions, which must be accounted for in the extrapolation exercise. The difference arises because the standard approach takes switching costs as fixed when other payoffs vary, whereas the predispositions in our model vary as parameters of the environment change.

A major difficulty in solving the model and obtaining the dynamic logit characterization arises due to the influence of current information acquisition on future beliefs. One

[1] This result extends the static logit result of Matějka and McKay (2015) to the dynamic setting.

approach would be to reduce the problem to a collection of static problems using the Bellman equation, with payoffs equal to current payoffs plus continuation values that depend on posterior beliefs. However, the resulting collection of problems cannot be solved by directly applying techniques developed for static rational inattention problems (henceforth, RI problems). In static RI problems, not including information costs, expected payoffs are linear in beliefs; in the Bellman equation, continuation values are not linear in probabilities. Nevertheless, we show that the solution can be obtained in a similar fashion by ignoring the effect of information acquisition on future beliefs: one can define continuation values as a function only of the histories of actions and states, and treat those values as fixed when optimizing at each history. Because of this property, we can characterize the solution to the dynamic RI problem in terms of solutions to static RI problems that are well understood.

The key step behind the reduction to static problems is to show that the dynamic RI problem can be reformulated as a control problem with observable states. In this reformulation, the agent first chooses a default choice rule that specifies a distribution of actions at each history independent of which states are realized. Then, after observing the realized state in each period, she chooses her actual distribution of actions, and incurs a cost according to how much she deviates from the default choice rule.[2] Because states are observable in the control problem, beliefs do not depend on choice behavior; as a result, this reformulation circumvents the main difficulty in the original problem of accounting for the effect of the current strategy on future beliefs.

We illustrate the general solution in three applications. In the first, the agent seeks to match her action to the state in each of two periods. We show that positive correlation between the states can lead to an apparent status quo bias: the agent never switches her action from one period to the next, and her choice is, on average, better in the first period than in the second. The correlation between the states creates a relatively strong incentive to learn in the first period because the information she obtains will be useful in both periods. Acquiring more information in the first period in turn reduces the agent's incentive to acquire information in the second, making her more inclined to choose the same action.[3] This effect arises because of the endogeneity of information: in a standard model with exogenous signals, the agent would have superior information in the second

[2]Mattsson and Weibull (2002) study essentially the same problem for a fixed default rule in a static setting.

[3]As in Baliga and Ely (2011), the agent's second-period beliefs are directly linked to her earlier decision, although the effect here arises due to costly information acquisition rather than forgetting.

period, and thus her choices would improve over time. Similarly, the effect would not arise if information acquisition was not dynamic: if the agent could acquire information about the state in both periods before choosing her first action, then her actions would be, on average, equally good across the two periods.

Our second application can be interpreted as a simple model of lagged adjustments to shocks. The state, which is either good or bad, follows a Markov chain. The agent chooses in each period whether or not to invest, preferring to invest if and only if the current state is good. When state transitions are rare, adjustment to shocks is slow and the expected reaction lags are proportional to the time between transitions; high persistence discourages the agent from closely monitoring the state. As volatility increases and transitions become more frequent, the speed of adjustment also increases. Relative adjustment speeds are driven by underlying incentives: if the incentive to disinvest when the state is bad is stronger than the incentive to invest when it is good, then lags in adjustment are shorter after bad shocks than after good ones.

The final application concerns a classic question in psychology, namely, the relationship between response times and accuracy of decisions. The state is binary and fixed over time. The agent chooses when to take one of two actions with the goal of matching her action to the state. Delaying is costly, but gives her the opportunity to acquire more information. We focus on a variant of the model in which the cost of information is replaced with a capacity constraint on how much information she can acquire in each period. The solution of the problem gives the joint distribution of the decision time and the chosen action. We find that delay is associated with better decisions. In addition, the expected delay time is non-monotone in the agent's capacity, with intermediate levels being associated with the longest delays.

We focus throughout the paper on information costs that are proportional to the reduction in entropy of beliefs. There are two related reasons for this choice. The first is tractability. With entropy-based costs, it is not necessary to consider all possible information acquisition strategies; we show that one can restrict attention to strategies that associate at most one signal realization to each action (and hence each action history is associated with just one belief). This substantially reduces the dimensionality of the problem in that it permits direct optimization over distributions of actions instead of over signals.[4] Entropy-based costs are also important for the reformulation of the dynamic RI problem

[4]In the static case, this one-to-one association of actions and signals holds under much weaker conditions on the cost function; see the discussion in Section 2.2.

as a control problem with observable states.

The second reason for using this cost function is that it abstracts from incentives to smooth or bunch information acquisition because of the shape of the cost function. In a problem involving a one-time action choice, the cost function we use has the feature that the number of opportunities to acquire information before the choice of action is irrelevant: the cost of multiple signals spread over many periods is identical to the cost of a single signal conveying the same information (Hobson, 1969). Although varying the shape of the cost function could generate interesting and significant effects, our goal is to first understand the problem in which we abstract away from these issues.[5]

This paper fits into the RI literature. This literature originated in the study of macroeconomic adjustment processes (Sims, 1998, 2003). More recently, Mackowiak and Wiederholt (2009, 2010) and Matějka (2010) study sluggish adjustment in dynamic RI models. Luo (2008) and Tutino (2013) consider dynamic consumption problems with RI. Each of these papers focuses either on an environment involving linear-quadratic payoffs and Gaussian shocks or on numerical solutions. A notable exception is Ravid (2014), who analyzes a class of RI stopping problems motivated by dynamic bargaining. In general static RI problems, Matějka and McKay (2015) show that the solution generates static logit behavior with an endogenous payoff bias. Our dynamic extension of this result links it back to the original motivation for the RI literature.

Although optimal behavior in our model fits the dynamic logit framework, the foundation is quite different from that of Rust (1987). He derives the dynamic logit rule in a complete information model with i.i.d. taste shocks that are unobservable to the econometrician. Our model has no such shocks and focuses on the agent's information acquisition. This difference accounts for the additional payoff term in our dynamic logit result.

While information acquisition dynamics appear to be central to many economic problems, they are rarely modeled explicitly in settings with repeated action choices. Exceptions outside of the RI literature include Compte and Jehiel (2007), who study information acquisition in sequential auctions, and Liu (2011), who considers information acquisition in a reputation model. In both cases, players acquire information at most once, in the former because information is fully revealing and in the latter because the players are short-lived. Their focus is on strategic effects, whereas we study single-agent problems with repeated

[5] Moscarini and Smith (2001) focus on information costs that are convex in the volume of information and study delay in decision-making resulting from the incentive to smooth information acquisition over time. Sundaresan and Turban (2014) study a different model with a similar incentive.

information acquisition. In a single-agent setting, Moscarini and Smith (2001) analyze a model of optimal experimentation with explicit information costs of learning about a fixed state of the world.

As described above, a key step in proving our results is to reformulate the problem as a control problem. This reformulation connects logit behavior in RI to that found by Mattsson and Weibull (2002), who solve a problem with observable states in which the agent pays an entropy-based control cost for deviating from an exogenous default action distribution. We show that, in both static and dynamic settings, each RI problem is equivalent to a two-stage optimization problem that combines Mattsson and Weibull's control problem with optimization of the default distribution. Like us, Fudenberg and Strzalecki (2015) derive dynamic logit choice as a solution to a control problem. They focus on preferences over flexibility, while we focus on incomplete information and optimization of the default choice rule.

2 Model

A single agent chooses an action a_t from a finite set A in each period $t = 1, 2, \dots$. For any sequence $(y_\tau)_\tau$, let $y^t = (y_1, \dots, y_t)$ for each t. We refer to the action history a^{t-1} as the *decision node at* t. A payoff-relevant state θ_t follows a stochastic process on a finite set Θ with probability measure $\pi \in \Delta(\Theta^{\mathbb{N}})$.[6] Before choosing an action in any period t, the agent can acquire costly information about the history of states, θ^t. There is a fixed signal space X satisfying $|A| \leq |X| < \infty$. At time t, the agent can choose *any* signal about the history θ^t with realizations in X. Accordingly, a strategy $s = (f, \sigma)$ is a pair of

1. an *information strategy* f consisting of a system of signal distributions $f_t(x_t \mid \theta^t, x^{t-1})$, one for each θ^t and x^{t-1}, with the signal x_t conditionally independent of future states $\theta_{t'}$ for all $t' > t$, and

2. an *action strategy* σ consisting of a system of mappings $\sigma_t : X^t \longrightarrow A$, where $\sigma_t(x^t)$ indicates the choice of action at time t for each history x^t of signals.

Given an action strategy σ, we denote by $\sigma^t(x^t)$ the history of actions up to time t given the realized signals.

[6]The restriction to finite action and state spaces is for technical convenience; we conjecture that our results would extend to standard continuous models.

The agent receives flow utilities $u_t(a^t, \theta^t)$ that are uniformly bounded across all t. We refer to u_t as *gross utilities* to indicate that they do not include information costs. The agent discounts payoffs received at time t by a factor $\delta^{(t)} := \prod_{t'=1}^{t} \delta_{t'}$, where $\delta_{t'} \in [0,1]$ and $\limsup_t \delta_t < 1$. This form of discounting accommodates both finite and infinite time horizons.

As is standard in the RI literature, we focus throughout this paper on entropy-based information costs. Consider a random variable Y with finite support S distributed according to $p \in \Delta(S)$. Recall that the entropy

$$H(Y) = -\sum_{y \in S} p(y) \log p(y)$$

of Y is a measure of uncertainty about Y (with the convention that $0 \log 0 = 0$). At any signal history x^{t-1}, we assume that the cost of signal x_t is proportional to the conditional mutual information

$$I\left(\theta^t; x_t \mid x^{t-1}\right) = H\left(\theta^t \mid x^{t-1}\right) - E_{x_t}\left[H\left(\theta^t \mid x^t\right)\right] \tag{1}$$

between x_t and the history of states θ^t.[7] The conditional mutual information captures the difference in the agent's uncertainty about θ^t before and after she receives the signal x_t. Before, her uncertainty can be measured by $H\left(\theta^t \mid x^{t-1}\right)$. After, her level of uncertainty becomes $H\left(\theta^t \mid x^t\right)$. The mutual information is the expected reduction in uncertainty averaged across all realizations of x_t.

The agent solves the following problem.

Definition 1. *The* dynamic rational inattention problem *(henceforth dynamic RI problem) is*

$$\max_{f,\sigma} E\left[\sum_{t=1}^{\infty} \delta^{(t)}\left(u_t(\sigma^t(x^t), \theta^t) - \lambda I\left(\theta^t; x_t \mid x^{t-1}\right)\right)\right], \tag{2}$$

where $\lambda > 0$ is an information cost parameter, and the expectation is taken with respect to the distribution over sequences $(\theta_t, x_t)_t$ induced by the prior π together with the information strategy f.

[7]When x^t is attained with 0 probability, the value of $H\left(\theta^t \mid x^t\right)$ is defined arbitrarily and has no effect on the mutual information.

To simplify notation, we normalize the information cost parameter λ to 1. Although we assume the information cost parameter is fixed over time, one could allow for varying cost by adjusting the discount factors and correspondingly rescaling the flow utilities (as long as doing so does not violate the restrictions on $\delta^{(t)}$ or the uniform boundedness of the utilities).

Note that the sum in (2) converges because the gross flow payoffs are bounded, and the mutual information is bounded (since the signal space is finite).

Since the strategy depends only on the history of signals, we are implicitly assuming that the agent does not observe the realized payoffs (from which she could infer information about the states) during the decision process; the agent must pay a cost to process *any* information, even information pertaining to her own experience. Since she can learn directly about the history of states, it makes no difference whether she could also obtain costly information about past payoffs. If the realized payoffs were freely or cheaply observable, the agent's actions would be driven in part by the information they reveal. The current setting abstracts from such experimentation motives. However, we conjecture that characterizations analogous to those of Theorem 1 and Proposition 3 would hold in settings with free information about payoffs.

2.1 Applications

The following are examples that fit into the general framework. We solve these examples in Section 4.

Example 1 (Status quo bias). The agent chooses an action $a_t \in \{0, 1\}$ in each period $t = 1, 2$. In both periods, the gross flow payoff u_t is 1 if the action a_t matches the current state $\theta_t \in \{0, 1\}$, and is 0 otherwise. The two states are correlated across periods. There is no discounting. In this setting, we analyze the correlation between choices in the two periods. In particular, if the agent chooses not to acquire any information in the second period, then her behavior exhibits an apparent status quo bias insofar as she never reverses her decision.

Example 2 (Inertia). The agent chooses an action $a_t \in \{0, 1\}$ in each period $t = 1, 2, \ldots$ with the goal of matching the current state. The state θ_t follows a time-homogeneous Markov chain on the set $\{0, 1\}$. In each period $t \in \mathbb{N}$, the gross flow payoff $u(a_t, \theta_t)$ is equal to $u_a > 0$ if $a_t = \theta_t = a$, and is 0 if $a_t \neq \theta_t$. Payoffs are discounted exponentially. The solution to this problem illustrates how the speed of adjustment depends on incentives and on the persistence of states.

Example 3 (Response times). The state $\theta \in \{0,1\}$ is fixed over time. The agent has a uniform prior belief. In each period $t = 1, \ldots, T$, she chooses among taking a terminal action 0 or 1, or waiting until the next period (denoted by w). She receives a benefit of 1 if her terminal action matches the state, and a benefit of 0 otherwise. In addition, she pays a cost $c \in (0,1)$ for each period that she waits. We use this example to study the tradeoff between speed and accuracy of decision making.

2.2 Preliminaries

Our main goal is to characterize the agent's behavior, i.e. the distribution of actions along each history of states. A *(stochastic) choice rule* p is a system of distributions $p_t\left(a_t \mid \theta^t, a^{t-1}\right)$ over A, one for each θ^t and a^{t-1}, interpreted as the probability of choosing a_t conditional on histories θ^t and a^{t-1}. We say that a strategy $s = (f, \sigma)$ *generates the choice rule* p if

$$p_t\left(a_t \mid \theta^t, a^{t-1}\right) \equiv \Pr\left(\sigma\left(x^t\right) = a_t \mid \theta^t, \sigma^{t-1}\left(x^{t-1}\right) = a^{t-1}\right),$$

where the probability is evaluated with respect to the joint distribution of sequences of states and signals generated according to f. To simplify notation, we drop the t subscript on $p_t(a_t \mid \theta^t, a^{t-1})$ and write $p(a_t \mid \theta^t, a^{t-1})$.

Conversely, a choice rule p can be associated (non-uniquely) with a strategy (f, σ). Roughly speaking, one can choose a particular signal realization for each action, and then match the probability of each of those signal realizations with the probability the choice rule assigns to its associated action.[8] If $\overline{p}$ is a strategy obtained in this way from a choice rule p, we say that p *induces* $\overline{p}$.

The following lemma simplifies the analysis considerably by allowing us to focus on a special class of information strategies in which signals correspond directly to actions. See also Ravid (2014), who has independently proved the corresponding result in a related dynamic model.

[8] Formally, fix any injection $\phi : A \longrightarrow X$ and, by a slight abuse of notation, for any t, let ϕ also denote the mapping from A^t to X^t obtained by applying ϕ coordinate-by-coordinate. Given any choice rule p, let $\overline{p} = (f, \sigma)$ be such that $f_t(\phi(a_t) \mid \theta^t, \phi(a^{t-1})) \equiv p(a^t \mid \theta^t, a^{t-1})$ and $\sigma(\phi(a^t)) \equiv a_t$.

Lemma 1. *Any strategy s solving the dynamic RI problem generates a choice rule p solving*

$$\max_p E\left[\sum_{t=1}^{\infty} \delta^{(t)} \Big(u_t\big(a^t, \theta^t\big) - I\big(\theta^t; a_t \mid a^{t-1}\big)\Big)\right], \tag{3}$$

where the expectation is with respect to the distribution over sequences $(\theta_t, a_t)_t$ *induced by p and the prior,* π*. Conversely, any choice rule p solving* (3) *induces a strategy solving the dynamic RI problem.*

Accordingly, we abuse terminology slightly by calling (3) the dynamic RI problem, and any rule p solving (3) a solution to the dynamic RI problem. Proofs are in the appendix.

To understand why the lemma holds, consider for contradiction a strategy s such that, at some decision node, two distinct signal realizations (generating distinct posterior beliefs) map to the same action. In that case, the strategy s acquires more information at that node than is required for the current choice. One can then coarsen the signal to correspond directly to the current action choice and recover all lost information by enriching the signal in the following period. Doing so has no effect on behavior. Nor does it affect the mutual information across those two periods, and since the agent discounts future costs, it therefore cannot increase the total information cost. By recursively delaying all excess information in this way, one is left with a strategy that associates to each action a unique signal.

In static models, the conclusion of Lemma 1 holds as long as the cost of signals is nondecreasing in Blackwell informativeness. In dynamic problems, more structure is needed. For example, if the cost was concave in the mutual information then the agent could have an incentive to acquire more information than what is necessary for her choice in a given period if she plans to use that information in a later period where the marginal cost of acquiring it would be higher. If costs are proportional to reduction in entropy (and discounted costs do not increase over time), there is no incentive for the agent to acquire information any earlier than necessary.

Proposition 1. *There exists a solution to the dynamic RI problem.*

3 Solution

In this section, we first describe two characterizations of the solution to the dynamic RI problem—the first in relation to dynamic logit behavior, and the second in relation to static

RI problems. Both characterizations rely on a reformulation of the problem as a control problem with observable states described in Section 3.3.

3.1 Dynamic logit

Our main result states that the solution of the dynamic RI problem is a dynamic logit rule with a bias. We begin by recalling the definition of dynamic logit.[9]

Definition 2 (Rust (1987))**.** *A choice rule r is a* dynamic logit rule under payoffs $(u_t)_t$ *if*

$$r_t\left(a_t \mid \theta^t, a^{t-1}\right) = \frac{e^{\hat{u}_t\left(a^t,\theta^t\right)}}{\sum_{a_t'} e^{\hat{u}_t\left(\left(a^{t-1},a_t'\right),\theta^t\right)}},$$

where

$$\hat{u}_t\left(a^t,\theta^t\right) = u_t\left(a^t,\theta^t\right) + \delta_{t+1} E_{\theta_{t+1}}\left[V_{t+1}\left(a^t,\theta^{t+1}\right) \mid \theta^t\right],$$

and the continuation values V_t satisfy

$$V_t\left(a^{t-1},\theta^t\right) = \log\left(\sum_{a_t} e^{\hat{u}_t\left(\left(a^{t-1},a_t\right),\theta^t\right)}\right). \tag{4}$$

The solution to the dynamic RI problem is a dynamic logit rule with an endogenous state-independent utility term. A *default rule* q is a system of conditional action distributions $q_t(a_t \mid a^{t-1})$, one for each decision node a^{t-1}. The difference between a default rule and a choice rule is that the latter conditions on states while the former does not. From this point on, we drop the subindex t from q_t.

Let $\mathcal{V}(u) = E_{\theta_1}[V_1(\theta_1)]$ denote the first-period expected value from (4) under the system of payoff functions $u = (u_t)_t$, and given any default rule q, write $u + \log q$ to represent the system of payoff functions

$$u_t\left(a^t,\theta^t\right) + \log q\left(a_t \mid a^{t-1}\right)$$

for all t. For any choice rule p, let $p\left(a_t \mid a^{t-1}\right)$ denote the probability of choosing action a_t conditional on reaching decision node a^{t-1}, that is,

$$p\left(a_t \mid a^{t-1}\right) = E_{\theta^t}\left[p\left(a_t \mid \theta^t, a^{t-1}\right) \mid a^{t-1}\right].$$

[9] Our definition is more restrictive than that of Rust (1987) in that we do not allow the agent's actions to affect the distribution of future states. Our model also differs in the form of discounting and in the state space.

We adopt the convention that $\log 0 = -\infty$ and $e^{-\infty} = 0$.

Theorem 1. *Let q be a default rule that solves*

$$\max_{\tilde{q}} \mathcal{V}(u + \log \tilde{q}).$$

Then the dynamic logit rule p under payoffs $u + \log q$ solves the dynamic RI problem. Moreover,

$$q\left(a_t \mid a^{t-1}\right) = p\left(a_t \mid a^{t-1}\right) \tag{5}$$

for every decision node a^{t-1} that is reached with positive probability according to p.

Given a default rule q, we refer to $q(a_t \mid a^{t-1})$ as the *predisposition* toward action a_t at the decision node a^{t-1}. According to the theorem, the optimal predispositions are identical to the average behavior at each decision node.

The $\log q$ term in the payoffs has a natural interpretation: the agent behaves *as if* she incurs a cost

$$c_t\left(a^{t-1}, a_t\right) \equiv -\log q\left(a_t \mid a^{t-1}\right) \tag{6}$$

whenever she chooses a_t after the action history a^{t-1}. This endogenous virtual cost is high when the action a_t is rarely chosen at a^{t-1}. The cost captures the cost of information that leads to the choice of action a_t; actions that are unappealing ex ante can only become appealing through costly updating of beliefs.

Theorem 1 may be relevant for identification of preferences in dynamic logit models. Suppose that, as in Rust (1987), an econometrician observes the states θ_t together with the choices a_t, and estimates the agent's utilities using the dynamic logit rule from Definition 2. If our model correctly describes the agent's behavior, then instead of estimating the utility u_t, the econometrician will in fact be estimating $u_t\left(a^t, \theta^t\right) - c_t\left(a^{t-1}, a_t\right)$—the utility less the virtual cost.

For a fixed decision problem, separately identifying u_t and c_t is not necessary to describe behavior: choice probabilities depend only on the difference $u_t - c_t$. However, the distinction is important when extrapolating to other decision problems. For example, Rust (1987) considers a bus company's demand for replacement engines. He estimates the replacement cost by fitting a dynamic logit model in which the agent trades off the replacement cost against the expected loss from engine failure. He then obtains the expected demand by extrapolating to different engine prices, keeping other components of the replacement cost fixed.

Our model suggests that, if costly information acquisition plays an important role, Rust's approach underestimates demand elasticity. Consider an increase in the engine price. *Ceteris paribus*, replacement becomes less common, leading to a decrease in the predisposition toward replacement (by (5)). This corresponds to an increase in the virtual cost c_t associated with replacement (by (6)), and hence to an additional decrease in demand relative to the model in which c_t is fixed. Intuitively, the price increase not only discourages the purchase of a new engine, it also discourages the agent from checking whether a new engine is needed.

Distinguishing the actual utility u_t from the virtual cost c_t is feasible using data on choices and states. As described above, one can estimate $u_t - c_t$ by fitting the dynamic logit rule from Definition 2. The virtual cost $c_t(a^{t-1}, a_t) = -\log p(a_t \mid a^{t-1})$ can be identified directly based on the frequency with which each action is chosen.

3.2 Reduction to static problems

In this section, we describe how the dynamic RI problem can be reduced to a collection of static RI problems. This reduction allows us to draw on well developed solution methods from the static RI literature. In particular, we obtain a system of equations describing necessary and sufficient conditions for the solution of the dynamic RI problem.

As noted in the introduction, the characterization in terms of static RI problems does not follow from the Bellman equation alone. Gross expected utilities in static RI problems are linear in beliefs, but the continuation value function is not. For the resulting problems to fit the static RI framework, we show that one can ignore the dependence of continuation values on beliefs and treat them simply as functions of histories. Doing so restores the linearity of the expected gross payoffs and ensures that the problem has the usual RI structure. We explain this step in detail in Section 3.3.

We begin with a brief description of existing results for the static case. Consider a fixed, finite action set A, a finite state space Θ, a prior $\pi \in \Delta(\Theta)$, and a payoff function $u(a, \theta)$. A static choice rule p is a collection of action distributions $p(a \mid \theta)$, one for each $\theta \in \Theta$. We abuse notation by writing $p(\theta \mid a)$ for the posterior belief after choosing action a given the choice rule p. Recall that $I(\theta; a)$ is the mutual information between θ and a.[10]

[10]The literature on static rational inattention is richer than Definition 3 suggests. We restrict to the definition provided here because it is sufficient for our characterization.

Definition 3. *The* static rational inattention problem *for a triple* (Θ, π, u) *is*

$$\max_p E\left[u(a,\theta) - I(\theta; a)\right].$$

Proposition 2 (Matějka and McKay, 2015; Caplin and Dean, 2013)**.** *The static RI problem with parameters* (Θ, π, u) *is solved by the choice rule*

$$p(a \mid \theta) = \frac{q(a)e^{u(a,\theta)}}{\sum_{a'} q(a')e^{u(a',\theta)}}, \tag{7}$$

where the default rule $q \in \Delta(A)$ *maximizes*

$$E_\pi\left[\log\left(\sum_a q(a)e^{u(a,\theta)}\right)\right]. \tag{8}$$

If action a *is chosen with positive probability under the rule* p*, then the posterior belief after choosing* a *is*

$$p(\theta \mid a) = \frac{\pi(\theta)e^{u(a,\theta)}}{\sum_{a'} q(a')e^{u(a',\theta)}}. \tag{9}$$

We show that the dynamic RI problem can be reduced to a collection of static RI problems, one for each decision node a^{t-1}. These static problems are interconnected in that the payoffs and prior in one generally depend on the solutions to the others. At each a^{t-1}, the gross payoff consists of the flow payoff plus a continuation value, and the prior belief is obtained by Bayesian updating given a^{t-1}.

We again abuse notation by writing $p(\theta^t \mid a^{t-1})$ for the agent's prior over θ^t at the decision node a^{t-1} and $p(\theta^t \mid a^t)$ for the posterior over θ^t after action history a^t.

We call a dynamic choice rule that reaches each action history with positive probability *interior*. For simplicity, we state here the result only for interior dynamic choice rules. We extend the result to the general case in Appendix B.

Proposition 3. *An interior dynamic choice rule* p *solves the dynamic RI problem if and only if, at each decision node* a^{t-1}*,* $p(a_t \mid \theta^t, a^{t-1})$ *solves the static RI problem with state space* Θ^t*, prior belief*

$$p\left(\theta^t \mid a^{t-1}\right) = \sum_{\theta_t} p\left(\theta^{t-1} \mid a^{t-1}\right)\pi\left(\theta_t \mid \theta^{t-1}\right), \tag{10}$$

and payoff function

$$\hat{u}_t\left(a^t, \theta^t\right) = u_t\left(a^t, \theta^t\right) + \delta_{t+1} E_{\theta_{t+1}}\left[V_{t+1}\left(a^t, \theta^{t+1}\right) \mid \theta^t\right], \tag{11}$$

where the posterior belief $p(\theta^t \mid a^t)$ *formed after taking action* a_t *at the decision node* a^{t-1} *complies with* (9)*, and the continuation values satisfy*

$$V_t\left(a^{t-1}, \theta^t\right) = \log\left(\sum_{a_t} p\left(a_t \mid a^{t-1}\right) e^{\hat{u}_t\left(a^t, \theta^t\right)}\right). \tag{12}$$

Perhaps surprisingly, this result indicates that when optimizing behavior at a particular node, we can treat the continuation value as fixed as a function of the history. To understand the role of the continuation values, we note that the solution can be interpreted as an equilibrium of a common interest game played by multiple players, one for each period. The player in each period observes the history a^{t-1} but not the distribution of actions chosen in the past. In equilibrium, each player forms beliefs according to the others' equilibrium strategies, corresponding to the updating rule described in the proposition. Since deviations in the distribution of actions are unobservable to future players, each treats the strategies of the others (and hence the continuation values) as fixed. Even though the agent in our model can recall her own past strategy, the proposition indicates that she can ignore the effect of deviations on future beliefs.

When combined with a result from Caplin and Dean (2013), Proposition 3 provides necessary and sufficient conditions for solutions to dynamic RI problems. Theorem 1 in Caplin and Dean (2013) describes necessary and sufficient first-order conditions characterizing the solutions of static RI problems. Therefore, satisfying Caplin and Dean's conditions in each of the static problems in Proposition 3 is necessary and sufficient for a choice rule to be a solution to the dynamic RI problem.

In finite horizon and in stationary problems, the proposition leads to a finite system of equations characterizing the solution to the dynamic RI problem. Section 4 illustrates this approach in several applications. The solution of the status quo bias example in Section 4.1 and Appendix C.1 is a particularly simple application of Proposition 3.

A complication arises for the characterization in Proposition 3 when the solution of the dynamic RI problem is not interior. If the choice rule assigns zero probability to some action at a decision node, then it is not immediately clear how to define the posterior belief following that action. This posterior is needed to pin down the optimal continuation play and value associated with taking the action, which in turn is needed to determine whether taking the action with zero probability is indeed optimal. Formula (24) in Appendix B extends the posteriors defined by (9) to histories reached with zero probability. We show in the appendix how the extended definition can be obtained by solving the problem in which the probability of each action is constrained to be at least some $\varepsilon > 0$, then taking the limit as $\varepsilon \to 0$.

3.3 The control problem

We now describe the key step of the proof that leads to the dynamic logit characterization and allows us to reduce the dynamic problem to a collection of static ones. The main idea is to establish an equivalence between the dynamic RI problem and a control problem with observable states in which the agent must pay a cost for deviating from a default choice rule.[11]

Reformulating the dynamic RI problem as a control problem addresses the difficulty described above involving the link between the current action distribution and the future beliefs. The control problem clarifies why this link can be disregarded and hence the continuation values associated with each history can be treated as fixed when optimizing the action distribution at each decision node.

Definition 4. *Given any default rule q, the* control problem for q *is*

$$\max_{p} E\left[\sum_{t=1}^{\infty} \delta^{(t)} \Big(u_t\left(a^t, \theta^t\right) + \log q\left(a_t \mid a^{t-1}\right) - \log p\left(a_t \mid \theta^t, a^{t-1}\right) \Big)\right], \tag{13}$$

where p is a stochastic choice rule, and the expectation is with respect to the joint distribution generated by π and p.

This definition is a dynamic extension of a static control problem studied by Mattsson and Weibull (2002). In the control problem, the agent has complete information about the history θ^t, but must trade off optimizing her flow utility u_t against a control cost: for each (θ^t, a^{t-1}), she pays a cost

$$\log p\left(a_t \mid \theta^t, a^{t-1}\right) - \log q\left(a_t \mid a^{t-1}\right)$$

[11] Control problems of this kind have been studied in game theory building on the trembling-hand perfection of Selten (1975). Van Damme (1983) offers an early version in which agents optimize the distributions of trembles. Stahl (1990) introduces entropy-based control costs.

for deviating from the default action distribution $q\left(a_t \mid a^{t-1}\right)$ to the action distribution $p\left(a_t \mid \theta^t, a^{t-1}\right)$.

The next result shows that the dynamic RI problem is equivalent to the control problem with the optimal default rule. In other words, the dynamic RI problem can be solved by first solving the control problem to find the optimal choice rule p for each default rule q, and then optimizing q.

Lemma 2. *A stochastic choice rule solves the dynamic RI problem if and only if it (together with some default rule) solves*

$$\max_{q,p} E\left[\sum_{t=1}^{\infty} \delta^{(t)}\Big(u_t\left(a^t, \theta^t\right) + \log q\left(a_t \mid a^{t-1}\right) - \log p\left(a_t \mid \theta^t, a^{t-1}\right)\Big)\right], \tag{14}$$

where the expectation is with respect to the joint distribution generated by π *and* p.

To see how Lemma 2 addresses the difficulty due to the link between the current action distribution and subsequent beliefs, note that for any fixed default rule q, optimizing the choice rule p in the control problem does not involve updating of beliefs since the agent observes θ^t in period t. Since q cannot depend on the history of states, the optimal q at each decision node a^{t-1} does depend on the posterior distribution $p(\theta^t \mid a^{t-1})$; however, for any fixed p, optimizing q does not require varying these posterior distributions because they are determined by p, not by q.

The proof of the lemma relies on two well-known properties of entropy:

Symmetry For any random variables X, Y, and Z, $I(X;Y \mid Z) = I(Y;X \mid Z)$.

Properness For any random variable X with finite support S and distribution $p(x) \in \Delta(S)$,

$$H(X) = -\max_{q\in\Delta(S)} E_p[\log q(x)]. \tag{15}$$

To interpret the latter property, consider an agent who believes that X is distributed according to p and is asked to report a distribution $q \in \Delta(S)$ before observing the realization of X, with the promise of a reward of $\log q(x)$ if the realized value is x. Properness states that the truthful report $q = p$ maximizes the expected reward.

Proof of Lemma 2. By the symmetry of mutual information, the dynamic RI problem is equivalent to

$$\max_p E\left[\sum_t \delta^{(t)}\Big(u_t\big(a^t,\theta^t\big) - I\big(a_t;\theta^t \mid a^{t-1}\big)\Big)\right]$$
$$= \max_p E\left[\sum_t \delta^{(t)}\Big(u_t\big(a^t,\theta^t\big) - H\big(a_t \mid a^{t-1}\big) + H\big(a_t \mid \theta^t, a^{t-1}\big)\Big)\right]. \quad (16)$$

By (15),

$$E\Big[-\sum_t \delta^{(t)} H\left(a_t \mid a^{t-1}\right)\Big] = \max_q E\Big[\sum_t \delta^{(t)} \log q\left(a_t \mid a^{t-1}\right)\Big].$$

Substituting this into (16) and recalling that

$$E\Big[H\big(a_t \mid \theta^t, a^{t-1}\big)\Big] = E\Big[-\log p\big(a_t \mid \theta^t, a^{t-1}\big)\Big]$$

gives the result. □

The dynamic logit result in Theorem 1 follows from solving problem (14). As the following lemma indicates, dynamic logit choice behavior (with biased payoffs) arises as the solution to the control problem for any fixed q. This lemma extends a result of Mattsson and Weibull to the dynamic case: they show that, in the static version of the control problem, the optimal action distribution is a logit rule with a bias toward actions that are relatively likely according to the default rule.

Lemma 3. *Given any default rule q, the dynamic logit rule under payoffs $u + \log q$ solves problem* (13).

4 Applications

In this section, we use Proposition 3 to analyze the examples described in Section 2.1.

4.1 Status quo bias

Recall that in Example 1, the agent chooses an action $a_t \in \{0,1\}$ at $t = 1,2$. In both periods, the gross flow payoff u_t is 1 if $a_t = \theta_t$, and is 0 otherwise. There is no discounting. The states are symmetrically distributed and positively correlated across time in the following

	Prob. of retaining decision across periods $\Pr(a_1 = a_2)$	Prob. of correct choice in period 1 $\Pr(a_1 = \theta_1)$	Prob. of correct choice in period 2 $\Pr(a_2 = \theta_2)$
Correlated states: $\Pr(\theta_1 = \theta_2) = 0.9$	1	0.86	0.79
Uncorrelated states: $\Pr(\theta_1 = \theta_2) = 0.5$	0.5	0.73	0.73

Table 1: Inertia and accuracy of choice in Example 1.

way: θ_1 is equally likely to be 0 or 1, and, whatever the realized value of θ_1, the probability that $\theta_2 = \theta_1$ is 0.9.

We show in Appendix C.1 that the optimal choice rule exhibits an apparent status quo bias: the agent never reverses her decision from one period to the next. The optimal strategy in this case acquires information only in the first period and then relies on that information for the action choices in both periods. Consequently, the agent performs better in the first period than in the second; see the first row of Table 1.

Which features of the model drive the status-quo-bias behavior? The superior performance in the first period arises because of the endogenous timing of information acquisition. In a variant of the model with exogenous conditionally i.i.d. signals, the agent would perform better in the second period than in the first since she obtains more precise information about θ_2 than about θ_1. When information is endogenous, the correlation between the two periods creates an incentive to acquire more information in the first period because that information can be used twice.

However, correlation does not generate the status quo bias on its own—the temporal structure also plays an important role in the sense that the effect would not arise if the agent could acquire information about both states in the first period. To see this, consider a static variant in which the agent simultaneously chooses a pair of actions (a_1, a_2) to maximize

$$E\left[u_1(a_1, \theta_1) + u_2(a_2, \theta_2) - I\big((\theta_1, \theta_2); (a_1, a_2)\big)\right].$$

In this case, as in the original example, the optimal strategy involves a single binary signal and identical actions in the two periods. In the static variant, however, the expected performance is constant across the two periods. The asymmetric performance in the original example arises because it is impossible for the agent to learn directly about the second period in the first, when information is most valuable.

Finally, to illustrate the role of correlation in the state across periods, consider a benchmark in which θ_1 and θ_2 are independent and uniform on $\{0, 1\}$. In that case, any information obtained in the first period is useless in the second. The problem therefore reduces to a pair of unconnected static RI problems (one for each period). The solution involves switching actions with probability 1/2 and constant performance across the two periods; see the second row of Table 1. In general, the predisposition towards switching the action in the second period decreases with the probability that the state remains the same, up to a critical level of correlation beyond which the agent never switches.[12]

Although the solution when the states are correlated may appear as if the agent has a preference against switching her action, the independent case highlights the difference between such a preference and the effect of information acquisition; if the "status quo bias" behavior were driven by preferences, it would not depend on correlation between the states.

4.2 Inertia

Recall that, in Example 2, the agent chooses an action $a_t \in \{0, 1\}$ in each period $t = 1, 2, \ldots$ with the goal of matching the current state. The state θ_t follows a Markov chain on the set $\{0, 1\}$ with time-homogeneous transition probabilities $\gamma(\theta, \theta')$ from state θ to state θ'. In each period $t \in \mathbb{N}$, the gross flow payoff $u(a_t, \theta_t)$ is equal to $u_a > 0$ if $a_t = \theta_t = a$, and is 0 if $a_t \neq \theta_t$. Payoffs are discounted exponentially with discount factor $\delta \in (0, 1)$. The solution in this case illustrates the method in a stationary environment in which the agent acquires information in every period.

This example can be viewed as a stylized model of a wide range of economic phenomena. For example, the action could represent an investor's choice of whether to hold a particular asset, a firm's choice of how many employees to have on staff, or a worker's choice of whether to participate in the labor market. An important question in each of these settings is how the timing of adjustment to shocks is shaped by the environment. Do the lengths of adjustment lags differ between booms and busts? How does volatility influence behavioral inertia?

Comparative statics of adjustment patterns with respect to the stochastic properties of the agent's environment are a central question in the RI literature. Existing studies, such as Sims (2003), Moscarini (2004), Luo (2008), and Mackowiak and Wiederholt (2009), provide results for quadratic payoffs and normally distributed shocks. Our framework provides an

[12]Details behind this computation are available upon request.

alternative approach suitable for general payoffs and distributions in discrete environments.

We start by noting that the long-run behavior is Markovian. After a finite number of periods, the choice rule, continuation values, and predispositions in any period t depend on the last action a_{t-1}, but not on any earlier actions. This implies that the long-run behavior is characterized by a finite set of equations; see Lemma 4 in Appendix C.2 for details. This Markov property of the solution holds for arbitrary finite sets of actions and states, general time-homogeneous Markov processes, and general utilities as long as all actions are chosen with positive probability at all decision nodes.[13] This feature highlights the relative simplicity of the rationally inattentive solution compared to that of similar decision problems with exogenous conditionally i.i.d. signals. In the exogenous case, the optimal strategy is not Markov: actions depend on the entire history of signals, the probabilities of which in turn depend on the entire history of states. Characterizing the distribution of actions is therefore complicated even in the simplest cases.

We focus here on the limit in which states become increasingly persistent, which allows for a simple analytical solution. Intuitively, if the state θ_t rarely changes, then the ex ante probability of an action switch between two consecutive periods is low, and hence the agent's predisposition goes against switching. By Theorem 1, she follows the dynamic logit choice rule under her true payoff function plus virtual switching costs, and these lead to delayed reactions to payoff shocks.

The agent's attention strategy is dynamically sophisticated. She largely relies in each period on her information from the previous period since it is likely that the state has not changed. However, she also acquires a small amount of information in each period to avoid prolonged stretches of suboptimal behavior. When deciding how much information to acquire, she takes into account her immediate incentives and the future value of any information she acquires.

To study the persistent-state case, let $\gamma(\theta,\theta') = \overline{\gamma}(\theta,\theta')\varepsilon$ for $\theta \neq \theta'$, and consider the limit as ε vanishes. It turns out that, for small ε, the predisposition to switch actions is of order ε. Accordingly, for $a' \neq a$ we normalize the predispositions by ε and define the limit predispositions to be

$$q^*(a,a') := \lim_{\varepsilon\to 0} \frac{q(a_t = a' \mid a_{t-1} = a)}{\varepsilon},$$

[13]The structure of our solution resembles that of the bounded memory model of Wilson (2014). Each action in our model can be viewed as a "memory state," with the agent's strategy describing stochastic transitions among them. As in Wilson's model, beliefs in each memory state depend on the agent's entire strategy.

and the limit adjustment rates to be

$$\alpha(a,a') := \lim_{\varepsilon\to 0} \frac{p(a_t = a' \mid a_{t-1} = a, \theta_t = a')}{\varepsilon}.$$

Thus $\varepsilon q^*(a,a')$ is (approximately) the probability that the agent switches to action a' after choosing $a \neq a'$ in the previous period, and $\varepsilon\alpha(a,a')$ is the probability that the agent switches from a to a' if this switch adjusts the action to match the current state.

Let $U_a = \exp \frac{u_a}{1-\delta}$, and assume that

$$\overline{\gamma}(a,a')\frac{U_a}{U_a - 1} - \frac{1}{U_{a'} - 1}\overline{\gamma}(a',a) \tag{17}$$

is positive whenever $a \neq a'$. If this is not the case, then one of the actions is absorbing: there exists a period after which the agent almost surely chooses the same action in all subsequent periods.

Proposition 4. *Let $a' \neq a$. The adjustment rate $\alpha(a,a')$ from a to a' increases with the incentive $u_{a'}$ to match state a', decreases with the incentive u_a to match state a, increases with the transition rate $\overline{\gamma}(a,a')$, and decreases with $\overline{\gamma}(a',a)$. The limit predispositions $q^*(a,a')$ are given by* (17)*, and the limit adjustment rates are*

$$\alpha(a,a') = q^*(a,a')U_{a'}. \tag{18}$$

Optimal adjustment to shocks involves significant delays in a persistent environment. For example, consider a symmetric setting with $\gamma(0,1) = \gamma(1,0) = \varepsilon$, and $u_0 = u_1 = u$. The expected time between adjacent switches of the state is $1/\varepsilon$. According to the proposition, if the agent's action is not aligned with the state, she switches her action with probability $e^{u/(1-\delta)}\varepsilon$ per period. This probability corresponds to an expected lag time of $e^{-u/(1-\delta)}/\varepsilon$, which is of the same order as the time between switches of the state. In particular, greater persistence in the state corresponds to longer adjustment lags. In the symmetric case, this monotonicity result also holds outside of the limit: increasing volatility reduces adjustment lags.[14] Intuitively, past actions are not reliable predictors of the current optimal action in volatile environments, which reduces the optimal predisposition towards repeating the last action.

Asymmetries in incentives have an intuitive impact on adjustment rates. Consider again

[14] Numerical results suggest that the same result holds in asymmetric settings.

a Markov chain with $\gamma(0,1) = \gamma(1,0) = \varepsilon$, and suppose now that $u_0 > u_1$. Interpreting state 1 as the good state in an investment problem, this corresponds to the loss from investing during a crisis exceeding the gain from investing during a boom. Then $\alpha(0,1) < \alpha(1,0)$, meaning that the agent reacts more quickly to negative shocks than to positive ones.

4.3 Response times

In this section, we study a simple model of response times in decision-making. The study of response times has a long tradition in psychology, and has more recently become the subject of a growing literature in economics based on the idea that the timing of choice may reveal useful information beyond what is revealed by the choice itself (e.g., see Rubinstein, 2007). We discuss below how an outside observer may exploit decision times to better understand the decision maker's choices.

An important methodological question in this area is whether choice procedures should be modeled explicitly or in reduced form. Sims (2003) argues that the RI framework is a promising tool for incorporating response times into traditional economic models that treat decision-making as a black box. Our model, with its focus on sequential choice, is a step in this direction. Woodford (2014) studies delayed decisions in an RI model that focuses on neurological decision procedures.[15]

Recall that in Example 3, the state $\theta \in \{0,1\}$ is uniformly distributed and fixed over time. In each period $t = 1, \dots, T$, the agent chooses among taking a terminal action 0 or 1, or waiting until the next period (denoted by w). She incurs a cost $c \in (0,1)$ for each period that she waits. The agent's total gross payoff is the undiscounted sum of the flow payoffs

$$u_t(a^t, \theta) = \begin{cases} 1 & \text{if } a^t = (w, \dots, w, \theta), \\ 0 & \text{if } a^t = (w, \dots, w, 1-\theta), \\ -c & \text{if } a^t = (w, \dots, w), \\ 0 & \text{otherwise.} \end{cases}$$

This formulation is similar to the model of Arrow, Blackwell, and Girshick (1949) except that information is endogenous; see also Fudenberg et al. (2015).

With the information cost function in our general model, the solution to this problem is

[15] See Spiliopoulos and Ortmann (2014) for a review of psychological and economic research on decision times, and of the methodological differences across the two fields.

trivial: since delay is costly, any strategy that involves delayed decisions is dominated by a strategy that generates the same distribution of terminal actions in the first period. Hence there is no delay if the marginal cost of information is constant across time. However, delay can be optimal in a closely related variation of the model in which—as in much of the RI literature—there is an upper bound on how much information the agent can process in each period. Accordingly, the agent solves

$$\begin{gathered}\max_p E\left[\sum_{t=1}^{T} u_t\left(a^t,\theta\right)\right] \\ \text{s.t. } E\left[I\left(\theta;a^t \mid a^{t-1}\right)\right] \leq \kappa \text{ for all } t=1,\ldots,T,\end{gathered} \tag{19}$$

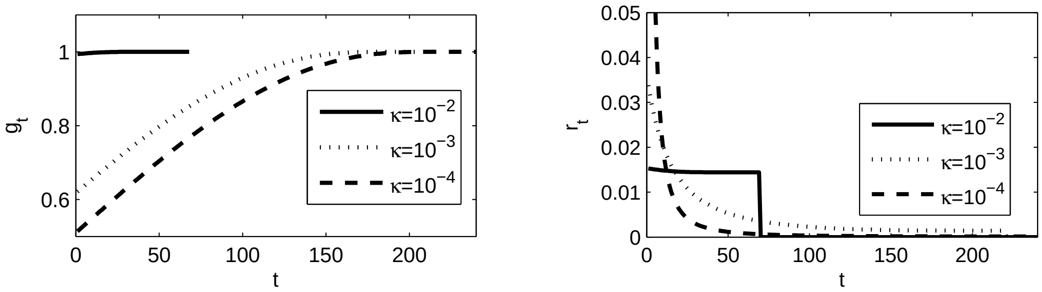
where $\kappa > 0$ is the capacity constraint on the information acquired per period, and $p(a_t \mid \theta, a^{t-1})$ is the choice rule.[16] Note that the expectation in the constraint is taken ex ante, capturing the idea that taking earlier decisions in some problems can free up capacity to be used in other problems that the agent may face. The ex ante form of the constraint is in the spirit of the standard constraint used in static RI models based on the expected reduction in entropy over all posteriors. This formulation ensures that the dual problem lies within the class of dynamic RI problems that we study.

The first-order conditions for this problem are closely related to those of our general model with information costs. The solution of (19) also solves a problem in which the capacity constraint is replaced by time-varying marginal costs of information λ_t, where λ_t is the shadow price of the capacity constraint for period t.

Delay costs introduce a trade-off between the speed and accuracy of decision-making: increasing the likelihood of a terminal decision early on decreases delay costs but also uses up the information capacity in early periods, thereby decreasing the accuracy of the early decisions.

Because of the symmetry of the model, the optimal posterior belief after waiting for any number of periods is always uniform; the agent learns nothing until the period in which she takes a final decision. The purpose of delay is not to accumulate increasingly precise information over time, but rather to reduce the information capacity required to achieve the desired accuracy in the terminal decisions.

[16]This formulation abstracts from explicit signal acquisition by imposing the constraint directly on the joint distribution of actions and states. As in the general model, we could allow the agent to choose signal distributions together with mappings from signals to actions. One can show that the result of Lemma 1 applies in this context, and therefore the present formulation is without loss of generality.



(A) (B)

Figure 1: (A) accuracy g_t as a function of time, and (B) the probability r_t that the decision is taken at time t for $\kappa \in \{10^{-4}, 10^{-3}, 10^{-2}\}$ when $T = 1000$ and the delay cost c is 10^{-3}.

The accuracy of the decision-maker's choice is correlated with the decision time. An outside observer interested in whether the decision maker made the correct choice could exploit this correlation. Accordingly, let g_t be the probability that the correct decision $a_t = \theta$ is made conditional on terminating at t. How is the timing related to accuracy? For a fixed capacity κ, the accuracy of decisions increases with time. Because the capacity is uniform over time, the delay cost makes it relatively more valuable early on. The shadow price of capacity is therefore decreasing over time, which in turn leads to increasing accuracy of decisions over time.

The optimal behavior exhibits an intuitive relationship between the distribution of decision times and the capacity κ. This relationship can, in principle, be used by an outside observer who would like to identify the agent's capacity in order to predict her behavior in other problems. This relationship can also be viewed as describing how the decision times vary with the difficulty of the problem. Let r_t denote the probability that the agent makes the terminal decision at round t. The speed of decision-making is not monotone in the capacity: Figure 1 shows that decisions are fastest when the capacity is high or low, and slowest for intermediate capacities.[17] If the capacity is low, there is little incentive to delay the decision since the cost of delay is large relative to the value of the additional information that can be acquired. If the capacity is high, the agent can acquire precise information quickly and then has little incentive to delay in order to acquire additional information.[18]

[17]The computations for Figure 1 in Appendix C.3 are analytical except for numerical solution of one unknown.

[18]This pattern may be familiar to course instructors who observe that the earliest exams to be turned in often combine both tails of the distribution of abilities.

If individual subjects can be treated as having a fixed capacity across problems in an experiment, this suggests that we should expect significant differences in the correlation between accuracy and decision times depending on whether the data is within or across subjects.

5 Summary

We solve a general dynamic decision problem in which an agent repeatedly acquires information, facing entropy-based information costs. The optimal behavior is stochastic—the action distribution at each decision node complies with a logit choice rule—and biased—compared to the standard dynamic logit model, the agent behaves as if she incurs a cost for choosing actions that are unlikely ex ante. When incentives are serially correlated, the agent exhibits an endogenous conservative bias that results in stickiness in her actions. The distinction between real and informational frictions is a central topic of the RI literature that has been studied in particular settings. This paper formalizes, in a general setting, an equivalence between the two frictions within any given decision problem, while showing that they lead to distinct predictions when extrapolating to different problems.

As a tool for solving the problem, we show that the RI model with incomplete information and learning is behaviorally equivalent to a complete information control problem. The agent behaves as if she faces a cost of deviating from a default choice rule, but also engages in a second layer of optimization: at the ex ante stage, she optimizes the default rule, which is independent of the state of the world, and ex post, the agent chooses an optimal deviation from the default rule given the incentives in the realized state and the control cost.

Appendix

A Proofs for Section 2.2

Proof of Proposition 1. Consider the space of strategies $\Pi = \prod_t P_t^{A^{t-1}}$, where P_t denotes the set of feasible joint distributions of a_t and θ^t. By Tychonoff's Theorem, the space Π is compact in the product topology, and because u_t is uniformly bounded, the objective function is continuous. Therefore, an optimum exists. $\square$

Proof of Lemma 1. Let Ψ denote the set of all strategies, and for $s \in \Psi$, let $U(s)$ denote the ex ante expected payoff from strategy s, that is, for $s = (f, \sigma)$, $U(s)$ is the objective function in (2).

Given any strategy $s = (f, \sigma)$ and any x^{t-1}, let $X(s, x^{t-1}) = \{x \in X : f_t(x \mid \theta^t, x^{t-1}) > 0 \text{ for some } \theta^t\}$. Let $\tilde{\sigma}_t(\cdot; x^{t-1}, s) : X(s, x^{t-1}) \longrightarrow A$ be such that $\tilde{\sigma}_t(x_t; x^{t-1}, s) \equiv \sigma_t((x^{t-1}, x_t))$.

The main idea of the proof is to take any information that is acquired but not used at time t and postpone its acquisition to time $t+1$. However, doing so may not be possible if all available signals are already being used at time $t+1$. Accordingly, for the purpose of the construction, we expand the signal spaces, and then note that, following an infinite recursion, the strategy we construct is feasible with the original signal spaces.

Let $\overline{X}_t = X^t$ and $\overline{X}^t = \prod_{t' \leq t} \overline{X}_{t'}$, and let $\overline{\Psi}$ denote the set of all strategies when the space of available signals in period t is $\overline{X}_t$. Given any strategy $s = (f, \sigma) \in \Psi$, we will construct a sequence $(s^\tau)_{\tau=0}^\infty = (f^\tau, \sigma^\tau)_{\tau=0}^\infty$ with $s^\tau \in \overline{\Psi}$ such that $U(s^0) = U(s)$ and, for every τ,

1. $\tilde{\sigma}_t^\tau(\cdot; x^{t-1}, s^\tau)$ is one-to-one for every $t \leq \tau$ and every x^{t-1},
2. $(f_t^\tau(\cdot|x^{t-1}, \theta^t), \sigma_t^\tau) = (f_t^{\tau-1}(\cdot|x^{t-1}, \theta^t), \sigma_t^{\tau-1})$ whenever $t \leq \tau - 1$, and
3. $U(s^\tau) \geq U(s^{\tau-1})$.

Endowing $\overline{\Psi}$ with the product topology, the sequence s^τ converges to a strategy s^* that is identical up to relabeling of signals to the strategy induced by the choice rule generated by s. Moreover, since s^* and s^τ are identical in periods 1 through τ, flow payoffs are uniformly bounded, and $\sum_t \delta^{(t)} < \infty$, we have $U(s^*) = \lim_\tau U(s^\tau)$. In particular, if s is optimal then so is s^*, proving the lemma.

The sequence $(s^\tau)_{\tau=0}^\infty$ is constructed recursively as follows. First we define s^0 by "embedding" s into the expanded signal space $\overline{X}_t$. Formally, for $x_t \in X$ and $\overline{x}^{t-1} = (\overline{x}_1, \ldots, \overline{x}_{t-1}) \in \overline{X}^{t-1}$, let

$$f_t^0((x_1, \ldots, x_t)|\overline{x}^{t-1}, \theta_t) = \begin{cases} f_t(x_t|\overline{x}_{t-1}, \theta_t) & \text{if } \overline{x}_{t-1} = (x_1, \ldots, x_{t-1}), \\ 0 & \text{otherwise,} \end{cases}$$

and $\sigma_t^0(\overline{x}^t) \equiv \sigma_t(\overline{x}_t)$. By construction, s and s^0 generate the same joint distribution of histories of actions and states and the same information costs in each period; hence we have $U(s^0) = U(s)$.

For $\tau > 0$, the idea is to construct s^τ by coarsening $s^{\tau-1}$ in period τ so that signals that occur with positive probability map one-to-one to actions and then restore the lost information in period $\tau+1$. Accordingly, if $\tilde{\sigma}_\tau^{\tau-1}(\cdot;\overline{x}^{\tau-1})$ is one-to-one for every $\overline{x}^{\tau-1}$ then let $s^\tau = s^{\tau-1}$. Otherwise, for each t, associate to each action $a \in A$ a signal $\overline{x}_t^a \in \overline{X}_t$ (chosen arbitrarily) such that $\overline{x}_t^a \neq \overline{x}_t^{a'}$ whenever $a \neq a'$. Let

$$f_t^\tau(\overline{x}_t|\overline{x}^{t-1},\theta^t) = \begin{cases} f_t^{\tau-1}(\overline{x}_t|\overline{x}^{t-1},\theta^t) & \text{if } t \leq \tau-1, \\ \sum_{\overline{x}\in\overline{X}_t:\tilde{\sigma}_t^{\tau-1}(\overline{x};\overline{x}^{t-1})=a} f_t^{\tau-1}(\overline{x}|\overline{x}^{t-1},\theta^t) & \text{if } t=\tau \text{ and } \overline{x}_t = \overline{x}_t^a, \\ 0 & \text{if } t=\tau \text{ and } \overline{x}_t \neq \overline{x}_t^a \text{ for any } a \in A, \\ \Pr_{f_t^{\tau-1}}(\overline{x}_t \mid \overline{\mu}_{t-1}^{\tau-1}(\overline{x}^{t-1}),\theta^t) & \text{otherwise,} \end{cases}$$

where $\overline{\mu}_t^\tau(\overline{x}^t) := (\tilde{\sigma}_1^\tau(\overline{x}_1),\ldots,\tilde{\sigma}_t^\tau(\overline{x}_t;\overline{x}_{t-1}))$, and

$$\sigma_t^\tau(\overline{x}^t) = \begin{cases} a & \text{if } t=\tau \text{ and } \overline{x}_t = \overline{x}_t^a, \\ \sigma_t^{\tau-1}(\overline{x}^t) & \text{otherwise.} \end{cases}$$

It is clear by construction that the sequence $(s^\tau)_\tau$ satisfies properties 1 and 2 above. All that remains is to show that it satisfies property 3.

First note that for every $\tau \geq 1$, s^τ induces the same distribution over sequences of action-state pairs as $s^{\tau-1}$. Hence $U(s^\tau) \geq U(s^{\tau-1})$ if and only if the total discounted expected information cost from s^τ is no more than that from $s^{\tau-1}$. Letting $x^0 := \emptyset$, for any $t \neq \tau$, the mutual information $I\left(\theta^t; x^t \mid x^0\right)$ is identical under $s^{\tau-1}$ and s^τ, and for $t=\tau$ it is (weakly) lower under s^τ. Since $\delta^{(\tau)} \geq \delta^{(\tau+1)}$ and, from the definition of the mutual information in (1),

$$I\left(\theta^t; x^t \mid x^0\right) = \sum_{t'=1}^{t} I\left(\theta^t; x^{t'} \mid x^{t'-1}\right),$$

it follows that the information cost is at least as high under $s^{\tau-1}$ as under s^τ. $\square$

B Proofs for Section 3

Proof of Lemma 3. Let $\overline{u}_t(a^t,\theta^t) = u_t(a^t,\theta^t) + \log q(a_t \mid a^{t-1})$. For each a^{t-1} and θ^t such that $\pi(\theta^t) > 0$ and $p(a^{t-1}) > 0$, let

$$V_t(a^{t-1},\theta^t) = \frac{1}{\delta^{(t)}} \max_{\{p_\tau(\cdot|a^{\tau-1},\theta^\tau)\}_{\tau=t}^{\infty}} E\left[\sum_{\tau=t}^{\infty} \delta^{(\tau)} \left(\overline{u}_\tau(a^\tau,\theta^\tau) - \log p_\tau(a_\tau \mid a^{\tau-1},\theta^\tau)\right) \mid \theta^t, a^{t-1}\right];$$

thus $V_t(a^{t-1},\theta^t)$ is the continuation value in the control problem for q. Note that V_t does not depend on the agent's strategy in earlier periods.

The value V_t satisfies the recursion

$$V_t(a^{t-1},\theta^t) = \max_{p(\cdot|a^{t-1},\theta^t)} E_{\theta_{t+1}}\left[\overline{u}_t(a^t,\theta^t) - \log p(a_t \mid a^{t-1},\theta^t) + \delta_{t+1} V_{t+1}(a^t,\theta^{t+1}) \mid \theta^t\right] \tag{20}$$

(recall that $\delta_{t+1} = \delta^{(t+1)}/\delta^{(t)}$).

To solve the maximization problem on the right-hand side of (20), note first that, since $\overline{u}_t(a^t,\theta^t) = u_t(a^t,\theta^t) + \log q(a_t \mid a^{t-1})$, if $q(a_t \mid a^{t-1}) = 0$ (and hence $\log(q(a_t \mid a^{t-1}) = -\infty$) for some a_t, then we must have $p\left(a_t \mid a^{t-1},(\theta^{t-1},\theta_t)\right) = 0$ for every θ_t satisfying $\pi(\theta^{t-1},\theta_t) > 0$.[19] Accordingly let $A(a^{t-1}) = \{\tilde{a}_t \in A_t : q(\tilde{a}_t \mid a^{t-1}) > 0\}$, and suppose $a_t \in A(a^{t-1})$ and $\pi(\theta^{t-1},\theta_t) > 0$. If $A(a^{t-1})$ is a singleton, then $p\left(a_t \mid a^{t-1},(\theta^{t-1},\theta_t)\right) = 1$. Otherwise, the first-order condition for (20) with respect to $p(a_t \mid a^{t-1},\theta^t)$ is

$$\overline{u}_t(a^t,\theta^t) - \left(\log p(a_t \mid a^{t-1},\theta^t) + 1\right) + \delta_{t+1} E_{\theta_{t+1}}\left[V_{t+1}\left(a^t,\theta^{t+1}\right) \mid \theta^t\right] = \mu_t(a^{t-1},\theta^t), \tag{21}$$

where $\mu_t(a^{t-1},\theta^t)$ is the Lagrange multiplier associated with the constraint $\sum_{a'_t} p(a'_t \mid a^{t-1},\theta^t) = 1$.

Rearranging the first-order condition gives

$$p(a_t \mid a^{t-1},\theta^t) = \exp\left(\overline{u}_t(a^t,\theta^t) - 1 + \delta_{t+1}\overline{V}_{t+1}\left(a^t,\theta^t\right) - \mu_t(a^{t-1},\theta^t)\right),$$

where, for brevity, $\overline{V}_{t+1}\left(a^t,\theta^t\right) := E_{\theta_{t+1}}\left[V_{t+1}\left(a^t,\theta^{t+1}\right) \mid \theta^t\right]$. Since $\sum_{a'_t \in A(a^{t-1})} p(a'_t \mid$

[19] If $\pi(\theta^{t-1},\theta_t) = 0$ then $p\left(a_t \mid a^{t-1},(\theta^{t-1},\theta_t)\right)$ has no effect on the value and can be chosen arbitrarily.

$a^{t-1}, \theta^t) = 1$, it follows that

$$p(a_t \mid a^{t-1}, \theta^t) = \frac{\exp\left(\overline{u}_t(a^t, \theta^t) - 1 + \delta_{t+1}\overline{V}_{t+1}\left(a^t, \theta^t\right) - \mu_t(a^{t-1}, \theta^t)\right)}{\sum_{a'_t \in A(a^{t-1})} \exp\left(\overline{u}_t((a^{t-1}, a'_t), \theta^t) - 1 + \delta_{t+1}\overline{V}_{t+1}\left((a^{t-1}, a'_t), \theta^t\right) - \mu_t(a^{t-1}, \theta^t)\right)}$$
$$= \frac{\exp\left(\overline{u}_t(a^t, \theta^t) + \delta_{t+1}\overline{V}_{t+1}\left(a^t, \theta^t\right)\right)}{\sum_{a'_t \in A(a^{t-1})} \exp\left(\overline{u}_t((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}\overline{V}_{t+1}\left((a^{t-1}, a'_t), \theta^t\right)\right)}.$$

Substituting into (20) gives the recursion

$$\begin{aligned}
&\overline{V}_t(a^{t-1}, \theta^{t-1}) \\
&= E\Bigg[- \delta_{t+1}\overline{V}_{t+1}(a^t, \theta^t) + \log\left(\sum_{a'_t \in A(a^{t-1})} \exp\left(\overline{u}((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}\overline{V}_{t+1}((a^{t-1}, a'_t), \theta^t)\right)\right) \\
&\qquad\qquad + \delta_{t+1}\overline{V}_{t+1}(a^t, \theta^t) \,\Bigg|\, \theta^{t-1} \Bigg],
\end{aligned}$$

and therefore,

$$\begin{aligned}
\overline{V}_t(a^{t-1}, \theta^{t-1}) &= E\left[\log\left(\sum_{a'_t \in A(a^{t-1})} \exp\left(\overline{u}((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}\overline{V}_{t+1}((a^{t-1}, a'_t), \theta^t)\right)\right) \,\middle|\, \theta^{t-1} \right] \\
&= E\left[\log\left(\sum_{a'_t \in A_t} q(a'_t | a^{t-1}) \exp\left(u((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}\overline{V}_{t+1}((a^{t-1}, a'_t), \theta^t)\right)\right) \,\middle|\, \theta^{t-1} \right],
\end{aligned}$$

as needed. □

Proof of Theorem 1. The first assertion follows immediately from Lemmas 2 and 3. For the second assertion, fixing p, if a^{t-1} is reached with positive probability, properness implies that $q(a_t \mid a^{t-1}) = p(a_t \mid a^{t-1})$ maximizes the objective in problem (14). □

Proof of Proposition 3. Given a^{t-1} and continuation values $V_{t+1}(a^t, \theta^{t+1})$, we refer to the static problem described in the proposition as the *static RI problem at* a^{t-1}. Each of these static problems is a special case of our general model; in particular, Lemma 2 implies that $p(a_t \mid a^{t-1}, \theta_t)$ solves the static RI problem at a^{t-1} if and only if it, together with some

$q(\cdot \mid a^{t-1})$ solves the corresponding control problem

$$\max_{q(\cdot|a^{t-1}),\{p(\cdot|\theta^t,a^{t-1})\}_{\theta^t}} E\left[\hat{u}_t\left((a^{t-1},a_t),\theta^t\right)+\log q\left(a_t \mid a^{t-1}\right)-\log p\left(a_t \mid \theta^t,a^{t-1}\right)\right], \quad (22)$$

where the expectation is with respect to the joint distribution of a_t and θ^t generated by the prior $p(\theta^t \mid a^{t-1})$ together with $\{p(\cdot \mid \theta^t, a^{t-1})\}_{\theta^t}$. We call (22) the *control problem at* a^{t-1}.

By Lemma 2, it suffices to prove that any solution of the control problem (problem (14)) coincides at every a^{t-1} with a solution of the control problem at a^{t-1}. By Lemma 3, for any given q, the objective function in (14) is maximized by p satisfying

$$p(a_t \mid a^{t-1},\theta^t)=\frac{q(a_t \mid a^{t-1})\exp\left(\hat{u}(a^t,\theta^t)\right)}{\sum_{a'_t} q(a'_t \mid a^{t-1})\exp\left(\hat{u}((a^{t-1},a'_t),\theta^t)\right)}. \quad (23)$$

Similarly, for each a^{t-1}, given $q(\cdot \mid a^{t-1})$, this p maximizes the objective function in the control problem at a^{t-1}.

Let q be a solution to (14) (together with p given by (23)). Since p is interior, it follows from (23) that $p(a_t \mid a^{t-1},\theta^t) > 0$ for every a^{t-1} and θ^t. The result now follows from the Principle of Optimality: the control problem at a^{t-1} corresponds to the Bellman equation at that decision node, and hence $q(\cdot \mid a^{t-1})$ and $p(\cdot \mid a^{t-1},\theta^t)$ also solve the control problem at a^{t-1} (and conversely). □

We now extend Proposition 3 to cases in which the solution to the dynamic RI problem is not interior. We define the posterior belief in a static RI problem after an action a is taken with zero probability to be

$$p(\theta \mid a)=\frac{1}{\sum_{\theta'}\pi(\theta')\frac{e^{u(a,\theta')}}{\sum_{a'} q(a')e^{u(a',\theta')}}}\frac{\pi(\theta)e^{u(a,\theta)}}{\sum_{a'} q(a')e^{u(a',\theta)}}. \quad (24)$$

This expression coincides with (9) when a is chosen with positive probability. Otherwise, it differs from (9) only by a renormalization. We show below that the posteriors in (24) arise in a modified dynamic RI problem in which the probability of each action is constrained to be at least some $\varepsilon > 0$. We then prove that solutions of the constrained problems converge to a solution of the unconstrained problem as $\varepsilon \to 0$.

Proposition 5. *There exists a dynamic choice rule p solving the dynamic RI problem such that, at each decision node a^{t-1}, $p(a_t \mid \theta^t, a^{t-1})$ solves the static RI problem with state space*

Θ^t, *prior belief* $p(\theta^t \mid a^{t-1})$ *satisfying* (10), *and payoff function* $\hat{u}_t$ *given by* (11), *where the posterior belief* $p(\theta^t \mid a^t)$ *formed after taking action* a_t *at the decision node* a^{t-1} *complies with* (24), *and the continuation values satisfy* (12).

Proof. This proof extends the proof of Proposition 3 to include the case in which $q(a_t \mid a^{t-1})$ may be 0 for some a^{t-1} and a_t, where we must ensure that prior beliefs in the static problems are defined appropriately to generate the correct continuation values. The main idea is to add an additional constraint placing a lower bound on every $q(a_t \mid a^{t-1})$ in both the original control problem and the control problem at a^{t-1}, and then examine the limit as this lower bound vanishes. The same argument as in the proof of Proposition 3 applies to the problems with the lower bound, and continuity yields the desired result in the limit.

Now consider, for $\varepsilon \in (0, 1/|A|)$, the ε-*control problem*, which is identical to the original control problem (problem (14)) except that for each a_t and a^{t-1}, there is a constraint that $q(a_t \mid a^{t-1}) \geq \varepsilon$. Define the ε-*control problem at* a^{t-1} analogously. The argument in the proof of Proposition 3 applies here to show that, for each ε, the solutions to the ε-control problem coincide with those of the full collection of ε-control problems at decision nodes a^{t-1}. Moreover, essentially the same argument as in the proof of Proposition 1 establishes that a solution to each ε-control problem exists.

By Lemma 3, the solution to (and value of) the control problem for q is continuous in q (with respect to the product topology). Therefore, any limit point—as ε vanishes—of the set of solutions to the ε-control problem is a solution to the original control problem. An analogous argument applies to the ε-control problem at each a^{t-1} provided that the continuation values and priors approach those described in the proposition as ε vanishes.

For each a^{t-1}, θ^t, and ε, let $V_t^\varepsilon(a^{t-1}, \theta^t)$ denote the continuation value in the ε-control problem. Consider the ε-control problem at a^{t-1}, and write π for the prior and $\hat{u}_t^\varepsilon$ for the analogue of $\hat{u}_t$ with continuation values V_t^ε in place of V_t. We have

$$V_t^\varepsilon(a^{t-1}, \theta^t) = \log\left(\sum_{a_t} q(a_t \mid a^{t-1}) \exp(\hat{u}_t^\varepsilon(a^t, \theta^t))\right),$$

which converges to the expression in (12) since $p(a_t \mid a^{t-1}) = q(a_t \mid a^{t-1})$ at an optimum.

For the priors, note that the first-order condition with respect to $q(a_t \mid a^{t-1})$ for a solution of the ε-control problem at a^{t-1} with $q(a_t \mid a^{t-1}) \in (\varepsilon, 1)$ is

$$\sum_{\theta} \frac{\pi(\theta^t) e^{\hat{u}_\varepsilon(a^t,\theta^t)}}{\sum_{a'_t} q(a'_t \mid a^{t-1}) e^{\hat{u}_\varepsilon((a^{t-1},a'_t),\theta^t)}} = \mu, \tag{25}$$

where μ is the Lagrange multiplier associated with the constraint $\sum_{a'_t} q(a'_t \mid a^{t-1}) = 1$. Note that there must exist some a_t for which $q(a_t \mid a^{t-1}) \in (\varepsilon, 1)$. For this action a_t, we have $p(a_t \mid a^{t-1}) = q(a_t \mid a^{t-1})$, and hence the left-hand side of (25) is the sum of posterior beliefs, which must be equal to 1.

Now consider a_t for which the solution is $q(a_t \mid a^{t-1}) = \varepsilon$. Then we must have

$$\sum_{\theta} \frac{\pi(\theta^t) e^{\hat{u}_\varepsilon(a^t,\theta^t)}}{\sum_{a'_t} q(a'_t \mid a^{t-1}) e^{\hat{u}_\varepsilon((a^{t-1},a'_t),\theta^t)}} \leq \mu = 1.$$

In this case, the posterior beliefs satisfy

$$\begin{aligned}
p(\theta^t \mid a^t) &= \frac{\pi(\theta^t)}{p(a^t \mid a^{t-1})} p(a^t \mid \theta^t, a^{t-1}) \\
&= \frac{q(a^t \mid a^{t-1})}{p(a^t \mid a^{t-1})} \frac{\pi(\theta^t) e^{\hat{u}_\varepsilon(a^t,\theta^t)}}{\sum_{a'_t} q(a'_t \mid a^{t-1}) e^{\hat{u}_\varepsilon((a^{t-1},a'_t),\theta^t)}} \\
&= \frac{1}{\sum_{\tilde{\theta}^t} \pi(\tilde{\theta}^t) \frac{\exp(\hat{u}_\varepsilon(\theta^t,a^t))}{\sum_{a'_t} q(a'_t \mid a^{t-1}) \exp(\hat{u}_\varepsilon((a^{t-1},a'_t),\tilde{\theta}^t))}} \frac{\pi(\theta^t) \exp(\hat{u}_\varepsilon(a^t,\theta^t))}{\sum_{a'_t} q(a'_t \mid a^{t-1}) \exp(\hat{u}_\varepsilon((a^{t-1},a'_t),\theta^t))}.
\end{aligned}$$

Therefore, as ε vanishes, the posteriors indeed approach those given by (24). □

C Proofs and computations for Section 4

C.1 Status quo bias

The symmetry of this example implies that there is a symmetric solution. The predisposition $q_1(a_1)$ toward action $a_1 \in \{0, 1\}$ in the first period is 1/2, the predisposition $s := q_2(a_2 = a \mid a_1 = a)$ toward maintaining the same action in the second period is independent of a, and the continuation value function attains only two values,

$$V_2\left(a_1, \theta^2\right) = \begin{cases} V_c & \text{if } a_1 = \theta_2, \\ V_w & \text{if } a_1 \neq \theta_2. \end{cases}$$

One may interpret V_c as the expected payoff in period 2, including the information cost, when the action a_1 suggests the correct choice of a_2, and V_w as the corresponding payoff when a_1 suggests the wrong choice. By (12), the continuation payoffs satisfy $V_c = \log(se + (1-s))$ and $V_w = \log(s + (1-s)e)$.

Proposition 3 states that the choice rule in each period is a solution to a static RI problem.[20] The first-stage static RI problem involves gross payoffs

$$\hat{u}_1(a_1, \theta_1) = \begin{cases} 1 + 0.9V_c + 0.1V_w & \text{if } a_1 = \theta_1, \\ 0.9V_w + 0.1V_c & \text{if } a_1 \neq \theta_1, \end{cases}$$

and a uniform prior on θ_1. The agent assigns posterior probability p_1 to her choice in the first period being correct, where, according to (9),

$$p_1 = p(\theta_1 = a_1 \mid a_1) = \frac{\frac{1}{2}e^{\hat{u}_1(a_1,a_1)}}{\frac{1}{2}e^{\hat{u}_1(a_1,a_1)} + \frac{1}{2}e^{\hat{u}_1(1-a_1,a_1)}}. \tag{26}$$

Note that p_1 is independent of $a_1 \in \{0, 1\}$, and that it is equal to the probability that the first-period decision is correct.

The second-stage static RI problem involves gross payoffs

$$\hat{u}_2(a_2, \theta_2) = \begin{cases} 1 & \text{if } a_2 = \theta_2, \\ 0 & \text{if } a_2 \neq \theta_2, \end{cases}$$

and a prior on θ_2 that depends on the choice of action in period 1. Specifically, at the beginning of period 2, the prior belief that $\theta_2 = a_1$ is $p_2 = p(\theta_2 = a_1 \mid a_1) = 0.9p_1 + 0.1(1 - p_1)$.

We want to show that under the optimal choice rule, $a_1 = a_2$ almost surely. This involves conjecturing that the predisposition s is equal to 1, computing the optimal choice rule p given this conjecture, and verifying that the predisposition $s = 1$ maximizes the expected value in the second period, where the expectation is computed under beliefs induced by the choice rule p. Accordingly, suppose the predisposition s is equal to 1. Then $V_c = 1$ and $V_w = 0$. It follows from (26) that $p_1 \approx 0.86$. Then $p_2 \approx 0.79$, and since $s = 1$, this is also

[20]The extension of Proposition 3 to non-interior choice rules is immediate in this case because the optimal choice rule never assigns zero probability to any action in the first period, only in the last period. Therefore, posterior beliefs following zero-probability histories are irrelevant for continuation values.

the probability that the second decision is correct. To verify that $s = 1$, we need to check that the predisposition $s = 1$ maximizes the expected value

$$p_2 \log (se + (1 - s)) + (1 - p_2) \log (s + (1 - s)e),$$

in the second period given the belief p_2, which is indeed the case.

C.2 Inertia

We say that a solution to Example 2 is *eventually interior* if there exists t' such that each action is chosen with positive probability in every period $t > t'$. The next lemma follows directly from Proposition 3.

Lemma 4. *Suppose there is an eventually interior solution to the model in Example 2. Then there exists t' such that for $t > t'$, conditional on a_{t-1} and θ_t, a_t is independent of θ^{t-1} and a^{t-2}. Moreover, there is an optimal choice rule for which, in each period $t > t'$,*

$$p(a_t \mid \theta_t, a_{t-1}) = \frac{q(a_t \mid a_{t-1}) \exp \left(u(a_t, \theta_t) + \delta E \left[V(a_t, \theta_{t+1}) \mid \theta_t\right]\right)}{\sum_{a'} q(a' \mid a_{t-1}) \exp \left(u(a', \theta_t) + \delta E \left[V(a', \theta_{t+1}) \mid \theta_t\right]\right)}, \tag{27}$$

where the continuation payoffs solve

$$V(a_{t-1}, \theta_t) = \log \left(\sum_a q(a \mid a_{t-1}) \exp \left(u(a, \theta_t) + \delta E \left[V(a, \theta_{t+1}) \mid \theta_t\right]\right) \right), \tag{28}$$

the predispositions $q(a_t \mid a_{t-1})$ solve

$$\sum_{\theta_{t-1}} p(\theta_{t-1} \mid a_{t-1}) \gamma(\theta_{t-1}, \theta_t) = \sum_{a_t} q(a_t \mid a_{t-1}) p(\theta_t \mid a_t) \tag{29}$$

for all θ_t and a_{t-1}, and the posteriors $p(\theta_t \mid a_t)$ satisfy

$$\frac{p(\theta_t \mid a_t)}{p(\theta_t \mid a'_t)} = \frac{\exp \left(u(a_t, \theta_t) + \delta E[V(a_t, \theta_{t+1}) \mid \theta_t]\right)}{\exp \left(u(a'_t, \theta_t) + \delta E[V(a'_t, \theta_{t+1}) \mid \theta_t]\right)}. \tag{30}$$

One can check whether there is an eventually interior solution by solving the system of equations in the lemma. If the resulting predispositions are positive then there is indeed an eventually interior solution.[21]

[21]Lemma 4 describes long-run behavior. Actions in early periods depend on the prior. If the distribution

Equation (29) is a condition on the posterior beliefs. The left-hand side is the prior belief about θ_t at the beginning of period t obtained by applying the transition probabilities of the Markov chain to the posterior about θ_{t-1} at the end of period $t-1$. The right-hand side is the same prior written as the expectation of the posterior at the end of period t. Equation (30) follows from (27) together with Bayes' rule; a similar expression appears in Caplin and Dean (2013).

Lemma 4 follows from the recursive characterization in Proposition 3 together with a result from Caplin and Dean (2013). They show that in static RI problems, the optimal posteriors $p(\theta \mid a)$ are constant across priors lying within their convex hull. In the present setting, this implies that the agent's posterior after choosing a_t is independent of her prior at the beginning of period t, and hence constant across all a_{t-1}. This property gives rise to the Markovian structure of the optimal actions and beliefs. The same argument applies for a general version of Example 2 with arbitrary finite state and action spaces and general payoffs.

Proof of Proposition 4. First, we claim that the rescaled predispositions $q^*(a, a')$ are bounded above by some K. To see this, note that condition (29) implies that, for $a \neq a'$,

$$q(a' \mid a) = \frac{\Pr(\theta_t = a \mid a_t = a)\gamma(a, a') - \Pr(\theta_t = a' \mid a_t = a)\gamma(a', a)}{\Pr(\theta_t = a' \mid a_t = a') - \Pr(\theta_t = a' \mid a_t = a)}. \tag{31}$$

The numerator of this expression is of order ε. The denominator is bounded away from zero because the difference $p(\theta_t = a' \mid a_t = a') - p(\theta_t = a' \mid a_t = a)$ in posteriors is larger than in the static RI problem in which the continuation values are zero.

Using the bound on $q^*(a, a')$ and the fact that the continuation values are bounded, (28) implies that there exists some K' such that for sufficiently small ε,

$$u(a, \theta) + \delta V(a, \theta) - K'\varepsilon \leq V(a, \theta) \leq u(a, \theta) + \delta V(a, \theta) + K'\varepsilon.$$

Thus $\lim_{\varepsilon \to 0} V(a, \theta) = \frac{u(a,\theta)}{1-\delta}$. Combining this with (30) leads to

$$\lim_{\varepsilon \to 0} p(\theta_t = a \mid a_t = a) = \frac{U_a U_{a'} - U_a}{U_a U_{a'} - 1},$$

of θ_1 lies in the convex hull of the long-run stationary posteriors $p(\theta_t \mid a_t)$ for the two actions, then the choice rule (27) is optimal beginning in period 2. If not, the agent acquires no information until her belief enters the convex hull of the stationary posteriors, after which (27) applies.

where $a' \neq a$. Substituting this last expression into (31) gives (17). Finally, (18) follows from (27). The comparative statics can be checked by taking derivatives of the expressions for α. $\square$

C.3 Response times

Consider the problem

$$\begin{aligned} &\max_p E\left[\sum_{t'=1}^{T} u_{t'}\left(a^{t'},\theta\right)\right] \\ \text{s.t. } &E\left[\sum_{t'=1}^{t} I\left(\theta; a^{t'} \mid a^{t'-1}\right)\right] \leq \kappa t \text{ for all } t = 1,\ldots,T \end{aligned} \tag{32}$$

in which the constraint is relaxed relative to the original problem, (19). We will show that solution of (32) also solves (19).

For each $t = 1,\ldots,T$, the capacity constraint in (32) is equivalent to $I(\theta; a^t) \leq \kappa t$. Hence, the set of the feasible joint distributions $p(\theta, a^T)$ satisfying the constraints of (32) is convex. Since the objective of (32) is linear in $p(\theta, a^T)$, the first-order conditions are sufficient for a global optimum.

The solution of (32) also solves

$$\max_p E\left[\sum_{t'=1}^{T}\left(u_{t'}\left(a^{t'},\theta\right) - \lambda_{t'} I\left(\theta; a^{t'} \mid a^{t'-1}\right)\right)\right], \tag{33}$$

where $\lambda_{t'}$ are the shadow prices of the information capacity for $t' = 1,\ldots,T$. If λ_t is decreasing then Problem (33) is a particular case of the dynamic RI problem from Definition 1 (after rescaling discount factors and payoffs as described following Definition 1).

Assume that λ_t is indeed decreasing (we verify this below). Then we may solve (33) using Proposition 3. The only non-trivial decision node at period t is the one with $a^{t-1} = w^{t-1}$. By symmetry, the prior belief about θ at the decision node w^{t-1} is uniform on $\{0,1\}$. Symmetry also implies that the continuation value $V_t(w^{t-1},\theta)$ is independent of $\theta \in \{0,1\}$; accordingly, we omit the arguments of V_t. Hence at the node w^{t-1}, the agent solves a static

RI problem with a uniform prior over θ and payoffs

$$\hat{u}_t\left(a_t,\theta\right)=\begin{cases}1 & \text{if } a_t=\theta,\\ 0 & \text{if } a_t=1-\theta,\\ V_{t+1}-c & \text{if } a_t=w.\end{cases}$$

This static RI problem can be solved using Proposition 2. Symmetry implies that the predisposition $q(a_t \mid w^{t-1})$ is the same for the two terminal actions $a_t\in\{0,1\}$; we denote it by $s_t/2$, which makes s_t the probability that the agent takes a terminal action at t conditional on waiting in the previous periods. In this case, Proposition 2 implies that s_t solves

$$\max_{s_t\in[0,1]}\log\left(\frac{s_t}{2}\left(e^{1/\lambda_t}+1\right)+(1-s_t)e^{(V_{t+1}-c)/\lambda_t}\right). \tag{34}$$

By (12), the value associated with the static RI problem at time t is

$$V_t=\lambda_t\log\left(\frac{s_t}{2}\left(e^{1/\lambda_t}+1\right)+(1-s_t)e^{(V_{t+1}-c)/\lambda_t}\right). \tag{35}$$

Using the first-order condition for problem (34) together with (35), it is easy to check that if the solution to (34) satisfies $s_t\in(0,1)$ then $V_{t+1}=V_t+c$, and that

$$V_t=\lambda_t\log\left(\frac{1}{2}\left(e^{1/\lambda_t}+1\right)\right) \tag{36}$$

whenever $s_t\in(0,1]$.

Since it can never be optimal to delay with probability one, there exists some $t^*\in\{1,\ldots,T\}$ such that $s_t\in(0,1)$ for all $t<t^*$, and $s_{t^*}=1$, meaning that the agent always makes a terminal decision by time t^*. In addition, there exists some V_1 such that $V_t=V_1+c(t-1)$ for each $t=1,\ldots,t^*$. Substituting into (36), we see that λ_t is decreasing in time, as claimed.

Recall that $g_t=\Pr\left(a_t=\theta \mid a_t\in\{0,1\}\right)$ denotes the accuracy of the terminal decision when it is made at time t. From (7), we obtain $g_t=\frac{e^{1/\lambda_t}}{1+e^{1/\lambda_t}}$. Solving for λ_t and substituting into (36) gives a condition relating g_t to V_1, namely

$$\frac{\log\left(2(1-g_t)\right)}{\log\left(\frac{1-g_t}{g_t}\right)}=V_1+c(t-1). \tag{37}$$

The next step is to derive an expression for r_t in terms of g_t. Recall that r_t is the probability that a terminal decision is made in period t. For $t \leq t^*$, the capacity constraint binds, and hence r_t satisfies

$$r_t \left(\log 2 + g_t \log g_t + (1 - g_t) \log(1 - g_t)\right) = \kappa. \tag{38}$$

The expression in the parentheses on the left-hand side is the difference between the entropy of the prior belief at t and that of the posterior belief after taking a terminal action at t.

We determine the value $V_1 \in (0, 1)$ numerically. It is the maximal value for which there exists a natural number $t^* \leq T$ such that $\sum_{t=1}^{t^*} r_t = 1$.

Notice that the solution of the relaxed problem (32) also solves the original problem (19) because the constraints of (32) are binding for $t = \{1, \ldots, t^*\}$, and hence the solution of the relaxed problem satisfies the constraints of the original problem.

References

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica 17*, 213–244.

Baliga, S. and J. C. Ely (2011). Mnemonomics: the sunk cost fallacy as a memory kludge. *American Economic Journal: Microeconomics 3*(4), 35–67.

Caplin, A. and M. Dean (2013). The behavioral implications of rational inattention with Shannon entropy. Working paper.

Compte, O. and P. Jehiel (2007). Auctions and information acquisition: sealed bid or dynamic formats? *Rand Journal of Economics 38*(2), 355–372.

Fudenberg, D., P. Strack, and T. Strzalecki (2015). Stochastic choice and optimal sequential sampling. Working Paper.

Fudenberg, D. and T. Strzalecki (2015). Dynamic logit with choice aversion. *Econometrica 83*(2), 651–691.

Hobson, A. (1969). A new theorem of information theory. *Journal of Statistical Physics 1*(3), 383–391.

Liu, Q. (2011). Information acquisition and reputation dynamics. *Review of Economic Studies 78*(4), 1400–1425.

Luo, Y. (2008). Consumption dynamics under information processing constraints. *Review of Economic Dynamics 11*(2), 366–385.

Mackowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review 99*(3), 769–803.

Mackowiak, B. and M. Wiederholt (2010). Business cycle dynamics under rational inattention. Forthcoming in the *Review of Economic Studies*.

Mattsson, L.-G. and J. W. Weibull (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior 41*(1), 61–78.

Matějka, F. (2010). Rationally inattentive seller: Sales and discrete pricing. CERGE-EI Working Paper Series, 408.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–98.

Moscarini, G. (2004). Limited information capacity as a source of inertia. *Journal of Economic Dynamics and Control 28*(10), 2003–2035.

Moscarini, G. and L. Smith (2001). The optimal level of experimentation. *Econometrica 69*(6), 1629–1644.

Ravid, D. (2014). Bargaining with rational inattention. Working paper.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal 117*(523), 1243–1259.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica 55*(5), 999–1033.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory 4*(1), 25–55.

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

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Spiliopoulos, L. and A. Ortmann (2014). The BCD of response time analysis in experimental economics. Working paper.

Stahl, D. O. (1990). Entropy control costs and entropic equilibria. *International Journal of Game Theory 19*(2), 129–138.

Sundaresan, S. and S. Turban (2014). Inattentive valuation and belief polarization. Working paper.

Tutino, A. (2013). Rationally inattentive consumption choices. *Review of Economic Dynamics 16*(3), 421–439.

Van Damme, E. (1983). *Refinements of the Nash Equilibrium Concept*, Volume 219 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag.

Wilson, A. (2014). Bounded memory and biases in information processing. *Econometrica 82*(6), 2257–2294.

Woodford, M. (2014). Stochastic choice: An optimizing neuroeconomic model. *American Economic Review 104*(5), 495–500.