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

September 26, 2016

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 is independent of the realized state. The model provides the same fit to choice data as dynamic logit, but, because of the bias, yields different counterfactual predictions. 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.

*We thank Victor Aguirregabiria, Mehmet Ekmekci, Štěpán Jurajda, Jakub Kastl, Philipp Kircher, Alessandro Pavan, Marcin Pęski, Vasiliki Skreta, Joel Sobel, Balázs Szentes, three anonymous referees, and participants in various seminars and conferences for their comments. Maxim Goryunov provided excellent research assistance. Steiner was supported by the Purkyně fellowship of the Czech Academy of Sciences and by GAČR grants 13-34759S and 16-00703S. Stewart is grateful to SSHRC for financial support of this research.

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

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 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. At the end of each period, she observes a costless signal that may depend on her action choice; for example, she may observe her realized flow payoff. 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 agent's information at the corresponding time.

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 counterfactual extrapolation of behavior based on identification of utility parameters. An

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

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 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, costless signals, 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

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

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]

Our second application extends the first one to a stationary, infinite-horizon setting. A binary state follows a Markov chain. The agent chooses a binary action in each period with the goal of matching the state. This model can be viewed, for instance, as capturing an investor's choice between two asset classes, or a consumer's choice between two products. Unlike analogous models with exogenous information or restrictions on available signals, long-run behavior satisfies a simple Markov property: choice probabilities in each period depend only on the action in the preceding period and the current state. Inertia in states is reflected by inertia in actions. Provision of costless information, although beneficial for the agent's payoff, can have a perverse effect on behavior: relative to the case in which all information is costly, additional costless signals can make the agent's actions *less* likely to match the state. The expectation of free signals in the future crowds out the agent's incentive to acquire costly information, in some cases to the extent that the overall precision of information declines.

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, the state, and the chosen action. We find that, for a range of delay costs, the probability of choosing the correct action is constant over time, and so is the hazard rate at which decisions are made (up until the final period). In addition, the expected delay time is non-monotone in the agent's capacity, with intermediate levels being associated with the longest delays.

[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.

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. Since the agent in our model can choose any signal in each period, the set of possible information acquisition strategies is very large. With entropy-based costs, 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 without explicitly considering all information acquisition strategies.[4] Entropy-based costs are also important for the reformulation of the dynamic RI problem 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). Similarly, in dynamic contexts, it makes no difference whether there are multiple opportunities to acquire information between action choices or just one; in this sense, preferences over the timing of information acquisition across periods are driven by the payoffs in the decision problem (together with discounting of costs). 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, Maćkowiak and Wiederholt (2009, 2015); Maćkowiak, Matějka, and Wiederholt (2016); and Matějka (2015) 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

[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 2.1.

[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.

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 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 $t = 1, 2, \ldots$.[6] A payoff-relevant state θ_t is drawn in each period t from a finite set Θ_t according to a probability measure $\pi \in \Delta(\prod_t \Theta_t)$. Let $\Theta^t = \prod_{\tau \leq t} \Theta_\tau$, and, for any sequence $(b_\tau)_\tau$, let $b^t = (b_1, \ldots, b_t)$ for each t. 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 $|A| \leq |X| < \infty$.[7] At time t, the agent can choose *any* signal about the history θ^t with realizations x_t in X. In addition, after choosing her action in each period t, she observes a costless signal y_t from a finite set Y distributed according to a given $g_t(y_t \mid \theta^t, a^t, y^{t-1}) \in \Delta(Y)$; we denote by g the complete family of these distributions across all histories (θ^t, a^t, y^{t-1}). The signal y_t incorporates all of the costless information the agent receives. For example, y_t may correspond to observation of the payoff she receives in period t, or observation of the realized state (either perfectly or with noise in each case). One important special case—which has been the focus of the previous dynamic RI literature—is when there is no costless information, corresponding to $|Y| = 1$. Let $z_t = (x_t, y_t)$ and $Z = X \times Y$. We refer to z^{t-1} as the *decision node at* t.

We assume that, for each a^t, y_t is independent of x^t conditional on θ^t and y^{t-1}; while the agent's choice of actions may affect the distribution of y_t, her choice of information does not.

A strategy $s = (f, \sigma)$ is a pair comprised of

1. an *information strategy* f consisting of a system of signal distributions $f_t(x_t \mid \theta^t, z^{t-1})$, one for each θ^t and z^{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 $\sigma_t : Z^{t-1} \times X \longrightarrow A$, where $\sigma_t(z^{t-1}, x_t)$ indicates the choice of action at time t for each history z^{t-1} and current costly signal x_t.[8]

[6] Although the action set is constant across time, the model can implicitly allow for varying action sets by making certain choices payoff-irrelevant.

[7] The restriction to finite action, signal, and state spaces is technically convenient in that it allows us to work with discrete distributions, avoiding issues of measurability, and the need to condition on zero probability histories along the realized action path. We conjecture that Lemma 1, Theorem 1, and Proposition 3 would extend to standard continuous models.

[8] This formulation implicitly assumes that the agent perfectly recalls all past information. In contrast, Woodford (2009) analyzes a model in which *all* information is costly, even observation of the current time.

Given an action strategy σ, we denote by $\sigma^t(z^{t-1}, x_t)$ the history of actions up to time t given the realized signals.

The agent receives flow utilities $u_t(a^t, \theta^t)$ that depend on the entire history of actions and states, and 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 W with finite support S distributed according to $p \in \Delta(S)$. Recall that the entropy

$$H(W) = -\sum_{w \in S} p(w) \log p(w)$$

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

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

between x_t and the history of states θ^t. 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 $H\left(\theta^t \mid z^{t-1}\right)$. After, her level of uncertainty becomes $H\left(\theta^t \mid z^{t-1}, x_t\right)$. The mutual information is the expected reduction in uncertainty averaged across all realizations of x_t.[9]

The agent solves the following problem.

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

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

where $\lambda > 0$ is an information cost parameter, and the expectation is taken with

[9]Note that $I\left(\theta^t; x_t \mid z^{t-1}\right)$ depends on the realization of z^{t-1}. Other authors sometimes use this notation to refer to the expectation of this quantity across z^{t-1}.

respect to the distribution over sequences $(\theta_t, z_t)_t$ induced by the prior π together with the strategy (f, σ) and the distributions g of costless signals.

To simplify notation, we normalize 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).

One implicit assumption of our model is that the agent's actions do not affect the distribution of future states. However, since flow payoffs depend on the entire histories of actions and states, any problem having this feature is equivalent to one with larger state spaces that fits within our framework. The idea is to make each state $\tilde{\theta}_t$ in the equivalent problem correspond to a vector of states in the original problem, one for each history a^{t-1} of actions. Payoffs in state $\tilde{\theta}_t$ are equal to those in the original problem for the component of $\tilde{\theta}_t$ associated with a^{t-1}. Similarly, the distribution of $\tilde{\theta}_t$ conditional on $\tilde{\theta}^{t-1}$ is constructed so as to ensure that, for each a^{t-1}, the marginal distribution of the component associated with a^{t-1} matches the distribution of states in the original problem.

To make this precise, suppose for simplicity that $\Theta_t = \Theta$ for all t, and let $\pi_t(\theta \mid \theta^{t-1}, a^{t-1}) \in \Delta(\Theta)$ denote the probability of state θ in period t following the history (θ^{t-1}, a^{t-1}). Let $\tilde{\Theta}_t = \Theta^{A^{t-1}}$, with typical element $\tilde{\theta}_t : A^{t-1} \longrightarrow \Theta$, where $A^0 := \{\emptyset\}$. An equivalent problem with state spaces $\tilde{\Theta}_t$ and distributions that are independent of actions may be obtained by taking gross utilities to be

$$\tilde{u}_t(a^t, \tilde{\theta}^t) \equiv u_t\big(a^t, \big(\tilde{\theta}_1(\emptyset), \ldots, \tilde{\theta}_t(a^{t-1})\big)\big)$$

and the distribution of states in each period t to be

$$\tilde{\pi}_t(\tilde{\theta}_t \mid \tilde{\theta}^{t-1}) \equiv \prod_{a^{t-1} \in A^{t-1}} \pi_t\big(\tilde{\theta}_t(a^{t-1}) \mid \big(\tilde{\theta}_1(\emptyset), \ldots, \tilde{\theta}_{t-1}(a^{t-2})\big), a^{t-1}\big).$$

Thus following each history $(\tilde{\theta}^{t-1}, a^{t-1})$, for $\theta^{t-1} = \big(\tilde{\theta}_1(\emptyset), \ldots, \tilde{\theta}_{t-1}(a^{t-2})\big)$, there is probability equal to $\pi_t(\theta_t \mid \theta^{t-1})$ of reaching a state $\tilde{\theta}_t$ in period t satisfying $\tilde{\theta}_t(a^{t-1}) = \theta_t$, ensuring that gross payoffs from each action correspond to $u_t(a^t, \theta^t)$.

2.1 Preliminaries

Our main goal is to characterize the agent's action choices. Let $\omega^t = (\theta^t, a^{t-1}, y^{t-1})$. A *(stochastic) choice rule* p is a system of distributions $p_t(a_t \mid \omega^t)$ over A, one for each ω^t, interpreted as the probability of choosing a_t at the history ω^t. We say that a strategy $s = (f, \sigma)$ *generates the choice rule* p if

$$p_t\left(a_t \mid \omega^t\right) \equiv \Pr\left(\sigma_t\left(z^{t-1}, x_t\right) = a_t \,\middle|\, \theta^t, \sigma^{t-1}\left(z^{t-2}, x_{t-1}\right) = a^{t-1}, y^{t-1}\right),$$

for all a_t and $\omega^t = (\theta^t, a^{t-1}, y^{t-1})$, where the probability is evaluated with respect to the joint distribution of sequences of states and signals generated according to f, σ, and g.

Conversely, a choice rule p can be associated (non-uniquely) with a strategy (f, σ). Roughly speaking, one can choose a particular signal realization in X 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.[10] If s is a strategy obtained in this way from a choice rule p, we say that p *induces* s.

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 independently proved the corresponding result in a related model with a continuum of states, and without costless signals.

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

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

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

Accordingly, we henceforth dispense with the signals x_t, replacing them with actions a_t, so that $z_t = (a_t, y_t)$, and 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.

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

In static models, the conclusion of Lemma 1 holds as long as the cost of signals is nondecreasing in Blackwell informativeness and all signal structures are feasible. To see why, consider a strategy in which two distinct signal realizations (each occurring with positive probability) lead to the same action. Combining these into a single realization has no effect on the distribution of actions and (weakly) reduces the information cost.

In dynamic problems, more structure is needed because information that is acquired but not used in a given period may be used later on. For the lemma to hold, it must be that delaying the acquisition of information until the time when it is used never increases the information cost relative to acquiring it earlier. For example, if the cost were a nonlinear function of 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, given the other information she acquires, the marginal cost of acquiring it would be higher. When the cost is linear in mutual information and the agent (weakly) discounts the future, the additive property of entropy ensures that delaying information acquisition never increases the cost, regardless of other information acquired by the agent.

Lemma 1 also relies on several other assumptions of our model. The result would not necessarily hold if $|X| < |A|$, or if there were restrictions on what information strategies are feasible. For example, it fails if past states are payoff-relevant but the agent can only learn about the current state. This lemma also relies on the conditional independence of x^t and y_t, for otherwise the choice of costly signal could affect the value of the free information conveyed by y_t directly (as opposed to indirectly through the choice of action).

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

Proposition 1 makes use of the finiteness of A, Y, and each Θ_t to ensure that the strategy space is compact, and the boundedness of payoffs together with discounting to ensure that the agent's objective function is continuous in her strategy.

3 Solution

This section presents 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 for general payoffs with states (θ^t, y^{t-1}) in period t. We denote the payoff function in period t by v_t to distinguish it from the payoff function u_t in the rational inattention problem (which does not depend on y^{t-1}). Recall that $\omega^t = (\theta^t, a^{t-1}, y^{t-1})$.

Definition 2 (Rust (1987))**.** *A choice rule r is a* dynamic logit rule under payoffs $(v_t(a_t, \omega^t))_t$ *if*

$$r_t\left(a_t \mid \omega^t\right) = \frac{\exp\left(\hat{v}_t\left(a_t, \omega^t\right)\right)}{\sum_{a_t'} \exp\left(\hat{v}_t\left(a_t', \omega^t\right)\right)},$$

where

$$\hat{v}_t\left(a_t, \omega^t\right) = v_t\left(a_t, \omega^t\right) + \delta_{t+1} E\left[V_{t+1}\left(\omega^{t+1}\right) \mid a_t, \omega^t\right],$$

and the continuation values V_t satisfy

$$V_t\left(\omega^t\right) = \log\Bigl(\sum_{a_t} \exp\left(\hat{v}_t\left(a_t, \omega^t\right)\right)\Bigr). \tag{4}$$

The solution to the dynamic RI problem is a dynamic logit rule under payoffs that are modified by an endogenous state-independent term. A *default rule* q is a system of conditional action distributions $q_t(a_t \mid z^{t-1})$, one for each decision node $z^{t-1} = (a^{t-1}, y^{t-1})$. The difference between a default rule and a choice rule is that the latter conditions on θ^t while the former does not.

Given any default rule q, write $u + \log q$ to represent the system of payoff functions

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

for all t. Let $\mathcal{V}(v) = E_{\theta_1}[V_1(\theta_1, \emptyset, \emptyset)]$ denote the first-period expected value from (4) under the system of payoff functions $v = (v_t)_t$. For any choice rule p, let $\overline{p}_t\left(a_t \mid z^{t-1}\right)$ denote the probability of choosing action a_t conditional on reaching decision node z^{t-1}, that is,

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

We adopt the convention that $\log 0 = -\infty$ and $\exp(-\infty) = 0$.

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,

$$q_t \left(a_t \mid z^{t-1}\right) = \overline{p}_t \left(a_t \mid z^{t-1}\right) \tag{5}$$

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

A converse to Theorem 1 also holds, with a minor caveat: for any solution p to the dynamic RI problem, and for q satisfying (5), p coincides with the dynamic logit rule under payoffs $u + \log q$ at every ω^t that is reached with positive probability.

Given a default rule q, we refer to $q_t(a_t \mid z^{t-1})$ as the *predisposition* toward action a_t at the decision node z^{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. To keep the discussion simple, consider the case in which there is no costless signal y_t. The agent behaves *as if* she incurs a cost

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

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 $u_t - c_t$. Put differently, the two models provide an equally good fit for the data. However, the distinction is important when extrapolating to other decision problems (as in counterfactual experiments). 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.[11]

Kennan and Walker (2011) estimate a dynamic model of migration decisions. Each agent chooses a location to maximize her expected income less the cost of moving. Estimated moving costs are surprisingly large. If, as in our model, agents can acquire costly information about wages at other locations, estimates of the moving costs would decrease: since moves are relatively rare, the log-predisposition term in our model creates a virtual cost of moving, which the cost identified by Kennan and Walker (2011) combines with the true moving cost. In addition, Kennan and Walker (2011) consider a counterfactual policy experiment involving a subsidy for moving. In our model, the effect of the subsidy would be larger than their estimates. Not only does the subsidy have a direct effect on payoffs, it also increases the predisposition toward moving, thereby lowering the associated virtual cost; the information acquisition induced by the subsidy reinforces the increase in migration.

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 $c_t(a^{t-1}, a_t) = -\log \overline{p}_t(a_t \mid a^{t-1})$ can be identified directly based on the frequency with which each action is chosen.

[11] Similar comments apply to other work estimating demands using dynamic logit models. For example, Gowrisankaran and Rysman (2012) study demand for durable goods, while Schiraldi (2011) focuses on automobiles.

3.2 Reduction to static problems

The dynamic logit characterization of Theorem 1 is related to a reduction of the dynamic RI problem 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 static 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 write $\pi^p(\theta \mid a)$ for the posterior belief after choosing action a given the choice rule p.[12]

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*

$$p(a \mid \theta) = \frac{q(a) \exp\left(u(a, \theta)\right)}{\sum_{a'} q(a') \exp\left(u(a', \theta)\right)}, \tag{7}$$

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

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

[12] 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.

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

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

We show that the dynamic RI problem can be reduced to a collection of static RI problems, one for each decision node z^{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 z^{t-1}, the gross payoff consists of the flow payoff plus a continuation value, and the prior belief is obtained by Bayesian updating given z^{t-1}.

We write $\pi_p(\theta^t \mid z^{t-1})$ for the agent's prior over θ^t at the decision node z^{t-1} given a choice rule p, and $\pi^p(\theta^t \mid z^t)$ for the posterior over θ^t after z^t.

We say that a dynamic choice rule is *interior* if, at every decision node, it chooses each action with positive probability. For simplicity, we state here the result only for interior dynamic choice rules. We extend the result to the general case in Proposition 7 in Appendix B.

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

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

and payoff function

$$\hat{u}_t\left(a_t, \theta^t; z^{t-1}\right) = u_t\left(a^t, \theta^t\right) + \delta_{t+1} E\left[V_{t+1}\left(\theta^{t+1}, z^t\right) \mid a_t, \theta^t, z^{t-1}\right], \tag{11}$$

where the posterior belief $\pi^p(\theta^t \mid z^t)$ satisfies Bayes' rule with respect to the prior $\pi_p\left(\theta^t \mid z^{t-1}\right)$, and the continuation values satisfy

$$V_t\left(\theta^t, z^{t-1}\right) = \log\Big(\sum_{a_t} p_t\left(a_t \mid z^{t-1}\right)\exp\left(\hat{u}_t\left(a_t, \theta^t; z^{t-1}\right)\right)\Big). \tag{12}$$

As for Theorem 1, the converse to this proposition holds at all decision nodes reached with positive probability.

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}, y^{t-1}) but not the choice rule used 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 choice rule 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. Sections 4.2 and 4.3 illustrate this approach.

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.

Reformulating the dynamic RI problem as a control problem with observable states 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*

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

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

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 ω^t (and in particular about θ^t), but must trade off optimizing her flow utility u_t against a control cost: for each $\omega^t=(\theta^t,z^{t-1})$, she pays a cost

$$E_{p_t}\left[\log p_t\left(a_t\mid\omega^t\right)-\log q_t\left(a_t\mid z^{t-1}\right)\mid\omega^t\right]$$

for deviating from the default action distribution $q_t\left(a_t\mid z^{t-1}\right)$ to the action distribution $p_t\left(a_t\mid\omega^t\right)$.

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*

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

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

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 z^{t-1} does depend on the distribution $\pi_p(\theta^t \mid z^{t-1})$; however, for any fixed p, optimizing q does not require varying these 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 $p(x) \in \Delta(S)$,

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

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.

The use of properness in the proof also relies on the information cost being proportional to the reduction in entropy; the result would not hold for costs that are nonlinear functions of the mutual information.

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

$$\max_p E\left[\sum_t \delta^{(t)}\Big(u_t\big(a^t,\theta^t\big) - I\big(a_t;\theta^t \mid z^{t-1}\big)\Big)\right]$$
$$= \max_p E\left[\sum_t \delta^{(t)}\Big(u_t\big(a^t,\theta^t\big) - H\big(a_t \mid z^{t-1}\big) + H\big(a_t \mid \omega^t\big)\Big)\right]. \quad (15)$$

By properness,

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

Substituting this into (15) and recalling that

$$E\Big[H\big(a_t \mid \omega^t\big)\Big] = E\Big[-\log p_t\big(a_t \mid \omega^t\big)\Big]$$

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 apply our results in three examples. The first illustrates the mechanics of the model in a particularly simple setting. The second shows how allowing for unrestricted information choice can generate a simpler solution than one would obtain with exogenous information or standard restricted classes of signals. The last example demonstrates that Proposition 3 can be useful in problems with a constraint on information acquisition in each period instead of a cost.

4.1 Status quo bias

Our first application uses a particularly simple instance of our model to illustrate intertemporal incentives to acquire information. In doing so, we show that the dynamics of choice by a rationally inattentive agent may resemble status-quo-bias behavior (see, e.g., Samuelson and Zeckhauser (1988)). The agent chooses an action $a_t \in \{0, 1\}$ at $t = 1, 2$. 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 way: θ_1 is equally likely to be 0 or 1, and, whatever the realized value of θ_1, the probability that $\theta_2 \neq \theta_1$ is $\gamma < 1/2$. The agent receives no costless signal.

This example can be interpreted as a stylized model of investment in a risky asset. The agent prefers to invest (corresponding to $a_t = 1$) if and only if the return from the asset exceeds the risk-free rate (corresponding to $\theta_t = 1$). Learning about the quality of the asset is costly, as is monitoring its performance.

We analyze the correlation between actions across the two periods. 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, even if the state changes; an outside observer who sees the realized states and the agent's actions might conclude that she has a preference for maintaining the same choice. The following proposition shows that the optimal strategy has this feature whenever the serial correlation in the state is sufficiently high.

Proposition 4. *There exists* $\gamma^* \in (0, 1/2)$ *such that, under the optimal choice rule,* $\Pr(a_2 = a_1) = 1$ *whenever* $\gamma < \gamma^*$.

The proposition holds for $\gamma^* \approx 0.16$; thus if the probability that the state changes is less than 0.16 then the agent acquires information only in the first period, and relies on that information in both periods. Since the state may change in between the periods, the agent performs better in the first period than in the second (in the sense that her action is more likely to match the state).

The superior performance in the first period illustrates the importance of the endogenous timing of information acquisition in our model. 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 if γ is sufficiently small. 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.

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

Our second application consists of a stationary infinite-horizon environment in which the state follows a Markov chain and the agent chooses an action in each period with the goal of matching the state. This example can be viewed as a stylized model of a wide range of economic phenomena. The action could represent a consumer's choice of what product to purchase, an investor's choice of whether to hold a particular asset, or a worker's choice of whether to participate in the labor market. We start by analyzing a model in which all information acquired by the agent is costly. For example, in product choice, one can think of the state as capturing which product offers a larger surplus, which is costly to monitor.

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 Maćkowiak and Wiederholt (2009), provide results for quadratic payoffs and normally distributed shocks. Our framework provides an alternative approach suitable for general payoffs and distributions in discrete environments.

The agent chooses an action $a_t \in \{0,1\}$ in each period $t \in \mathbb{N}$. The state θ_t follows a Markov chain on $\{0,1\}$ with time-homogeneous transition probabilities $\gamma(\theta,\theta')$ from θ to θ'. In each period t, the gross flow payoff $u(a_t,\theta_t)$ is $u_a > 0$ if $a_t = \theta_t = a$, and 0 if $a_t \neq \theta_t$. Payoffs are discounted exponentially with discount factor $\delta \in (0,1)$. The agent receives no costless signal.

In contrast to the analogous model with exogenous information, behavior in this framework is Markovian: the choice rule, continuation values, and predispositions in any period t depend on the last action a_{t-1} and the current state θ_t, but not on any earlier actions or states. Moreover, after a finite number of periods, the choice rule is stationary. This implies that the long-run behavior is characterized by a finite set of equations, making it amenable to numerical computation. 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.

The Markov property of the solution follows from Proposition 3 together with a result of Caplin and Dean (2013). They show that in static RI problems, the set of posterior beliefs that arises from the optimal choice of signal is constant across priors lying within its convex hull. By Proposition 3, the same result holds in dynamic problems. In the present setting, it follows that the agent's posterior after choosing a_t is independent of her prior at the beginning of period t whenever she acquires a nontrivial signal. In particular, this posterior is independent of a^{t-1}.

Given an optimal choice rule p, we denote by $\hat{p}(a_t \mid \theta_t, a_{t-1})$ the long-run stationary choice rule; thus $\hat{p}(a_t \mid \theta_t, a_{t-1}) \equiv p_t(a_t \mid \theta^t, a^{t-1})$ for sufficiently large t.

We say that a solution is *eventually interior* if there exists t' such that, for every $t > t'$, each action is chosen with positive probability at each a^{t-1}. Lemma 6 in the appendix translates Proposition 3 to characterize the long-run solution in the present

[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.

setting in terms of a system of equations, provided that it is eventually interior. One can check whether this is the case by solving the system of equations in the lemma. If all of the resulting predispositions lie in $(0, 1)$ then there is indeed an eventually interior solution.

The next result shows that the model generates intuitive comparative statics.[14] We say that the states θ_t have *positive persistence* if $\gamma(0,0) + \gamma(1,1) > 1$; similarly, actions a_t eventually have positive persistence if $\Pr(a_t = 0 \mid a_{t-1} = 0) + \Pr(a_t = 1 \mid a_{t-1} = 1) > 1$ for all t sufficiently large. Positive persistence captures inertia in the process: it says that the state one period earlier is more likely to be the same as the current state than different.

Proposition 5. *Suppose states have positive persistence and the solution is eventually interior. Then*

1. *actions eventually have positive persistence, and moreover, the choice rule satisfies* $\hat{p}(a_t \mid \theta_t, a_{t-1}) > \hat{p}(a_t \mid \theta_t, a'_{t-1})$ *whenever* $a_t = a_{t-1} \neq a'_{t-1}$*; and*

2. *the posterior probability* $\pi^p(\theta_t = a \mid a_t = a')$ *is nonincreasing in the payoff* u_a *for all* $a, a' \in \{0, 1\}$ *and all* t*.*

The first part of the proposition says that inertia in the state will be reflected by inertia in behavior. The second part says that if u_a increases, the agent adjusts her information in such a way that her degree of certainty when choosing a falls, while her degree of certainty when choosing the other action rises. Both results follow from analyzing the system of equations described in Lemma 6.

We now extend the model to allow for the agent to receive costless signals—in the form of observation of flow payoffs—at the end of each period. How does the provision of free information affect choice behavior? We show that costless signals crowd out information acquisition. In some cases, the crowding-out effect is so strong that the agent is *less* likely to choose the optimal action with costless signals than without them. This result implies that providing free information can lower the agent's gross payoffs. Her net payoffs, however, cannot decrease: whatever loss she incurs from choosing suboptimal actions is compensated by a lower cost of information acquisition.

[14] Additional comparative statics results may be found in an earlier version of this paper (Steiner, Stewart, and Matějka, 2015).

As a benchmark, consider a static RI problem (Θ, π, u) with solution p. Suppose the agent, before acquiring information, receives a costless signal, y. We focus on the case in which y is less informative than the signal the agent acquires in the original problem; that is, we suppose that for each realization of y, the belief $\pi(\theta \mid y)$ lies in the convex hull of the posteriors arising from p.[15] Upon observing y, the agent solves the static RI problem $(\Theta, \pi(\cdot \mid y), u)$. We denote the optimal choice after observing y by $\tilde{p}(a \mid \theta, y)$. We are interested in the distribution of the agent's actions in each state; accordingly, define the *average choice rule* $\tilde{p}(a \mid \theta) := \sum_y \pi(y \mid \theta)\tilde{p}(a \mid \theta, y)$.

In this case, the average choice rule $\tilde{p}(a \mid \theta)$ is identical to $p(a \mid \theta)$; whether or not the agent receives free information has no effect on her behavior in each state. This observation follows from the result in Caplin and Dean (2013) that the optimal posteriors in static RI problems are the same across all priors within their convex hull: since the prior belief and the sets of posteriors are the same in the two problems, so is the distribution of actions in each state.

In dynamic decision problems, provision of free information has an additional effect. In choosing what information to acquire, the agent must consider its value not only in the current period but also in the future. If the agent did not expect to receive costless signals in future periods, the effect of a (not-too-strong) costless signal before the current period would be to exactly crowd out information acquisition, as in the static case. However, receiving additional signals in the future tends to lower the value of acquiring information today, leading to a *reduction* in the overall precision of information in each period.

To illustrate, suppose the realized payoff given a_t and θ_t is stochastic. More specifically, if $a_t = \theta_t$, the agent receives—and freely observes—a flow payoff of $\frac{1}{2\lambda-1}$ with probability λ and of 0 with probability $1-\lambda$, where $\lambda \in (1/2, 1]$. If $a_t \neq \theta_t$, the agent receives a flow payoff of $\frac{1}{2\lambda-1}$ with probability $1-\lambda$ and of 0 with probability λ. Flow payoffs are independent across periods conditional on the history of states. Let $u(a_t, \theta_t)$ denote the expected flow payoff given (a_t, θ_t), and note that, for every λ, the payoff difference $u(\theta, \theta) - u(1-\theta, \theta)$ is 1 for each θ.

Higher values of λ correspond to more precise costless signals. When $\lambda = 1$, the flow payoff perfectly reveals the state θ_t at the end of each period. In this case, continuation values are independent of the current action. The agent therefore acquires the same signal as if the choice for the current period were a static problem. At the

[15] Whenever $\pi(\theta \mid y)$ lies outside this convex hull, the agent acquires no additional information.

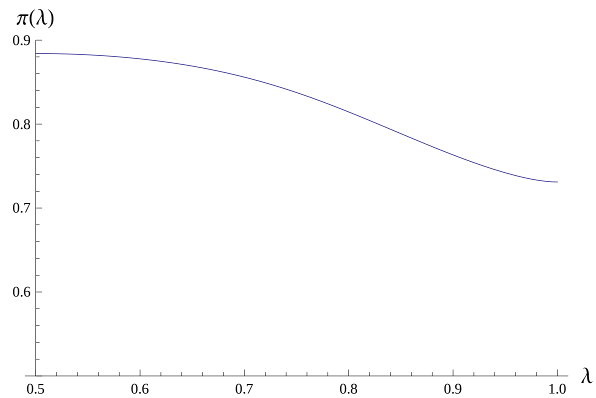

Figure 1: Precision, $\pi(\lambda)$, as a function of the informativeness of flow payoffs for discount factor $\delta = 0.9$ and transition probabilities $\gamma(0,1) = \gamma(1,0) = 0.9$.

other extreme, for λ close to $1/2$, this model approximates the one with no costless signal.

Suppose $\gamma(0,1) = \gamma(1,0)$. Symmetry implies that the solution is characterized by a *precision* $\pi(\lambda) \in (1/2, 1)$ such that, in every period t for which the agent's prior belief that $\theta_t = 0$ lies in $[1-\pi(\lambda), \pi(\lambda)]$, her posterior after observing x_t is either $\pi(\lambda)$ or $1 - \pi(\lambda)$.

From the characterization in Proposition 3, it follows that the optimal precision is the value of π that maximizes

$$\pi + H(\pi) - \delta E_y \left[H \left(\pi_p \left(\pi, y \right) \right) \right], \tag{16}$$

where $\pi_p(\pi, y)$ denotes the prior belief assigned to a given state in period $t+1$ if π is the belief in period t at the time the action is chosen, and the agent receives flow payoff y in period t. The first term in (16) captures the impact of π on the expected gross payoff in period t. The second term captures the impact on the information cost in period t. The third term captures the impact of the belief in period t on the information cost in period $t+1$ through its effect on the prior belief in the latter period.

Figure 1 depicts the optimal precision, obtained by maximizing (16) numerically. Precision decreases in the informativeness of the flow payoffs. Since the precision is equal to the probability that the agent's action matches the state in each period, the agent's gross payoff also decreases with λ.

4.3 Response times

A large body of research in psychology—and more recently in economics—has examined response times in decision-making (e.g., see Rubinstein, 2007). 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.[16]

We focus on the following simple model. The state $\theta \in \{0, 1\}$ is uniformly distributed and fixed over time. In each period $t = 1, \dots, T$, the agent chooses a terminal action 0 or 1, or waiting (denoted by w). She receives a benefit of 1 if her terminal action matches the state, and 0 otherwise. In addition, she incurs a cost $c \in (0, 1)$ for each period that she waits. Letting $w^t = (w, \dots, w)$ (t times), the agent's total gross payoff is the undiscounted sum of the flow payoffs

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

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

With the information cost function as in our general model, the solution to this problem is 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. However, some delay is 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 a given amount of time; thus delaying allows her to process more information. We view this formulation as natural for capturing perceptual experiments that take place over a short time. Accordingly, the

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

agent solves

$$\max_{f,\sigma} E\Big[\sum_{t=1}^{T} u_t\left(\sigma^t(x^t),\theta\right)\Big], \tag{17}$$

$$\text{s.t. } I\left(\theta; x_t \mid x^{t-1}\right) \leq \kappa \text{ for all } x^{t-1} \text{ and } t = 1,\dots,T,$$

where $\kappa > 0$ is the capacity constraint on the information acquired per period, $f = (f(x_t \mid x^{t-1}))_t$ is the information strategy, and $\sigma = (\sigma_t(x^t))_t$ is the action strategy. We assume that there are at least four signal realizations (i.e., $|X| \geq 4$), and that $\kappa T < \log 2$, which means that the capacity is not large enough for the agent to learn the state perfectly within the T periods.

Given a strategy for the agent, let $r_t = \Pr\left(a_t \neq w \mid a^{t-1} = w^{t-1}\right)$ and $g_t = \Pr(a_t = \theta \mid a^{t-1} = w^{t-1}$ and $a_t \neq w)$. We refer to r_t as the (hazard) *rate* and g_t as the *accuracy* of terminal actions at time t.

Proposition 6. *For each κ, there exist $\underline{c}$ and $\overline{c}$ with $\overline{c} > \underline{c} > 0$ such that if $c \in (\underline{c},\overline{c})$, then* (17) *has a solution in which the rate r_t is positive and constant across $t < T$, $r_T = 1$, and $g_t = \frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}$ for every t.*

This result indicates that the solution involves constant accuracy across periods, and a constant rate until the final period (at which point taking a terminal action is always optimal). This solution reflects two tradeoffs. First, for given rates r_t, the agent can trade off accuracy across periods: decreasing accuracy in one period frees up capacity that can be used to acquire information that improves the accuracy of future decisions. The marginal value of capacity must be equalized across periods, which occurs when accuracy is constant. The second tradeoff is between speed and accuracy. Increasing the rate of terminal actions lowers the expected waiting cost but requires a corresponding decrease in accuracy so as not to violate the capacity constraint. The optimal accuracy therefore depends on both the capacity, κ, and the waiting cost, c.

Although problem (17) does not fit directly into our general model, we show in the proof of Proposition 6 that it can be solved by transforming it into a problem that does. We first consider a relaxed problem in which capacity is storable; the agent therefore faces a cumulative capacity constraint at each decision node. Because of the additivity of mutual information, the behavior in any solution to this relaxed problem can be replicated in the original problem (although the timing of information acquisition may differ). The Lagrangian for this relaxed problem is a special case of the objective function in our general model; accordingly, we find the solution using Proposition 3.

Unlike our general model, problem (17) cannot be solved by a strategy in which signal realizations map one-to-one to actions (as in Lemma 1): with constant accuracy and binding constraints, the only way to achieve a higher rate in the final period is to acquire information even when choosing to wait. With four possible signal realizations, an optimal strategy has two realizations leading to posteriors g_t and $1 - g_t$—at which the agent chooses a terminal action—and two realizations leading to posteriors closer to $1/2$, at which the agent waits. Lemma 1 does hold in the relaxed problem with storable capacity; in that problem, any information obtained while waiting can be delayed until it is used for a terminal decision.

How do response times vary with the capacity κ? Higher values of κ can be interpreted as describing an agent with higher ability, or a decision problem that is easier to solve. Given the solution in Proposition 6 together with the fact that the constraints in (17) bind, one can compute the rate r_t numerically. We find that the rate is not monotone in the capacity: decisions are fastest when the capacity is high or low, and slowest for intermediate capacities. 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. 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 2.1

The next two lemmas are used to prove Lemma 1. The first relies on the conditional independence of x^t and y_t, the additive property of entropy, and the symmetry of mutual information.

Lemma 4. *Let $a^t = \sigma^t(z^{t-1}, x_t)$. The total discounted information cost associated with any strategy (f, σ) satisfies*

$$
\begin{aligned}
&E\left[\sum_{t=1}^{\infty} \delta^{(t)} I(\theta^t; x_t \mid z^{t-1})\right] \\
&\qquad = E\Bigg[\sum_{t=1}^{\infty} \Big(\left(-\delta^{(t)} + \delta^{(t+1)}\right) H(\theta^t \mid x^t, y^t) - \delta^{(t)} H(y_t \mid x^t, y^{t-1}) \\
&\qquad\qquad + \delta^{(t)} H(y_t \mid \theta^t, a^t, y^{t-1}) + \delta^{(t)} H(\theta_t \mid \theta^{t-1})\Big)\Bigg]. \qquad (18)
\end{aligned}
$$

Proof. Recall that

$$E\left[\sum_{t=1}^{\infty}\delta^{(t)}I(\theta^t;x_t\mid z^{t-1})\right]=E\left[\sum_{t=1}^{\infty}\delta^{(t)}\left(H(\theta^t\mid x^{t-1},y^{t-1})-H(\theta^t\mid x^t,y^{t-1})\right)\right]. \tag{19}$$

By the symmetry of mutual information, for $a^t=\sigma^t(z^{t-1},x_t)$,

$$\begin{aligned}E\left[H(\theta^t\mid x^t,y^{t-1})-H(\theta^t\mid x^t,y^t)\right]&=E\left[I(\theta^t;y_t\mid x^t,y^{t-1})\right]\\&=E\left[I(y_t;\theta^t\mid x^t,y^{t-1})\right]\\&=E\left[H(y_t\mid x^t,y^{t-1})-H(y_t\mid \theta^t,x^t,y^{t-1})\right]\\&=E\left[H(y_t\mid x^t,y^{t-1})-H(y_t\mid \theta^t,a^t,y^{t-1})\right],\end{aligned}$$

where the last step follows from the independence of x^t and y_t conditional on (θ^t,y^{t-1},a^t).

In addition, by the additive property of entropy and the independence of θ_t and (x^{t-1},y^{t-1}) conditional on θ^{t-1},

$$\begin{aligned}E\left[H(\theta^t\mid x^{t-1},y^{t-1})\right]&=E\left[H(\theta^{t-1}\mid x^{t-1},y^{t-1})+H(\theta_t\mid\theta^{t-1},x^{t-1},y^{t-1})\right]\\&=E\left[H(\theta^{t-1}\mid x^{t-1},y^{t-1})+H(\theta_t\mid\theta^{t-1})\right].\end{aligned}$$

Substituting the last two identities into the right-hand side of (19) gives

$$E\left[\sum_{t=1}^{\infty}\delta^{(t)}\left(H(\theta^{t-1}\mid x^{t-1},y^{t-1})-H(\theta^t\mid x^t,y^t)-H(y_t\mid x^t,y^{t-1})\right.\right.$$
$$\left.\left.+H(y_t\mid\theta^t,a^t,y^{t-1})+H(\theta_t\mid\theta^{t-1})\right)\right].$$

Rearranging terms gives the result. □

Lemma 5. *Let χ, ξ, and ζ be finite random variables such that ζ is measurable with respect to ξ. Then $E[H(\chi\mid\xi)]\le E[H(\chi\mid\zeta)]$.*

Proof. Since ζ is measurable with respect to ξ, $\Pr(\chi\mid\zeta)\equiv\sum_{\xi}\Pr(\chi\mid\xi)\Pr(\xi\mid\zeta)$. Thus $\Pr(\chi\mid\zeta)$ is a convex combination of the distributions $\Pr(\chi\mid\xi)$ (as ξ varies). The result follows from the concavity of entropy. □

Proof of Lemma 1. Let s be a strategy and p the choice rule generated by s. By construction, s and p give the same stream of expected gross payoffs. We claim that the information cost

$$E\left[\sum_{t=1}^{\infty} I(\theta^t; a_t \mid a^{t-1}, y^{t-1})\right]$$

associated with p is no larger than that associated with s, which, by Lemma 4, is equal to the right-hand side of (18). The information cost associated with p can be expressed in the exactly same way except with a^t in place of each x^t. These two expressions can be compared term-by-term. By Lemma 5, $E\left[H(\theta^t \mid x^t, y^t)\right] \leq E\left[H(\theta^t \mid a^t, y^t)\right]$ and $E\left[H(y_t \mid x^t, y^{t-1})\right] \leq E\left[H(y_t \mid a^t, y^{t-1})\right]$ for every t. Since $\delta^{(t+1)} \leq \delta^{(t)}$, this implies that the first two terms of the sum on the right-hand of (18) are at least as large as the corresponding terms in the expression associated with p. Since the last two terms of the sum are identical in the two cases, the claim follows.

We have shown that the discounted expected payoff from any strategy s is no larger than the value of the objective function in (3) given the choice rule generated by s. Conversely, the discounted expected payoff from any strategy induced by a choice rule p is identical to the value of the objective function in (3) given p. Together, these two relationships imply the result. □

Proof of Proposition 1. Consider the space of strategies $\Pi = \prod_t \Delta(A)^{\Theta^t \times A^{t-1} \times Y^{t-1}}$. 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. □

B Proofs for Section 3

Proof of Lemma 3. Given q, let $v_t(a_t, \omega^t) = u_t(a^t, \theta^t) + \log q_t(a_t \mid a^{t-1}, y^{t-1})$ for all $\omega^t = (\theta^t, a^{t-1}, y^{t-1})$. For each ω^t such that $\Pr(\omega^t) > 0$ (where the probability is with respect to π, q, and g), let

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

thus $V_t(\omega^t)$ is the continuation value in the control problem for q. If $\Pr(\omega^t) = 0$, we define $V_t(\omega^t)$ arbitrarily.

When $\Pr(\omega^t) > 0$, the value V_t satisfies the recursion

$$V_t(\omega^t) = \max_{p_t(\cdot|\omega^t)} E\left[v_t(a_t, \omega^t) - \log p_t(a_t \mid \omega^t) + \delta_{t+1} V_{t+1}(\omega^{t+1}) \mid \omega^t\right] \tag{20}$$

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

To solve the maximization problem in (20), note first that, since $v_t(a_t, \omega^t) = u_t(a^t, \theta^t) + \log q_t(a_t \mid z^{t-1})$ (for $z^{t-1} = (a^{t-1}, y^{t-1})$), if $q_t(a_t \mid z^{t-1}) = 0$—and hence $\log q_t(a_t \mid z^{t-1}) = -\infty$—for some a_t, then we must have $p_t\left(a_t \mid (\theta^{t-1}, \theta_t), z^{t-1}\right) = 0$ for every θ_t satisfying $\pi(\theta^{t-1}, \theta_t) > 0$.[17] Accordingly, let $A(z^{t-1}) = \{a_t \in A : q_t(a_t \mid z^{t-1}) > 0\}$, and suppose $a_t \in A(z^{t-1})$ and $\pi(\theta^{t-1}, \theta_t) > 0$. If $A(z^{t-1})$ is a singleton, then $p_t\left(a_t \mid (\theta^{t-1}, \theta_t), z^{t-1}\right) = 1$. Otherwise, the first-order condition for the optimization problem in (20) with respect to $p_t(a_t \mid \omega^t)$ is

$$v_t(a_t, \omega^t) - \left(\log p_t(a_t \mid \omega^t) + 1\right) + \delta_{t+1} E\left[V_{t+1}\left(\omega^{t+1}\right) \mid \omega^t, a_t\right] = \mu_t(\omega^t), \tag{21}$$

where $\mu_t(\omega^t)$ is the Lagrange multiplier associated with the constraint $\sum_{a_t'} p_t(a_t' \mid \omega^t) = 1$.

Rearranging the first-order condition gives

$$p_t(a_t \mid \omega^t) = \exp\left(v_t(a_t, \omega^t) - 1 + \delta_{t+1} \overline{V}_{t+1}\left(a_t, \omega^t\right) - \mu_t(\omega^t)\right),$$

where $\overline{V}_{t+1}\left(a_t, \omega^t\right) := E\left[V_{t+1}\left(\omega^{t+1}\right) \mid \omega^t, a_t\right]$. Since $\sum_{a_t' \in A(z^{t-1})} p_t(a_t' \mid \omega^t) = 1$, it follows that

$$\begin{aligned} p_t(a_t \mid \omega^t) &= \frac{\exp\left(v_t(a_t, \omega^t) - 1 + \delta_{t+1} \overline{V}_{t+1}\left(a_t, \omega^t\right) - \mu_t(\omega^t)\right)}{\sum_{a_t' \in A(z^{t-1})} \exp\left(v_t(a_t', \omega^t) - 1 + \delta_{t+1} \overline{V}_{t+1}\left(a_t', \omega^t\right) - \mu_t(\omega^t)\right)} \\ &= \frac{\exp\left(v_t(a_t, \omega^t) + \delta_{t+1} \overline{V}_{t+1}\left(a_t, \omega^t\right)\right)}{\sum_{a_t' \in A(z^{t-1})} \exp\left(v_t(a_t', \omega^t) + \delta_{t+1} \overline{V}_{t+1}\left(a_t', \omega^t\right)\right)}. \end{aligned}$$

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

Substituting into (20) gives the recursion

$$\overline{V}_t(a_{t-1}, \omega^{t-1}) = E\Bigg[- \delta_{t+1}\overline{V}_{t+1}(a_t, \omega^t)$$

$$+ \log\left(\sum_{a'_t \in A(z^{t-1})} \exp\left(v_t(a'_t, \omega^t) + \delta_{t+1}\overline{V}_{t+1}(a'_t, \omega^t)\right) \right) + \delta_{t+1}\overline{V}_{t+1}(a_t, \omega^t) \,\Bigg|\, \omega^{t-1}, a_{t-1} \Bigg],$$

and therefore,

$$\overline{V}_t(a_{t-1}, \omega^{t-1})$$
$$= E\left[\log\left(\sum_{a'_t \in A(z^{t-1})} \exp\left(v_t(a'_t, \omega^t) + \delta_{t+1}\overline{V}_{t+1}(a'_t, \omega^t)\right) \right) \,\middle|\, \omega^{t-1}, a_{t-1} \right]$$
$$= E\left[\log\left(\sum_{a'_t \in A} q(a'_t | z^{t-1}) \exp\left(u_t((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}\overline{V}_{t+1}(a'_t, \omega^t)\right) \right) \,\middle|\, \omega^{t-1}, a_{t-1} \right],$$

as needed. □

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

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

$$\max_{q_t(\cdot|z^{t-1}), \{p_t(\cdot|\theta^t, z^{t-1})\}_{\theta^t}} E\left[\hat{u}_t\left(a_t, \theta^t; z^{t-1}\right) + \log q_t\left(a_t \mid z^{t-1}\right) - \log p_t\left(a_t \mid \theta^t, z^{t-1}\right) \mid z^{t-1}\right], \tag{22}$$

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

By Lemma 2, it suffices to prove that any solution of the control problem (problem

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

$$p_t(a_t \mid \theta^t, z^{t-1}) = \frac{q_t(a_t \mid z^{t-1}) \exp\left(\hat{u}(a_t, \theta^t; z^{t-1})\right)}{\sum_{a'_t} q_t(a'_t \mid z^{t-1}) \exp\left(\hat{u}(a'_t, \theta^t; z^{t-1})\right)}. \tag{23}$$

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

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

We now extend Proposition 3 to cases in which the solution to the dynamic RI problem is not interior. To do this, we must ensure that prior beliefs in the static problems are defined appropriately to generate the correct continuation values. We define the posterior belief in a static RI problem after an action a is taken with zero probability to be

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

This expression coincides with (9) when a is chosen with positive probability. Otherwise, it differs from (9) only by a renormalization. The main idea of the proof is to add an additional constraint placing a lower bound on every $q_t(a_t \mid z^{t-1})$ in both the original control problem and the control problem at z^{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.

Proposition 7. *There exists a dynamic choice rule p solving the dynamic RI problem such that, at each decision node z^{t-1}, $p_t(a_t \mid \theta^t, z^{t-1})$ solves the static RI problem with state space Θ^t, prior belief $\pi_p(\theta^t \mid z^{t-1})$ satisfying* (10)*, and payoff function $\hat{u}_t$ given by* (11)*, where the posterior belief $\pi^p(\theta^t \mid a_t, z^{t-1})$ formed after taking action a_t at the decision node z^{t-1} complies with* (24)*, and the continuation values satisfy* (12)*.*

Proof. Consider, for $\varepsilon \in (0, 1/|A|)$, the *$\varepsilon$-control problem*, which is identical to the original control problem (problem (14)) except that for each a_t and z^{t-1}, there is a constraint that $q_t(a_t \mid z^{t-1}) \geq \varepsilon$. Define the *$\varepsilon$-control problem at z^{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 z^{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 z^{t-1} provided that the continuation values and priors approach those described in the proposition as ε vanishes.

For each $\omega^t = (\theta^t, a^{t-1}, y^{t-1})$ and ε, let $V_t^\varepsilon(\omega^t)$ denote the continuation value in the ε-control problem. Consider the ε-control problem at $z^{t-1} = (a^{t-1}, y^{t-1})$, and write π_p 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

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

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

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

$$\sum_{\theta^t} \frac{\pi_p(\theta^t) \exp \hat{u}_\varepsilon(a_t, \omega^t)}{\sum_{a_t'} q_t(a_t' \mid z^{t-1}) \exp \hat{u}_\varepsilon(a_t', \omega^t)} = \mu, \tag{25}$$

where μ is the Lagrange multiplier associated with the constraint $\sum_{a_t'} q_t(a_t' \mid z^{t-1}) = 1$. Note that there must exist some a_t for which $q_t(a_t \mid z^{t-1}) \in (\varepsilon, 1)$. For this action a_t, we have $\overline{p}_t(a_t \mid z^{t-1}) = q_t(a_t \mid z^{t-1})$, 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 $q_t(a_t \mid z^{t-1}) = \varepsilon$ (if such an a_t exists).

Then we must have

$$\sum_{\theta^t} \frac{\pi_p(\theta^t) \exp \hat{u}_\varepsilon(a_t, \omega^t)}{\sum_{a'_t} q_t(a'_t \mid z^{t-1}) \exp \hat{u}_\varepsilon(a'_t, \omega^t)} \leq \mu = 1.$$

In this case, the posterior beliefs satisfy

$$\begin{aligned}
\pi^p(\theta^t \mid a_t, z^{t-1}) &= \frac{\pi_p(\theta^t)}{\overline{p}_t(a_t \mid z^{t-1})} p_t(a_t \mid \omega^t) \\
&= \frac{q_t(a_t \mid z^{t-1})}{\overline{p}_t(a_t \mid z^{t-1})} \frac{\pi_p(\theta^t) \exp(\hat{u}_\varepsilon(a_t, \omega^t))}{\sum_{a'_t} q_t(a'_t \mid z^{t-1}) \exp(\hat{u}_\varepsilon(a'_t, \omega^t))} \\
&= \frac{1}{\sum_{\tilde{\theta}^t} \pi_p(\tilde{\theta}^t) \frac{\exp(\hat{u}_\varepsilon(a_t, \omega^t))}{\sum_{a'_t} q_t(a'_t \mid z^{t-1}) \exp(\hat{u}_\varepsilon(a'_t, \tilde{\omega}^t))}} \frac{\pi_p(\theta^t) \exp(\hat{u}_\varepsilon(a_t, \omega^t))}{\sum_{a'_t} q_t(a'_t \mid z^{t-1}) \exp(\hat{u}_\varepsilon(a'_t, \omega^t))},
\end{aligned}$$

where $\tilde{\omega}^t = (\tilde{\theta}^t, a^{t-1}, y^{t-1})$. Therefore, as ε vanishes, the posteriors indeed approach those given by (24). □

C Proofs and computations for Section 4

C.1 Status quo bias

Proof of Proposition 4. By symmetry, the predispositions in the first period are given by $q_1(0) = q_1(1) = 1/2$, and in the second period, $q_2(0 \mid 0) = q_2(1 \mid 1)$. Denote the latter by s. Also by symmetry, the continuation value function attains only two values:

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

where V_c is the expected payoff in period 2 (including the information cost) when $a_1 = \theta_2$, and V_w is the corresponding payoff when $a_1 \neq \theta_2$. By Theorem 1 and (4), the continuation values satisfy $V_w = \log(s + (1-s)e)$ and $V_c = \log(se + (1-s))$.

Applying Theorem 1 and (4) together with symmetry in the first period gives the expected payoff

$$\log\left(\frac{1}{2}\exp\left(1 + (1-\gamma)V_c + \gamma V_w\right) + \frac{1}{2}\exp\left((1-\gamma)V_w + \gamma V_c\right)\right),$$

where the predisposition s maximizes this expression (subject to $0 \leq s \leq 1$). This is equivalent to maximizing

$$\begin{aligned} W(s;\gamma) &:= \exp\left(1 + (1-\gamma)V_c + \gamma V_w\right) + \exp\left((1-\gamma)V_w + \gamma V_c\right) \\ &= e(se + (1-s))^{1-\gamma}(s + (1-s)e)^{\gamma} + (s + (1-s)e)^{1-\gamma}(se + (1-s))^{\gamma}. \end{aligned}$$

Note that, because $\gamma \in (0, 1/2)$, $W(s;\gamma) \geq W(1-s;\gamma)$ for $s > 1/2$, and thus the maximand is at least $1/2$.

It is straightforward to verify that $W(s;\gamma)$ is concave in s, and its derivative (with respect to s) at $s = 1$ is positive when $\gamma = 0$. It follows that the optimal value of s is 1 when $\gamma = 0$, and therefore, by continuity, there exists $\gamma^* \in (0, 1/2)$ such that this is true whenever $\gamma < \gamma^*$. □

C.2 Inertia

Lemma 6. *Suppose there is an eventually interior solution. 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 $p_t(a_t \mid \theta^t, a^{t-1}) \equiv \hat{p}(a_t \mid \theta_t, a_{t-1})$ in each period $t > t'$, where*

$$\hat{p}(a_t \mid \theta_t, a_{t-1}) = \frac{\hat{q}(a_t \mid a_{t-1}) \exp\left(u(a_t, \theta_t) + \delta E\left[V(a_t, \theta_{t+1}) \mid \theta_t\right]\right)}{\sum_{a'_t} \hat{q}(a'_t \mid a_{t-1}) \exp\left(u(a'_t, \theta_t) + \delta E\left[V(a'_t, \theta_{t+1}) \mid \theta_t\right]\right)}, \tag{26}$$

where the continuation payoffs solve

$$V(a_{t-1}, \theta_t) = \log\left(\sum_{a_t} \hat{q}(a_t \mid a_{t-1}) \exp\left(u(a_t, \theta_t) + \delta E\left[V(a_t, \theta_{t+1}) \mid \theta_t\right]\right)\right), \tag{27}$$

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

$$\sum_{\theta_{t-1}} \pi^p(\theta_{t-1} \mid a_{t-1})\gamma(\theta_{t-1}, \theta_t) = \sum_{a_t} \hat{q}(a_t \mid a_{t-1})\pi^p(\theta_t \mid a_t) \tag{28}$$

for all θ_t and a_{t-1}, and the posteriors $\hat{\pi}^p(\theta_t \mid a_t) \equiv \sum_{\theta^{t-1}} \pi^p((\theta^{t-1}, \theta_t) \mid (a^{t-1}, a_t))$ satisfy

$$\frac{\hat{\pi}^p(\theta_t \mid a_t)}{\hat{\pi}^p(\theta_t \mid a'_t)} = \frac{\exp\left(u(a_t, \theta_t) + \delta E[V(a_t, \theta_{t+1}) \mid \theta_t]\right)}{\exp\left(u(a'_t, \theta_t) + \delta E[V(a'_t, \theta_{t+1}) \mid \theta_t]\right)}. \tag{29}$$

Proof. The result follows directly from Proposition 3. It suffices to verify that $\hat{p}(a_t \mid \theta_t, a_{t-1})$ and $\hat{q}(a_t \mid a_{t-1})$ solve the static RI problem at each a_{t-1}. □

Proof of Proposition 5. For each $a \in \{0,1\}$, let π^a denote the posterior $\hat{\pi}^p(\theta_t = 1 \mid a_t = a)$, and π_a the prior $\hat{\pi}_p(\theta_t = 1 \mid a_{t-1} = a) \equiv \sum_{\theta^{t-1}} \pi_p((\theta^{t-1}, 1) \mid (a^{t-2}, a))$ associated with the stationary solution when t is sufficiently large. Since the expected posterior is equal to the prior, we have $\hat{q}(a|a)\pi^a + \hat{q}(1-a|a)\pi^{1-a} = \pi_a$ for each $a \in \{0,1\}$. Together with $\hat{q}(a \mid a) + \hat{q}(1-a \mid a) = 1$, these two equations imply

$$\hat{q}(a \mid a) - \hat{q}(a \mid 1-a) = \frac{\pi_a - \pi_{1-a}}{\pi^a - \pi^{1-a}}$$

for each a. Substituting $\pi_a = \pi^a\gamma(1,1) + (1-\pi^a)\gamma(0,1)$ and the analogous expression for π_{1-a} leads to $\hat{q}(a \mid a) - \hat{q}(a \mid 1-a) = \gamma(1,1) - \gamma(0,1) = \gamma(1,1) + \gamma(0,0) - 1$, which is positive since states have positive persistence. That $\hat{p}(a \mid \theta, a) > \hat{p}(a \mid \theta, 1-a)$ for each a follows from this last result together with (26).

For part 2, suppose without loss of generality that $a = 1$. Consider a static RI problem with $\Theta = A = \{0,1\}$ and payoffs $u(a,\theta)$ satisfying $u(a,a) \equiv u_a > 0$ and $u(1-a,a) = 0$ for each $a \in \{0,1\}$. Suppose moreover that the solution is interior. By (9),

$$\frac{\pi_s^1}{\pi_s^0} = \exp u_1 \quad \text{and} \quad \frac{1-\pi_s^1}{1-\pi_s^0} = \exp(-u_0),$$

where π_s^a denotes the posterior $\pi^p(\theta = 1 \mid a)$. It is straightforward to verify that the posteriors π_s^1 and π_s^0 solving these two equations decrease in u_1 and increase in u_0.

Now consider the dynamic problem. By Proposition 3 and Lemma 6, the solution is the same as in a static RI problem with payoffs $\hat{u}(a_t, \theta_t) = u_{a_t} 1_{a_t = \theta_t} + \delta E\left[V(a_t, \theta_{t+1}; u_1) \mid \theta_t\right]$, where $V(a_t, \theta_{t+1}; u_1)$ solves (27) (given u_1). By the previous paragraph, it suffices to prove that the payoff difference $\hat{u}(1,\theta) - \hat{u}(0,\theta)$ increases in u_1 for each $\theta \in \{0,1\}$, which follows if $V(1,\theta;u_1) - V(0,\theta;u_1)$ increases in u_1 for each $\theta \in \{0,1\}$.

Differentiating (27) gives

$$\frac{\partial}{\partial u_1} V(a_{t-1}, \theta_t; u_1) = \hat{p}(1 \mid \theta_t, a_{t-1}) + \delta E\left[\frac{\partial}{\partial u_1} V(a_t, \theta_{t+1}; u_1) \mid a_{t-1}, \theta_t\right]. \qquad (30)$$

Letting $d(a_t, \theta_t) := \hat{p}(a_t \mid \theta_t, 1) - \hat{p}(a_t \mid \theta_t, 0)$ and $\Delta(\theta_t; u_1) := V(1, \theta_t; u_1) - V(0, \theta_t; u_1)$,

(30) implies

$$\begin{aligned}\frac{\partial}{\partial u_1}\Delta\left(\theta_t;u_1\right) &= d(1,\theta_t)+\delta E\left[d(1,\theta_t)\frac{\partial}{\partial u_1}V\left(1,\theta_{t+1};u_1\right)\mid\theta_t\right] \\ &\qquad\qquad + \delta E\left[d(0,\theta_t)\frac{\partial}{\partial u_1}V\left(0,\theta_{t+1};u_1\right)\mid\theta_t\right] \\ &= d(1,\theta_t)+\delta E\left[d(1,\theta_t)\frac{\partial}{\partial u_1}\Delta\left(\theta_{t+1};u_1\right)\mid\theta_t\right],\end{aligned}$$

where the last equality follows from the identity $d(a,\theta_t)\equiv -d(1-a,\theta_t)$. Iterating gives

$$\frac{\partial}{\partial u_1}\Delta\left(\theta_t;u_1\right)=\sum_{t'=t}^{\infty}\delta^{t'-t}E\left[\prod_{t''=t}^{t'}d(1,\theta_{t''})\mid\theta_t\right].$$

By part 1 of the proposition, $d(1,\theta_t)>0$ for each θ_t, and hence $\frac{\partial}{\partial u_1}\Delta\left(\theta_t;u_1\right)>0$. □

C.3 Response times

Proof of Proposition 6. Since actions following any $a^{t-1}\neq w^{t-1}$ are payoff-irrelevant, the problem has a solution in which the agent acquires information only before the first terminal action; that is, if $\sigma^{t-1}(x^{t-1})\neq w^{t-1}$ then $I(\theta;x_t\mid x^{t-1})=0$. We restrict attention to solutions of this form.

Note that the system of constraints in (17) together with the preceding paragraph imply

$$E\left[I\left(\theta;x_t\mid x^{t-1}\right)\right]\leq\kappa\Pr\left(\sigma^{t-1}\left(x^{t-1}\right)=w^{t-1}\right)\text{ for all }t=1,\ldots,T.$$

Taking partial sums gives

$$\sum_{\tau=1}^{t}E\left[I\left(\theta;x_\tau\mid x^{\tau-1}\right)\right]\leq\kappa\sum_{\tau=1}^{t}\Pr\left(\sigma^{\tau-1}\left(x^{\tau-1}\right)=w^{\tau-1}\right)\text{ for all }t=1,\ldots,T.$$

In addition, if $a^t=\sigma^t(x^t)$, then

$$\sum_{\tau=1}^{t}E\left[I\left(\theta;a_\tau\mid a^{\tau-1}\right)\right]=I(\theta;a^t)\leq I(\theta;x^t)=\sum_{\tau=1}^{t}E\left[I\left(\theta;x_\tau\mid x^{\tau-1}\right)\right].$$

Therefore, the value of the problem

$$\max_p E\Big[\sum_{t=1}^{T} u_t\left(a^t,\theta\right)\Big] \tag{31}$$

$$\text{s.t.}\ \sum_{\tau=1}^{t} E\left[I\left(\theta;a_\tau \mid a^{\tau-1}\right)\right] \leq \kappa \sum_{\tau=1}^{t} \Pr\left(a^{\tau-1}=w^{\tau-1}\right)\ \text{ for all } t=1,\ldots,T$$

is an upper bound on the value of problem (17). Thus if we find a choice rule p^* solving (31), and construct a strategy (f,σ) feasible in (17) that generates p^*, then (f,σ) solves (17).

Since the set of choice rules satisfying the constraints in (31) is convex and the objective is linear in p, the first-order conditions are sufficient for a global optimum. For each t, let $\tilde{\lambda}_t \geq 0$ denote the shadow price of the constraint in (31) for t. Consider the problem

$$\max_p E\Big[\sum_{t=1}^{T}\left(u_t(a^t,\theta)+\lambda_{t+1}\kappa\mathbf{1}_{a^t=w^t}-\lambda_t I(\theta;a_t\mid a^{t-1})\right)\Big], \tag{32}$$

where $\lambda_{T+1}=0$, and $\lambda_t=\sum_{\tau=t}^{T}\tilde{\lambda}_\tau$ for each $t=1,\ldots,T$. We will use Proposition 3 to find a solution to (32) with $\lambda_1=\lambda_2=\cdots=\lambda_T=c/\kappa$, and then show that, for a range of values of c, this solution satisfies the first-order conditions for (31).

The only non-trivial decision node in each period t is $a^{t-1}=w^{t-1}$. Each of these nodes is associated with a unique belief about θ, which, by symmetry, is the uniform belief. Symmetry also implies that, for each t, the continuation value $V_t(w^{t-1},\theta)$ is independent of θ; accordingly, we omit the arguments of V_t. Multiplying the objective by κ/c to eliminate the λ_t coefficient on the mutual information term and applying Proposition 3 implies that, at each node w^{t-1}, the solution corresponds to that of the static RI problem with a uniform prior over θ and payoffs

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

where $V_{T+1}:=0$. Note that the last expression simplifies to $\hat{u}_t(w,\theta)=\kappa V_{t+1}/c$.

We solve this static RI problem using Proposition 2. By symmetry, the rate r_t

satisfies $r_t/2 = q_t(0 \mid w^{t-1}) = q_t(1 \mid w^{t-1})$ for each t. By (7), the accuracy satisfies $g_t = \frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}$ for each t.

Since action w is dominated at T by a uniform mixture of 0 and 1, $r_T = 1$. By (12), the value associated with the static RI problem at time T (including the rescaling by κ/c) is $\kappa V_T/c = \log\left(\frac{1}{2}\left(\exp(\kappa/c)+1\right)\right)$. Proposition 2 implies that, for each t, r_t solves

$$\max_{r_t \in [0,1]} \log\left(\frac{r_t}{2}\left(\exp(\kappa/c)+1\right) + (1-r_t)\exp(\kappa V_{t+1}/c)\right). \tag{33}$$

This problem is solved by any $r_t \in [0,1]$ if and only if $\kappa V_{t+1}/c = \log\left(\frac{1}{2}\left(\exp(\kappa/c)+1\right)\right)$, in which case $V_t = V_{t+1}$. Proceeding recursively from period T back to period 1, it follows that, for each $r \in (0,1)$, there is a solution with $r_1 = \cdots = r_{T-1} = r$.

For this to solve (31), r must be such that the constraints are satisfied. Note that, since $g_t = \frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}$ for each t,

$$\sum_{\tau=1}^{T} E\left[I\left(\theta; a_\tau \mid a^{\tau-1}\right)\right] = I(\theta; a^T) = h\left(1/2\right) - h\left(\frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}\right),$$

where $h(p) := -p\log p - (1-p)\log(1-p)$. Since the constraint for period T in (31) is binding, r must satisfy

$$h\left(1/2\right) - h\left(\frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}\right) = \kappa \sum_{\tau=1}^{T}(1-r)^{\tau-1}. \tag{34}$$

The left-hand side of this equation is decreasing in c, while the right-hand side ranges from κ to κT as r ranges from 1 to 0. By assumption, $h(1/2) > \kappa T$, and hence there exist $\underline{c}$ and $\overline{c}$ (whose values depend on κ) such that (34) has a solution $r \in (0,1)$ whenever $c \in (\underline{c}, \overline{c})$. Since g_t is constant, $I(\theta; a^t) = \Pr(a^t \neq w^t)\kappa \sum_{\tau=1}^{T}(1-r)^{\tau-1}$ for each t, from which it is straightforward to verify that the constraints in (31) hold for each period $t < T$. In addition, because $\lambda_t = \lambda_T$ for each $t < T$, the shadow price $\tilde{\lambda}_t$ of the constraint for t is 0, and that for the binding constraint at T is positive. Therefore, the choice rule corresponding to $g_t \equiv \frac{\exp(\kappa/c)}{1+\exp(\kappa/c)}$ and $r_t \equiv r$ satisfying (34) solves (31).

All that remains is to construct a strategy (f, σ) satisfying the system of constraints in (17) that generates this choice rule. Without loss of generality, let 0, 1, w_0, and w_1 be distinct elements of X, and let $W = \{w_0, w_1\}$. Let σ satisfy

$\sigma_t(x^t) = x_t$ if $x_t \in \{0,1\}$ and $\sigma_t(x^t) = w$ if $x_t \in W$. Let the information strategy f satisfy $\Pr(x_t = 0 \mid x^{t-1} \in W^{t-1}) = \Pr(x_t = 1 \mid x^{t-1} \in W^{t-1}) = r_t/2$ and $\Pr(x_t = w_0 \mid x^{t-1} \in W^{t-1}) = \Pr(x_t = w_1 \mid x^{t-1} \in W^{t-1}) = (1-r_t)/2$ for every t, and generate the following posteriors: if $x^{t-1} \in W^{t-1}$ and $x_t = a \in \{0,1\}$ then $\Pr(\theta = a \mid x^t) = g$; and if $x^{t-1} \in W^{t-1}$ and $x_t = w_a$ for $a \in \{0,1\}$ then $\Pr(\theta = a \mid x^t) = \tilde{g}_t$, where $\tilde{g}_t$ satisfies

$$h\left(1/2\right) - h(\tilde{g}_t) = \kappa \sum_{\tau=T-t+1}^{T} (1-r)^{\tau-1}. \tag{35}$$

Since $\tilde{g}_t$ is increasing and smaller than g for all $t < T$, there exists an information strategy generating these posteriors. Equation (35) is equivalent to

$$\left(h\left(1/2\right) - h\left(\tilde{g}_t\right)\right)(1-r)^t + \left(1-(1-r)^t\right)\kappa \sum_{\tau=1}^{T}(1-r)^{\tau-1} = \kappa \sum_{\tau=1}^{t}(1-r)^{\tau-1}.$$

For $t = 1$, this together with (34) implies that the capacity constraint is binding. Proceeding inductively, the constraint binds at every w^t. □

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.

Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy 120*(6), 1173–1219.

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

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

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.

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

Maćkowiak, B. and M. Wiederholt (2015). Business cycle dynamics under rational inattention. *Review of Economic Studies 82*(4), 1502–1532.

Maćkowiak, B. A., F. Matějka, and M. Wiederholt (2016). The rational inattention filter. CEPR Discussion Paper 11237.

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

Matějka, F. (2015). Rationally inattentive seller: Sales and discrete pricing. Forthcoming in the *Review of Economic Studies*.

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.

Samuelson, W. and R. Zeckhauser (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty 1*(1), 7–59.

Schiraldi, P. (2011). Automobile replacement: a dynamic structural approach. *The RAND Journal of Economics 42*(2), 266–291.

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.

Steiner, J., C. Stewart, and F. Matějka (2015). Rational inattention dynamics: Inertia and delay in decision-making. CEPR Discussion Paper 10720.

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.

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

Woodford, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics 56*, 100–124.

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