Rational inattention dynamics

Citation for published version:
Steiner, J, Stewart, C & Matejka, F 2016 ‘Rational inattention dynamics: Inertia and delay in decision-making’ Edinburgh School of Economics, University of Edinburgh.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Rational Inattention Dynamics: Inertia and Delay in Decision-Making*

Jakub Steiner†  Colin Stewart
CERGE-EI and University of Edinburgh  University of Toronto
Filip Matějka
CERGE-EI

December 22, 2014

PRELIMINARY AND INCOMPLETE

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. Optimal behavior resembles that in a dynamic logit model. The distribution of actions at each decision node depends on the expected payoffs (not including information costs) and on a system of endogenous switching costs that do not depend on the realized history of states. Ex ante unlikely action paths are associated with high virtual switching costs, reflecting that the agent undertakes them only if she receives a high volume of costly information contradicting her prior. We apply the general solution to the study of (i) the sunk-cost fallacy; (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, Jakub Kastl, Philipp Kircher, Alessandro Pavan, Balázs Szentes and participants in seminars at CERGE-EI, Edinburgh, IESE, and Surrey university for their comments. Maxim Goryunov provided excellent research assistance.

†email: jakub.steiner@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 both what and how much information to acquire. In each period, the agent can choose an arbitrary signal about a payoff-relevant state of the world before taking an action. The state follows an arbitrary stochastic process, and the agent’s flow payoff is a function of the histories of actions and states. Following Sims (2003), the agent pays a cost to acquire information that is proportional to the reduction in her uncertainty as measured by the entropy of her beliefs. We characterize the stochastic behavior that maximizes the sum of the agent’s expected discounted utilities less the cost of the information she acquires.

We find that the optimal choice rule coincides with dynamic logit behavior (Rust, 1987) with respect to payoffs that differ from the agent’s true payoffs by an endogenous additive term. This additional term, which we refer to as a predisposition, depends on the history of actions but does not depend directly on the states. Relative to dynamic logit behavior with the agent’s true payoffs, the predisposition tends to increase the probability assigned to actions that perform well on average across all states of the world given the history of actions up to that point. For example, if states are positively serially correlated, the influence of predispositions can resemble switching costs; because learning whether the state has changed is costly, the agent relies in part on her past behavior to inform her current decision, and is therefore predisposed toward repeating her previous action.

Our results provide a new foundation for the use of dynamic logit in empirical research with an important caveat: the presence of predispositions affects extrapolation of behavior based on identification of utility parameters from data. An econometrician applying standard dynamic logit techniques to the agent in our model would correctly identify her behavior in repetitions of the same 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.

\footnote{This result extends the static logit result of Matějka and McKay (2014) and Caplin and Dean (2013b) to the dynamic setting.}
As a tool for solving the model, we characterize the optimal behavior as the solution of a collection of interconnected static rational inattention problems. These static problems resemble those corresponding to a Bellman equation in that they involve utilities given by flow payoffs plus continuation values; however, they differ in that the continuation values for each action are fixed as we vary the agent’s information. This difference is what allows us to use static rational inattention techniques: the static problems that arise under a standard dynamic programming approach do not fit into the rational inattention framework.

The key step behind the reduction to static problems is to show that the dynamic rational inattention problem can be reformulated as a control problem with observable states. In the control problem, the agent first chooses her predispositions at the ex ante stage. Then, after observing the realized state in each period, she chooses her distribution of actions, and incurs a cost according to how much she deviates from her chosen predisposition. The control problem is simpler than the original problem insofar as it does not require updating of beliefs. This feature allows us to optimize at each stage without accounting for the effect on subsequent beliefs.

We illustrate the general solution in three applications. In the first, the agent seeks to match her action to the state in each of two periods. We show that positive correlation between the states can lead to an apparent sunk cost fallacy: 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 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 her incentive to acquire information in the second.

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

\(^2\)Mattsson and Weibull (2002) study essentially the same problem for fixed predispositions.

\(^3\)As in Baliga and Ely (2011) agent derives information from her own earlier decision, though in our setting the effect arises due to costly information acquisition rather than forgetting.
stronger than the incentive to invest when it is good, then lags in adjustment are shorter for bad shocks than for good ones.

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

We focus throughout the paper on information costs that are proportional to the reduction in entropy of beliefs. There are two related reasons for this choice. The first is tractability. With entropy-based costs, it is not necessary to consider all possible signals; we can restrict attention to signals that associate at most one signal to each action. In particular, each action history is associated with a unique belief, thereby avoiding a substantial complication that arises from tracking beliefs in solving many dynamic models with exogenous information (such as hidden Markov models). Entropy-based costs are also important for the control problem reformulation, and hence for the reduction to static rational inattention problems.

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


---

4In the static case, this property holds under much weaker conditions on the cost function; see the discussion in Section 2.2.
consumption problems with RI. Each of these papers focuses either on an environment involving linear-quadratic payoffs and Gaussian shocks or on numerical solutions. A notable exception is Ravid (2014), who analyzes a class of RI stopping problems motivated by dynamic bargaining. In general static RI models, Matějka and McKay (2014) and Caplin and Dean (2013b) 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. This difference accounts for the additional payoff term in our dynamic logit result.

Work on dynamic information acquisition outside of the RI literature includes 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.

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 the RI problem is equivalent to a two-stage optimization problem that combines Mattsson and Weibull’s control problem with optimization of the default distribution. In this sense, RI can be viewed as capturing a dual process involving the interaction between heuristic and analytical decision making.\footnote{Tversky and Kahneman (1974) and Evans (1984) are seminal psychological studies of dual decision processes.}

2 Model

A single agent chooses an action \(a_t\) from a finite set \(A\) in each period \(t = 1, 2, \ldots\). We denote the action history \((a_1, \ldots, a_t)\) by \(a^t\), and refer to \(a^{t-1}\) as the decision node at \(t\). A payoff-relevant state \(\theta_t\) follows a stochastic process on a finite set \(\Theta\) with probability
measure $\pi \in \Delta(\mathbb{N}^\mathbb{N})$. Before choosing an action in any period $t$, the agent can acquire costly information about the history of states to date. There is a fixed signal space $X$ satisfying $|A| \leq |X| < \infty$. At time $t$, the agent can choose any signal about the history $\theta^t$ with realizations in $X$. Accordingly, a strategy $s = (f, \sigma)$ is a pair of

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

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

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

The agent receives flow utilities $u_t(a_t, \theta^t)$ that 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 rational inattention literature, we focus throughout this paper on entropy-based information costs. Consider a random variable $Y$ with finite support $S$ distributed according to $p(y) \in \Delta(S)$. Recall that the entropy

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

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

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

between $x_t$ and the history of states $\theta^t$.\footnote{When $x^t$ is attained with 0 probability, the value of $H(\theta^t \mid x^t)$ is chosen arbitrarily.} The conditional mutual information captures the difference in the agent’s uncertainty about $\theta^t$ before and after she receives the signal $x_t$. Before, her uncertainty can be measured by $H(\theta^t \mid x^{t-1})$. After, her uncertainty is
changed to $H(\theta^t | x^t)$. The mutual information is the expected reduction in uncertainty averaged across all realizations of $x_t$.

The agent solves the following problem.

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

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

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

To simplify notation, we normalize the information cost parameter $\lambda$ to 1.\(^7\)

The objective in (2) is well defined because the gross flow payoffs are bounded, and the mutual information is bounded (since the signal space is finite). Therefore, the sum converges.

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

### 2.1 Applications

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

**Example 1** (Sunk cost fallacy). The agent chooses an action $a_t \in \{0, 1\}$ in each period $t = 1, 2$. In both periods, the gross flow payoff $u_t$ is 1 if the action $a_t$ matches the current

---

\(^7\)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).
state \( \theta_t \in \{0, 1\} \), and is 0 otherwise. The two states are correlated across periods. There is no discounting. The agent therefore maximizes the sum of the expected gross flow payoffs less the information cost.

In this setting, we analyze the correlation between choices in the two periods. In particular, if the agent chooses not to acquire any information in the second period, then her behavior can exhibit an apparent sunk cost fallacy insofar as she never reverses her decision.

**Example 2 (Inertia).** The agent chooses an action \( a_t \in \{0, 1\} \) in each period \( t = 1, 2, \ldots \) with the goal of matching the current state. The state \( \theta_t \) follows a Markov chain on the set \( \{0, 1\} \). In each period \( t \in \mathbb{N} \), the gross flow payoff \( u(a_t, \theta_t) \) is equal to \( u_a > 0 \) if \( a_t = \theta_t = a \), and is 0 if \( a_t \neq \theta_t \). Payoffs are discounted exponentially.

The solution of this problem illustrates how the speed of adjustment depends on incentives and on the persistence of states.

**Example 3 (Response times).** The state \( \theta \in \{0, 1\} \) is fixed over time. The agent has a uniform prior belief. In each period \( t = 1, \ldots, T \), she chooses among taking a terminal action 0 or 1, or waiting until the next period (denoted by \( w \)). She receives a benefit of 1 if her terminal action matches the state, and a benefit of 0 otherwise. In addition, she pays a cost \( c \in (0, 1) \) for each period that she waits. The agent’s total gross payoff is the undiscounted sum of the flow payoffs

\[
u(a_t, \theta) = \begin{cases} 
1 & \text{if } a_t = (w, \ldots, w, \theta), \\
0 & \text{if } a_t = (w, \ldots, w, 1 - \theta), \\
-c & \text{if } a_t = (w, \ldots, w), \\
0 & \text{otherwise.}
\end{cases}
\]

We use this example to study the tradeoff between speed and accuracy of decision making.

### 2.2 Preliminaries

Our main goal is to characterize the agent’s observable behavior, i.e. her action choices. A *(stochastic)* choice rule \( p \) is a system of distributions \( p_t(a_t \mid \theta^t, a^{t-1}) \) over \( A \), one for each \( \theta^t \) and \( a^{t-1} \), interpreted as the probability of choosing \( a_t \) conditional on histories \( \theta^t \) and

a^{t-1}. We say that a strategy \( s = (f, \sigma) \) generates the choice rule \( p \) if

\[
p(a_t | \theta^t, a^{t-1}) \equiv \Pr \left( \sigma(x^t) = a_t | \theta^t, \sigma^{t-1}(x^{t-1}) = a^{t-1} \right),
\]

where the probability is evaluated with respect to the joint distribution of states and sequences of signals generated according to \( f \).

Conversely, a choice rule \( p \) can be associated with a strategy \( (f, \sigma) \). Roughly speaking, one can choose a particular signal for each action, and then match the probability of each of those signals with the probability the choice rule assigns to its associated action. Formally, fix any injection \( \phi : A \rightarrow X \) and, by a slight abuse of notation, for any \( t \), let \( \phi \) also denote the mapping from \( A^t \) to \( X^t \) obtained by applying \( \phi \) coordinate-by-coordinate. Given any choice rule \( p \), let \( \overline{p} = (f, \sigma) \) be such that \( f(\phi(a_t) | \theta^t, \phi(a^{t-1})) = p(a_t | \theta^t, a^{t-1}) \) and \( \sigma(\phi(a^t)) = a_t \). Call \( \overline{p} \) the strategy induced by \( p \).

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

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

\[
\max_{p} E \left[ \sum_{t=1}^{\infty} \delta^{(t)} \left( u_t(a^{t}, \theta^t) - I(\theta^{t}; a_t | a^{t-1}) \right) \right], \tag{3}
\]

where the expectation is with respect to the joint distribution of \((\theta_t, a_t)\) induced by \( \pi \) and \( p \). Conversely, any choice rule \( p \) solving (3) induces a strategy solving the dynamic RI problem.

Accordingly, we call any rule \( p \) solving (3) a solution to the dynamic RI problem.

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

In static models, the conclusion of Lemma 1 holds as long as the cost of signals is non-decreasing in Blackwell informativeness. In dynamic problems, more structure is needed. For example, if cost was concave in the mutual information then the agent to acquire more information than what is necessary for her choice in a given period if she plans to use that information in a later period where the marginal cost of acquiring it would be higher. If costs are proportional to reduction in entropy, there is no incentive for the agent to acquire information any earlier than necessary.

**Proposition 1.** There exists a solution to (3).

Proofs are in the Appendix.

## 3 Solution

### 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.\(^8\)

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

\[
r_t(a_t \mid \theta^t, a_{t-1}) = \frac{e^{\tilde{u}_t(a_t, \theta^t)}}{\sum_{a_t'} e^{\tilde{u}_t(a_t', \theta^t)}},
\]

where

\[
\tilde{u}_t(a_t^t, \theta^t) = u_t(a_t^t, \theta^t) + \delta_{t+1}E_{\theta_{t+1}}[V_{t+1}(a_t^{t+1} \mid \theta^{t+1}) \mid \theta^t],
\]

and the continuation values \(V_t\) satisfy

\[
V_t(a_{t-1}^t, \theta^t) = \log \left( \sum_{a_t} e^{\tilde{u}_t((a_{t-1}, a_t^t), \theta^t)} \right).
\]

The solution to the dynamic RI problem is a dynamic logit rule with an endogenous state-independent utility term. A default rule \(q\) is a system of conditional action distributions \(q_t(a_t \mid a_{t-1}^t)\), one for each decision node \(a_{t-1}^t\). We refer to \(q_t(a_t \mid a_{t-1}^t)\) as the

---

\(^8\)Our definition is more restrictive than that of Rust (1987) in that we do not allow the agent actions to affect the realization of future states. Our model also differs in the form of discounting and in the state spaces.
predisposition toward action \(a_t\) at the decision node \(a_{t-1}\). The difference between a default rule and a choice rule is that the latter conditions on states while the former does not. Below, we interpret the default rule \(q_t(a_t \mid a_{t-1})\) as a system of probabilistic hints that suggest a good decision in period \(t\) conditional on the action history \(a_{t-1}\), formed at the ex ante stage.

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

\[
u_t (a^t, \theta^t) + \log q_t (a_t \mid a_{t-1}) \quad \text{for} \ t \in \mathbb{N}.
\]

For any choice rule \(p\), let \(p_t (a_t \mid a_{t-1})\) denote the probability of choosing action \(a_t\) conditional on reaching decision node \(a_{t-1}\), that is,

\[
p_t (a_t \mid a_{t-1}) = E_{\theta_t} [p (a_t \mid \theta_t, a_{t-1}) \mid a_{t-1}] .
\]

Finally, we adopt the convention that \(\log 0 = -\infty\) and \(e^{-\infty} = 0\).

**Theorem 1.** The dynamic RI problem under payoffs \(u\) is solved by a dynamic logit rule \(p\) under payoffs \(u + \log q\), where \(q\) is a default rule that solves

\[
\max_{\tilde{q}} V(u + \log \tilde{q}).
\]

Moreover,

\[
q_t (a_t \mid a_{t-1}) = p_t (a_t \mid a_{t-1})
\]

for every decision node \(a_{t-1}\) that is reached with positive probability according to \(p\).

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

\[
c_t (a_{t-1}, a_t) \equiv - \log q_t (a_t \mid a_{t-1})
\]

whenever she chooses \(a_t\) after the action history \(a_{t-1}\). This “switching 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 \(\theta_t\) 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 \( u_t (a^t, \theta^t) - c_t (a^{t-1}, a_t) \)—the utility less the virtual switching cost.

For a fixed decision problem, separately identifying \( u_t \) and \( c_t \) is not necessary to describe behavior; choice probabilities depend only on the difference \( u_t - c_t \). However, the distinction can be important when extrapolating to other decision problems. For example, Rust (1987) considers a bus company’s demand for replacement engines. He estimates the replacement cost by fitting a dynamic logit in which the agent trades that cost off against the expected cost of 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 could underestimate 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 switching cost \( c_t \) associated with replacement (by (6)), and hence to an additional decrease in demand relative to the model in which \( c_t \) is fixed. Intuitively, the price increase not only discourages the purchase of a new engine, it also discourages the agent from checking whether a new engine is needed.

Distinguishing the actual utility \( u_t \) from the virtual switching 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 switching cost \( c_t(a^{t-1}, a_t) = -\log p(a_t | a^{t-1}) \) can be estimated directly from the agent’s choice data.

### 3.2 Reduction to static problems

While Theorem 1 is useful for understanding behavior, it is less helpful when it comes to computing the default rule \( q \). In this section, we show that the dynamic RI problem can be reduced to a collection of static RI problems, one for each decision node, that can be used to solve for the predispositions \( q_t(a_t | a^{t-1}) \).

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

**Definition 3.** The static rational inattention problem for a triple \( (\Theta, \pi, u) \) is

\[
\max_p E_p [u(a, \theta) - I(\theta; a)] .
\]

**Proposition 2** (Matějka and McKay, 2014; Caplin and Dean, 2013a). The static rational inattention problem with parameters \( (\Theta, \pi, u) \) is solved by the choice rule

\[
p(a \mid \theta) = \frac{q(a) e^{u(a, \theta)}}{\sum_{a'} q(a') e^{u(a', \theta)}} ,
\]

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

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

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

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

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

One complication that arises for this characterization is that, 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 (which is needed to determine whether choosing that action with zero probability is optimal). To cover this case, we define the posterior belief after an action \( a \) is taken with zero probability to be

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

This expression coincides with (9) when \( a \) is chosen with a positive probability. Otherwise,

---

9The literature on static rational inattention is richer than Definition 3 suggests. We state here the minimal definition sufficient for our purpose.
it differs from (9) only by a renormalization. We show in the proof of Proposition 3 that the belief (10) arises from solving the problem in which the probability of each action is constrained to be at least some \(\varepsilon > 0\), then taking the limit as \(\varepsilon \to 0\).

As in the static case, we abuse notation by writing \(p\) to denote the joint distribution of \(\theta_t\) and \(a_t\), which depends on both the stochastic process \(\pi\) governing \(\theta_t\) and the choice rule \(p(a_t \mid \theta^t, a^{t-1})\). We interpret \(p(\theta^t \mid a^{t-1})\) as the agent's prior over \(\theta^t\) held at the beginning of period \(t\) at the decision node \(a^{t-1}\), and \(p(\theta^t \mid a^t)\) as the posterior over \(\theta^t\) held at the end of period \(t\) after action history \(a^t\). Posterior beliefs at action histories that are reached with zero probability are obtained by applying (10) recursively.

**Proposition 3.** There exists a dynamic choice rule \(p\) solving the dynamic RI problem and a default rule \(q\) such that, at each decision node \(a^{t-1}\), \(p_t(a_t \mid \theta^t, a^{t-1})\) and \(q_t(a_t \mid a^{t-1})\) solve the static RI problem with state space \(\Theta^t\), prior belief

\[
p(\theta^t \mid a^{t-1}) = \sum_{\theta_t \in \Theta} p(\theta^{t-1} \mid a^{t-1}) \pi(\theta_t \mid \theta^{t-1}),
\]

and payoff

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

where the posterior belief \(p(\theta^t \mid a^t)\) formed after taking action \(a_t\) at the decision node \(a^{t-1}\) complies with (10), and the continuation values satisfy

\[
V_t(a^{t-1}, \theta^t) = \log \left( \sum_{a_t} q_t(a_t \mid a^{t-1}) e^{\hat{u}_t(a^t, \theta^t)} \right).
\]

At any given decision node, the static problem in Proposition 3 depends on past behavior through the prior belief and on future behavior through the continuation values. Nonetheless, when optimizing behavior at a particular node, we can treat the continuation values as fixed. In finite horizon and stationary problems, the proposition leads to a finite system of equations characterizing the solution to the dynamic RI problem. Section 4 illustrates this approach in several applications.

### 3.3 The control problem

In this section, we describe the key step of the proof that allows us to reduce the dynamic problem to a collection of static ones. The main idea is to connect the rational inattention
problem to a control problem with observable states in which the agent must pay a cost for deviating from a default choice rule.\footnote{Control problems of this kind have been studied in game theory building on the trembling-hand perfection of Selten (1975). van Damme (1983) offers an early version in which agents optimize the distributions of trembles. Stahl (1990) introduces entropy-based control costs.}

Reformulating the dynamic RI problem as a control problem addresses a crucial difficulty in the analysis. According to Proposition 3, the solution at node $a^{t-1}$ also solves a static RI problem with payoffs

\[ u_t(a^t, \theta^t) + \delta_{t+1} E_{\theta_{t+1}} [V_{t+1}(a^t, \theta^{t+1}) | \theta^t]. \]

The difficulty is that the choice of action distribution $p(a_t | \theta^t, a^{t-1})$ affects the posterior beliefs in subsequent periods, which in turn may affect the continuation value function. The control problem clarifies why the continuation values can be treated as fixed when optimizing the distribution of $a_t$.

**Definition 4.** The control problem for default rule $q$ is

\[ \max_p E_p \left[ \sum_{t=1}^{\infty} \delta_t \left( u_t(a^t, \theta^t) + \log q_t(a_t | a^{t-1}) - \log p_t(a_t | \theta^t, a^{t-1}) \right) \right], \tag{12} \]

where $p$ is a stochastic choice rule.

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 of states, but must trade off optimizing her flow utility $u_t$ against a control cost: for each $(\theta^t, a^{t-1})$, she pays a cost

\[ \log p_t(a_t | \theta^t, a^{t-1}) - \log q_t(a_t | a^{t-1}) \]

for deviating from the default action distribution $q_t(a_t | a^{t-1})$ to the action distribution $p_t(a_t | \theta^t, a^{t-1})$. Mattsson and Weibull show that, in the static problem, the optimal action distribution is a logit rule with a bias toward actions that are relatively likely under the default rule.

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 \( p \) solves the dynamic RI problem if and only if it, together with a default rule \( q \), solves

\[
\max_{q,p} E_p \left[ \sum_{t=1}^{\infty} \delta^t \left( u_t(a_t, \theta_t) + \log q_t(a_t | a^{t-1}) - \log p_t(a_t | \theta, a^{t-1}) \right) \right]. \tag{13}
\]

To see how Lemma 2 addresses the difficulty described at the beginning of this section, 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 \( \theta_t \) in period \( t \). Similarly, for any fixed \( p \), optimizing the default rule \( q \) does not require varying posterior beliefs because those are determined by \( p \), not by \( q \).

## 4 Applications

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

### 4.1 Sunk cost fallacy

We now revisit Example 1. Recall that the agent chooses an action \( a_t \in \{0, 1\} \) at \( t = 1, 2 \). In both periods, the gross flow payoff \( u_t \) is 1 if the action \( a_t \) matches the current state \( \theta_t \in \{0, 1\} \), and is 0 otherwise. There is no discounting.

Suppose the states are symmetrically distributed and positively correlated across time in the following way: \( \theta_1 \) is equally likely to be 0 or 1, and, whatever the realized value of \( \theta \), the probability that \( \theta_2 = \theta_1 \) is 0.9.

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

Which features of the model drive the sunk-cost-fallacy behavior? The superior performance in the first period arises because of the endogenous timing of information acquisition. In a variant of the model with exogenous conditionally i.i.d. signals, the agent would perform better in the second period since she has obtained more precise information. When
Prob. of retaining decision across periods
\[ \Pr(a_1 = a_2) \]
Prob. of the correct choice in period 1
\[ \Pr(a_1 = \theta_1) \]
Prob. of correct choice in period 2
\[ \Pr(a_2 = \theta_2) \]

<table>
<thead>
<tr>
<th>Correlated states</th>
<th>Uncorrelated states</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.86</td>
</tr>
<tr>
<td>0.86</td>
<td>0.79</td>
</tr>
<tr>
<td>0.73</td>
<td>0.73</td>
</tr>
</tbody>
</table>

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

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 sunk-cost effect 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

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

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

Finally, to illustrate the role of correlation across periods, consider a benchmark in which \(\theta_1\) and \(\theta_2\) 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 expected payoffs across the two periods; see the second row of Table 1.

### 4.2 Inertia

Recall that, in Example 2, the agent chooses an action \(a_t \in \{0, 1\}\) in each period \(t = 1, 2, \ldots\) with the goal of matching the current state. The state \(\theta_t\) follows a Markov chain on the set \([0, 1]\) with time-homogeneous transition probabilities \(\gamma(\theta, \theta')\) from state \(\theta\) to state \(\theta'\). In each period \(t \in \mathbb{N}\), the gross flow payoff \(u(a_t, \theta_t)\) is equal to \(u_a > 0\) if \(a_t = \theta_t = a\), and is 0 if \(a_t \neq \theta_t\). Payoffs are discounted exponentially with discount factor \(\delta \in (0, 1)\).
This example can be viewed as a stylized model of a wide range of economic phenomena. For example, the action could represent the consumption level of an agent who responds to macroeconomic variables (captured by $\theta$) that affect her permanent income. Similarly, the action could be thought of as a pricing decision by a firm facing varying demand, or an investor’s choice of whether to hold a particular asset. An important question in each of these settings is how the timing of adjustment to shocks is shaped by the environment. Do the length of adjustment lags differ between booms and busts? How does volatility influence behavioral inertia?

We start with an implicit characterization of long-run behavior. The following proposition, which is an application of Proposition 3, states that the choice rule, continuation values, and predispositions in any period $t$ depend on the last action $a_{t-1}$, but not on any earlier actions. This implies that the long-run behavior is characterized by a finite set of equations, which can be solved numerically.

We say that a solution to this Markovian model is interior if there exists $t'$ such that each action is chosen with positive probability in every period $t > t'$. If the solution is not interior, then there is some $t'$ for which a particular action is chosen almost surely in every period $t > t'$, in which case we call that action absorbing.

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

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

(14)

where the continuation payoffs solve

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

(15)

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

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

(16)
for all $\theta_t$ and $a_{t-1}$, and the posteriors $p(\theta_t \mid a_t)$ satisfy

$$
p(\theta_t \mid a_t) = \frac{\exp (u(a_t, \theta_t) + \delta E[V(a_t, \theta_{t+1}) \mid \theta_t])}{\exp (u(a_t, \theta_t) + \delta E[V(a_t, \theta_{t+1}) \mid \theta_t])}.
$$

(17)

One can check whether there is an interior solution by solving the system of equations in Proposition 4. If the resulting predispositions are positive then there is an interior solution and the result applies. Otherwise, one of the actions is absorbing.\textsuperscript{11}

Equation (16) is condition for the posterior beliefs. The left-hand side is the prior belief about $\theta_t$ at the beginning of the period $t$ obtained by applying the transition probabilities of the Markov chain to the posterior about $\theta_{t-1}$ at the end of period $t-1$. The right-hand side is the same prior written as the expectation of the posterior at the end of period $t$.

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

The system of equations in Proposition 4 can not be solved analytically in general. There is, however, a simple analytical solution in the limit as states become perfectly persistent. Accordingly, let $\gamma(\theta, \theta') = \tilde{\gamma}(\theta, \theta')\varepsilon$ for $\theta \neq \theta'$, and consider the limit as $\varepsilon$ vanishes. For $a' \neq a$, define the limit predispositions $q^*(a, a') := \lim_{\varepsilon \to 0} q(a_t = a' \mid a_{t-1} = a)/\varepsilon$, and the limit adjustment rates $\alpha(a, a') := \lim_{\varepsilon \to 0} p(a_t = a' \mid a_{t-1} = a, \theta_t = a')/\varepsilon$. In words, $q^*(a, a')$ is the probability, scaled by $\varepsilon$, that the agent switches to action $a'$ after she has chosen $a \neq a'$ in the previous period, and $\alpha(a, a')$ is the rescaled probability that the agent switches from $a$ to $a'$ if this switch aligns the action with the current state. Let $U_a = \exp u_a$.

Proposition 5. Let $a' \neq a$. The adjustment rate $\alpha(a, a')$ increases in $u_{a'}$ and $\tau(a, a')$, and decreases in $u_a$ and $\tau(a', a)$. The limit predispositions are

$$
q^*(a, a') = \frac{\tau(a, a') U_{a'} - 1}{U_a - 1} \frac{1}{\tau(a', a)}
$$

(18)

\textsuperscript{11}Proposition 4 describes long-run behavior. Actions in early periods depend on the prior. If the distribution of $\theta_1$ lies in the convex hull of the long-run stationary posteriors $p(\theta_t \mid a_t)$ for the two actions, then the choice rule (14) is optimal beginning in period 2. If not, the agent acquires no information until her belief enters the convex hull of the stationary posteriors, after which (14) applies.
and the limit adjustment rates are

\[ \alpha(a, a') = q^*(a, a')U_{a'}. \] (19)

Adjustment to shocks involves significant delays in a persistent environment. For example, consider a symmetric setting with \( \gamma(0, 1) = \gamma(1, 0) = \varepsilon \), and \( u_0 = u_1 = u \). The expected time between adjacent switches of the state is \( 1/\varepsilon \). According to the proposition, if the agent’s action is not aligned with the state, she switches her action with probability \( e^{u/(1-\delta)}\varepsilon \) per period. This probability corresponds to an expected lag time of \( e^{-u/(1-\delta)}/\varepsilon \), which is of the same order as the time between switches of the state. In particular, greater persistence corresponds to longer adjustment lags. In the symmetric case, this monotonicity result also holds outside of the limit: increasing volatility reduces adjustment lags.\cite{12}

Asymmetries in incentives have an intuitive impact on adjustment rates. Consider the same Markov chain with \( \gamma(0, 1) = \gamma(1, 0) = \varepsilon \), and suppose \( u_0 > u_1 \). Interpreting state 1 as the good state in an investment problem, this corresponds to the loss from investing during a crisis exceeding the profit from investing during a boom. Proposition 5 implies that \( \alpha(0, 1) < \alpha(1, 0) \), meaning that the agent reacts more quickly to negative shocks than to positive ones.

### 4.3 Response times

In this section, we study a simple model of response times in decision-making. The study of response times has a long tradition in psychology, and has more recently become the subject of a growing literature in economics (for example, see Rubinstein, 2007). Spiliopoulos and Ortmann (2014) attribute the historical difference in interest between the two fields to the fact that psychologists focus on procedures of decision-making whereas economists traditionally treat decision-making as a black box. Our approach follows Sims (2003), who argues for RI as a way to incorporate decision times into economic models in the black-box fashion. Woodford (2014) studies similar issues in a model that merges neurological evidence with an RI framework.

Recall that in Example 3, the state \( \theta \in \{0, 1\} \) is uniformly distributed and fixed over time. In each period \( t = 1, \ldots, T \), the agent chooses among taking a terminal action 0 or 1, or waiting until the next period (denoted by \( w \)). The agent’s total gross payoff is the

\[^{12}\text{Numerical results suggest that the same result holds in asymmetric settings.}\]
undiscounted sum of the flow payoffs

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

This formulation is similar to the model of Arrow, Blackwell, and Girshick (1949) except that information is endogenous.

With the information cost function 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. Hence there is no delay if the marginal cost of information is constant across time. However, delay can be optimal in a closely related variation of the model in which—as in much of the rational inattention literature—there is an upper bound on how much information the agent can process in each period. Accordingly, the agent solves

\[
\max_p E \left[ \sum_{t'=1}^T u_{t'}(a_{t'}, \theta) \right]
\]

s.t. \( E \left[ I(\theta; a^t | a^{t-1}) \right] \leq \kappa \) for all \( t = 1, \ldots, T, \)

where \( \kappa > 0 \) is the capacity constraint on the information per period, and \( p(a_t | \theta, a^{t-1}) \) is the choice rule.\(^{13}\)

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

The first-order conditions for this problem are closely related to those of the general model with information costs. The solution of (20) also solves a problem in which the capacity constraint is replaced by time-varying marginal costs of information \( \lambda_t \), where \( \lambda_t \)

\(^{13}\)This formulation abstracts from explicit signal acquisition by imposing the constraint directly on the joint distribution of actions and states. As in the general model, we could allow the agent to choose signal distributions together with mappings from signals to actions. One can show that the result of Lemma 1 applies in this context, and therefore the present formulation is without loss of generality.
Figure 1: (A) accuracy as a function of time, (B) distributions of the decision time for three values of the capacity $\kappa$.

is the shadow price of the capacity constraint for period $t$.

How is the timing of terminal decisions related to their accuracy? For a fixed capacity $\kappa$, the accuracy of decisions increases with time. Because the capacity is uniform over time, the delay cost makes it relatively more valuable early on. The shadow price of capacity is therefore decreasing over time, which in turn leads to increasing accuracy of decisions over time. However, the speed of decision-making is not monotone in the capacity: numerical results show that 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 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. Computations are in the appendix.

Let $r_t$ denote the probability that the agent makes the terminal decision at round $t$, and let $p_t$ be the probability that the correct decision $a_t = \theta$ is made conditional on terminating at $t$. Figure 1 depicts the values of $r_t$ and $p_t$ when $T = 1000$, the per period capacity is $\kappa \in \{10^{-4}, 10^{-3}, 10^{-2}\}$, and the delay cost $c$ is $10^{-3}$. 


Appendix

A Proofs for Section 2.2

Proof of Proposition 1. Consider the space of strategies $\Pi = \prod_t \prod_{a^{t-1}} P(a_t, \theta_t; a^{t-1})$, where $P(a_t, \theta_t; a^{t-1})$ denotes the set of feasible joint distributions of $a_t$ and $\theta_t$ given $a^{t-1}$. By Tychonoff’s Theorem, the space $\Pi$ is compact in the product topology, and because $u_t$ is uniformly bounded, the objective function is continuous. Therefore, an optimum exists.

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

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

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

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

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

Endowing $\overline{\Psi}$ with the product topology, the sequence $s^\tau$ converges to a strategy $s^*$ that is identical up to relabeling of signals to the strategy induced by the choice rule generated by $s$. Moreover, since $s^*$ and $s^\tau$ are identical in periods 1 through $\tau$, flow payoffs are uniformly bounded, and $\sum_t \delta^{(t)} < \infty$, we have $U(s^*) = \lim_\tau U(s^\tau)$. In particular, if $s$ is optimal then so is $s^*$, proving the lemma.
The sequence \((s^\tau)_{\tau=0}^\infty\) is constructed recursively as follows. First we define \(s^0\) by “embedding” \(s\) into the expanded signal space \(\overline{X}_t\). Formally, for \(x_t \in X\) and \(\overline{x}_{t-1} = (\overline{x}_1, \ldots, \overline{x}_{t-1}) \in \overline{X}_{t-1}\), let

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

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

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

\[
f^\tau_t(\overline{x}_t|\overline{x}^{\tau-1}, \theta_t) = \begin{cases} f^{\tau-1}_t(\overline{x}_t|\overline{x}^{\tau-1}, \theta_t) & \text{if } t \leq \tau - 1, \\ \sum_{\overline{x} \in X_t} \tilde{\mu}_t^{\tau-1}(\overline{x};\overline{x}^{\tau-1}) a f^{\tau-1}_t(\overline{x}|\overline{x}^{\tau-1}, \theta_t) & \text{if } t = \tau \text{ and } \overline{x}_t = \overline{x}_t^a, \\ 0 & \text{if } t = \tau \text{ and } \overline{x}_t \neq \overline{x}_t^a \text{ for any } a \in A, \\ \Pr_{\tilde{f}^{\tau-1}_t} (\overline{x}_t | \tilde{\mu}_t^{\tau-1}(\overline{x}^{\tau-1}), \theta_t) & \text{otherwise,} \end{cases}
\]

where \(\tilde{\mu}^t(\overline{x}^t) := (\tilde{\mu}^t_1(\overline{x}_1), \ldots, \tilde{\mu}^t_t(\overline{x}_t;\overline{x}_{t-1}))\), and

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

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

First note that for every \(\tau \geq 1\), \(s^\tau\) induces the same joint distribution of actions and states as \(s^{\tau-1}\). Hence \(U(s^\tau) \geq U(s^{\tau-1})\) if and only if the total discounted expected information cost from \(s^\tau\) is no more than that from \(s^{\tau-1}\). Recall that for signals \(x_t\), the
total discounted expected information cost is
\[ E \left[ \sum_{t=1}^{\infty} \delta(t) I (\theta^t; x^t \mid x^{t-1}) \right]. \]

Letting \( x^0 = \emptyset \), for any \( t \neq T \), the mutual information \( I (\theta^t; x^t \mid x^0) \) is identical under \( s^{T-1} \) and \( s^T \), and for \( t = T \) it is (weakly) lower under \( s^T \). Since
\[ I (\theta^T; x^T \mid x^0) = \sum_{t'=1}^{T} I (\theta^t; x^{t'} \mid x^{t'-1}) \]
and \( \delta(T) \geq \delta(T+1) \), it follows that the information cost is at least as high under \( s^{T-1} \) as under \( s^T \). \( \square \)

B Proofs for Section 3

Proof of Lemma 2. First we show that the optimization problem (3) from Lemma 1 is equivalent to
\[
\max_p E \left[ \sum_{t=1}^{\infty} \delta(t) \left( u_t(a^t, \theta^t) + \log p_t(a_t \mid a^{t-1}) - \log p_t(a_t \mid a^{t-1}, \theta^t) \right) \right]. \tag{21}
\]

Using the definition of mutual information, we can rewrite the objective in (3) as
\[
E \left[ \sum_{t=1}^{\infty} \delta(t) \left( u_t(a^t, \theta^t) - \log p(\theta^t \mid a^t) + \log p(\theta^t \mid a^{t-1}) \right) \right].
\]

When \( a^t \) is attained with 0 probability, define \( \log p(\theta^t \mid a^t) \) as an arbitrary constant, and note that the choice of the constant does not affect the expectation.

Using the identities \( p(\theta^t \mid a^s) = p(a^s \mid \theta^t)p(\theta^t)/p(a^s) \) and \( p(\theta^t \mid a^{t-1}) = p(\theta^t \mid \theta^{t-1})p(\theta^{t-1} \mid a^{t-1}) \), and dropping terms such as \(-\log p(\theta^t)\) that do not depend on actions gives the equivalent objective function
\[
E \left[ \sum_{t=1}^{\infty} \delta(t) \left( u_t(a^t, \theta^t) - \log p(a^t \mid \theta^t) + \log p(a^t) + \log p(a^{t-1} \mid \theta^{t-1}) - \log p(a^{t-1}) \right) \right].
\]
which can be rearranged to give

\[ E \left[ \sum_{t=1}^{\infty} \left( \delta^{(t)}(a_t', \theta_t') + (\delta^{(t+1)} - \delta^{(t)}) \left( \log p(a_t' | \theta_t') - \log p(a_t') \right) \right) \right]. \]

Using \( \log p(a_t) = \sum_{t'=1}^{t} \log p(a_t' | a_t'-1) \) and

\[ \log p(a_t' | \theta_t') = \sum_{t'=1}^{t} \log p(a_t' | a_t'-1, \theta_t') = \sum_{t'=1}^{t} \log p(a_t' | a_t'-1, \theta_t') \]

and rearranging shows that the problem is equivalent to (21).

Consider problem (13). For any given \( p \), the optimal \( q \) solves

\[ \max_q E \left[ \sum_{t=1}^{\infty} \delta^{(t)} \log q_t(a_t | a_t^{t-1}) \right], \]

where the distribution of action paths is governed by \( p \). This problem can be solved separately for each \( t \) and \( a_t-1 \), and it is straightforward to show that the solution is

\[ q_t(a_t | a_t^{t-1}) = p_t(a_t | a_t^{t-1}). \]

In particular, a choice rule \( p \) solves (13) if and only if it solves (21).

**Proof of Theorem 1.** Consider the control problem given some \( q \) (i.e. the problem where we choose \( p \) to maximize the objective of (13) for fixed \( q \)). Let \( \pi_t(a_t', \theta_t') = u_t(a_t', \theta_t') + \log q_t(a_t | a_t^{t-1}) \). For each \( a_t^{t-1} \) and \( \theta_t' \) such that \( \pi(\theta_t') > 0 \), let

\[ V_t(a_t^{t-1}, \theta_t') = \frac{1}{\delta^{(t)}} \max_{\{p_t(a_t | a_t^{t-1}, \theta_t')\}_{t'=t}} E \left[ \sum_{t'=t}^{\infty} \delta^{(t')} \left( \pi_t(a_t', \theta_t') - \log p_t(a_t | a_t^{t-1}, \theta_t') \right) \right]. \]

Note that \( V_t \) satisfies the recursion

\[ V_t(a_t^{t-1}, \theta_t') = \max_{\{p_t(a_t | a_t^{t-1}, \theta_t')\}} E \left[ \pi_t(a_t', \theta_t') - \log p_t(a_t | a_t^{t-1}, \theta_t') + \delta_{t+1} V_{t+1}(a_t, \theta_t') \right] \]

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

To solve the maximization problem on the right-hand side of (22), note first that, since
\( \pi_t(a^t, \theta^t) = u_t(a^t, \theta^t) + \log q(a_t \mid a^{t-1}), \) if \( q(a_t \mid a^{t-1}) = 0 \) for some \( a_t \), then we must have \( p_t(a_t \mid a^{t-1}, (\theta^{t-1}, \theta_t)) = 0 \) for every \( \theta_t \) satisfying \( \pi(\theta^{t-1}, \theta_t) > 0 \).\(^{14}\) Accordingly let
\( \tilde{A}(a^{t-1}) = \{ \tilde{a}_t \in A_t : q(\tilde{a}_t \mid a^{t-1}) > 0 \} \), and suppose \( a_t \in \tilde{A}(a^{t-1}) \) and \( \pi(\theta^{t-1}, \theta_t) > 0 \). If \( \tilde{A}(a^{t-1}) \) is a singleton, then \( p_t(a_t \mid a^{t-1}, (\theta^{t-1}, \theta_t)) = 1 \). Otherwise, the first-order condition for (22) with respect to \( p_t(a_t \mid a^{t-1}, \theta^t) \) is
\[
\pi(\theta^t \mid \theta^{t-1}) (\pi_t(a^t, \theta^t) - (\log p_t(a_t \mid a^{t-1}, \theta^t) + 1) + \delta_{t+1}E_{\theta_{t+1}}[V_{t+1}(a^t, \theta^{t+1})] \) \leq \mu_t(a^{t-1}, \theta^t), \tag{23}
\]
where \( \mu_t(a^{t-1}, \theta^t) \) is the Lagrange multiplier associated with the constraint \( \sum a'_t p_t(a'_t \mid a^{t-1}, \theta^t) = 1 \). Moreover, (23) holds with equality if \( p_t(a_t \mid a^{t-1}, \theta^t) \in (0, 1) \), which must be the case to ensure that the left-hand side of (23) is finite for all \( a_t \in \tilde{A}(a^{t-1}) \).

For \( a_t \in \tilde{A}(a^{t-1}) \), rearranging the first-order condition gives
\[
p_t(a_t \mid a^{t-1}, \theta^t) = \exp \left( \pi_t(a^t, \theta^t) - 1 + \delta_{t+1}V_{t+1}(a^t, \theta^t) - \frac{\mu_t(a^{t-1}, \theta^t)}{\pi(\theta^t \mid \theta^{t-1})} \right),
\]
where \( V_{t+1}(a^t, \theta^t) := E_{\theta_{t+1}}[V_{t+1}(a^t, \theta^{t+1})] \). Since \( \sum a'_t \in \tilde{A}(a^{t-1}) p_t(a'_t \mid a^{t-1}, \theta^t) = 1 \), it follows that
\[
p_t(a_t \mid a^{t-1}, \theta^t) = \frac{\exp \left( \pi_t(a^t, \theta^t) - 1 + \delta_{t+1}V_{t+1}(a^t, \theta^t) - \frac{\mu_t(a^{t-1}, \theta^t)}{\pi(\theta^t \mid \theta^{t-1})} \right)}{\sum a'_t \in \tilde{A}(a^{t-1}) \exp \left( \pi_t((a^{t-1}, a'_t), \theta^t) - 1 + \delta_{t+1}V_{t+1}((a^{t-1}, a'_t), \theta^t) - \frac{\mu_t(a^{t-1}, \theta^t)}{\pi(\theta^t \mid \theta^{t-1})} \right)} \]
\[
= \frac{\exp \left( \pi_t(a^t, \theta^t) + \delta_{t+1}V_{t+1}(a^t, \theta^t) \right)}{\sum a'_t \in \tilde{A}(a^{t-1}) \exp \left( \pi_t((a^{t-1}, a'_t), \theta^t) + \delta_{t+1}V_{t+1}((a^{t-1}, a'_t), \theta^t) \right)} .
\]

Substituting into (22) gives the recursion
\[
V_t(a^{t-1}, \theta^{t-1})
\]
\[
= E \left[ -\delta_{t+1}V_{t+1}(a^t, \theta^t) + \log \left( \sum_{a_t \in \tilde{A}(a^{t-1})} \exp \left( \pi((a^{t-1}, a_t), \theta^t) + \delta_{t+1}V_{t+1}((a^{t-1}, a_t), \theta^t) \right) \right) \right] + \delta_{t+1}V_{t+1}(a^t, \theta^t) | \theta^{t-1} ,
\]
\[
\]
\[
^{14}\text{If } \pi(\theta^{t-1}, \theta_t) = 0 \text{ then } p_t(a_t \mid a^{t-1}, (\theta^{t-1}, \theta_t)) \text{ has no effect on the value and can be chosen arbitrarily.}
\]

27
and therefore,

\[
V_t(a^{t-1}, \theta^t) = E \left[ \log \left( \sum_{a_t' \in A(a^{t-1})} \exp \left( u((a^{t-1}, a_t'), \theta^t) + \delta_{t+1} V_{t+1}((a^{t-1}, a_t'), \theta^t) \right) \right) \theta^{t-1} \right]
\]

\[
= E \left[ \log \left( \sum_{a_t' \in A_t} q(a_t' | a^{t-1}) \exp \left( u((a^{t-1}, a_t'), \theta^t) + \delta_{t+1} V_{t+1}((a^{t-1}, a_t'), \theta^t) \right) \right) \theta^{t-1} \right].
\]

(24)

The result now follows from Lemma 2. \(\square\)

**Proof of Proposition 3.** Note first that in both the control problem and the recursive problem, for any given \(q\), there exists an optimal \(p\) satisfying

\[
p_t(a_t | a^{t-1}, \theta^t) = \frac{q_t(a_t | a^{t-1}) \exp \left( \hat{u}(a_t, \theta^t) \right)}{\sum_{a_t'} q_t(a_t' | a^{t-1}) \exp \left( \hat{u}(a^{t-1}, a_t') \right)}.
\]

(25)

Thus it suffices to consider, in each case, the problem of optimizing with respect to \(q\) under the assumption that \(p\) is given by (25).

If the optimal \(q\) in problem (13) is fully interior (i.e. \(q(a_t | a^{t-1}) > 0\) for every \(a_t\) and \(a^{t-1}\)), then the result is straightforward to verify. In particular, the solutions coincide if we add to each problem additional constraints that \(q(a_t | a^{t-1}) \geq \varepsilon/|A_t|\) for every \(a^{t-1}\) and \(a_t\), where \(\varepsilon \in (0, 1)\) (and take the posterior beliefs in the recursive problem based on (25)).

Note that for every \(a^{t-1}\) and \(\theta^t\), the difference between the continuation value \(V_t^\varepsilon(a^{t-1}, \theta^t)\) in this problem and the continuation value \(V_t(a^{t-1}, \theta^t)\) in the original problem is bounded by a quantity proportional to \(\varepsilon\). Moreover, the values vary continuously in \(q\) (with the product topology). Hence, as \(\varepsilon\) vanishes, any convergent subsequence of solutions to the problems bounded by \(\varepsilon\) converges to a solution of the original problem. All that remains is to show that the same is true of the recursive problem, that is, that, as \(\varepsilon\) vanishes, some sequence of solutions to the recursive problems with bounds \(\varepsilon\) converges to a solution of the recursive problem with no such bounds.

Consider the static control problem in any period of the recursion, and write \(\pi\) for the prior in that period and \(\hat{u}_\varepsilon\) for the analogue of \(\hat{u}\) with continuation values \(V_\varepsilon\). The optimal
default rule \( q \) for the problem with bounds \( \varepsilon \) solves

\[
\max_q E_\theta \left[ \ln \left( \sum_a q(a) e^{\hat{u}(\theta,a)} \right) \right]
\]

s.t. \( \varepsilon \leq q(a) \leq 1, \sum_a q(a) = 1. \)  

(26)

By the Maximum Theorem, the set of solutions to this problem is upper hemicontinuous in \( \hat{u} \) and \( \pi \). Since the continuation values approach the true values as \( \varepsilon \) vanishes, all that remains is to show that, at each history, the prior \( \pi \) approaches that of the recursive problem.

The first-order condition for a solution of (26) with \( q(a) \in (\varepsilon,1) \) is

\[
\sum_\theta \frac{\pi(\theta)e^{\hat{u}_\varepsilon(\theta,a)}}{\sum_{a'} q(a') e^{\hat{u}_\varepsilon(\theta,a')}} = \mu,
\]

(27)

where \( \mu \) is the Lagrange multiplier associated with the constraint \( \sum_{a'} q(a') = 1 \). Note that there must exist some \( a \) for which \( q(a) \in (\varepsilon,1) \). For that action \( a \), we have \( p(a) = q(a) \), and hence the left-hand side of (27) is the sum of posterior beliefs, which must be equal to 1.

Now consider \( a \) for which the solution is \( q(a) = \varepsilon \). Then we must have

\[
\sum_\theta \frac{\pi(\theta)e^{\hat{u}_\varepsilon(\theta,a)}}{\sum_{a'} q(a') e^{\hat{u}_\varepsilon(\theta,a')}} \leq \mu = 1.
\]

In this case, the posterior beliefs satisfy

\[
p(\theta | a) = \frac{\pi(\theta)}{p(a)} p(a | \theta) = \frac{q(a)}{p(a)} \frac{\pi(\theta)e^{\hat{u}_\varepsilon(\theta,a')}}{\sum_{a'} q(a') e^{\hat{u}_\varepsilon(\theta,a')}} = \frac{1}{\sum_{\theta'} \pi(\theta')} \frac{e^{\hat{u}_\varepsilon(a',\theta')}}{\sum_{a'} q(a') e^{\hat{u}_\varepsilon(a',\theta')}} \frac{\pi(\theta)e^{\hat{u}_\varepsilon(\theta,a')}}{\sum_{a'} q(a') e^{\hat{u}_\varepsilon(\theta,a')}}.
\]

Therefore, as \( \varepsilon \) vanishes, the posteriors indeed approach those given by (10).\(^\text{15}\)

\(^{15}\)Although we have only shown this within a single period of the recursion, simple induction implies that it holds across all histories, giving convergence of the solutions in the product topology.
C Proofs and computations for Section 4

C.1 Sunk cost fallacy

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

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

One may interpret \( V_c \) as the expected payoff in period 2, including the information cost, when the action \( a_1 \) suggests the correct choice of \( a_2 \), and \( V_w \) as the corresponding payoff when \( a_1 \) suggests the wrong choice. By (11), the continuation payoffs satisfy \( V_c = \log(se + (1 - s)) \) and \( V_w = \log(s + (1 - s)e) \).

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

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

and a uniform prior on \( \theta_1 \). If the agent chooses action \( a_1 \), she assigns posterior probability \( p_1 \) to her choice being correct, where, according to (9),

\[
p_1 = p(\theta_1 = a_1 \mid a_1) = \frac{\frac{1}{2}e^{\hat{u}_1(a_1, a_1)}}{\frac{1}{2}e^{\hat{u}_1(a_1, a_1)} + \frac{1}{2}e^{\hat{u}_1(1-a_1, a_1)}}. \tag{28}
\]

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

The second-stage static RI problem involves gross payoffs

\[
u_2(a_2, \theta_2) = \begin{cases} 
1 & \text{if } a_2 = \theta_2, \\
0 & \text{if } a_2 \neq \theta_2,
\end{cases}
\]

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

We want to show that under the optimal choice rule, $a_1 = a_2$ almost surely. Suppose the predisposition $s$ is equal to 1. Then $V_c = 1$ and $V_w = 0$. It follows from (28) that $p_1 \approx 0.86$. Then $p_2 \approx 0.79$, and since $s = 1$, this is also the probability that the second decision is correct. To verify that $s = 1$, we need to check that the predisposition $s = 1$ maximizes

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

which is indeed the case.

### C.2 Inertia

As noted in the text, Proposition 4 follows directly from Proposition 3 together with a result from Caplin and Dean (2013b).

**Proof of Proposition 5.** First, note that the rescaled predispositions $q^*(a, a')$ are bounded above by some $K$. Condition (16) implies that, for $a \neq a'$,

$$q(a' \mid a) = \frac{\Pr(\theta_t = a \mid a_t = a) \gamma(a, a') - \Pr(\theta_t = a' \mid a_t = a) \gamma(a', a)}{\Pr(\theta_t = a' \mid a_t = a') - \Pr(\theta_t = a' \mid a_t = a)}.$$  \(29\)

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

Using the bound on $q^*(a, a')$ and the fact that the continuation values are bounded, (15) implies that there exists some $K'$ such that for sufficiently small $\varepsilon$

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

Thus $\lim_{\varepsilon \to 0} V(a, \theta) = \frac{u(a, \theta)}{1 - \sigma}$. Condition (17) therefore implies that

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

where $a' \neq a$. Substituting this last expression into (29) gives (18). Finally, (19) follows from (14). The comparative statics can be checked by taking derivatives of the expressions for $\alpha$ and $q^*$. \(\square\)
C.3 Response times

Consider a problem with a relaxed constraint

\[
\max_p E \left[ \sum_{t'=1}^T u_t \left( a_t', \theta \right) \right] \quad (30)
\]

s.t. \( E \left[ \sum_{t'=1}^t I \left( \theta; a_{t'} \mid a_{t'-1} \right) \right] \leq \kappa t \) for all \( t = 1, \ldots, T. \) \quad (31)

We will show that solution of (30) also solves the original problem (20).

For each \( t = 1, \ldots, T, \) the capacity constraint in (30) is equivalent to \( E[I(\theta; a_t)] \leq \kappa t. \) Hence, in (30), the set of the feasible joint distributions \( p(\theta, a_t) \) is convex. Since the objective (20) is linear in \( p(\theta, a_t), \) the first-order conditions are sufficient for a global optimum.

The solution of (30) also solves

\[
\max_p E \left[ \sum_{t'=1}^T \left( u_{t'} \left( a_{t'}, \theta \right) - \lambda_{t'} I \left( \theta; a_{t'} \mid a_{t'-1} \right) \right) \right], \quad (32)
\]

where \( \lambda_{t'} \) are the shadow prices of the information capacity for \( t' = 1, \ldots, T. \) If \( \lambda_t \) is decreasing then Problem (32) is a particular case of the dynamic RI problem from Definition 1 (after rescaling discount factors and payoffs as described in footnote 7).

Assume that \( \lambda_t \) is indeed decreasing (we verify this below). Then we may solve (32) using Proposition 3. The only non-trivial decision node at period \( t \) is the one with \( a_{t-1} = w_{t-1}. \) By symmetry, the prior belief about \( \theta \) at the decision node \( w_{t-1} \) is uniform on \{0, 1\}. Symmetry also implies that the continuation value \( V_t(w_{t-1}, \theta) \) is independent of \( \theta \in \{0, 1\}; \) accordingly, we omit the arguments of \( V_t. \) Hence at the node \( w_{t-1}, \) we can treat the problem as a static RI problem with a uniform prior over \( \theta \) and payoffs

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

This static RI problem can be solved using Proposition 2. Symmetry implies that the predisposition \( q_t(a_t \mid w_{t-1}) \) is the same for the two terminal actions \( a_t \in \{0, 1\}; \) we denote
it by \( s_t / 2 \), which makes \( s_t \) the probability that the agent takes a terminal action at \( t \) conditional on waiting in the previous periods. In this case, (8) implies that \( s_t \) solves

\[
\max_{s_t \in [0, 1]} \log \left( \frac{s_t}{2} \left( e^{1/\lambda_t} + 1 \right) + (1 - s_t)e^{(V_{t+1}-c)/\lambda_t} \right).
\] (33)

By (11), the value associated with the static RI problem at time \( t \) is

\[
V_t = \lambda_t \log \left( \frac{s_t}{2} \left( e^{1/\lambda_t} + 1 \right) + (1 - s_t)e^{(V_{t+1}-c)/\lambda_t} \right).
\] (34)

Using the first-order condition together with (34), it is easy to check that if the solution to (33) satisfies \( s_t \in (0, 1) \) then \( V_{t+1} = V_t + c \), and that

\[
V_t = \lambda_t \log \left( \frac{1}{2} \left( e^{1/\lambda_t} + 1 \right) \right)
\] (35)

whenever \( s_t \in (0, 1) \).

Since it can never be optimal to delay with probability one, there exists some \( t^* \in \{1, \ldots, T\} \) such that \( s_t \in (0, 1) \) for all \( t < t^* \), and \( s_{t^*} = 1 \), meaning that the agent always makes a terminal decision by time \( t^* \). In addition, there exists some \( V_1 \) such that \( V_t = V_1 + c(t-1) \) for each \( t = 1, \ldots, t^* \). Substituting into (35), we see that \( \lambda_t \) is decreasing in time, as claimed.

Recall that \( p_t = \Pr(a_t = \theta \mid a_t \in \{0, 1\}) \) denotes the accuracy of the terminal decision when it is made at time \( t \). From (7), we obtain \( p_t = \frac{e^{1/\lambda_t}}{1 + e^{1/\lambda_t}} \). Solving for \( \lambda_t \) and substituting into (35) gives

\[
\frac{\log (2(1 - p_t))}{\log \left( \frac{1 - p_t}{p_t} \right)} = V_1 + c(t - 1).
\] (36)

Next, recall that \( r_t = \Pr(a_{t-1} = w^{t-1} \text{ and } a_t \in \{0, 1\}) \) is the probability that a terminal decision is made in period \( t \). The value of \( r_t \) can be computed using the capacity constraints. For \( t \leq t^* \), the capacity constraint binds, and hence \( r_t \) satisfies

\[
r_t \left( \log 2 + p_t \log p_t + (1 - p_t) \log(1 - p_t) \right) = \kappa.
\] (37)

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

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

Notice that the solution of (30) also solves the original problem (20) because the constraints (31) are binding for $t = \{1, \ldots, t^*\}$, and hence the solution of (30) is also a solution of (30).

References


