Rational inattention model under pushed information

Yujue Wang[1] and Rui Yao*[1]

[1]Civil and Environmental Engineering, Technion

Short summary

This paper develops a rational inattention (RI) model that separates exogenous information pushed by external platforms from travelers' endogenous costly information processing. A platform first publishes a signal that shifts the traveler's prior into an interim belief; the traveler then solves an RI problem conditional on this interim belief. The RI value function is convex in the interim belief, generating non-learning regions where the traveler optimally ignores additional information. Following Caplin et al. (2019); Matyskova & Montes (2023), we derive closed-form threshold beliefs that partition the belief space into learning and non-learning regimes, and show that pushed signals improve welfare only when they shift interim beliefs across these thresholds, demonstrating an information paradox. In an illustrative example with garbled signals parameterized by noise level ϵ, we obtain analytical conditions on ϵ for welfare improvement. We further show that the sensitivity of unconditional choice probabilities to ϵ varies with the information cost λ, implying that systematic variation in signal precision can screen travelers' information costs.

Keywords: rational inattention; exogenous information; Bayesian updating; cognitive constraints; information design

1 Introduction

Digital navigation platforms have made traffic information an active input to travel decisions. High-resolution trajectories, roadside sensing, and connected vehicles support near real-time inference of congestion and incidents, and platforms routinely *push* signals through route recommendations, alerts, and warnings. Yet travelers cannot fully process what they receive, often related to their limited cognitive capacity. Empirical evidence shows that drivers often ignore traffic messages, inconsistently follow navigation recommendations, or selectively attend to certain types of information depending on the context (Imants et al. (2021); Harms (2023)). These patterns indicate selective attention rather than full information use.

This setting is not well captured by the standard assumption in random utility models that travelers have full knowledge of all relevant attributes and process this information without cost (McFadden, 1972, 1981; Train, 2009). Rational Inattention (RI) theory (Sims, 2006) provides an alternative: decision-makers optimally choose how much information to acquire by trading off expected utility gains against information-processing costs, typically measured by Shannon mutual information (Matějka & McKay, 2015; Maćkowiak et al., 2023). In transportation, RI has been used in route choice models to capture limited attention, generating stochastic and state-dependent behavior under travel time variability and uncertainty (Fosgerau & Jiang, 2019; Jiang et al., 2020; Zhou & Liu, 2024; Chen et al., 2025).

A remaining gap is that the standard RI formulation represents all information relevant for the decision through a single, endogenous information structure chosen by the decision-maker, and imposes costs only for its informativeness (Sims, 2006; Matějka & McKay, 2015). In this formulation, externally provided platform messages are not modeled as a distinct stage of belief updating separate from the traveler's subsequent, costly information processing. We address this by modeling travel decisions as a two-stage belief updating process. A platform first publishes an exogenous (pushed) signal that shifts the traveler's prior into an interim belief. Conditional on this interim belief, the traveler then decides whether to acquire additional information subject to RI costs

*Corresponding author: rui.yao@technion.ac.il

and chooses an action based on the resulting posterior. This separation lets us characterize how pushed information changes optimal attention through its effect on interim beliefs. In particular, interim beliefs define a *learning* region where the traveler acquires additional information and a *non-learning* region where she optimally acquires none. We use an illustrative example to show how pushed signals shift interim beliefs across these regions and how this drives welfare and choice behaviors.

Our analysis is related to Bayesian persuasion with costly information acquisition (Matyskova & Montes, 2023). Under Shannon mutual information costs, our model is a special case of their framework. We do not study optimal sender design; instead, we analyze how exogenous belief shifts interact with endogenous attention choices, and how this interaction shapes behavior and welfare.

2 Preliminary: Classic RI Model

We briefly review the classic rational inattention (RI) framework and fix notation. The next section extends this setup by introducing a pushed (exogenous) signal prior to the traveler's endogenous information acquisition.

Rational Inattention (RI), introduced by Sims (2006), models decision-making when information processing is costly. Under Shannon mutual information costs, Matějka & McKay (2015) show that optimal (state-contingent) choice probabilities take the form of a generalized logit form, and can be characterized without explicitly constructing the underlying signal structure or posterior beliefs.

In our context, consider a traveler choosing an action $j \in \mathcal{J}$ (e.g., choosing bus) under an uncertain state $x \in \mathcal{X}$, where $\mathcal{J}$ and $\mathcal{X}$ are finite. Each action yields a state-dependent utility $u_j(x) \in \mathbb{R}$. The traveler starts from a prior belief $\mu(x)$ and selects an information structure $p(s \mid x)$ that maps states to an internal signal $s \in S$. After observing s, she forms the posterior $\mu(x \mid s)$ via Bayes' rule and chooses accordingly a decision rule $p(j \mid s)$.

In the classic RI formulation, the information cost is proportional to mutual information, $\lambda I(\mathcal{X}; S) = \lambda(H(\mathcal{X}) - H(\mathcal{X}|S))$, where $\lambda > 0$ is the unit cost of information and $H(\cdot)$ is Shannon entropy. Following Matějka & McKay (2015), this cost can be written as:

$$C(p;\mu) = \lambda \left(\sum_{s \in S} \pi(s) H\big(p(\cdot \mid s)\big) - H\big(\mu(\cdot)\big) \right), \tag{1}$$

where $\pi(s) = \sum_{x \in \mathcal{X}} \mu(x) p(s \mid x)$ is the unconditional probability, and Shannon entropy function $H(q(\cdot)) = -\sum_x q(x) \log q(x)$. Accordingly, the traveler chooses information structure $p(s \mid x)$ to maximize expected utility net of information costs:

$$\max_{p(S|\mathcal{X})} U(p) \coloneqq \sum_s \sum_x \mu(x) p(s \mid x) \sum_j p(j \mid s) u_j(x) - C(p;\mu) = \widetilde{U}(p;\mu) - C(p;\mu), \tag{2}$$

where $\widetilde{U}(p;\mu)$ is the expected payoff.

A useful implication of optimality in Eq. (2) is that attention is not wasted on signals that do not affect behavior: one can restrict attention to information that is only relevant for distinguishing actions Matějka & McKay (2015). Consequently, this leads to an equivalent formulation in terms of (state-contingent) choice probabilities $P(j \mid x)$ alone:

$$\max_{P(\mathcal{J}|\mathcal{X})} U(P) \coloneqq \sum_x \mu(x) \sum_j u_j(x) P(j \mid x) - C\big(P(\mathcal{J} \mid \mathcal{X}); P(\mathcal{J})\big), \tag{3}$$

where $P(j) = \sum_x P(j \mid x)\mu(x)$ is the unconditional probability of choosing action j. Under mutual information costs, the optimal solution takes the generalized logit form (Matějka & McKay, 2015):

$$P(j \mid x) = \frac{P(j) \exp\left(\frac{u_j(x)}{\lambda}\right)}{\sum_{j'} P(j') \exp\left(\frac{u_{j'}(x)}{\lambda}\right)}. \tag{4}$$

Fosgerau et al. (2020) show that a similar choice form exists for generalized entropy functions.

We note that, in Eq. (4), choices depend on both state-dependent utilities $u_j(x)$ and the endogenous weights $P(j)$, which capture the ex ante action frequencies and reflect optimal attention allocation. The parameter λ governs responsiveness to utility differences: as $\lambda \to 0$, choices converge to a deterministic best response, while larger λ leads to weaker state dependence and more dispersed choice probabilities. Thus, the RI model can be interpreted as a logit choice model with endogenous choice weights, where stochasticity arises from optimal endogenous information processing rather than random shocks.

3 Rational Inattention Under Pushed Information

We extend the classic RI model by introducing a pushed (exogenous) signal released by a platform before the traveler undertakes any costly information processing. The traveler's decision problem therefore has two stages. In the first stage, the traveler receives a pushed signal from the platform and updates her prior belief into an interim belief. In the second stage, conditional on this interim belief, she endogenously decides additional information acquisition subject to RI costs and makes choices based on the posterior belief. The overall decision process is illustrated in Fig. 1. We refer detailed derivations of our model to the full paper.

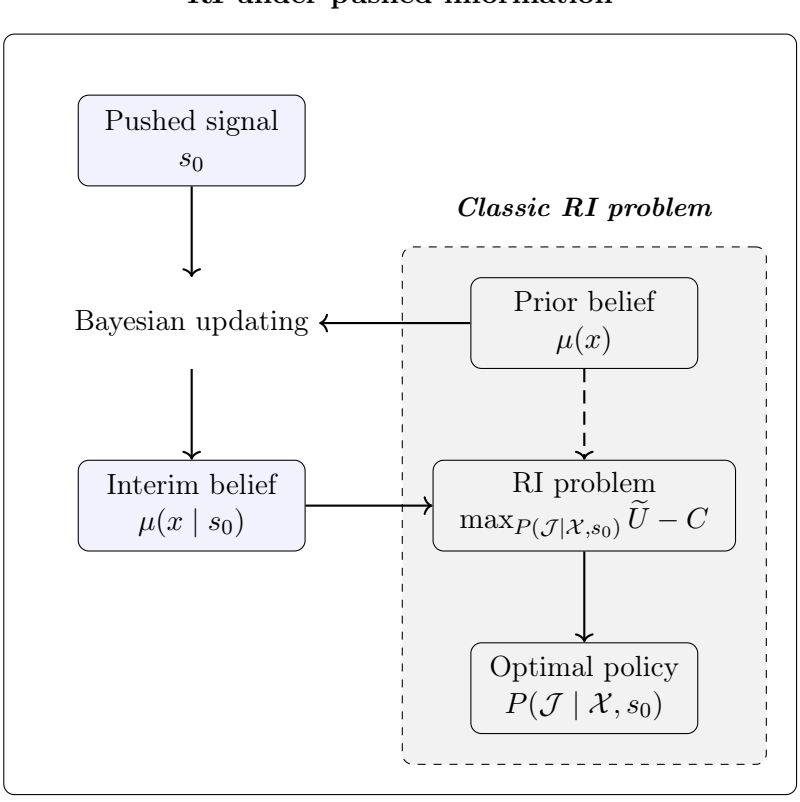


Figure 1: Decision process in the RI model under pushed information.

In our extended framework, before choosing, the traveler receives a pushed signal $s_0 \in S_0$ generated according to a publishing rule $\kappa(s_0 \mid x)$. The signal s_0 induces the interim belief:

$$\mu(x \mid s_0) = \frac{\mu(x)\kappa(s_0 \mid x)}{\pi(s_0)}, \qquad \pi(s_0) = \sum_{x \in \mathcal{X}} \mu(x)\kappa(s_0 \mid x). \tag{5}$$

where $\pi(s_0)$ is the ex ante (unconditional) probability that the platform sends pushed signal s_0.

In the second stage, conditional on the pushed signal s_0, the traveler chooses an endogenous information structure $p(s \mid x, s_0)$ over internal signals $s \in S$. Upon observing s, she forms the

posterior $\mu(\cdot \mid s, s_0)$ via Bayes' rule and selects an action j according to a decision rule $a(j \mid s, s_0)$. Her conditional expected utility

$$\sum_{j \in \mathcal{J}} a(j \mid s, s_0) \sum_{x \in \mathcal{X}} \mu(x \mid s, s_0)\, u_j(x)$$

is linear in $a(\cdot \mid s, s_0)$, so an optimum assigns probability one to a payoff-maximizing action. Hence, the maximized posterior expected payoff is

$$A(s \mid s_0) \;\coloneqq\; \max_{j \in \mathcal{J}} \sum_{x \in \mathcal{X}} \mu(x \mid s, s_0)\, u_j(x). \tag{6}$$

The ex ante expected payoff, taken over all signal realizations, is

$$\widetilde{U}\big(p(\cdot \mid \cdot, s_0); \mu(\cdot \mid s_0)\big) \;=\; \sum_{s \in S} \pi(s \mid s_0)\, A(s \mid s_0), \qquad \pi(s \mid s_0) \;=\; \sum_{x \in \mathcal{X}} \mu(x \mid s_0)\, p(s \mid x, s_0). \tag{7}$$

The traveler chooses $p(s \mid x, s_0)$, the endogenous information structure, to maximize expected payoff net of information costs:

$$\max_{p(\mathcal{S} \mid \mathcal{X}, s_0)} \; \widetilde{U}\big(p(\cdot \mid \cdot, s_0); \mu(\cdot \mid s_0)\big) \;-\; C\big(p(\cdot \mid \cdot, s_0); \mu(\cdot \mid s_0)\big), \tag{8}$$

where $C(\cdot)$ is the mutual information cost as in the classic model, now evaluated at the interim belief $\mu(\cdot \mid s_0)$. This is the key structural consequence of pushed information: the platform's signal enters the traveler's RI problem entirely through its effect on the interim belief.

As in the classic RI model, one can work with an equivalent action-based formulation. Under an optimal policy, information that does not affect actions is irrelevant, so the mutual information between states and internal signals can be replaced by the mutual information between states and actions conditional on s_0, namely, at optimal

$$I(\mathcal{X}; S \mid s_0) = I(\mathcal{X}; \mathcal{J} \mid s_0).$$

Let $P(j \mid x, s_0)$ denote the state-contingent choice probabilities after observing s_0, and let $P(j \mid s_0) = \sum_x \mu(x \mid s_0) P(j \mid x, s_0)$ be the corresponding unconditional probabilities. Invoking the chain rule of the mutual information yields the following equivalent action-based formulation:

$$\max_{P(\mathcal{J} \mid \mathcal{X}, s_0)} U(P; \mu) := \sum_x \sum_j \mu(x \mid s_0) P(j \mid x, s_0) u_j(x) - C(P(\mathcal{J} \mid \mathcal{X}, s_0); P(\mathcal{J} \mid s_0)), \tag{9}$$

subject to:

$$\sum_j P(j \mid x, s_0) = 1, \; \forall x, \tag{10}$$

$$P(j \mid x, s_0) \geq 0, \; \forall j, x. \tag{11}$$

With Shannon costs,

$$\begin{aligned} C\big(P(\mathcal{J} \mid \mathcal{X}, s_0); P(\mathcal{J} \mid s_0)\big) &= \lambda\Big(H(\mathcal{J} \mid s_0) - H(\mathcal{J} \mid \mathcal{X}, s_0)\Big) \\ &= \lambda\left(\sum_x \mu(x \mid s_0) \sum_j P(j \mid x, s_0) \ln P(j \mid x, s_0) - \sum_j P(j \mid s_0) \ln P(j \mid s_0)\right). \end{aligned} \tag{12}$$

We note that Eq. (9) shares a similar structure compared with the classic RI model. However, the prior belief $\mu(x)$ is now replaced by the interim belief $\mu(x \mid s_0)$ that is induced by the platform's publishing strategy κ.

Remark 1 (Reduction to classic RI)**.** *If the pushed signal is uninformative, i.e., $\kappa(s_0 \mid x)$ does not depend on x, then $\mu(\cdot \mid s_0) = \mu(\cdot)$, interim belief reduces to the prior, and* (9) *reduces to the classic RI problem* (3)*.*

Proposition 1 (Existence)**.** *Define $0 \ln 0 = 0$, RI problem* (9) *admits at least one solution.*

Proof sketch. We prove existence by showing the objective $U(\cdot;\mu)$ is continuous on the simplex, which is compact. Existence then follows immediately. □

Proposition 2 (Uniqueness)**.** *Suppose there exists $x \in \mathcal{X}, j \in \mathcal{J}$ such that $P(j \mid x, s_0) \neq P(j \mid s_0)$ and assumptions in Prop. 1 hold, there exists a unique solution to the RI problem* (9)*.*

Proof sketch. Our proof follows similarly to (Fosgerau et al., 2020). The payoff term in program (9) is linear in $P(j \mid x, s_0)$, so uniqueness reduces to strict convexity of the information cost. By a standard Cauchy–Schwarz argument on the second-order variation of mutual information, C is strictly convex whenever $P(j \mid x, s_0) \neq P(j \mid s_0)$ for some x, j, i.e., whenever some action is state-dependent. □

Solution algorithm

A direct analysis on the KKT conditions of RI problem (9) yields the optimal choice probabilities:

$$P(j \mid x, s_0) = \frac{P(j \mid s_0) \exp\left(\frac{u_j(x)}{\lambda}\right)}{\sum_{j'} P(j' \mid s_0) \exp\left(\frac{u'_j(x)}{\lambda}\right)}. \tag{13}$$

This is a generalized logit with the same structure as the classic RI solution (4), but the unconditional weights $P(j \mid s_0)$ now depend on the interim belief $\mu(\cdot \mid s_0)$.

Let $q_j \equiv P(j|s_0)$ and denote $q \coloneqq (q_j)_{j\in\mathcal{J}} \in \Delta(\mathcal{J})$, where $\Delta(\mathcal{J})$ is the probability simplex. Substituting (13) into the Bayesian plausibility condition

$$q_j = \sum_x \mu(x \mid s_0) P(j \mid x, s_0)$$

yields the fixed point condition:

$$q = Tq \tag{14}$$

where the operator $T : \Delta(\mathcal{J}) \to \mathbb{R}_+^{|\mathcal{J}|}$ is defined component-wise by

$$Tq(j) \;=\; q_j \sum_x \frac{\mu(x \mid s_0)\, \exp\left(\frac{u_j(x)}{\lambda}\right)}{\sum_{j'} q_{j'}\, \exp\left(\frac{u_{j'}(x)}{\lambda}\right)}, \quad \forall j \in \mathcal{J}. \tag{15}$$

Similar to Caplin et al. (2019), a sufficient and necessary condition for optimality is

$$Tq^* \leq 1, \tag{16}$$

with equality holds if $q_j > 0$. The solution method is summarized in Algorithm 1. We provide detail proof on convergence in the full paper.

Algorithm 1 RI Fixed-Point Iteration with Pushed Information

Require: Prior $\mu(x)$, pushed signal kernel $\kappa(s_0 \mid x)$, utility $u_j(x)$, unit cost λ, tolerance $\delta > 0$

Update interim belief $\mu(x \mid s_0)$ via Eq. (5) (Bayesian update)
Initialize $q_j^{(0)} = 1/|\mathcal{J}|, \forall j$ (uniform initialization)
for $n = 0, 1, 2, \ldots$ **do**
  $q^{(n+1)} \leftarrow Tq^{(n)}$ (fixed-point iteration)
  if $\|q^{(n+1)} - q^{(n)}\|_\infty < \delta$ **and** Eq. (16) holds **then**
    break
  end if
end for
return $q^{(n+1)}$

4 Illustrative examples

To illustrate the properties of the proposed model, we consider a simple mode choice problem with two actions under two states. A traveler chooses between two modes: car (j_1) and metro (j_2). The optimal choice depends on the congestion level of the road network, which can be either *low traffic* (x_1) or *high traffic* (x_2). Driving is attractive when traffic is light, but becomes less desirable when congestion is severe, in which case taking the metro is preferable.

The prior belief over states is given by $[\mu(x_1), \mu(x_2)] = [0.4, 0.6]$, indicating that high traffic is considered more likely a prior. The information cost parameter is set to $\lambda = 1$. The state-dependent utilities are given as

$$u_{j_1}(x_1) = -1;\ u_{j_1}(x_2) = -5;\ u_{j_2}(x_1) = -3;\ u_{j_2}(x_2) = -4.$$

Under this specification, each mode is optimal in a particular state: driving is preferable when traffic is low, while the metro performs better under congestion. Let

$$\Delta_1 = u_{j_1}(x_1) - u_{j_2}(x_1) > 0, \quad \Delta_2 = u_{j_2}(x_2) - u_{j_1}(x_2) > 0$$

denote the utility advantages of each action in its favorable state. In this case, $\Delta_1 = 2$ and $\Delta_2 = 1$. Since the traveler does not directly observe the state, she faces a trade-off between acquiring costly information and deciding under uncertainty.

The platform publishes binary signals $s_0 \in S_0$, $|S_0| = 2$, according to a publishing strategy $\kappa(s_0 \mid x)$. For a given $\kappa(s_0 \mid x)$, let $V(\mu_\kappa) \coloneqq \max_{P(J|\mathcal{X},s_0)} U(P;\mu_\kappa)$ denote the RI value function under interim belief $\mu_\kappa(\cdot|s_0)$ induced by κ, the traveler's expected welfare is

$$\mathcal{W}(\kappa) = \sum_{s_0} \pi_\kappa(s_0) V(\mu_\kappa). \tag{17}$$

Benchmark cases: uninformative and perfect information

We begin by establishing two benchmarks that bound the welfare effects of pushed information: the case of no information and the case of perfect information.

When pushed signals are uninformative, i.e., $\kappa(s_0 \mid x) = 1/2$ for all x, the interim belief coincides with the prior and the model reduces to the classic RI problem. Solving the traveler's optimization yields the choice probabilities $P(j_1) = 0.4611$, $P(j_2) = 0.5389$, and welfare $\mathcal{W} = -3.2534$. The traveler slightly favors the metro, reflecting that high traffic is considered more likely under the prior, but the choice remains stochastic due to costly information processing.

At the other extreme, suppose pushed signals perfectly reveal the true state, i.e., $\kappa(s_0 \mid x) = \mathbb{I}$ is the identity matrix. The traveler then learns the traffic state before choosing and no longer needs to acquire additional information. The optimal action is chosen deterministically in each state: the ex ante choice probabilities coincide with the prior state probabilities $P(j_1) = \mu(x_1) \times 1 + \mu(x_2) \times 0 = 0.4$, $P(j_2) = \mu(x_1) \times 0 + \mu(x_2) \times 1 = 0.6$, and welfare improves to $\mathcal{W} = -2.800$, which serves as an upper bound on the value of information.

The natural question is what happens between these two extremes, i.e., when signals are partially informative. To answer this, we first characterize how the RI value function depends on the interim belief.

Value function structure and non-learning regions

Since the pushed signal shifts the prior into an interim belief $\mu(\cdot \mid s_0)$, the welfare under any publishing strategy depends on the shape of the RI value function $V(\mu_\kappa)$. Hence, we start by studying the impact of interim belief on value function $V(\mu_\kappa)$. Fig. 2 shows $V(\mu_\kappa)$ with respect to varying $\mu(x_1 \mid s_0)$.

A notable feature is that the value function becomes locally linear near extreme beliefs, i.e., when $\mu(x_1 \mid s_0) \to 0$ or $\mu(x_1 \mid s_0) \to 1$. In these regions, the traveler is sufficiently confident about the optimal action that the marginal benefit of acquiring information falls below its cost. She

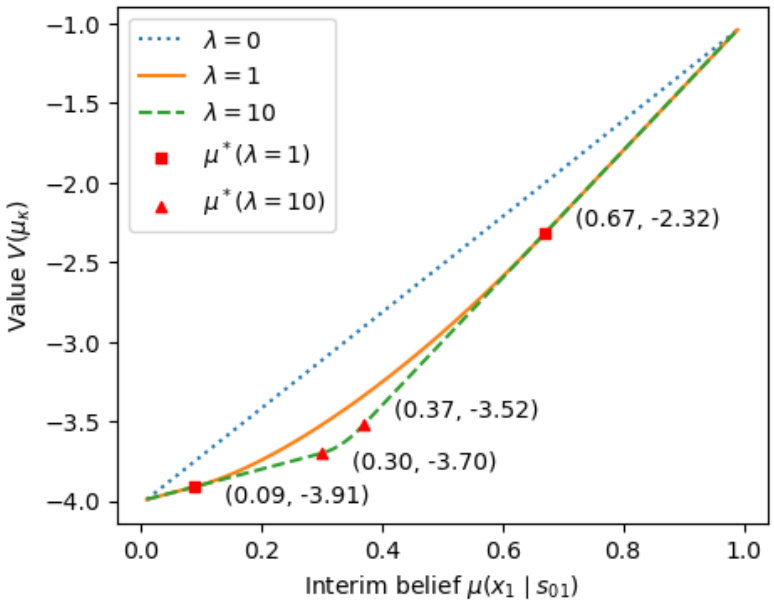


Figure 2: RI value function of the interim belief $\mu(x_1 \mid s_0)$.

optimally commits to a single action without learning, and the mutual information between actions and states is zero. We refer to such regions as *non-learning regions*. Under Shannon costs, Caplin et al. (2019); Matyskova & Montes (2023) show that, by using the fixed likelihood ratio property, the non-learning region of action j can be characterized by

$$\sum_x \mu(x|s_0) \exp\left(\frac{u_{j'}(x) - u_j(x)}{\lambda}\right) \leq 1, \quad \forall j' \neq j. \tag{18}$$

The boundary at which (18) holds with equality for some $j' \neq j$ defines the *threshold interim belief* μ_j^* for action j. In the two-state, two-action setting, the thresholds are then given by

$$\mu_{j_1}^* = \frac{e^{\Delta_2/\lambda} - 1}{e^{\Delta_2/\lambda} - e^{-\Delta_1/\lambda}}, \quad \mu_{j_2}^* = \frac{1 - e^{-\Delta_2/\lambda}}{e^{\Delta_1/\lambda} - e^{-\Delta_2/\lambda}}. \tag{19}$$

Since $\mu_{j_1}^*/\mu_{j_2}^* = e^{\Delta_1/\lambda} > 1$, we have $\mu_{j_1}^* > \mu_{j_2}^*$. The belief space is thus partitioned into a learning region $(\mu_{j_2}^*, \mu_{j_1}^*)$, where the traveler acquires information and implements a state-contingent strategy, and non-learning regions $[0, \mu_{j_2}^*] \cup [\mu_{j_1}^*, 1]$, where she acts without learning. This partition determines whether pushed information can improve welfare: a signal is valuable only if it shifts the interim belief across a threshold, a key insight similar to Matyskova & Montes (2023).

Garbling, information paradox, and screening

With the threshold structure in place, we now analyze how signal quality affects welfare and choice behavior. To do so, we parameterize the publishing strategy as a *garbling* of the perfectly informative signal $\tilde{s}_0$ through a binary symmetric channel with parameter $\epsilon \in (0, 1/2]$ given by

$$\kappa(\tilde{s}_0 \mid x) = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}.$$

The garbled signals represent noisy or imperfect information, such as delayed congestion reports, inaccurate travel time predictions, or probabilistic route recommendations (e.g., "possible congestion ahead"), which only partially reflect the true state. As ϵ increases, the signal becomes less informative and the interim beliefs are pulled toward the prior. When $\epsilon = 1/2$, the signal is completely uninformative and the model collapses to the benchmark case. The induced interim beliefs are

$$\mu(x_1 \mid \tilde{s}_{01}) = \frac{(1-\epsilon)\mu(x_1)}{(1-\epsilon)\mu(x_1) + \epsilon\mu(x_2)}, \quad \mu(x_1 \mid \tilde{s}_{02}) = \frac{\epsilon\mu(x_1)}{\epsilon\mu(x_1) + (1-\epsilon)\mu(x_2)}. \tag{20}$$

Information paradox. The non-learning region structure reveals that the relationship between signal quality and welfare is more subtle. When the prior lies in a non-learning region, pushed

information improves welfare *only if* it shifts the interim belief across a threshold boundary. Since $\mu(x_1 \mid s_{01}) \geq \mu(x_1) \geq \mu(x_1 \mid s_{02})$, signal s_{01} shifts beliefs upward and s_{02} shifts them downward. For a prior $\mu \in [0, \mu^*_{j_2}]$, welfare improves only if $\mu(x_1 \mid s_{01}) > \mu^*_{j_2}$. The following conditions characterize the maximum admissible noise level. For $\mu \in [0, \mu^*_{j_2}]$, welfare improves if

$$\epsilon < \frac{\mu\big(e^{\Delta_1/\lambda} - 1\big)}{\mu\big(e^{\Delta_1/\lambda} - 1\big) + (1-\mu)\big(1 - e^{-\Delta_2/\lambda}\big)}. \tag{21}$$

For $\mu \in [\mu^*_{j_1}, 1]$, welfare improves if

$$\epsilon < \frac{(1-\mu)\big(e^{\Delta_2/\lambda} - 1\big)}{(1-\mu)\big(e^{\Delta_2/\lambda} - 1\big) + \mu\big(1 - e^{-\Delta_1/\lambda}\big)}. \tag{22}$$

These bounds are governed by a signal-to-noise ratio determined by the relative utility advantages across states. The numerator captures the strength of the action that is optimal in the relevant state, while the denominator aggregates contributions from both states. The maximum admissible noise level ϵ increases when the utility gap in the relevant state becomes larger, and decreases when the competing state becomes more influential in the prior.

When the prior is far from a threshold, the platform must provide a highly informative signal (low ϵ) to shift the interim belief across the boundary. When the prior is close to a threshold, even a relatively noisy signal can induce a regime switch. If both the prior and the induced interim beliefs remain within the same non-learning region, the pushed signal has no value and may even reduce welfare. This is the information paradox: more information does not necessarily improve welfare; it is beneficial only when it changes decisions. Fig. 3 illustrates the joint dependence of welfare

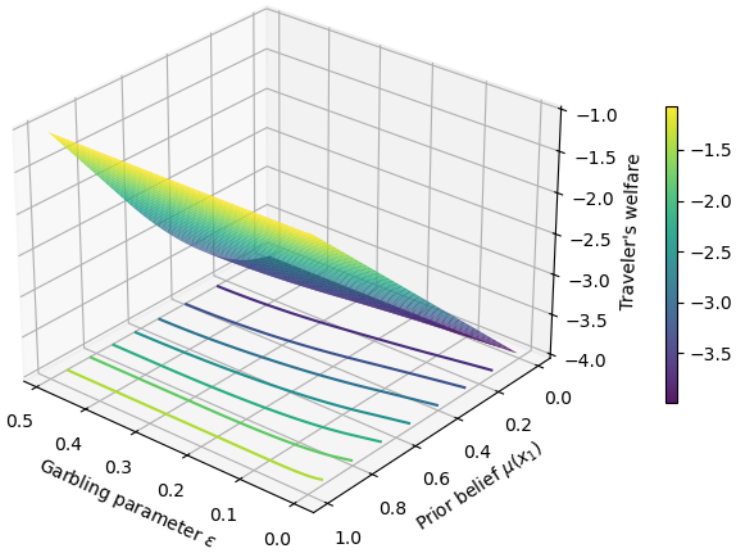


Figure 3: Traveler's welfare as a function of ϵ and prior belief $\mu(x_1)$.

on ϵ and the prior $\mu(x_1)$. For extreme priors (near 1), welfare is relatively high and insensitive to ϵ, because the traveler is already confident about the optimal action. For intermediate priors, lowering ϵ yields substantial welfare gains (indicated by crossing contours), as more precise signals help the traveler distinguish between states. The flat regions at extreme priors correspond to the non-learning regime, where pushed signals of moderate quality fail to shift the interim belief across a threshold.

Screening. The response of choice behavior to signal quality also depends on the traveler's information cost, suggesting a role for pushed signals as a screening tool. Fig. 4 shows the unconditional choice probability $P(j_1)$ (the overall frequency of choosing car) as a function of ϵ for different values of λ. In screening/identification applications, this is the quantity a platform would observe in aggregate travel data.

When λ is small, $P(j_1)$ responds smoothly and monotonically to changes in ϵ: the traveler processes all available information and adjusts behavior even when signals are noisy. When λ is large, $P(j_1)$

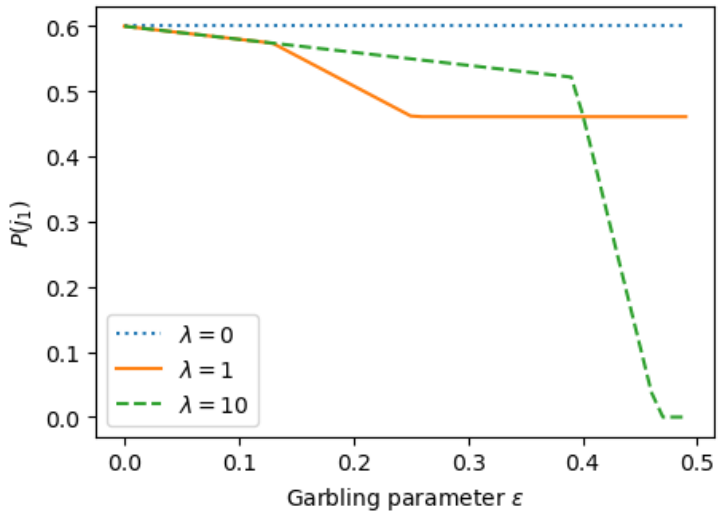


Figure 4: Unconditional choice probability $P(j_1)$ as a function of ϵ for $\lambda \in \{0, 1, 10\}$ while prior belief $[\mu(x_1), \mu(x_2)] = [0.4, 0.6]$.

remains (relatively) flat over a wide range of ϵ, meaning that the traveler is in the non-learning regime and ignores the pushed signal until the signal becomes sufficiently informative to cross a threshold where behavior shifts abruptly. This heterogeneity implies that the same variation in signal precision produces different behavioral responses across travelers with different information costs. A platform could exploit this by systematically varying ϵ and observing the resulting changes in aggregate mode shares, thereby inferring the distribution of information-processing costs in the population and designing targeted publishing strategies. This also implies one interesting application direction for the proposed model.

5 Conclusions

This paper extends the classic RI framework by separating exogenous pushed information from endogenous costly information acquisition. The platform's signal enters the traveler's RI problem entirely through its effect on the interim belief, and the resulting choice probabilities retain the generalized logit form of Matějka & McKay (2015). We establish existence and uniqueness of the solution and provide a fixed-point algorithm for computation.

The illustrative example reveals three main findings. First, the RI value function is convex in the interim belief, with non-learning regions where the traveler optimally ignores additional information, in line with Caplin et al. (2019) and Matyskova & Montes (2023). The threshold beliefs that separate learning from non-learning regimes are derived in closed form for the two-state, two-action case. Second, pushed signals improve welfare only when they shift the interim belief across a threshold, illustrating one mechanism underlying the information paradox. We derive analytical conditions on the garbling parameter ϵ that characterize when this occurs in terms of maximum admissible noise. Third, unconditional choice probabilities respond differently to signal quality depending on the information cost λ: low-λ travelers adjust smoothly, while high-λ travelers exhibit flat responses followed by abrupt shifts. This observation implies that systematic variation in signal precision can screen travelers' cognitive costs, providing a basis for personalized information provision or estimation.

Several limitations remain. Pushed information is treated as costless and fully attended; in practice, travelers may rationally ignore exogenous signals. Future work could consider alternative belief updating mechanism on pushed signals and other information cost structures beyond Shannon entropy.

References

Caplin, A., Dean, M., & Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, *86*(3), 1061–1094.

Chen, X., Jiang, G., & Jiang, Y. (2025). Multivariate discrete choice with rational inattention: Model development, application, and calibration. *Transportation Research Part E: Logistics and Transportation Review*, *194*, 103899.

Fosgerau, M., & Jiang, G. (2019). Travel time variability and rational inattention. *Transportation Research Part B: Methodological*, *120*, 1–14.

Fosgerau, M., Melo, E., de Palma, A., & Shum, M. (2020). Discrete choice and rational inattention: A general equivalence result. *International Economic Review*, *61*(4), 1569-1589. doi: https://doi.org/10.1111/iere.12469

Harms, I. M. (2023). Distracted by familiarity: Implications of 'autopilot'as a default cognitive mode. *Transportation research part F: traffic psychology and behaviour*, *99*, 274–288.

Imants, P., Theeuwes, J., Bronkhorst, A., & Martens, M. (2021). Effect of multiple traffic information sources on route choice: A driving simulator study. *Transportation research part F: traffic psychology and behaviour*, *81*, 1–13.

Jiang, G., Fosgerau, M., & Lo, H. K. (2020). Route choice, travel time variability, and rational inattention. *Transportation Research Part B: Methodological*, *132*, 188–207.

Matějka, F., & McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, *105*(1), 272–298.

Matyskova, L., & Montes, A. (2023). Bayesian persuasion with costly information acquisition. *Journal of Economic Theory*, *211*, 105678.

Maćkowiak, B., Matějka, F., & Wiederholt, M. (2023). Rational inattention: A review. *Journal of Economic Literature*, *61*(1), 226–273.

McFadden, D. (1972). Conditional logit analysis of qualitative choice behavior.

McFadden, D. (1981). Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, *198272*.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, *96*(2), 158–163.

Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.

Zhou, B., & Liu, R. (2024). A generalized rationally inattentive route choice model with non-uniform marginal information costs. *Transportation Research Part B: Methodological*, *189*, 102993.