Paying for Inattention

Ala Avoyan* and Giorgia Romagnoli†

August 27, 2019

Abstract

We extend the discrete-choice rational inattention model to the case in which the decision maker can influence the payoff distribution across states. By reducing the gap between payoffs in different states, the decision maker is able to affect her own incentives to pay attention. The smaller the gap, the less attentive the decision maker needs to be. This new framework with endogenous incentives allows to derive a novel method for eliciting the attention level solely by observing the decision maker's incentive redistribution choice. As a result, we have two methods for observing the same variable of interest—targeted success probability: (i) through actual performance (method used in the literature); and (ii) through our model estimation, using incentive redistribution choices. Having two ways of identifying the targeted probability of success allows us to test rational inattention models without making any assumptions on the cost of attention function. Furthermore, the new framework allows us to identify and test the properties of the attention cost function.

JEL Classification: C72, C91, C92, D83; **Keywords:** testing rational inattention, experiment, endogenous incentives.

*Department of Economics, Indiana University. E-mail: aavoyan@iu.edu

†Department of Economics, University of Amsterdam. E-mail: g.romagnoli@uva.nl

1 Introduction

The seminal work of Sims (1998, 2003) built the foundations of the concept of rational inattention (RI), the idea that the decision maker optimally chooses the level of attention based on a trade-off between attention costs and expected monetary gains. This concept has been explored theoretically and applied to various areas of economics.[1] In the typical setup of RI, the decision maker chooses among actions with state-dependent payoffs and she is uncertain about the actual state realization (Matejka and McKay (2014), Caplin and Dean (2015)). By being more attentive the agent can refine her posterior distribution over states and improve her choice of the payoff-maximizing action. Models in this literature have typically assumed that the agent takes the payoff structure as exogenously given. That is, while attention is endogenous, incentives to be attentive have been kept outside of the control of the decision maker.

In this paper, we extend the discrete-choice RI model to the case in which the decision maker can influence the payoff distribution across states. There is an interesting interplay between the choice of attention levels and the choice of payoff distribution. The extreme example of full equalization of payoffs over all states leaves the agent with no incentive to exert costly attention and learn the true state. In general, the more the payoffs are smoothed out across states, the lower the incentives to pay attention and refine beliefs about the true state. By observing how the decision maker chooses to redistribute payoffs across states, we can identify the underlying attention choices and infer the underlying unobservable attentional costs.

The motivation for this work is twofold. First, the model seeks to capture the real-life circumstances in which the decision maker chooses his own incentives to exert costly attention. Examples include insurance choices (with full-coverage contracts reducing the incentives to pay attention, compared to partial coverage), and financial choices (certain portfolios requiring more monitoring than

[1] See, for example, Woodford (2008), Bartoš et al. (2016), Mondria et al. (2010), Matějka (2015), Matějka and Tabellini (2017), Martinelli (2006), Martin (2017), Mackowiak and Wiederholt (2009), Kacperczyk et al. (2016), Gaballo (2016), De Oliveira et al. (2017), Caplin et al. (2014), Andrade and Le Bihan (2013).

others), among others. Second, our new framework provides a more stringent test of the rational inattention model. The existing tests of rational inattention can be categorized into two strands. On the one hand, there are tests concerned with special versions of the rational inattention model where certain parametric assumptions are made about the attentional cost function. These tests posit one functional form against the other, or derive, from the chosen functional form, additional predictions that allow to compare rational inattention to other paradigms, such as the random utility models (see Dean and Neligh (2017)). These tests are well fit to compare various theories; however, parametric assumption are needed and testing general predictions are not feasible with this method.

There are studies concerned with the general rational inattention model, making minimal assumptions about the shape of the cost function. Caplin and Dean (2015) show that, under these minimal assumptions, as long as attention is increasing in incentives, the RI model holds in that a cost function can be found which rationalizes the data. These assumptions assess monotonic response of attention to incentives and therefore corroborate RI. However, the generality and unobservability of the cost function makes these types of tests rather weak. Attention may be monotonically responsive to incentives, while systematic departures from the assumption of optimal trade-off, which is at the heart of the RI model, could still be in place. As a proof of concept, imagine a decision maker who, faced with two tasks, say A and B, systematically underestimates the attention required to succeed in task A and overestimate the attention required by task B. As a result, this decision maker may allocate less attention than it is optimal to task A and more attention to task B. This would violate the optimality of the trade-off inherent in the RI paradigm. However, so long as the decision maker monotonically responds to incentives, her non-optimal behavior would still be rationalized within the RI paradigm.

By endogenizing payoffs, in this paper we elicit an additional piece of information that we use to make predictions about the level of precision that the decision maker is planning to obtain from her chosen level of attention (as well as an estimation of the attentional costs associated to that targeted level of precision). We then compare these estimates to the actual level of precision attained during the execution of the task. This comparison allows for a stringent test of the RI model without making restrictive assumptions about the parametric form of the attentional costs function. This exercise enables the detection of behavioral biases (i.e. overconfidence) within the attentional framework. Besides testing the premises of the RI paradigm, our model also allows us to estimate, in a parameter-free way, the shape of the cost function and therefore provide a test for the most commonly used functional forms in the attention literature, such as the Shannon capacity (Matejka and McKay (2014)).

In this paper, we present the extension of the RI model with endogenous payoffs followed by the identification of attentional choices and costs that this model enables. Next, we conduct a lab experiment to validate the method. We find that 94% of subjects choose to smooth out payoffs at least once. Moreover, in 87% of all possible cases payoff redistribution was strictly positive. We conclude that subjects are willing to pay to avoid paying more attention. Some even fully insure by completely equalizing the payoffs in both states, and thus, avoiding the need to pay attention to the task at all. Moreover, we investigate attentional choices during the task, after payoffs have been decided. We observe that attention is monotonically increasing in incentives, confirming that our data passes the standard tests of the general RI model currently in use in the literature (Caplin and Dean (2015)).

Finally, we provide an exploration of our test of RI. We measure attention in three different ways. First, as it is standard, we observe actual performance in the task. Second, we use insurance choices and estimate the targeted probability of success through our model. Finally, we ask subjects to provide their own estimate of how well they think they did, which provides us with an ex-post subjective measure of the targeted probability of success. We find that model-estimated probability has a significant and sizable prediction power for performance. At the same time, we observe that targeted and observed probabilities do not fully align.

2 The Model

In the typical setup of the costly information acquisition model (Sims (2003)), the decision maker chooses among actions with state-dependent payoffs and she is uncertain about the actual state realization (Matejka and McKay (2014), Caplin and Dean (2015)).

Our model builds on the costly information acquisition framework of Matejka and McKay (2014) and Caplin and Dean (2015) and endogenizes the payoff structure. Here the decision maker maximizes utility by choosing an information structure and a payoff redistribution function at the same time. The decision maker's action set is $A = \{1, ..., N\}$. The state of nature is a vector $\mathbf{v} \in S \subset \mathbb{R}^N$, where v_i is the payoff of action $i \in A$ (let $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_M\}$) and $\mu \in \Gamma$ is the prior belief with Ω^μ support and $\gamma \in \Gamma$ posterior belief. Π^μ is the set of feasible attention strategies of all mappings $\pi : \Omega^\mu \to \Delta(\Gamma)$ with finite support and that satisfy Bayes' law. Let $\Pi = \cup_\mu \Pi^\mu$. Let $G : \Gamma \times \Pi \to \mathbb{R}$ be the gross payoff of using some information structure:

$$G(\mu, \pi) = \sum_{\gamma \in \Gamma(\pi)} \left[\sum_{j=1}^{M} \mu(\mathbf{v}_j)\pi(\gamma|\mathbf{v}_j) \right] g(\gamma) \tag{1}$$

where

$$g(\gamma) = \max_{i \in A} \sum_{i=1}^{M} \mathbf{v}_i \, \gamma(\mathbf{v}_i)$$

The gross expected payoff in (1) can be rewritten in terms of induced probability of selecting an action given a state. That is, let $\mathcal{P}_i(\mathbf{v})$ be

$$\mathcal{P}_i(\mathbf{v}) := \sum_{\gamma \in S_i} \pi(\gamma|\mathbf{v})$$

where S_i is a set of all signals for which optimal action is i (formally, $S_i := \{\gamma \in \mathbb{R}^N : \arg\max_{j \in A} \mathbb{E}_{\pi(\gamma|\mathbf{v})}[v_j] = i\}$). Again, $\mathcal{P}_i(\mathbf{v})$ is the induced probability of selecting action i conditional on state $\mathbf{v}$. Expected payoff in terms of $\mathcal{P}_i(\mathbf{v})$ is as follows

$$G(\mu, \mathcal{P}) = \sum_{i=1}^{N} \sum_{j=1}^{M} v_j^i \mathcal{P}_i(\mathbf{v_j})\mu(\mathbf{v_j})$$

and $\mathcal{P} := \{\mathcal{P}_i(\mathbf{v})\}_{i=1}^N$. Optimal choice given cost K:

$$\hat{\mathcal{P}}(\mu, K) = \arg\max_{\mathcal{P}}\{G(\mu, \mathcal{P}) - K(\mu, \mathcal{P})\}$$

Introducing Redistribution. Now, let us introduce the payoff redistribution function $X : \mathbb{R}^{N\times M} \to \mathbb{R}^{N\times M}$. Let V be $(N \times M)$ matrix such that V_{ij} is agent's payoff for action i and state $\mathbf{v}_j$. Function X redistributes V to $\tilde{V}$, $X(V) = \tilde{V}$.

$$G(\mu, \mathcal{P}, X) = \sum_{i=1}^{N}\sum_{j=1}^{M} \tilde{V_{ij}}\mathcal{P}_i(\mathbf{v_j})\mu(\mathbf{v_j})$$

Optimal information structure and optimal redistribution given attention cost and redistribution cost:

$$(\hat{\mathcal{P}}, \hat{X}) := \arg\max_{(\mathcal{P},X)} \{G(\mu, \mathcal{P}, X) - K(\mu, \mathcal{P}) - \mathcal{Q}(X)\} \tag{2}$$

2.1 Matching-the-state task with endogenous payoffs

In this section, we focus on the simplest environment that delivers our identification results. This is also the setup that most closely matches our experimental design. The setup is a symmetric two-state and two-action environment, with matching-the-state task. In the Appendix A, we show how our method can be easily extended to the case with N states, generic prior distribution and potentially asymmetric attention costs.

Consider the case in which there are two states of the world $S = \{\mathbf{w}, \mathbf{b}\}$ to which the agent's prior assigns *equal probability*. The agent can take two possible actions $A = \{W, B\}$. The initial payoff structure describes a typical "matching-the-state" task: An agent receives payoff Y if she matches the state (that is, she picks action W in state $\mathbf{w}$ and action B in state $\mathbf{b}$) and 0 otherwise. An agent can choose to change the payoff structure by purchasing insurance that guarantees a portion x of total payoff Y (that is, a dollar amount equal to xY) in the event she fails to match the state. Insurance costs are given by the function $q(x)$. Hence, the initial payoff structure is given by the matrix V, where the element v_{ij} is the payoff of playing action $i \in \{W, B\}$ in state $j \in \{\mathbf{w}, \mathbf{b}\}$.

$$V = \begin{bmatrix} Y & 0 \\ 0 & Y \end{bmatrix}$$

By purchasing insurance $x \in [0, 1]$ at cost $q(x)$, the agent has the ability to alter the initial payoff matrix and replace it with any of the following final payoff matrices $\tilde{V}(x)$ net of insurance costs:

$$\tilde{V}(x) = \begin{bmatrix} Y - q(x) & xY - q(x) \\ xY - q(x) & Y - q(x) \end{bmatrix}$$

By paying attention, the agent targets a certain posterior probability of matching the state $\{\mathcal{P}_W, \mathcal{P}_B\}$, where $\mathcal{P}_W$ is the probability of taking action W when the state is $\mathbf{w}$ and $\mathcal{P}_B$ is the probability of taking action B when the state is $\mathbf{b}$. We assume that (i) only symmetric posterior probabilities can be selected (that is, agents can only choose posterior probabilities such that $\mathcal{P}_W = \mathcal{P}_B = \mathcal{P}$ and (ii) posterior probabilities $\mathcal{P}$ that can be chosen by the agent vary in the continuum, $\mathcal{P} \in [0, 1]$. Let $K(\mu, \mathcal{P})$ be the *mental cost function*, which is the cost of the attention required to target a posterior probability $\mathcal{P}$ of matching the state[2]; $K(\mu, \mathcal{P})$ is assumed to be differentiable.

The agent has utility u defined over monetary outcomes and is rationally inattentive, that is, she chooses x and $\mathcal{P}$ by taking into account both the monetary rewards and her mental costs. The objective function is given by the expected utility of the monetary payoffs as given by matrix $\tilde{V}(x)$, weighted by the probability of success $\mathcal{P}$, minus the mental costs of targeting $\mathcal{P}$.

[2] Mental effort is conceived of as the minimum utility loss required to target a probability of success of $\mathcal{P}$ or larger.

Identification of $\mathcal{P}$

An external observer cannot identify the targeted probability of success $\mathcal{P}$ or the mental cost function $K(\mu, \mathcal{P})$ from the standard maximization problem with exogenous payoffs.[3] Here we show how, in the augmented maximization problem, where the payoff distribution is made endogenous, $\mathcal{P}$ and $K'_{\mathcal{P}}(\mu, \hat{\mathcal{P}})$ can be inferred from observing x and $q(x)$.

The agent's maximization problem 12 with 2 states, 2 actions and risk-neutral utility function leads to the first order conditions:[4]

$$(1 - \mathcal{P})Y - q'(x) = 0 \tag{3}$$

$$Y - xY - K'_{\mathcal{P}}(\mu, \mathcal{P}) = 0 \tag{4}$$

Let $(x^*_q, \hat{\mathcal{P}}_q)$ be the solution to the FOC conditions given by 3 and 4 and let the pair $\left(x^*_q, q'(x^*_q)\right)$ be the observable output of the decision problem. The FOC can be rewritten in terms of observables as follows:

$$\hat{\mathcal{P}}_q = 1 - q'(x^*_q)Y^{-1} \tag{5}$$

$$K'_{\mathcal{P}}(\mu, 1 - q'(x^*_q)Y^{-1}) = Y(1 - x^*_q) \tag{6}$$

Hence, the optimal probability of success, $\hat{\mathcal{P}}_q = 1 - q'(x^*_q)Y^{-1}$, and the value of the derivative of the mental cost function at that point, $K'_{\mathcal{P}} = Y(1 - x^*_q)$, can all be expressed in terms of observables. By adopting a family of insurance cost functions $q(x) \in \mathbb{Q}$ and observing one choice of insurance level for each cost function in set $\mathbb{Q}$, we obtain values of the derivative of the cost function $K'_{\mathcal{P}}(\mu, \hat{\mathcal{P}})$

[3] In the standard problem the agent chooses $\mathcal{P}$ to maximize revenues minus the cost of mental effort. The maximization is $\max_{\mathcal{P}} \mathcal{P}Y - K(\mu, \mathcal{P})$ with FOC $Y = K'_{\mathcal{P}}(\mu, \hat{\mathcal{P}})$. The only flexible variable available to the experimenter is Y. However, varying Y does not lead to identification because both $\mathcal{P}$ and $K(\mu, \mathcal{P})$ remain unobserved.

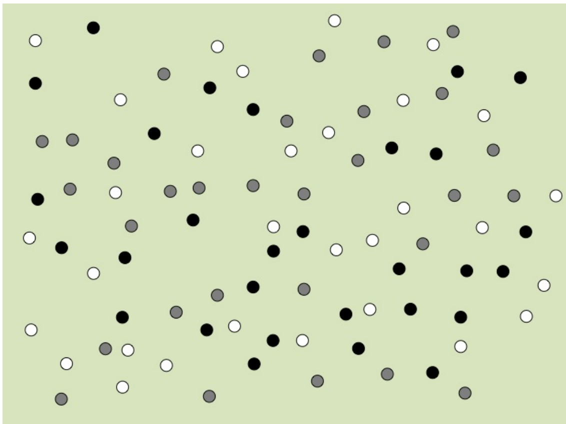
[4] For the sake of clean illustration we assume risk neutrality. However, the model generalizes to the case of risk aversion. In this case identification is reached provided that we specify a utility function and separately estimate the coefficient of risk aversion.

for several values of $\hat{\mathcal{P}}$. An interpolation method could be used to obtain an estimate of the derivative of the cost function, called $\hat{K}'_{\mathcal{P}}$. The domain $[\underline{\mathcal{P}}, \overline{\mathcal{P}}]$, on which the derivative of the cost function can be interpolated, depends upon the specific task being executed and the number of alternatives available for choice. For example, in multiple-choice tasks with 2 options, a utility-maximizing agent can secure a probability of $\underline{\mathcal{P}} = 1/2$ by exerting no attention and choosing one option randomly. Hence, in this case the mental cost function can be recovered over a range contained in the interval $(\frac{1}{2}, 1]$. (In general, in multiple choice tasks with uniform prior and n options, the mental cost function can be interpolated over a range contained in the interval $(\frac{1}{n}, 1]$.) Notice also that all interior solutions must satisfy $K'_{\mathcal{P}}(\mu, \hat{\mathcal{P}}) \leq Y$, which implies that the mental cost function cannot be estimated for values of $\mathcal{P}$ such that $K'_{\mathcal{P}}(\mu, \hat{\mathcal{P}}) > Y$.

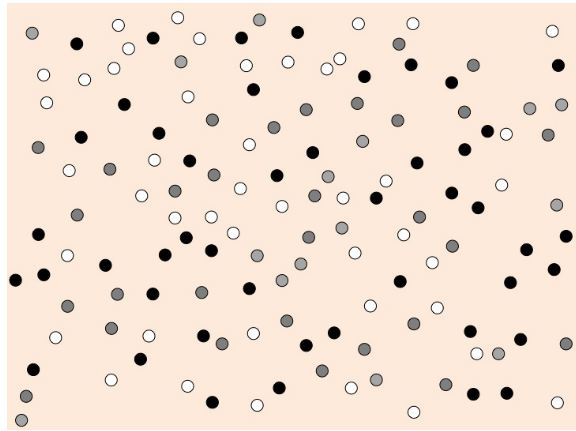
3 Experimental Design

The experiment consists of four phases. In Phase 1, the incentives are provided exogenously and subjects are asked to perform four match-the-state tasks with two levels of task difficulty and two incentive. In Phase 2, after subjects have been familiarized with the task and the difficulty of performing it, we ask them to choose their insurance for six tasks they will perform in the next phase. The first three questions are for one level of difficulty, and the second three questions are for the other level of difficulty. The level of difficulty for the first set of questions is randomized at the session level. In Phase 3, subjects perform the tasks with the incentives they chose in Phase 2. The order of task execution is randomized at the session level. In Phase 4, we elicit subjects' risk and ambiguity attitudes using well-established measures in the literature.

In the experiment, subjects solve a matching-the-state task in 10 consecutive rounds. The task requires subjects to look at screens filled with white, gray and black balls and determine which of the three colors is the most numerous. Subjects earn a larger payoff for a correct answer and a lower payoff for an incorrect answer. Tasks alternate between two difficulty levels, *easy* and *difficult*, whereby the screens of the *easy* task contain 90 balls and the screens of the *difficult* task contain 135 balls.[5]



(a) Easy task



(b) Difficult task

Figure 1

Tasks also differ in terms of how payoffs are determined. In the first four rounds, payoffs for succeeding and failing the task are exogenously determined. In the last six rounds payoffs are determined by the subjects themselves by means of a transfer choice. Starting off with an extreme payoff distribution that awards \$20 if they succeed and \$0 if they fail the task, subjects can choose to transfer any percentage of their payoffs from a successful to a failed outcome, up to the point at which payoffs are completely equalized. An underlying transfer function determines the cost of operating such a transfer. Subjects do not see the transfer function but are given a calculator which they can use by entering any transfer level to obtain the resulting success and failure payoffs. Subjects can familiarize with the transfer function for as long as they wish to. Once they identify their preferred transfer level, their choice is finalized.

Transfer functions vary across rounds and are associated with three cost levels: low, medium and high. The three transfer functions $q_i(x)$ with $i \in \{low,\ medium,\ high\}$, are summarized in Figure 2. Upon choosing a level of transfer $x \in [0, 1]$, payoffs are given by $20 - q_i(x)$ in the event of success, and by $20x - q_i(x)$ in the event of failure. The left panel of Figure 2 shows the accumulated costs

[5] In the instructions the two types of task were not referred to as difficult and easy, but rather as the 90-task and the 135-task

for increasing transfer levels for each of the three transfer functions. Notice that moving from low to medium costs, and similarly from medium to high costs, is associated with an increase in both the total and the marginal costs for each transfer level. The right panel of Figure 2 shows all possible payoffs of success and failure for each transfer level.

(a) Cost of insurance

(b) Success and failure incentives

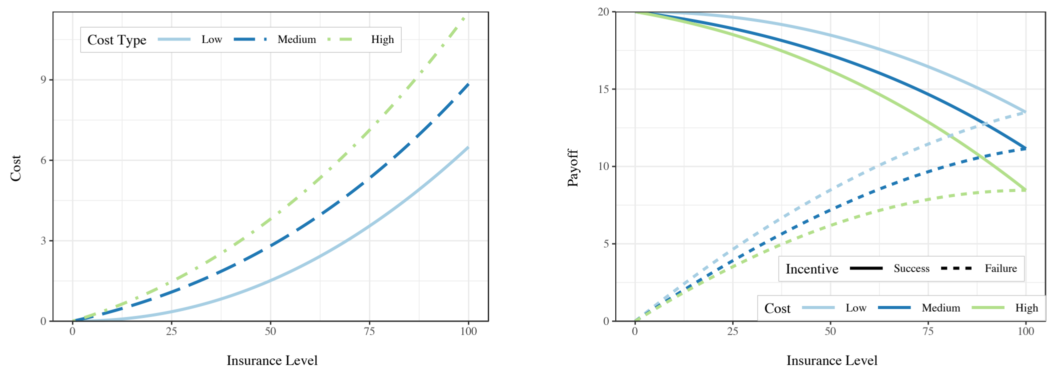


Figure 2: Transfer functions and payoff redistribution

At the end of the experiment, we elicit subjects' attitudes toward risk, ambiguity, and compound lotteries. We first elicit their attitudes towards risk using two procedures: the classic procedure developed by Holt and Laury (2002), and a second developed by Gneezy and Potters (1997) and Charness and Gneezy (2010). We then build on these tests to elicit subjects' attitudes towards ambiguity and compound lotteries. In the first task, developed by Gneezy and Potters (1997) and Charness and Gneezy (2010), subjects are endowed with 100 tokens, any number of which can be invested in a risky asset that has a 0.5 chance of success and, if successful, a payoff of 2.5 times the investment. Whatever is not invested is kept. The number of tokens not invested is proportional to the degree of risk- aversion of a subject who obeys expected utility. The second investment task was identical to the first, except that the success of the investment was determined by a compound lottery which reduces to a 0.5 chance of success. A subject with no aversion nor attraction to compound lotteries should invest the same number of tokens in both questions, while a subject who dislikes (likes) compound lotteries might decide to invest less (more) in the second rather than in the first investment task.

The Holt and Laury (2002) method for eliciting risk aversion elicitation presents subjects with a list of choices between sure payments and a lottery performed on a bag with a 50-50 composition of blue and red chips. Subjects choose the color they want to bet on. The lottery pays \$4 if their chosen color is extracted from the bag and \$0 otherwise. Risk aversion is measured by the certainty equivalent of the lottery. We further incorporate the elicitation of ambiguity aversion in this setting by adopting a procedure similar to that of Halevy (2007) and Dean and Ortoleva (2015). Subjects are presented with a second bag containing blue and red chips in unknown proportions. Subjects choose the color they want to bet on and receive \$4 if the chosen color is extracted from the bag with unknown composition and \$0 otherwise. The certainty equivalent of this gamble is elicited with the same choice list approach. Under subjective expected utility, the certainty equivalent for the ambiguous bet should be at least as high as the certainty equivalent for the risky lottery. Ambiguity aversion is therefore measured as the difference between the certainty equivalent of the risky bet and the certainty equivalent of the ambiguous bet.

All experimental sessions were conducted at the Center for Experimental Social Science (CESS) laboratory at New York University (NYU), using the software z-Tree (Fischbacher (2007)). All participants were NYU students. The experiment lasted about 90 minutes, and subjects earned, on average, \$22, which included a \$7 show-up fee. The experiment consists of four sessions with 51 subjects. In each session, written instructions were distributed to the subjects and also read aloud.[6]

4 Preliminaries and Predictions

In this section, we provide some definitions and details for clear exposition of comparative statics and theoretical predictions.

Definition 1 *An Assignment is a triple $\{T, Y, q(x)\}$ consisting of a task T, a payoff Y for the successful completion of the task, and an insurance cost function $q(x)$ specifying the cost of each level of coverage x.*

[6] Experimental instructions will be provided upon request.

Definition 2 *A Sequence is a set $S = \{T, Y, \{q_1(x), ...q_k(x)\}\}$ consisting of a task T, a payoff Y for the successful completion of the task, and a set of k insurance cost functions $q_j(x)$, $j \in \{1, ..., k\}$.* ***In*** *other words, a sequence is a series of k assignments, all sharing the same task and payoff structure.*

Next, let x^*_{hj} denote the choice of insurance performed by a subject facing sequence h and insurance function $q_j(x)$. Similarly, let $\hat{\mathcal{P}}_{hj}$ denote the predicted probability of success in the same assignment. $\hat{\mathcal{P}}_{hj}$ is obtained by plugging x^*_{hj} in the first order condition, that is $\hat{\mathcal{P}}_{hj} = 1 - q_j\left(x^*_{hj}\right)$.

A subject's answer to sequence h, is a k-tuple of insurance choices $\{x_1^{h*}, ..., x_k^{h*}\}$, one for each insurance function. These points are used for the interpolation of the cost of attention function $\hat{K}_i^h(\mu, \mathcal{P})$, which is estimated according to FOCs. $\hat{K}_i^h(\mu, \mathcal{P})$ is interpreted as the estimated attention cost function for individual i and sequence h.[7]

4.1 Comparative statics

The following comparative statics results can be derived from the first order conditions and provide a testing ground for our elicitation technique.

1. **Variation of the cost of insurance.** Fixing a sequence (e.g. for fixed task and payoff structure) consider two insurance costs functions $q_1(x)$ and $q_2(x)$, such that insurance 1 is more expensive than 2, that is $q_1(x) > q_2(x), \forall \in X$. Then the model predicts that: (i) Subjects buy more insurance when it is cheaper, $x_2^* > x_1^*$, and (ii) Subjects target a lower probability of success for the assignment with cheaper insurance, that is $\hat{\mathcal{P}}_1 > \hat{\mathcal{P}}_2$. This prediction can be tested by comparing answers given by an individual subject and for a given sequence.

2. **Variation of the degree of complexity of the task.** Consider two sequences, A and B, which have the same payoff structure Y, the same set of insurance functions $\{q_1(x), ..., q_k(x)\}$, and such that the task in sequence A is more difficult, that is $K_A(\mu, \hat{\mathcal{P}}) > K_B(\mu, \hat{\mathcal{P}})$ for all $\hat{\mathcal{P}}$. Then the model predicts that for each insurance cost functions $q_j(x)$: (i) Subjects buy more

[7] The subject index is omitted whenever possible.

insurance for the sequence with larger attention costs, $x_j^{A*} > x_j^{B*}$, and (ii) Agents target a lower probability of success for the sequence with larger attention costs, that is $\hat{\mathcal{P}}_j^A < \hat{\mathcal{P}}_j^B$.

Given that the two comparative statics results are supported by the data, we can then move to test the following prediction.

3. **Correspondence between targeted and observed probability of success**. The probability of success derived from the first order conditions, $\hat{\mathcal{P}} = 1 - q'(x^*)$, should have predictive power for the success rate observed in the actual execution of the assignment.

5 Results

In this section we investigate whether insurance choices and attention respond to incentives, task difficulty, and the cost of insurance in the expected manner. After presenting the determinants of insurance choices and their effect on attention, we proceed to probability measures and we find significant explanatory power of model estimated targeted probability of success on performance. We end the section with a discussion and steps of our ongoing exploration.

5.1 Insurance and attention

Let us first examine whether attention responds to incentives as predicted by the theory. First, we determine the effect of task difficulty and the insurance cost on the choice of insurance and, subsequently, the effect on success probability.

We state the theoretical predictions and examine the insurance choices first.

Prediction 1 *As the underlying task becomes more difficult, the decision maker buys more insurance and targets a lower probability of success.*

Prediction 2 *As insurance gets more expensive, the decision maker buys less insurance and targets a higher probability of success.*

To test these predictions, let us first consider the average insurance bought when the task is easy and when it is hard. Subjects on average choose lower insurance, 37.3%, when the task is easy than when it is hard, 54.2% (Mann-Whitney U (MWU) test with $p < 0.001$). The averages provide support for the prediction, and when we examine insurance distribution we find more support for it. Insurance bought under a difficult task first-order stochastically dominates the easy task: Subjects insure themselves more in the harder task (Figure 3b).

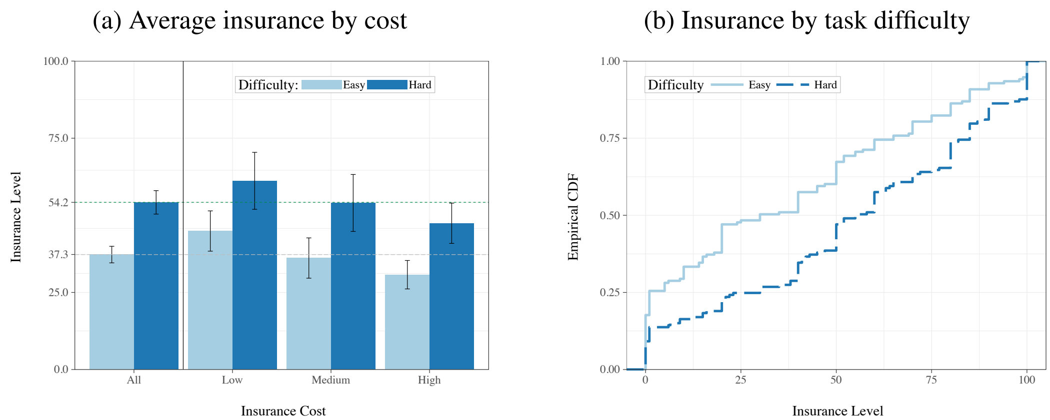


Figure 3: Effect of the task type and cost on insurance

The comparative statics results about the difficulty of the task are averaged over all levels of insurance costs (in the experiment, buying insurance—i.e., transferring a fraction of a winning prize to a fail prize—has low, medium, or high cost). We first break down the insurance choice by how expensive each was (Figure 3a). For each insurance cost, the same pattern emerges for task difficulty of the task: The harder the task, the more insurance subjects buy. Now, what happens when we make insurance more expensive? As expected, subjects buy less insurance when it is more expensive. To better understand subject' insurance choices, we run a regression that includes the task difficulty and insurance cost, as well as two types of subjects' personal characteristics. We include economic characteristics elicited during the experiment and information about the subjects provided in the exit survey.

The estimation provides further support for Propositions 1 and 2. In addition to the effects of

Table 1: Results

	Dependent variable: Insurance		
	(1)	(2)	(3)
Difficulty: Easy	-16.938^{***}	-16.938^{***}	-16.938^{***}
	(2.658)	(2.658)	(2.663)
Cost: Low	13.956^{***}	13.956^{***}	13.956^{***}
	(3.213)	(3.213)	(3.219)
Cost: Medium	5.985^{**}	5.985^{**}	5.985^{**}
	(2.369)	(2.369)	(2.373)
Ambiguity Attitude	16.620^{**}	15.873^{**}	16.553^{**}
	(8.122)	(7.939)	(7.944)
Risk Aversion HL	4.033^{**}		4.463^{**}
	(2.029)		(1.816)
Risky Investment		0.173	0.211
		(0.147)	(0.146)
Compound Lottery	7.125	8.015	7.315
	(9.509)	(9.570)	(9.359)
GPA	19.317	14.478	16.876
	(15.883)	(16.022)	(15.606)
Gender	−4.565	−0.035	−1.057
	(7.794)	(8.988)	(8.790)
GameTheory	−11.974	−11.783	−13.931
	(9.106)	(9.643)	(8.918)
Constant	−53.228	−24.362	−59.327
	(60.585)	(57.564)	(58.653)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$; errors are clustered at the subject level.

task type and cost level, the regression reveals a significant effect of risk and ambiguity attitudes of people on their insurance choices. The more risk averse a person is (measured by a commonly used procedure developed by Holt and Laury (2002)), the more they insure themselves and consequently, lower probability of success they target. In appendix A, we impose CES utility function in our model and show that, as the risk aversion coefficient goes up, agents should choose higher insurance levels, smoothing over their risks. Hence, the result on risk aversion leading to higher insurance choices is in line with the model predictions. An interesting implication of this prediction is that risk averse individuals do worse in this tasks.

If a subject is ambiguity averse, they buy about 15% more insurance than subjects that have neutral ambiguity attitudes. This is a significant effect both statistically and economically. The quantitative effect is similar to the impact of task difficulty going from easy to hard. A possible interpretation of this result is that exerting attention may be perceived as an ambiguous act, with a degree of uncertainty surrounding the mapping from paying attention to success probabilities. If this is true, then an ambiguity averse person will express a desire to hedge against such ambiguity by choosing a larger payoff redistribution. In appendix A we show that our setup can be modeled within the framework of Snow (2011) and Alary et al. (2013) which indeed show a positive link between ambiguity aversion and insurance choices. Finally, we notice that payoff redistribution and attentional choices do not correlate with attitudes towards compound lotteries, GPA, Gender or familiarity with Game Theory.

5.2 Attention and performance

The second part of the propositions concerns the targeted probabilities. Let us start by looking at the probability of success of all subjects by difficulty. On average, in 60.8% of the time subjects are successful at the hard task, while they reach on average 79.7% success rate with an easy task. Table 2 presents regression analyses for the variables of interest—performance. We incorporate in the analysis a series of control variables, including gender, GPA, and measures of risk and ambiguity aversion. We find that the two predictions of the rational inattention model with endogenous incentives are confirmed: (1) when the task is more difficult, subjects buy more insurance and perform worse; and (2) as insurance becomes more expensive, subjects buy less of it and target a higher probability of success.

The theory is silent on the effect of time on the performance in a task; however, it is fairly intuitive to think that successful performance in a task would related to longer time spent on it. We do analysis similar to the results summarized in Tables 1 and 2, and we combine all the results in the Table 3.

5.3 Attention Measures

Since the data strongly supports the theoretical predictions on incentive reactions we can proceed and investigate the targeted probability. First, let us look at the relationship between model estimated

Table 2: Results

	Dependent variable: Performance		
	(1)	(2)	(3)
Incentives	0.022^{***}	0.017^{***}	0.018^{***}
	(0.004)	(0.004)	(0.004)
Difficulty: Easy	0.179^{***}	0.209^{***}	0.206^{***}
	(0.040)	(0.039)	(0.040)
Exogenous	-0.218^{***}	-0.183^{***}	-0.189^{***}
	(0.041)	(0.041)	(0.042)
Response Time		0.001^{***}	0.001^{***}
		(0.000)	(0.000)
Ambiguity Attitude			-0.029
			(0.043)
Risk Aversion HL			0.023
			(0.015)
Risky Investment			0.0001
			(0.001)
Compound Lottery			0.001
			(0.001)
GPA			0.090
			(0.063)
Gender			0.097^{*}
			(0.051)
GameTheory			0.058
			(0.053)
Constant	0.374^{***}	0.279^{***}	-0.356
	(0.063)	(0.074)	(0.288)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$; errors are clustered at the subject level.

probability and subjects' actual performance and forecast (see Figure 4). Figure 4 shows that model estimated probability is a good predictor of performance and subject's own forecast. To further explore this relationship we run a regression using performance as a dependent variable and model prediction as an explanatory variable.

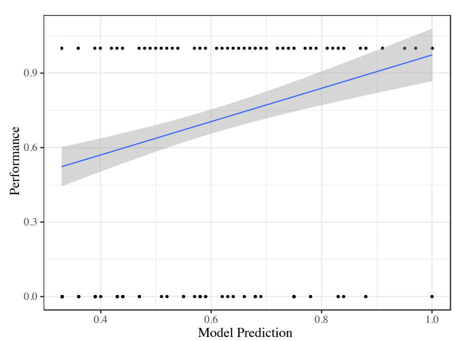
Given our experimental design, for given incentives set by the subjects we observe one repetition of the task. Hence, we know whether a subject was successful or not in that particular setting. To estimate the frequency of successes we use the standard method in the literature and aggregate the data

Table 3: Regression Analysis

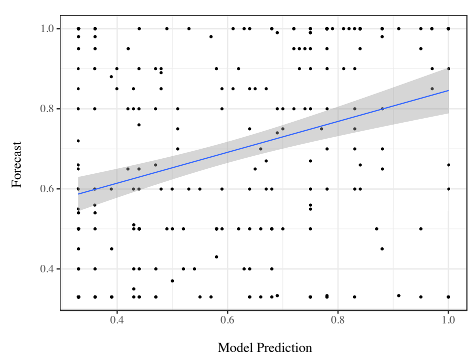
	(1) Insurance	(2) Performance	(3) Performance	(4) Time	(5) Time
RiskAversionHL	4.463**	-0.000355	0.0207	0.894	4.639
	(2.215)	(0.0211)	(0.0165)	(4.302)	(3.739)
RiskTask	0.211	0.000369	0.00136	-0.134	0.0426
	(0.143)	(0.00136)	(0.00105)	(0.278)	(0.240)
CompoundTask	7.315	-0.0354	-0.000977	-14.56	-8.417
	(8.592)	(0.0819)	(0.0626)	(16.69)	(14.33)
Ambiguity	16.55**	-0.0400	0.0379	-25.93*	-12.04
	(7.870)	(0.0750)	(0.0588)	(15.29)	(13.30)
Gender	1.057	0.00827	0.0133	11.96	12.84
	(8.072)	(0.0769)	(0.0585)	(15.68)	(13.43)
GPA	16.88	0.0207	0.100	11.53	25.69
	(14.84)	(0.141)	(0.109)	(28.83)	(24.81)
GameTheory	-13.93	0.183**	0.117*	15.21	3.519
	(9.480)	(0.0903)	(0.0697)	(18.41)	(15.90)
Ins_costs	-6.978***	0.0735***	0.0407	6.745*	0.889
	(1.286)	(0.0283)	(0.0291)	(3.620)	(3.632)
Difficulty	16.94***	-0.190***	-0.110**	8.922	23.14***
	(2.100)	(0.0462)	(0.0487)	(5.912)	(6.191)
Incentives			0.0235***		4.196***
			(0.00433)		(0.709)
Constant	-62.64	0.613	-0.153	82.79	-53.70
	(54.63)	(0.522)	(0.421)	(106.2)	(93.85)
Observations	306	306	306	306	306

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Model Prediction, Performance and Forecast



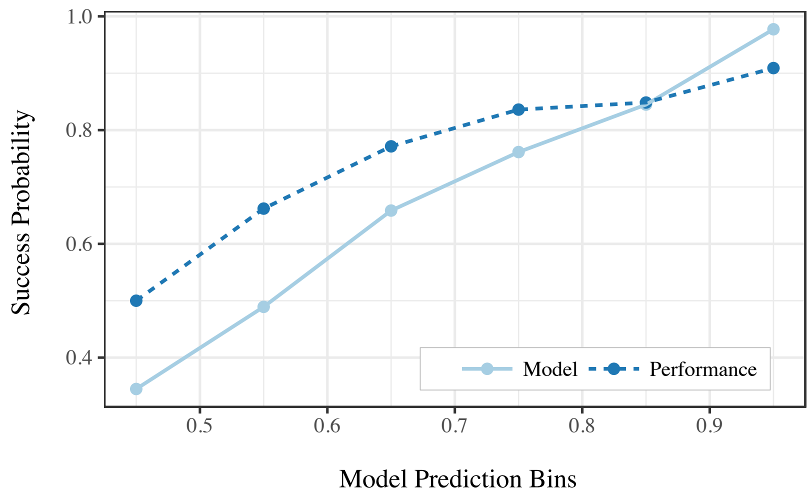
(a) Model Prediction and Performance



(b) Model Prediction and Forecast

in decile bins, in the following way. First, we take all the cases in which the model estimate of targeted probability was within a certain decile, for example, between 0.6 and 0.7. For all those instances, we

Figure 5: Model Predictions vs actual performance



calculate the frequency of actual successes. Figure 5 presents the results of this exercise. We can see the model average and performance in each interval starting from the minimal success probability of 1/3 (i.e., the success probability of a subject exerting no attention and choosing one of the three actions at random). We observe that the predicted probabilities by the model track the actual probabilities of success quite closely. We also observe the actual performance being almost everywhere higher than the model predicts, suggesting that the subjects may have been underconfident about their ability to succeed and therefore overinsured themselves, except at the highest levels of estimated and actual success probabilities.

6 Discussion and Conclusion

We extend the discrete choice rational inattention model to the case in which a decision maker can influence the payoff distribution across states. By reducing the gap between payoffs in different states, the decision maker is able to affect her own incentives to pay attention. The smaller the gap, the less attentive the decision maker needs to be, since the payoffs in different realized states are closer. We find that subjects have a demand for attention insurance as 94% of subjects buy insurance at least once. Moreover, in 87% of all possible cases insurance bought towards the failed state was strictly positive. Therefore, subjects are willing to pay to avoid paying more attention. Some even fully insure by completely equalizing the payoffs in both states, and thus, avoiding the need to pay attention to the task at all.

In our experiment, we measure attention in three different ways. First, we observe actual performance in the task. Second, we use insurance choices and estimate the targeted probability of success through our model. Finally, we ask subjects to provide their own estimate of how well they think they did, which provides us with a subjective measure of the targeted probability of success. We find that model-estimated probability has a significant and sizable prediction power, for both performance and own forecast.

In the theoretical framework or in the conditions of the experiment, we have not made any parametric restrictions on the attention cost function. Therefore, the condition we get from agent's maximization problem can be applied to any cost function and then the predictions can be tested using the existing data. Therefore, the data generated by this experiment can therefore be used to test any cost function that might be of theoretical or applied interest.

References

Alary, David, Christian Gollier, and Nicolas Treich, "The effect of ambiguity aversion on insurance and self-protection," *The Economic Journal*, 2013, *123* (573), 1188–1202.

Andrade, Philippe and Hervé Le Bihan, "Inattentive professional forecasters," *Journal of Monetary Economics*, 2013, *60* (8), 967–982.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention discrimination: Theory and field experiments with monitoring information acquisition," *American Economic Review*, 2016, *106* (6), 1437–75.

Caplin, Andrew and Mark Dean, "Revealed preference, rational inattention, and costly information acquisition," *The American Economic Review*, 2015, *105* (7), 2183–2203.

_ , John Leahy, and Filip Matejka, "Learning from market share when consumers are rationally inattentive," Technical Report, NYU Working Paper 2014.

Charness, Gary and Uri Gneezy, "Portfolio choice and risk attitudes: An experiment," *Economic Inquiry*, 2010, *48* (1), 133–146.

Dean, Mark and Nate Leigh Neligh, "Experimental tests of rational inattention," 2017.

_ and Pietro Ortoleva, "Is it all connected? A testing ground for unified theories of behavioral economics phenomena," *A Testing Ground for Unified Theories of Behavioral Economics Phenomena (August 13, 2015)*, 2015.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, 2007, *10* (2), 171–178.

Gaballo, Gaetano, "Rational inattention to news: the perils of forward guidance," *American Economic Journal: Macroeconomics*, 2016, *8* (1), 42–97.

Gneezy, Uri and Jan Potters, "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Halevy, Yoram, "Ellsberg revisited: An experimental study," *Econometrica*, 2007, *75* (2), 503–536.

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American economic review*, 2002, *92* (5), 1644–1655.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, "A rational theory of mutual funds' attention allocation," *Econometrica*, 2016, *84* (2), 571–626.

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

Martin, Daniel, "Strategic pricing with rational inattention to quality," *Games and Economic Behavior*, 2017, *104*, 131–145.

Martinelli, César, "Would rational voters acquire costly information?," *Journal of Economic Theory*, 2006, *129* (1), 225–251.

Matějka, Filip, "Rigid pricing and rationally inattentive consumer," *Journal of Economic Theory*, 2015, *158*, 656–678.

Matejka, Filip and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *The American Economic Review*, 2014, *105* (1), 272–298.

Matějka, Filip and Guido Tabellini, "Electoral competition with rationally inattentive voters," *Available at SSRN 3070204*, 2017.

Mondria, Jordi, Thomas Wu, and Yi Zhang, "The determinants of international investment and attention allocation: Using internet search query data," *Journal of International Economics*, 2010, *82* (1), 85–95.

Oliveira, Henrique De, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek, "Rationally inattentive preferences and hidden information costs," *Theoretical Economics*, 2017, *12* (2), 621–654.

Sims, Christopher A, "Stickiness," in "Carnegie-Rochester Conference Series on Public Policy," Vol. 49 Elsevier 1998, pp. 317–356.

_ , "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Snow, Arthur, "Ambiguity aversion and the propensities for self-insurance and self-protection," *Journal of risk and uncertainty*, 2011, *42* (1), 27–43.

Woodford, Michael, "Inattention as a source of randomized discrete adjustment," *unpublished, Columbia University*, 2008.

Appendices

A Proofs and Extension

A.1 Risk preferences

In what follows, we extend the model to the case of a decision maker with general risk attitude. The basic setup remains unchanged and the only addition is represented by the introduction of a possibly non-linear utility function over monetary payoffs. The new maximization is as follows:

$$\max_{\mathcal{P},x} \mathcal{P} \cdot u\left(Y - q(x)\right) + (1-\mathcal{P}) \cdot u\left(xY - q(x)\right) - c(\mathcal{P})$$

leading to the following first order conditions:

$$\begin{aligned}
&-\mathcal{P}u'(Y - q(x))q'(x) + (1-\mathcal{P})u'(xY - q(x))(Y - q'(x)) = 0 \\
&u(Y - q(X)) - u(xY - q(x)) - K'(\mathcal{P}) = 0
\end{aligned}$$

which, after rearrangment, lead to:

$$\mathcal{P} = \frac{u'(xY - q(x))(Y - q'(x))}{u'(Y - q(x))q'(x) + u'(xY - q(x))(Y - q'(x))} \tag{7}$$

$$K'(\mathcal{P}) = u(Y - q(x)) - u(xY - q(x)) \tag{8}$$

We can still obtain identification of p and $K'(p)$ if we make parametric assumptions about the utility function and separately estimate its parameters. For example, we can impose the CRRA utility function:

$$u(m) = \begin{cases} \frac{m^{1-\rho}}{1-\rho} & \text{for} \quad \rho \neq 1 \\ \ln(m) & \text{for} \quad \rho = 1 \end{cases}$$

and, after the appropriate substitutions, the identification equations in (7) and (8) become:

$$\mathcal{P} = \frac{(xY - q(x))^{-\rho}(Y - q'(x))}{(Y - q(x))^{-\rho} \cdot q'(x) + (xY - q(x))^{-\rho} \cdot (Y - q'(x))}$$

$$K'(\mathcal{P}) = \frac{(Y - q(x))^{1-\rho}}{1 - \rho} - \frac{(xY - q(x))^{1-\rho}}{1 - \rho}$$

The model is still identified, provided that we obtain a separate estimation for the risk parameter ρ.

A.2 Continuous payoffs

The method outlined in the main body of the text can be extended to settings where the decision maker is executing an assignment characterized by continuous actions and payoffs. Consider a DM that chooses one action a to play with $a \in \mathbb{R}$. Payoffs are determined as $-(a - a^*)^2$ where a^* is a normal random variable $a^* \sim N(0, \sigma_a)$ whose realization is unknown to the decision maker. The DM draws a signal $s = a^* + \epsilon$ with $\epsilon \sim N(0, \sigma_\epsilon)$ and chooses the precision of the signal.

There is a monotonically increasing relationship between σ_ϵ and the average loss $L = E((a - a*)^2)$, therefore we can re-parametrize the choice problem and say that the DM is choosing L. A signal with higher precision (lower variance) is more costly. It follows that we can express the attention cost function $k(\sigma_\epsilon)$ as a function of the average loss $c(-L)$. Finally the DM can purchase insurance x by paying a cost of $q(x)$. x is the percentage of L which is returned to the DM, i.e. when the DM buys x units of insurance his payoffs are given by $(x-1)L - q(x)$.

Under risk neutrality the maximization problem is as follows:[8]

$$\begin{aligned} \max_{x,L} \quad -L + xL - q(x) - c(-L) \\ \text{s.t.} \\ 0 \leq x \leq 1 \\ 0 \leq L \end{aligned}$$

The first order conditions are:

$$(1-x) = c'(x) \tag{9}$$

$$L = q'(x) \tag{10}$$

The second order conditions for a maximum are:

$$\begin{aligned} -q''(x) < 0 \\ c''(p)q''(x) > 1 \end{aligned}$$

Here again we can express the FOC only in terms of observables:

[8] In Appendix A we provide the extension of the method to general risk attitude.

$$c'(-q'(x)) = 1 - x \tag{11}$$

The derivation of the cost of attention can be obtained following the same procedure outlined in the main body of paper.

A.3 Generalization to N states, N actions and general prior distribution

Here we present the generalization of our identification strategy to the case with N states, N actions and general prior distribution. The basic idea is unchanged. We exploit the choice of payoff redistribution across actions and states to infer the costs of attention.

We assume, again, that the decision maker is faced with a matching-the-state task, this time with with N actions, N states and a prior distribution over states given by μ. The decision maker wins prize Y if she matches the state, that is, if she plays action j in state j. In all instances where there is a mismatch between actions and states, the decision maker can insure a portion x_{ij} of the payoff, that is when the action is i and the state is j, the decision maker obtains payoff $x_{ij} \cdot Y$. Hence, absent any insurance, the payoff matrix is given by:

$$V = \begin{bmatrix} Y & 0 & ... & 0 \\ 0 & Y & ... & 0 \\ . & . & . & . \\ 0 & 0 & ... & Y \end{bmatrix}$$

After payoff redistribution, the payoff matrix becomes:

$$\tilde{V} = \begin{bmatrix} Y - q(X) & ... & ... & x_{1N}Y - q(X) \\ x_{21}Y - q(X) & Y - q(X) & ... & x_{2N}Y - q(X) \\ . & . & . & . \\ x_{N1}Y - q(X) & x_{N2}Y - q(X) & ... & Y - q(X) \end{bmatrix}$$

with:

$$X = \begin{bmatrix} x_{11} & x_{12} & ... & x_{1M} \\ x_{21} & x_{22} & ... & x_{2M} \\ . & . & . & . \\ x_{N1} & x_{N2} & ... & x_{NM} \end{bmatrix}$$

The approriate generalization of the symmetric probability of matching the state $\mathcal{P}$ introduced in the main body of the paper, is given by $\mathcal{P}_i(\mathbf{v_j})$ that is the probability of choosing action i given state $\mathbf{v_j}$:

$$\mathcal{P} = \begin{bmatrix} p_{11} & p_{12} & ... & p_{1N} \\ p_{21} & p_{22} & ... & p_{2N} \\ . & . & . & . \\ p_{N1} & p_{N2} & ... & p_{NN} \end{bmatrix}$$

$$\sum_i p_{ij} = 1, \forall j$$

The expected monetary payoff given the payoff reditribution $\tilde{V}$ is:

$$G(\mu, \mathcal{P}, x) = \sum_{i=1}^{N}\sum_{j=1}^{N} \mu(\mathbf{v_j}) p_{ij} \tilde{V}_{ij} = \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} \mu(\mathbf{v_j}) p_{ij} x_{ij} Y + \sum_{j=1}^{N} \mu(\mathbf{v_j}) p_{ij} Y$$

Where we note that $p_{jj} = 1 - \sum_{\substack{i=1 \\ i\neq j}}^{N} p_{ij}$. The augmented maximization problem is:

$$(\hat{\mathcal{P}}, \hat{X}) := \arg\max_{(\mathcal{P},X)} \{G(\mu, \mathcal{P}, X) - K(\mu, \mathcal{P}) - \mathcal{Q}(X)\} \tag{12}$$

The $2 \times N \times (N-1)$ first order conditions with respect to x_{ij} and p_{ij} with $j \neq i$ are given by:

$$\mu(\mathbf{v_j}) p_{ij} Y = q_{ij}(X) \quad \text{for all} \quad i, j \neq i \tag{13}$$

$$\mu(\mathbf{v_j})(1 - x_{ij}) Y = K_{ij}(\mu, \mathcal{P}) \quad \text{for all} \quad i, j \neq i \tag{14}$$

where $q_{ij}(X)$ is the fully observable first derivative of the cost of insurance function with respect to component x_{ij} and evaluated at X, and $K_{ij}(\mu, \mathcal{P})$ is the derivative of the attention cost function with respect to component p_{ij}. The model is identified via a procedure similar to the one highlighted in the paper. The $N \times (N-1)$ targeted probabilities p_{ij} can be identified by an equal number of first order conditions in (13). Once $\mathcal{P}$ has been reconstructed, the equations in (14) provide the value of the first derivative of the cost function at $\mathcal{P}$ along dimension *ij*. Finally, the adoption of different insurance cost functions $q(X)$ allows to interpolate the attention cost function $K(\mu, \mathcal{P})$.

This generalization allows us to test additional properties of the cost function such as the assumption of .... which is a feature of the Shannon capacity.

A.4 N actions

Maximization problem

$$V = \begin{bmatrix} v_1 & 0 & ... & 0 \\ 0 & v_2 & ... & 0 \\ . & . & . & . \\ 0 & 0 & ... & v_N \end{bmatrix}$$

$$\tilde{V} = \begin{bmatrix} v_1 & xv_2 & ... & xv_N \\ xv_1 & v_2 & ... & xv_N \\ . & . & . & . \\ xv_1 & xv_2 & ... & v_N \end{bmatrix}$$

$$X = \begin{bmatrix} v_{11} & v_{12} & ... & v_{1M} \\ v_{21} & v_{22} & ... & v_{2M} \\ . & . & . & . \\ v_{N1} & v_{N2} & ... & v_{NM} \end{bmatrix}$$

$$\mu : \text{ prior probability of each state } \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} & ... & \mathbf{v_N} \\ p_1 & p_2 & ... & p_N \end{bmatrix}$$

$$\mathcal{P}_i(\mathbf{v_j}) : \text{ probability of choosing an action } i \text{ given the state } \mathbf{v_j}$$

$$\mathcal{P} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1M} \\ q_{21} & q_{22} & \cdots & q_{2M} \\ . & . & . & . \\ q_{N1} & q_{N2} & \cdots & q_{NM} \end{bmatrix}$$

$$\sum_{i} q_{ij} = 1, \forall j$$

$$G(\mu, \mathcal{P}, x) = \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{V}_{ij} \mathcal{P}_i(\mathbf{v_j}) \mu(\mathbf{v_j}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{V}_{ij} q_{ij} p_j = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} x v_j q_{ij} p_j + \sum_{i=1}^{N} v_i q_{ii} p_i$$

$$C(\mu, \mathcal{P}, x) = \kappa(\mu, \mathcal{P}) + q(x)$$