Paying for Inattention*

Ala Avoyan† and Giorgia Romagnoli‡

April 6, 2023

Abstract

We augment the standard rational inattention model by allowing the decision maker to alter the distribution of payoffs. The model captures real-life circumstances in which decision makers choose their incentives to pay attention, for instance, through the choice of insurance (full-coverage contracts reduce the incentives to pay attention compared to partial coverage). This new framework, specifically the ability to observe the decision maker's choice of payoff redistribution, allows us to elicit the decision maker's targeted attention level. This is a novel method of eliciting the object of interest—attention—typically obtained through performance (i.e., repetitions) in the literature. Furthermore, by manipulating the cost of payoff redistribution, the framework allows us to examine rational inattention models without making parametric assumptions on the cost of attention function. With a laboratory experiment, we validate novel comparative static predictions of our model. The subjects respond to the link between payoff redistribution decisions and attention in accordance to the theory.

JEL Classification: C72, C91, C92, D83;

Keywords: rational inattention, experiment, attention cost function, endogenous incentives.

*We would like to thank Andrew Schotter, Guillaume Frechette, Filip Matejka, Jakub Steiner, and Amnon Maltz, for their helpful comments and discussions. This paper also benefited from comments by seminar and conference participants of NRET at NYU, the 28th International Conference on Game Theory, the 2017 ESA North American Meetings, and the 2019 Sloan Nomis Workshop on the Cognitive Foundations of Economic Behavior. We gratefully acknowledge financial support from the Center for Experimental Social Science at New York University (CESS) and Indiana University.

†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 rational inattention (RI), the idea that a decision maker optimally chooses the level of attention based on a trade-off between attention costs and expected returns.[1] The decision maker chooses among actions with state-dependent payoffs and is uncertain about the actual state realization. Through costly attention, she can refine the posterior distribution over states and improve her choice. This literature has assumed that the agent takes the payoff structure as given. While attention is endogenous, incentives to be attentive have been kept outside of the decision maker's control.

In this paper, we endogenize the payoff structure. The model captures real-life circumstances in which decision makers choose their incentives to pay attention. Examples include financial choices (some investment portfolios requiring more monitoring than others), and insurance choices (full-coverage contracts reducing the incentives to pay attention compared to partial coverage), among others.

Allowing the decision maker to choose the payoff structure leads to two further results. First, a single endogenous payoff choice reveals the level of attention the decision maker intends to utilize. Intuitively, the more the payoffs are smoothed across states, the lower the incentives to pay attention and refine beliefs about the true state. The extreme example of full equalization of payoffs over all actions and states leaves the agent with no incentive to pay costly attention and learn the true state. In this paper we show that a single choice of payoff structure can reveal the attention level the decision maker is targeting. Secondly, this framework provides an evaluation of any theorized attention cost function. This is particularly useful since the methodology imposes no parametrization on the cost function and only makes minimal assumptions about it (e.g., differentiability).

We complement the theoretical analysis with a lab experiment to examine the validity of our model. In particular, we vary the underlying task's difficulty and the cost of payoff distribution across state-action combinations. We find strong support for the comparative static predictions implied by our model. The subjects choose lower transfers when the task is easy rather than hard. Similarly, the subjects choose lower

[1] Rational inattention has been theoretically explored and applied to various areas of economics (see, for examples, Martinelli (2006), Woodford (2008), Mackowiak and Wiederholt (2009), Mondria et al. (2010), Andrade and Le Bihan (2013), Bartoš et al. (2016), Matějka (2015), Martin (2017), Caplin et al. (2014), Kacperczyk et al. (2016), Gaballo (2016), De Oliveira et al. (2017), Matějka and Tabellini (2021), Luo and Tsang (2020), Zhang and Mu (2021) and Ilinov and Jann (2022)). Also, see Maćkowiak et al. (2023) for an excellent review of the literature.

transfers when the transferring becomes more expensive. Furthermore, lower transfers are associated with higher levels of attention. Strong empirical support for the comparative static predictions indicates that the subjects internalize the joint determination of payoff distribution and attention choice, which validates our approach of using observable payoff decisions to infer the implicit attention choices. In the existing literature, measuring an decision maker's attention at an individual level requires the said decision maker to repeat the task numerous times, so that attention can be measured via success frequencies.[2] With our novel methodology, similar results can be reached bypassing the repetition. Our method reveals the targeted attention level from a single decision of the payoff choice.

Related to our work are the theoretical studies of De Oliveira et al. (2017) and Lin (2022), showing how the choice of (or from) menus can reveal information regarding the shape of the cost of attention. Compared to these papers, our model provides a parsimonious measure of the targeted attention levels via a single choice of payoff distribution. This study derives a feasible experimental protocol that can be implemented in a lab setting. Notably, our experiment empirically validates the necessary condition for identification, i.e., the link between payoff distributions and attention decisions.

2 The Model

Our model builds on the costly information acquisition framework of Matejka and McKay (2014) and Caplin and Dean (2015). We first provide the standard setting and then augment it with an additional choice variable. The decision maker's action set is $A = \{1, ..., N\}$. There are M states of nature which are vectors $\mathbf{v} = (v_1, ..., v_N) \in R^N$. Let V be the $(N \times M)$ matrix whose element v_{ij} is agent's payoff for action i in state $\mathbf{v}_j$. Let $\mu \in \Gamma$ be the prior belief over states, with Ω^μ support, and $\gamma \in \Gamma$ be the 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' rule. Let $\Pi = \cup_\mu \Pi^\mu$ and $G : \Gamma \times \Pi \to \mathbb{R}$ be the gross payoff of using some information structure:

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

[2] It has been highlighted that the repetition method can suffer from learning effects and fatigue since the same setting is repeated numerous times to obtain accurate frequencies.

where

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

The gross expected payoff in (1) can be rewritten in terms of the induced probability of selecting an action given a state. That is, let

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

where S_i is a set of all signals for which optimal action is i. The expected payoff in terms of $\mathcal{P}_i(\mathbf{v})$ is

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

and $\mathcal{P} := \{\mathcal{P}_i(\mathbf{v})\}_{i=1}^{N}$. The decision maker maximizes utility by choosing a costly information structure. The optimal choice given cost K is:

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

Endogenizing payoffs In our augmented model, the decision maker maximizes utility by jointly choosing an information structure and a payoff redistribution function. The payoff redistribution function is $X : \mathbb{R}^{N \times M} \rightarrow \mathbb{R}^{N \times M}$. Function X transforms the outcome matrix V into $\tilde{V}$, $X(V) = \tilde{V}$, and carries a monetary cost of $\mathcal{Q}(X)$. The expected payoff given redistribution X is:

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

The optimal joint choice of an information structure *and* payoff redistribution, given attention cost and redistribution cost is:

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

Identification of attention Let us focus on the simplest environment that delivers the main results.[3] There are two states of the world $S = \{\mathbf{w}, \mathbf{b}\}$ with equal prior probability, $\mu = 1/2$. The agent can take two actions $A = \{W, B\}$. The initial payoff structure describes a typical "matching-the-state" task: decision maker receive payoff Y if they match the state (that is, they pick action W in state w and action B in state b) and 0 otherwise. Decision maker can change the payoff structure by guaranteeing a portion x of total payoff Y (that is, a dollar amount equal to xY) in the event she fails to match the state. Redistribution costs are given by $q : [0, 1] \to \mathbb{R}^+$. The initial payoff structure is given by the matrix V, where the element v_{ij} is the payoff of action $i \in \{W, B\}$ in state $j \in \{w, b\}$:

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

By redistributing a portion $x \in [0, 1]$ of Y to the event in which the decision maker fails to match the state, the resulting endogenous payoff matrix $\tilde{V}(x)$ is:

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

By exerting attention, the agent targets a certain posterior probability of matching the state $\{\mathcal{P}_W, \mathcal{P}_B\}$, where $\mathcal{P}_W$ ($\mathcal{P}_B$) is the probability of taking action W (B) when the state is w (b). We assume that (i) only symmetric posterior probabilities can be selected (that is $\mathcal{P}_W = \mathcal{P}_B = \mathcal{P}$) and (ii) the 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 cost of the attention required to target a posterior probability $\mathcal{P}$ of matching the state; $K(\mu, \mathcal{P})$ is assumed differentiable. The agent has utility u defined over monetary outcomes, is rationally inattentive, and simultaneously chooses x and $\mathcal{P}$. 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 attention costs of targeting $\mathcal{P}$.

The maximization problem is:

[3] In the online appendix, we show how our method can be extended to incorporate risk preferences and involve continuous actions and payoffs.

$$\max_{\mathcal{P},x} \mathcal{P}Y + (1-\mathcal{P})xY - q(x) - K(\mu, \mathcal{P}) \tag{3}$$

An external observer cannot identify the targeted probability of success $\mathcal{P}$ or the cost function $K(\mu, \mathcal{P})$ from the standard maximization problem with exogenous payoffs. Here we show how, in the augmented maximization problem, where the payoff distribution is made endogenous, $\mathcal{P}$ and $K'(\mu, \mathcal{P})$ can be inferred from observing x and associated $q(x)$. The first order conditions (FOC) derived from (3) are:

$$(1-\mathcal{P})Y - q'(x) = 0 \tag{4}$$

$$Y - xY - K'(\mu, \mathcal{P}) = 0 \tag{5}$$

Let $(x_q^*, \hat{\mathcal{P}}_q)$ be the solution to the FOC 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:

$$\hat{\mathcal{P}}_q = 1 - q'(x_q^*)Y^{-1}; \tag{6}$$

$$K'(\mu, 1 - q'(x_q^*)Y^{-1}) = Y(1 - x_q^*). \tag{7}$$

Identification of targeted attention From equations (6) and (7), we see that the optimal attention level $\hat{\mathcal{P}}_q = 1 - q'(x_q^*)Y^{-1}$, and the value of the derivative of the attention cost function at that point, $K'_{\mathcal{P}} = Y(1 - x_q^*)$, are all expressed in terms of observables. This is the core finding of the paper. Specifically, equation (6) reveals the optimal attention level targeted by an agent as a function of $q'(x_q^*)$ and Y, where the shape of $q'(x_q^*)$ is controlled by the experimenter and x_q^* is the observable choice of payoff redistribution taken by the agent.

Identification of the shape of the attention cost function Condition (7) reveals the value of the derivative of the cost function $K'(\cdot)$ at a given point. This can be used in two ways. Firstly, by adopting a set of redistribution functions $q(\cdot) \in \mathbb{Q}$ and observing one choice of payoff redistribution for each function in the set $\mathbb{Q}$, one can obtain values of the derivative of the cost function $K'(\mu, \hat{\mathcal{P}})$ for several values of $\hat{\mathcal{P}}$. An interpolation method could be used to obtain an estimate of the derivative of the cost function, $\hat{K}'$. This can be further integrated to obtain a parameter-free estimate of the attention cost function. Secondly, any theorized cost functions by researchers can be tested using this condition. That is, because we impose only minimal assumptions (e.g., differentiability and symmetry), any well-behaved function can be tested using existing data.

3 Model validation

In this section, we derive comparative static predictions, introduce the experimental design, and present experimental results to validate our model and the ensuing methodology for estimating attention levels. Our main objective is to demonstrate that decision makers comprehend the relationship between payoff redistribution and attention choices and can respond appropriately to changes in the task difficulty and the transfer costs. By doing so, we can validate the novel identification strategy that leverages payoff redistribution decisions to predict targeted attention.

3.1 Comparative static predictions

Examining the comparative static predictions will indicate whether decision makers internalize the joint determination of payoff distribution and attention choice. The literature has employed exogenous variations in incentives, task difficulty, and changing priors to examine aspects of the rational inattention model (see Dewan and Neligh (2020), Caplin et al. (2020), and Dean and Neligh (2022)). None of these papers study the impact and benefits of endogenous incentive choices.

We vary task difficulty and the transfer functions to derive four comparative static predictions from our model. Consider an agent facing two problems, A and B, which have the same transfer function $q(x)$ and the same payoff structure Y. However, the task in problem A is more difficult, that is $K_A(\mu, \hat{\mathcal{P}}) \geq K_B(\mu, \hat{\mathcal{P}})$ for all $\hat{\mathcal{P}}$. The model predicts that for each transfer function $q(x)$: (i) decision makers choose higher transfer levels in problem A, $x_A^* \geq x_B^*$, and (ii) decision makers target a lower probability of success in case A, that is $\hat{\mathcal{P}}_A \leq \hat{\mathcal{P}}_B$.

Let us consider now a given task and a payoff structure and introduce two transfer cost functions $q_1(x)$ and $q_2(x)$. Transfer function 1 is (weakly) more expensive than transfer function 2, that is, $q_1(x) \geq q_2(x), \forall x \in [0, 1]$. Then, the model predicts that: (i) decision makers choose higher transfers with $q_2(x)$, $x_2^* \geq x_1^*$, and (ii) decision makers target a lower probability of success with $q_2(x)$, that is $\hat{\mathcal{P}}_2 \leq \hat{\mathcal{P}}_1$. Combining the predictions above, we can write them together in the following manner.

Predictions *A more difficult attention task or less expensive transfer function leads to*

(a) higher payoff transfers, and

(b) lower targeted probability of success.

3.2 Experimental Design

The experiment was conducted at the Center for Experimental Social Science (CESS) laboratory at New York University (NYU) during the Spring of 2017, using the software z-Tree (Fischbacher (2007)), 51 participants were recruited from the general population of NYU students using *hroot* (Bock et al. (2014)). Additionally, four sessions of the identical experiment were conducted at Interdisciplinary Experimental Laboratory (IELAb) at Indiana University (IU) during the Spring of 2023. Fifty-five participants were recruited from the general population of IU students using the ORSEE recruitment system (Greiner (2015)).

Figure 1: Timing of the experiment

Phase 1	**Phase 2**	**Phase 3**	**Phase 4**
– Attention task with **exogenous** incentives – 4 cases (2 task difficulties × 2 incentive levels)	– Choice of **transfers** – 6 cases (2 task difficulties × 3 transfer costs)	– Attention task with **endogenous** incentives – 6 cases from Phase 2	– Risk-aversion elicitations – Ambiguity elicitation

The four main phases of the experiment are outlined in Figure 1. Phase 1 allows subjects to understand how strenuous the attention tasks are, that is, to form a mental mapping of the level of costly attention required to achieve various levels of performance. First, the subjects are shown what they would earn if they succeed or fail, with exogenous payoffs. Then, the subjects perform the task. Phase 1 allows subjects to make informed decisions on payoff transfers in Phase 2. In Phase 3, subjects repeat the task with endogenous payoffs. This phase incentivizes truthful choices in Phase 2, and provides data on success frequencies. Phase 4 elicits risk and ambiguity attitudes to be able to control for these characteristics in transfer and attention choices.[4]

[4] We elicit subjects' attitudes toward risk, compound lotteries, and ambiguity. We elicit subjects' 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 elicit subjects' attitudes towards compound lotteries (Agranov and Ortoleva (2017)). Finally, we elicit ambiguity aversion using a procedure similar to that of Halevy (2007) and Dean and Ortoleva (2015).

(a) Cost of transfer (b) Success and failure incentives

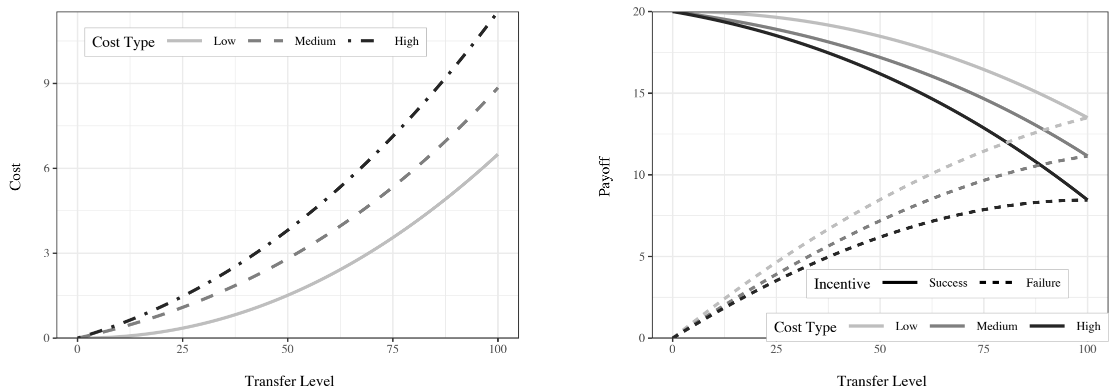


Figure 2: Transfer functions and payoff redistribution

During the attention task, participants view screens filled with white, gray, and black balls. [5] The subjects “succeed” if they correctly determine which of the three colors is the most numerous. *Exogenous incentives* have two incentive levels: (a) awarding \$19.5 if the subject succeeds and \$1.5 if they fail the task, and (b) awarding \$10.4 if the subject succeeds and \$8.4 if they fail the task.[6]

Endogenous incentives allow subjects 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 are provided with a menu consisting of 4 transfer levels (20, 40, 60, and 80%) and corresponding payoffs. In addition, subjects have a calculator to input any transfer level in the interval $[0, 100]$ and preview the implied payoff distribution.

The experiment provides 2 primary variations: (i) the complexity of the task, which can be *difficult* or *easy* (corresponding to having a total number of 90 and 135 balls, respectively); and (ii) the costs of payoff redistribution from the success to the failure states, which can be *high, medium* or *low* (see Figure 2). Note, Figure 2a shows that the same level of transfer level costs more under high-cost functions than medium and low-cost functions. Further, Figure 2b shows the corresponding incentive levels under success and failure for all three transfer cost functions and all transfer levels (0 to 100). A transfer level of 0% leads to incentive distribution of 20-0 under success and failure cases under all cost functions. A transfer level of 100% leads

[5] The literature usually employs screens with two colors only, leading to a 50% chance of success if a subject randomly picks a color. We chose to introduce an additional color to increase the range of targeted probability from $[.5, 1]$ to $[.(3), 1]$. This is because now a random choice of color leads to success only 1/3 of the time.

[6] The incentives correspond to 10% and 90% transfer levels in the case of the high transfer cost function.

to equalizing payoffs under success and failure cases, leading to about \$13.5, \$11, and \$8.5, respectively.

3.3 Experimental Results

We begin by evaluating the comparative static predictions from Section 3.1. We find that there is strong evidence supporting them (see Figures 3a and 3b).

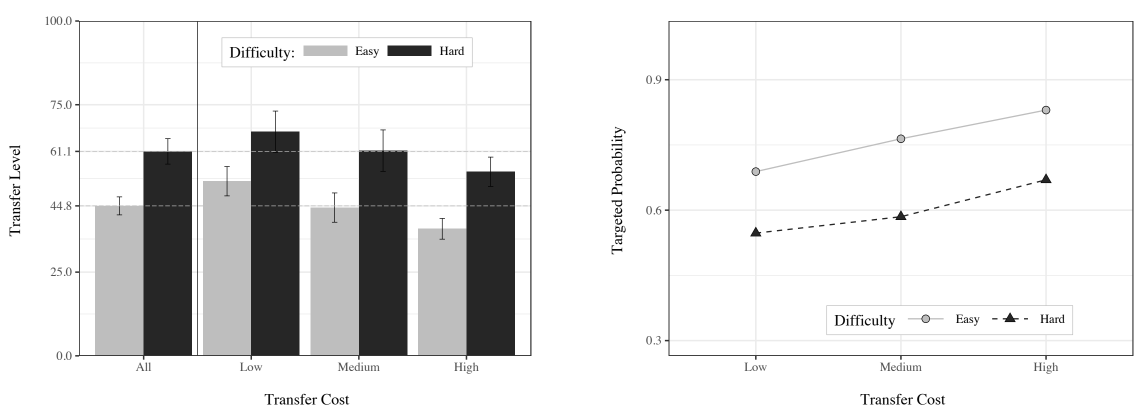
Subjects, on average, choose lower transfer, 44.8%, when the task is easy than when it is hard, 61.1% (Mann-Whitney U (MWU) test, $p-$value < 0.01). Furthermore, the hard task transfers first-order stochastically dominate those of the easy task. The result holds at the aggregate level and when we look at three different levels of transfer costs separately. For each transfer cost function (low, medium, and high), the hard task results in significantly higher transfers than the easy task.

Our analysis further reveals that subjects decrease their transfers as the cost of transfer increases. This finding holds true at both the aggregate level and for each level of task difficulty. Given the strong support for part (a) predictions, let us examine the prediction on the targeted probabilities of success (b). Here, we will use targeted probabilities calculated using subjects' performance, following the literature.[7] The probability of success grows with the increasing transfer cost, and it is lower for difficult than easy tasks at each transfer cost level, Figure 3b. Overall, the results depicted in Figure 3 provide substantial support for the comparative static predictions of our new framework.

Figure 3: Comparative statics: task type and cost of transfers

(a) Transfer levels

(b) Targeted Probability (performance)



[7] Note that since predictions in part (a) are supported, if we use our novel measure for targeted probability from expression (6), predictions in (b) will automatically follow. We examine the effects of task difficulty and transfer cost on performance for a more stringent test. See Appendix A for analysis comparing model targeted probability and performance measures.

Controlling for risk and ambiguity attitudes Let us examine the robustness of the above results controlling for subjects' risk and ambiguity attitudes. Table 1 reports results from a regression with transfer levels as dependent variable. The independent variables include the level of difficulty of the task and transfer cost dummy variables. Additionally we include risk and ambiguity attitudes as well as some demographic information.

Table 1: Payoff transfer Choice Regressions

	Dependent variable: Transfer		
	(1)	(2)	(3)
Difficulty: Easy	-16.272^{***}	-16.272^{***}	-16.272^{***}
	(2.487)	(2.497)	(2.463)
Cost: Low	13.078^{***}	13.078^{***}	13.078^{***}
	(3.046)	(3.059)	(3.017)
Cost: Medium	6.323^{**}	6.323^{**}	6.323^{**}
	(3.046)	(3.059)	(3.017)
Ambiguity Attitude	6.570^{***}	5.825^{**}	6.919^{***}
	(2.539)	(2.538)	(2.516)
Risk Aversion (HL)	2.890^{***}		3.340^{***}
	(0.777)		(0.780)
Risky Investment		0.138^{***}	0.169^{***}
		(0.047)	(0.047)
Compound Lottery	3.120	2.108	1.986
	(2.746)	(2.775)	(2.737)
Demographics	Yes	Yes	Yes
Observations	636	636	636

Note: Significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Both comparative static predictions are strongly supported even when we control for risk, ambiguity, and other observables. Additionally, this analysis reveals a significant effect of subjects' risk and ambiguity attitudes on their transfer levels. The more risk-averse a person is, the more they smooth the payoffs. This result implies that risk-averse individuals perform worse in the tasks.

Ambiguity-averse subjects transfer more than subjects with neutral ambiguity attitudes. A possible interpretation of this result is that paying attention may be perceived as an ambiguous act, with a degree of uncertainty surrounding the mapping from paying attention to success probabilities. Then, an ambiguity-averse person will desire to hedge against such ambiguity by choosing a more extensive transfer. This intriguing effect of ambiguity attitudes on attention presents an opportunity for further investigation in future work.

References

Agranov, Marina and Pietro Ortoleva, "Stochastic choice and preferences for randomization," *Journal of Political Economy*, 2017, *125* (1), 40–68.

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.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch, "hroot: Hamburg registration and organization online tool," *European Economic Review*, 2014, *71*, 117–120.

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

_ , Dániel Csaba, John Leahy, and Oded Nov, "Rational inattention, competitive supply, and psychometrics," *The Quarterly Journal of Economics*, 2020, *135* (3), 1681–1724.

_ , 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," *Working Paper*, 2022.

_ 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.

Dewan, Ambuj and Nathaniel Neligh, "Estimating information cost functions in models of rational inattention," *Journal of Economic Theory*, 2020, *187*, 105011.

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.

Greiner, Ben, "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.

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.

Ilinov, Pavel and Ole Jann, "An equivalence between rational inattention problems and complete-information conformity games," *Economics Letters*, 2022, *214*, 110444.

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

Lin, Yi-Hsuan, "Stochastic choice and rational inattention," *Journal of Economic Theory*, 2022, *202*, 105450.

Luo, Shaowen and Kwok Ping Tsang, "Elasticity of attention and optimal monetary policy," *Economics Letters*, 2020, *194*, 109393.

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

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt, "Rational inattention: A review," *Journal of Economic Literature*, 2023, *61* (1), 226–273.

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," *Journal of the European Economic Association*, 2021, *19* (3), 1899–1935.

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.

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

Zhang, Yuhua and Congming Mu, "Optimal ownership of entrepreneurial firms with rational inattention," *Economics Letters*, 2021, *209*, 110116.

Appendix (For Online Publication)

A Comparing measures of attention

In this appendix, we present a comparison between the estimation of attention levels provided by our model and the one existing in the literature (i.e., performance, or success frequency).

We start by premising that the goal of our experimental design is to validate the model by examining the comparative static results obtained by varying the task difficulty and the costs of payoff transfers. As such, and due to time constraints, the design did not include multiple repetitions of the task for the same combination of parameters. Thus, for each set of parameters, we can only observe whether the subject completed the task successfully (1) or not (0). This implies that, at the individual level, the comparison between success frequencies and the success probability predicted by the model cannot be derived in the current setting. We can nonetheless provide a coarse (aggregate level) comparison.

Let us aggregate the data in bins as follows. First, we take all the cases in which the model estimate of targeted probability is within a certain interval, for example, between 0.4 and 0.5. Then, we calculate the frequency of actual successes for all those instances. That is, we take all the cases for which the model estimate of targeted probability was within a bin and calculate the fraction of successful performance out of all those cases.

Figure 4: Comparison of model and performance estimates

(a) Bin size 0.05 (b) Bin size 0.025

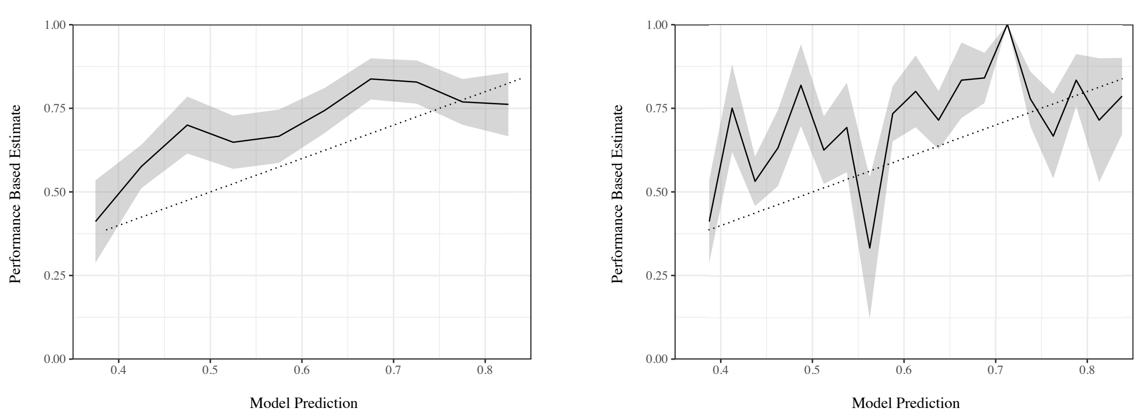


Figure 4 presents the results of this exercise with bin sizes of 0.05 and 0.025. Note that the smaller the bin size, the fewer cases will fall in each bin, leading to more erratic (and noisy) performance measure estimates. We can see the model average and performance in each interval starting from the 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. We also observe the actual performance being higher than the model predicts. With fewer data points per bin, i.e., with a bin size of 0.025, there is a fair amount of variation. We note that this is an exploratory comparison, with data aggregated over individuals that have presumably different attention costs and different risk preferences. We leave more robust exploration of the connection between the two measures of attention for future work. A different design, not varying the task difficulty and the transfer cost, and extending the number of repetitions within the same parameter set, could allow for a more precise measure of success frequencies at the individual level.

B Extensions

B.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 rearrangement, lead to:

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

$$K'(\mathcal{P}) = u(Y-q(x)) - u(xY-q(x)) \tag{9}$$

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 (8) and (9) 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.

B.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-parameterize 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:

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

The first order conditions are:

$$(1 - x) = c'(x) \tag{10}$$

$$L = q'(x) \tag{11}$$

The second order conditions for a maximum are:

$$-q''(x) < 0$$

$$c''(p)q''(x) > 1$$

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

$$c'(-q'(x)) = 1 - x \tag{12}$$

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