

Essays in Information and Behavior

Ambuj Dewan

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

© 2017
Ambuj Dewan
All rights reserved

ABSTRACT

Essays in Information and Behavior

Ambuj Dewan

This dissertation comprises three essays in behavioral and information economics. The first, “Estimating Information Cost Functions in Models of Rational Inattention,” uses laboratory data to analyze the properties of cost functions in models of rational inattention and determine their functional forms. The second, “Promises and Pronouncements,” uses a laboratory experiment to determine whether the propensity to tell monetarily advantageous lies depends on the ability to control the final outcome; in other words, whether renegeing on a commitment (breaking a promise) is more or less likely than lying about something out of one’s control (making a false pronouncement). The third, “Costly Information and Multiattribute Choice” provides an information-theoretic explanation for some commonly observed phenomena in consumer choice when goods are defined by multiple characteristics.

This page intentionally left blank.

Contents

List of Tables	iii
List of Figures	v
Acknowledgements	vii
1 Estimating Information Cost Functions in Models of Rational Inattention	1
1.1 Introduction	1
1.2 Related Literature	6
1.3 Theoretical Framework	8
1.4 Uniform Guess Tasks	22
1.5 Cost Functions	27
1.6 Experimental Design	37
1.7 Basic Results and Categorization	40
1.8 Model Selection	53
1.9 Additional Results	62
1.10 Application to the Delegation of Investment	70
1.11 Conclusion	78

2 Promises and Pronouncements	81
2.1 Introduction	81
2.2 Related Literature	83
2.3 The Game	87
2.4 Experimental Design	102
2.5 Results	104
2.6 Discussion and Conclusion	114
3 Costly Information and Multiattribute Choice	117
3.1 Introduction	117
3.2 The Model	119
3.3 Focusing	123
3.4 Salience	126
3.5 Conclusion	130
Bibliography	131
Appendices	139

List of Tables

1.1	Properties of cost functions	37
1.2	Laboratory Demographics	41
1.3	Regressions of correctness on incentive level and demographic covariates . .	44
1.4	Categorization of subjects	53
1.5	Performance functions estimated and their corresponding cost functions . .	54
1.6	Model Selection for Responsive Subjects	58
1.7	Average AIC and Rank for Estimated Models	60
1.8	Demographics and Categorization: Logit Regressions	63
1.9	Model Selection and Demographics	64
1.10	Linear regression of reaction time on incentive level	66
1.11	Linear regression of correctness on reaction time	66
2.1	Laboratory Demographics	105
2.2	Categorization of Receivers	107
2.3	Categorization of Senders (1)	109
2.4	Categorization of Senders (2)	110
2.5	Two-way contingency table for lying and beliefs	112
2.6	Regression of Lying on Scenario	113

B.1	Linear regression of correctness on incentive level and demographic covariates in the “angle” tasks	150
C.1	Online Demographics	152

List of Figures

1.1	Fixed cost for information acquisition.	30
1.2	Normal signals with cost of precision given by $K(\zeta) = 4\zeta^2$	34
1.3	Quadratic costs.	36
1.4	Incentive display for a task	38
1.5	Arrangement of dots for a task	39
1.6	Isotone nonparametric regression of correctness on incentive level	42
1.7	Categorization of subjects	43
1.8	Unrestricted cubic polynomial regression of correctness on incentive level for Subjects 1 and 19. The former rejects NIAC (and therefore rejects rational inattentiveness), and the latter fails to reject NIAC.	47
1.9	Linear regressions of correctness on incentive level for two subjects. The left panel shows an unresponsive subject, and the right panel shows a responsive one.	49
1.10	Concave costs	56
1.11	Best-fit binary performance	59
1.12	Best-fit logistic performance	60
1.13	Best-fit concave (normal) performance	61

1.14	Histogram of reaction times for all subjects	68
1.15	Histogram of reaction times for responsive subjects	69
1.16	Optimal payment as a function of potential net return, fixed costs	73
1.17	Principal's expected payoff as a function of payment for $x = 20$, mutual in- formation costs	74
1.18	Optimal payment as a function of potential net return, mutual information costs	75
1.19	Optimal payment as a function of potential net return, normally-distributed signals	76
1.20	Optimal payment as a function of potential net return, hybrid model with empirical proportions of types	77
1.21	Optimal payment as a function of potential net return, hybrid model with 50% fixed-cost types	77
2.1	The Promises scenario	90
2.2	The Pronouncements scenario	90
2.3	The full game	91
2.4	The rematching procedure	103
2.5	Histogram of first-order beliefs	106
2.6	Histogram of first-order beliefs	111
B.1	Angle display for a task	151

Acknowledgements

First and foremost, I would like to thank my advisors, Navin Kartik and Mark Dean, whose support and guidance miraculously transformed by scatterbrained thoughts into something approaching coherence.

Next, I must thank my coauthors, Nathaniel Neligh and Charles Maurin, who each coauthored a chapter of this dissertation. Working alongside them not only was rewarding, but also challenged me intellectually and pushed me to work my hardest.

I am grateful to the Columbia Experimental Laboratory in the Social Sciences (CELSS) and the microeconomic theory colloquium at Columbia University for providing research funding for the experiments in this dissertation.

I would also like to extend my thanks to the professors who advised me on various aspects of this dissertation: Alessandra Casella; Judd Kessler; José Luis Montiel Olea; Pietro Ortoleva; and Michael Woodford.

Finally, I thank all my colleagues who helped me test out my experiments at CELSS. There are too many of them to thank individually, but this dissertation would not have been possible without their remarkable ability to tolerate bugs, crashes, glitches, and all manner of technical nonsense.

To Mom, Dad, and Ambika, for their unwavering love and encouragement.

Chapter 1

Estimating Information Cost Functions in Models of Rational Inattention

This chapter is coauthored with Nathaniel Neligh.

1.1 Introduction

It has been observed in many settings that people have a limited capacity for attention, and this affects the decisions they make. For example, Chetty et al. (2009) demonstrate that consumers underreact to non-salient sales taxes; De los Santos et al. (2012) show that people only visit a small number of online retailers before making book purchases; and Allcott and Taubinsky (2015) provide evidence that people do not fully think about energy efficiency when making purchasing decisions about light bulbs.¹ Several laboratory experiments demonstrating limited attention have been conducted, including Gabaix et al. (2006), Caplin and Martin (2013), and Caplin and Dean (2014).

A common explanation for this phenomenon is the theory of *rational inattention* (Sims, 2003; Sims, 2006; Caplin and Dean, 2015; Matějka and McKay, 2015). This theory posits that people rationally choose the information to which they attend, trading off the costs

¹For a survey that discusses many more similar field studies, see DellaVigna (2009).

of paying more attention with the ensuing benefits of better decisions. This decision-making process occurs in two stages. In the first stage, the decision-maker chooses what information to acquire and pays costs accordingly. In the second stage, the decision-maker uses the information she acquired to make decisions.

The first-stage information costs are typically assumed to have some fixed functional form. Most common is mutual information (cf. Sims, 2003; Matějka and McKay, 2015), which measures the expected reduction in entropy from a decision-maker's prior beliefs to their posterior beliefs. Other cost functions, such as fixed costs for information acquisition (e.g. Grossman and Stiglitz, 1980; Barlevy and Veronesi, 2000; Hellwig et al., 2012) or costs for increasing the precision of normally distributed signals (e.g. Verrecchia, 1982; Van Nieuwerburgh and Veldkamp, 2010), have also been used. However, little is known about what form these costs take in reality, and different assumptions on these costs can lead to starkly different predictions.

In this paper, we use a laboratory experiment to characterize these information costs,² a crucial input for models of rational inattention. Subjects complete a series of perceptual tasks with fine-grained variation in the levels of potential rewards. For each reward level, we observe both the correct answer and the subject's response. We interrogate these data in two ways: (1) testing various properties of cost functions; (2) determining which models of information costs are consistent with observed behavior.

The cost function properties of greatest interest to us are continuity, convexity, and perceptual distance, the last of which refers to nearby states being harder to distinguish from each other than distant ones. The presence or absence of each of these properties

²We also conducted an online experiment, the results of which we report in Appendix C.

can have profound impacts on the predictions of a model of rational inattention.

Continuity and convexity are important characteristics of many cost functions, such as the aforementioned mutual-information cost function. The convexity of a cost function can greatly affect model predictions. For example, Van Nieuwerburgh and Veldkamp (2010) study a portfolio choice problem in which investors choose which assets to learn about and how much to learn about each of them. Depending on the convexity of the investor’s utility and cost functions, it can be optimal for the investor to learn about all available assets or to simply concentrate their attention on a single asset; utility and cost functions that imply concave objective functions result in generalized learning, whereas those that imply convex objective functions result in specialized learning. Convexity also has implications for comparative statics in models of rational inattention. As we prove in this paper, continuity and a specific form of convexity³ together imply that gross payoffs (excluding information costs) change continuously in incentives.

Whether or not an information cost function embeds some notion of perceptual distance can also have an effect on model predictions. Morris and Yang (2016) study a global game of regime change where players acquire information at a cost. If nearby states of the world are sufficiently costlier to distinguish from each other as compared to distant ones, then the game has a unique equilibrium. This result stands in contrast to the result of Yang (2015), who studies a model where information costs are such that nearby states are equally as easy to distinguish from each other as distant ones and finds that the game has multiple equilibria.

In our experiment, we find that roughly one-third of the subjects whose performance

³We call this form of convexity “almost strict convexity.” It is formally defined in Section 1.3.

on the tasks improves with increasing potential rewards (whom we call “responsive subjects”) have behavior that is consistent with “well-behaved” (i.e. continuous, convex) cost functions. Roughly 60% of the responsive subjects have behavior consistent with a cost function that embeds some notion of perceptual distance, in contrast to the mutual information cost function, which does not embed such a notion.

The second important set of analyses in our paper fits various classes of cost functions to our subjects’ data and selects the best fit for each subject. From subjects’ responses for each reward level, we infer how their performance in the experimental tasks changes with potential rewards; put differently, we estimate a *performance function* that traces out the relationship between the potential reward and the probability of success. Each class of cost functions implies a specific performance function, and so our model-fitting exercise involves determining which of these model-implied performance functions most closely reflects each subject’s performance data.

Of particular interest to us are cost functions with fixed costs for information acquisition, normal signals with linear precision costs, and the mutual information cost function, because of their usage in the economic literature. The first implies a binary performance function with two levels of performance, the second implies a concave performance function, and the third implies a logistic performance function. Of the set of models we estimate, we find that the data of the subjects who are responsive to incentives are best fit by one of these three models, with roughly a quarter of subjects best fit by the first model, one-seventh of subjects best fit by the second model, and two-thirds of subjects best fit by the third model. Thus, while there is some heterogeneity in the population with respect to which cost functions best reflect human behavior, the set of potential cost functions

that we need to consider can be reduced to three cost functions commonly found in the literature.

Finally, we apply the fixed-cost and mutual-information cost functions to a simple principal-agent model of investment delegation and find that they imply starkly different comparative statics results. Fixed costs for information acquisition imply payment structures that are discontinuous in potential returns, whereas mutual information implies continuous payment structures.

To our knowledge, our paper is the first to use an experiment with fine-grained variation in incentives to infer properties of information cost functions. This fine-grained variation is crucial for testing the continuity and convexity of cost functions and for estimating subjects' performance functions, which is crucial for our model-fitting exercise. Although several papers have examined competing hypotheses of dynamic evidence accumulation using perceptual data,⁴ ours is the first to run a "horse race" between a large number of types of cost functions in a static model of rational inattention.⁵

The remainder of the paper proceeds as follows. Section 1.2 reviews various experimental approaches to limited attention in the literature. Section 1.3 presents the theoretical framework that we use in this paper. Section 1.4 introduces the type of task that we implement in our experiment and situates it within our theoretical framework. Section 1.5 introduces various models of cost functions and applies them to the tasks of our experiment. Section 1.6 presents our experimental design and compares it to previous ex-

⁴We discuss some of this literature in the following section.

⁵Cheremukhin et al. (2015) use a "horse race" approach to select between probability weighting functions while assuming a specific but flexible functional form for information costs; their comparison is not between different information cost functions.

periments about limited attention. Section 1.7 presents and discusses basic experimental results and categorizes subjects according to the behaviors they exhibit. Section 1.8 fits various models of cost functions to the subjects' data and runs a "horse race" to determine which is the best fit for each subject. Section 1.9 presents additional results relating to demographics and reaction times. Section 1.10 presents an application of our results to the delegation of investment. Section 1.11 concludes. Most proofs are relegated to Appendix A. Additional experimental results are reported in Appendices B and C.

1.2 Related Literature

One approach that has been taken in the experimental literature to examine limited attention has been to give subjects a series of perceptual and/or cognitive tasks and observe their responses. Several experiments use this technique to test implications and fit parameters of models of limited attention. Caplin and Dean (2014) use a task involving the counting of differently-colored balls to test the necessary and sufficient conditions for a model of rational inattention with discrete choices where information acquisition is modeled as a static process. Caplin and Dean (2013) use that same task to estimate subjects' cost parameters in a mutual information cost function. Shaw and Shaw (1977) employ a protocol where stimuli are presented at random locations on a tachistoscope.⁶ Using the data they obtained, they estimate how much attention the subjects allocated to those locations on their visual field. Gabaix et al. (2006) use an experiment involving the comparison of multiattribute choices to calibrate a model of myopic search.

⁶A tachistoscope is a device used to present visual stimuli for a controlled duration. It has become much less common in behavioral research since the advent of personal computers.

This parameter-fitting approach has also been extended outside of the domain of visual perception to the domain of relative value assessment. Pinkovskiy (2009) and Chere-mukhin et al. (2015) fit experimental data on choice over gambles to models of limited attention. Krajbich et al. (2010) ask subjects to choose between pairs of snack food items that the subjects have rated on a numerical preference scale, and using choice and reaction-time data, they estimate the parameters of a drift-diffusion model (DDM), where information acquisition is modeled as a dynamic process.

There is a large literature that uses choice and reaction-time data to compare models of evidence accumulation. For example, Woodford (2014) presents a model of dynamic evidence accumulation with mutual-information costs and uses Krajbich et al.'s data to compare the fit of his optimizing model to the DDM, and Ratcliff and Smith (2004) use data from several experiments to compare the fits of four different dynamic evidence accumulation models.

Our experiment synthesizes these approaches. Using choice data from an experiment with perceptual tasks, we test hypotheses about behavior, including necessary and sufficient conditions for rational inattention, convexity and continuity of cost functions, and whether subjects are better at distinguishing nearby states from distant ones. We also estimate parameters of various models of rational inattention and then compare their fits to each other.

The tasks in our experiment involve the perception of numerosity, a long-standing area of research in perceptual psychology. Our tasks are similar to those implemented by Saltzman and Garner (1948) and Kaufman et al. (1949), who present subjects with fields of randomly arranged dots whose numerosity they have to judge. More recently, the ball-

counting tasks of Caplin and Dean (2014) have also involved the perception of numerosity.

1.3 Theoretical Framework

Various models of limited attention and imperfect perception have been proposed in the literature. These models can be classified along two axes: optimizing vs. non-optimizing; and static vs. dynamic.

In optimizing models, the information acquired by a decision-maker depends on an explicit choice that is made optimally given their cost and/or constraints. As a result, any “errors” observed in their decisions can be considered rational. By contrast, in non-optimizing models, errors in decisions or perception are the result of some exogenous process that is not chosen by the decision-maker; the choice of information is not explicitly modeled or is assumed to be out of the decision-maker’s control.

Models of limited attention and imperfect perception can also be classified according to whether they are static or dynamic. In static models, information acquisition is modeled as a one-time occurrence; either the decision-maker makes a single information-acquisition choice, or information is modeled as if it is delivered to the decision-maker all at once. On the other hand, in dynamic models, information is accumulated over time. That is not to say that static models cannot be applied to situations where information acquisition is a dynamic process; static models simply restrict their scope to the *outcome* of such a process, not the process itself.

We refer to the class of optimizing models as models of *rational inattention*. In particular, this paper considers a general static model of rational inattention with finite state and

action spaces and additively separable costs. Such a model is the focus of Caplin and Dean (2015) (henceforth CD15), who derive necessary and sufficient conditions for observed behavior to be consistent with it. Other papers have considered static rational inattention models with specific cost functions. For example, Matějka and McKay (2015) derive the implications of the mutual-information cost function, and Woodford (2012) considers the prior-independent channel-capacity cost function.

There are also several static, non-optimizing models of limited attention and imperfect perception. In consumer choice, one class of papers studies the imperfect perception of attributes of multi-attribute choices and introduces distortions to those attributes that depend on the set of available choices (e.g. Bordalo et al., 2013; Kőszegi and Szeidl, 2013; Bushong et al., 2016). These distortions are exogenously imposed; they are not chosen by the decision-maker. The size and extent of these distortions can be seen as a measure of the consumer's inattention. In perceptual psychology, signal detection theory (cf. Chapter 12 of Frisby and Stone, 2010) provides an important theoretical framework for studying imperfect perception. In this framework, states of the world are observed with exogenously given noise, and the decision-maker must determine what the most likely state was given their observations. We implement a task of this nature in our experiment; however, we model the noise process as an endogenous choice.

Dynamic models of evidence accumulation have a long tradition in mathematical psychology. In drift-diffusion models (DDMs) (e.g. Ratcliff, 1978; Diederich, 1997), evidence is modeled as a stochastic process that evolves according to a diffusion process (Smith, 2000), such as Brownian motion. The decision-maker stops gathering evidence and makes a decision when this process hits some (possibly time-dependent) boundary. This bound-

ary is often exogenously given, as in Ratcliff (1978), implying a non-optimizing modeling approach. However, under some conditions, the boundary can be derived as the result of an optimal stopping problem (e.g. Fudenberg et al., 2015; Tajima et al., 2016). Other optimizing approaches consider the optimal selection of the intensity of evidence accumulation when the stopping rule is exogenously given (e.g. Woodford, 2014) or the optimal selection of both evidence accumulation intensity and stopping rule (e.g. Moscarini and Smith, 2001). The vast majority of these dynamic evidence accumulation models restrict their focus to situations where the decision-maker must choose between two options, though Moscarini and Smith extend their model to consider situations with multiple discrete choice alternatives.

In this paper, we adopt a static framework that allows us to consider any finite number of options for the decision-maker as well as a flexible choice of information-acquisition technologies.

1.3.1 General Model

In remainder of this section, we present a general framework for analyzing problems of rational inattention. In this framework, there is an unknown state of the world about which an decision-maker (DM) can choose to acquire information. This information affects her beliefs about the state of the world. After obtaining this information, she makes a decision that maximizes her payoff given her beliefs.

We model information as a collection of probabilistic mappings from states of the world to a set of subjective signals. We define an *information structure* to be a set of

conditional distributions of signals given states. Observing a signal generates a corresponding posterior belief over states. Given this posterior belief, the DM maximizes her payoff by guessing the most likely state. Each information structure has a cost associated with it.

We remain agnostic about what the exact source of information costs is. Information costs could represent cognitive or physical effort exerted in learning about the true state, as well as the opportunity cost of time spent doing so.

This framework has several beneficial features. Firstly, it has the same behavioral implications as the model of CD15, which means we can apply their necessary and sufficient conditions for models of rational inattention to the problems we study. Secondly, it expresses information structures as stochastic matrices, which as demonstrated later in the paper, will permit us to easily compare information structures and to define a simple geometric notion of convexity of information costs.

Let $\Theta = \{\theta_i\}_{i=1}^{|\Theta|}$ be a finite state space, let $M = \{m_i\}_{i=1}^{|M|}$ be a finite signal space,⁷ and let $A = \{a_i\}_{i=1}^{|A|}$ be a finite action space, with $|M| \geq |A|$ so that there are at least as many signals as there are actions. Let $\pi = (\pi_i)_{i=1}^n \in \Delta(\Theta)$, where $n := |\Theta|$, be the DM's prior

⁷ Given that the state space is finite, the finiteness of the signal space is not a substantive restriction. In fact, if we assume that more informative information structures are costlier (our Restriction E, presented later in the paper), it can be shown that given a finite state space, a DM never need use more than a finite number of signals. This follows from Proposition 4 of Kamenica and Gentzkow (2010). They study a game where the information structure and the action are chosen by different players, but if we assume those players' preferences are perfectly aligned, then ignoring information costs, our framework maps onto theirs. By their Proposition 4, if a DM employs an information structure with an infinite number of signals, then ignoring information costs, she could have done at least as well with an information structure with a finite number of signals. Moreover, since the former information structure is more informative than the latter, it is costlier. Therefore, the DM will choose to use a finite number of signals.

over Θ . Each action-state pair (a, θ) has an associated utility $u(a, \theta)$. The DM maximizes:

$$\mathbb{E}_{\gamma_Q} \left[\mathbb{E}_{\langle \pi | m \rangle} [u(a, \theta)] \right] - C(\pi, Q) \quad (1.1)$$

where Q is an information structure (a collection of conditional signal probabilities, given states), γ_Q is the distribution of posterior beliefs it induces, and $\langle \pi | m \rangle$ is the posterior belief associated with signal m .

As explained above, the DM's problem has two stages. First, she selects an information structure Q . She then observes a signal m according to that information structure, which gives her a posterior belief $\langle \pi | m \rangle$. Second, given this posterior belief, she chooses an action a to maximize her expected payoff.

We can express this problem more formally using matrix notation. Let $\Pi = \text{diag}(\pi)$.⁸ Let $U \in M_{|A| \times |\Theta|}(\mathbb{R})$ be a matrix with entries $u_{i,j} := u(a_i, \theta_j)$, i.e. the utility of taking action i in state j . We refer to U as the *payoff matrix*.

Let \mathcal{Q} be the space of right-stochastic matrices of dimension $|\Theta| \times |M|$, and let \mathcal{D} be the space of right-stochastic matrices of dimension $|M| \times |A|$. $C : \Delta(\Theta) \times \mathcal{Q} \rightarrow \bar{\mathbb{R}}$ gives the cost⁹ of selecting an information structure from \mathcal{Q} , given a prior in $\Delta(\Theta)$.¹⁰

⁸ $\text{diag}(x)$ is the square matrix that has the entries of x in order on its diagonal and zeroes elsewhere.

⁹ $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ is the set of extended reals. If for some $\tilde{\pi}$ and \tilde{Q} , $C(\tilde{\pi}, \tilde{Q}) = \infty$, then the cost of the information structure \tilde{Q} given $\tilde{\pi}$ is infinite, and the DM will never select it, provided there is at least one information structure available at a finite cost.

¹⁰In principle, though the cost-function approach implies flexibility in the selection of information structures, it can accommodate restrictions on the space of available information structures as well. For example, if $\hat{\mathcal{Q}}$ is the set of admissible structures, then $C(\pi, Q)$ is finite if $Q \in \hat{\mathcal{Q}}$ and infinite otherwise. If a modeler wishes to impose an exogenous process of information acquisition, then he may simply set $\hat{\mathcal{Q}}$ to be a singleton.

The decision-maker's problem, then, is (cf. Leshno and Spector, 1992):¹¹

$$\max_{Q \in \mathcal{Q}, D \in \mathcal{D}} \text{tr}(QDU\Pi) - C(\pi, Q) \quad (1.2)$$

where the entries of Q are $q_{i,j} = \Pr(m_j|\theta_i)$, i.e. the probably of signal m_j in state θ_i , and the entries of D are $d_{i,j} = \Pr(a_j|m_i)$, i.e. the probability of selecting action a_j given signal m_i . We denote the distribution of posteriors (with finite support) that Q induces by $\gamma_Q \in \Delta(\Delta(\Theta))$. The i -th row of Q represents the conditional distribution of signals given state θ_i , and so Q can be seen as a collection of signal distributions given states. We refer to D as the *decision matrix*.

We refer to the maximand in (1.2) as the *net payoff* and its first component as the *ex-ante gross payoff*. Specific realizations of this payoff are called the *ex-post gross payoff*. Where it will not cause confusion, we will drop the “ex-ante” and “ex-post.”

This setup allows us to index decision problems of the form of (1.2) by (π, U) . In this paper, we will hold π fixed, and thus we will simply index decision problems by U where it will cause no confusion. If we give a DM a finite set of decision problems $\{U_i\}$, then we can observe the true state θ_i chosen by nature and action a_i chosen by the DM for each decision problem. Using the data set $\{(U_i, \theta_i, a_i)\}$ will allow us to infer the properties of $C(\cdot, \cdot)$. We refer to a data set of this type as *stochastic choice data*.

¹¹ $\text{tr}(X)$ denotes the trace of X , the sum of its diagonal entries.

At this point, it may be useful to consider a rewritten version of (1.2):

$$\max_{Q \in \mathcal{Q}, D \in \mathcal{D}} \sum_{i=1}^n \pi_i \left[\sum_{j=1}^{|A|} u_{j,i} \left[\sum_{k=1}^{|M|} q_{i,k} d_{k,j} \right] \right] - C(\pi, Q) \quad (1.3)$$

This version of the DM's problem allows us to see how the ex-ante gross payoff is constructed.

For each (i, j) , $\sum_{k=1}^{|M|} q_{i,k} d_{k,j} = \Pr(a_j | \theta_i)$ is the probability of taking action j in state i . Summing these probabilities and weighting by the utilities for each action gives $\sum_{j=1}^{|A|} u_{j,i} \left[\sum_{k=1}^{|M|} q_{i,k} d_{k,j} \right]$, the expected utility in state i , given the subject's choice of information structure and decision matrix. Summing these expected utilities and weighting by the probability of each state gives the total ex-ante gross payoff, $\sum_{i=1}^n \pi_i \left[\sum_{j=1}^{|A|} u_{j,i} \left[\sum_{k=1}^{|M|} q_{i,k} d_{k,j} \right] \right]$.

In this setup, for each $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, |A|\}$, π_i and $u_{j,i}$ are exogenous parameters. For each $i \in \{1, \dots, n\}$, $j \in \{1, \dots, |A|\}$, and $k \in \{1, \dots, |M|\}$, $q_{i,k}$ and $d_{k,j}$ are chosen by subjects. Though one cannot observe $q_{i,k}$ and $d_{k,j}$ separately, one can observe the products $q_{i,k} d_{k,j}$, i.e. if a DM solves the same decision problem repeatedly, one can observe how often each action is chosen in each state.

1.3.2 Cost Equivalence

As written, the model Subsection 1.3.1 generalizes the model of rational inattention studied by CD15, assuming a finite set of actions. In contrast to their model, our model allows for the cost of a distribution of posteriors to depend on the specific signals that generated each posterior in its support. Put differently, a version of their model with a finite number of actions is equivalent to ours with the following restriction.

Restriction A. Cost equivalence. For all priors π , $C(\pi, Q_1) = C(\pi, Q_2)$ whenever Q_1 and Q_2 induce the same distribution of posteriors.

However, as we show below, cost equivalence imposes no additional behavioral restrictions, and therefore, any behavior that is consistent with our model is also consistent with CD15. This result allows us to apply CD15’s necessary and sufficient conditions for rational inattention to our framework without imposing any restrictions.

Proposition 1. *Stochastic choice data are consistent with (1.2) iff they are consistent with CD15.*¹²

To outline the proof, the ‘if’ direction is obvious, since our model generalizes CD15. To see the ‘only if’ direction, suppose that $\{(U_i, \theta_i, a_i)\}$ can be rationalized by (1.2) with some cost function $C(\pi, Q)$. Define \mathcal{Q}_{γ_Q} to be set of information structures that induce the distribution γ_Q over posteriors, and define $\tilde{C}(\pi, Q) := \min_{R \in \mathcal{Q}_{\gamma_Q}} C(\pi, R)$. It is obvious that \tilde{C} satisfies cost equivalence. Moreover, since a given posterior distribution always induces the same ex-ante gross payoff, the DM should always choose the lowest-cost way of inducing that posterior distribution. Thus, the proof boils down to showing that this minimum is well-defined. Details are in Appendix A.

1.3.3 Testing for Rational Inattention

As CD15 demonstrate, observed behavior is consistent with their model if and only if it satisfies their “no improving attention cycles” (NIAC) and “no improving action switches”

¹²Though we have assumed a finite action space in our paper, the proof of Proposition 1 does not rely on this.

(NIAS) conditions.¹³ Their NIAC condition ensures that improvements to gross payoffs cannot be made by reallocating attention cyclically across decision problems, and their NIAS condition ensures that the DM's actions are optimal given the beliefs induced by her chosen information structure. Because our model is behaviorally equivalent to theirs, NIAC and NIAS are necessary and sufficient conditions for stochastic choice data to satisfy our model. Put differently, the DM fails to fulfill either of those two conditions if and only if there does not exist a cost function that rationalizes her stochastic choice data.

In our notation, the NIAC condition can be expressed as follows. Assume a fixed prior π , and let U_0, U_1, \dots, U_{J-1} be any set of two or more payoff matrices. Let Q_0, Q_1, \dots, Q_{J-1} and D_0, D_1, \dots, D_{J-1} be the corresponding information structures and decision matrices selected by the DM, and let D_i^j be a decision matrix that maximizes the gross payoff given payoff matrix U_i and information structure Q_j . Then the NIAC condition states:

$$\sum_{j=0}^{J-1} \text{tr} (Q_j D_j U_j \Pi) \geq \sum_{j=0}^{J-1} \text{tr} \left(Q_{(j+1) \bmod J} D_j^{(j+1) \bmod J} U_j \Pi \right) \quad (1.4)$$

The NIAS condition can be expressed as follows. Assume a fixed prior π . Then for any payoff matrix U , let Q^* be the information structure and D^* be the decision matrix chosen by the DM. Then the NIAS condition states that for any $k \in \{1, \dots, |A|\}$ such that the k -th

¹³The NIAS condition is due to Caplin and Martin (2015). It is the key condition that characterizes their Bayesian expected utility model.

column of D^* (denoted by $d_{\bullet,k}^*$) has at least one nonzero entry and any $l \in \{1, \dots, |A|\}$:

$$u_{k,\bullet} \Pi Q^* d_{\bullet,k}^* \geq u_{l,\bullet} \Pi Q^* d_{\bullet,k}^* \quad (1.5)$$

where $u_{k,\bullet}$ and $u_{l,\bullet}$ are the k -th and l -th rows of U , respectively.

Proposition 2. *NIAC and NIAS are necessary and sufficient conditions for stochastic choice data to satisfy (1.2).*

Proof. This follows directly from Proposition 1 of the present paper and Theorem 1 of CD15. □

1.3.4 Responsiveness

A set of behaviors that is consistent with rational inattention is one where the DM's behavior does not change across decision problems; regardless of the decision problem, she chooses the same information structure and decision matrix. This is consistent with models such as signal detection theory, where the DM's information structure is exogenously given. In those cases, the DM simply does not respond to changes in incentives across decision problems. More interesting are cases where the DM does modify her behavior in response to changes in incentives.

Definition 1. Suppose that a DM is given a set of decision problems $\mathcal{U} = \{U_1, U_2, \dots, U_J\}$.

Further suppose that $\exists U, \tilde{U} \in \mathcal{U}$ satisfying the following: for each $i \in \{1, \dots, n\}$, let

$$\tau_i \in \operatorname{argmax}_{j \in \{1, \dots, |A|\}} u_{i,j}; \quad \forall i \in \{1, \dots, n\}, \quad \tilde{u}_{i,j} \geq u_{i,j} \text{ if } j = \tau_i \text{ and } \tilde{u}_{i,j} \leq u_{i,j} \text{ if } j \neq \tau_i,$$

with at least one strict inequality. Then we say the DM is *responsive (to incentives)*

(or *exhibits responsiveness*) if her behavior is such that $\Pr \left(a \in \operatorname{argmax}_{z \in A} \tilde{u}(z, \theta) \right) > \Pr \left(a \in \operatorname{argmax}_{z \in A} u(z, \theta) \right)$.

Put differently, a DM is responsive to incentives if for some pair of decision problems, her probability of taking a (gross) payoff-maximizing action increases when the utility associated with payoff-maximizing actions increases and the utility associated with non-payoff-maximizing actions decreases.

Responsiveness is a fairly intuitive condition for human behavior to fulfill. Roughly speaking, it says that people perform better (by choosing the best option more often) when the stakes are higher.

1.3.5 Continuity and Convexity

In this subsection, we establish sufficient conditions for continuous gross payoffs in general rational inattention problems. Roughly speaking, continuity and convexity of the information cost function imply gross payoffs that are continuous in incentives.

Restriction B. *Continuity.* $C(\pi, Q)$ is continuous in its second argument.¹⁴

Continuity is a typical assumption in much of economic analysis. In this case, it implies that gathering a small amount of additional information increases the total cost of information by only a small amount. This may seem like a fairly innocuous assumption, but it precludes some plausible cost functions, such as those with fixed costs for information acquisition, as will be seen in Section 1.5.

¹⁴Because $\bar{\mathbb{R}}$ is not metrizable with the standard Euclidean topology, we are implicitly assuming here that C maps to \mathbb{R} , i.e. it is nowhere infinite.

Restriction C. Convexity. $\forall \pi \in \Delta(\Theta), \forall \lambda \in (0, 1), \forall Q_1, Q_2 \in \mathcal{Q}, C(\pi, \lambda Q_1 + (1 - \lambda)Q_2) \leq \lambda C(\pi, Q_1) + (1 - \lambda)C(\pi, Q_2).$

Restriction D. Almost strict convexity. $\forall \pi \in \Delta(\Theta), \forall \lambda \in (0, 1), \forall Q_1, Q_2 \in \mathcal{Q}, C(\pi, \lambda Q_1 + (1 - \lambda)Q_2) \leq \lambda C(\pi, Q_1) + (1 - \lambda)C(\pi, Q_2),$ where the inequality is strict except possibly if Q_1 and Q_2 induce the same distribution of posteriors.

These notions of convexity can be contrasted with CD15’s. CD15 define a notion of convexity over the space of distribution of posteriors called “mixture feasibility”;¹⁵ however it is not testable. In our framework, cost functions are defined over stochastic matrices instead of the distributions of posteriors they induce. Since the space of stochastic matrices can be identified with a subset of Euclidean space, Restrictions C and D give us easily interpretable “geometric” notions of convexity. Moreover, Restriction D has testable implications.

It is clear that Restriction D is a special case of Restriction C. Restriction D is also a slight relaxation of strict convexity. We employ this slight relaxation because strict convexity would be incompatible with Restriction A, cost equivalence. To see why, consider the 2×2 case where $Q_1 = \begin{pmatrix} 0.25 & 0.75 \\ 0.25 & 0.75 \end{pmatrix}, Q_2 = \begin{pmatrix} 0.75 & 0.25 \\ 0.75 & 0.25 \end{pmatrix},$ and $Q_3 = 0.5Q_1 + 0.5Q_2 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}.$ In all three of these information structures, a given signal is induced by each state with the same probability. Therefore, the distribu-

tion of posteriors each one generates is the degenerate distribution on the prior, and by

¹⁵Restriction C involves mixtures of conditional signal probabilities, which could yield posteriors not generated by either information structure in the mixture, whereas mixture feasibility involves mixtures of distributions of posteriors whose support is the union of the supports of the distributions in the mixture.

cost equivalence, all three information structures should have the same cost. However, Q_3 is a convex combination of Q_1 and Q_2 , and so if the cost function were strictly convex it would have to have a strictly lower cost than Q_1 or Q_2 . Thus, Restriction A is incompatible with strict convexity. However, under almost strict convexity, Q_3 would be permitted to have the same cost as Q_1 and Q_2 .

In order to ensure that continuity and almost strict convexity imply continuous gross payoffs, we require two additional conditions.

Restriction E. *Monotonicity of information.* *Let R be a right-stochastic matrix of dimension $|M| \times |M|$, which we refer to as a garbling matrix. Then for any $\pi \in \Delta(\Theta)$ and $Q \in \mathcal{Q}$, $C(\pi, Q) \geq C(\pi, QR)$.*

Restriction E is equivalent to Condition K1 of CD15. If Q is an information structure and R is a garbling matrix, then Q can be thought of as containing all the information contained in QR , i.e. QR simply adds noise to Q . In this case, we shall say that Q *Blackwell-dominates* QR . As shown by Blackwell (1953), Q yields a (weakly) higher gross payoff than QR for any decision problem, given an optimal selection of decision matrices. Therefore, Restriction E implies that if one information structure is more informative than another, then it is also costlier. This restriction does not provide a complete order on information costs, since it is possible that two experiments are not ranked in the Blackwell sense. In other words, if Q_1 and Q_2 are information structures of the same dimension, there does not necessarily exist R of appropriate dimension such that $Q_1R = Q_2$ or $Q_2R = Q_1$.

CD15 show that this restriction is not testable; any stochastic choice data set that is

consistent with some cost function C is also consistent with some cost function \tilde{C} that satisfies Restriction E. Therefore, requiring it does not eliminate any additional sets of stochastic choice data from being consistent with a model of rational inattention.

Restriction F. Cost symmetry. *Let R be a $|M| \times |M|$ permutation matrix. Then $\forall Q \in \mathcal{Q}$, $C(\pi, Q) = C(\pi, QR)$.*

It is easy to see that cost equivalence (Restriction A) implies cost symmetry. This restriction says that the cost of an information structure is invariant to the labeling of its signals; only the conditional probabilities of generating each signal matter.

We have now established a set of sufficient conditions that ensure that the DM's ex-ante gross payoff is continuous in incentives.

Proposition 3. *Suppose that π is fixed and C satisfies Restrictions B, D, E, and F. Then the ex-ante gross payoff is continuous in U .*

Monotonicity of information and cost symmetry ensure that the decision matrix chosen by the DM can be fixed, which in turn ensures the convexity of the problem. While almost strict convexity does not ensure a unique solution to the problem, it does ensure that the optimal ex-ante gross payoff is single-valued, which together with the continuity of the cost function implies the result.

At this point, a clarification is in order. Proposition 3 is a statement about what the properties of an information cost function imply about behavior. To obtain a statement about what behavior implies about the properties of cost functions, we invoke the contrapositive: if gross payoffs are discontinuous in incentives, then this behavior cannot

be rationalized by an information cost function that satisfies Restrictions B, D, E, and F simultaneously. However, as we explained earlier in this subsection, Restriction E is not testable, and furthermore, since Restriction F is implied by the untestable cost-equivalence Restriction A, it is also untestable. Therefore, given stochastic choice data, we can assume the cost function that rationalizes it satisfies Restrictions E and F, and so if we observe that ex-ante gross payoffs are discontinuous in incentives,¹⁶ then this implies that the DM's cost function either is discontinuous or fails almost strict convexity.

The restrictions necessary for 3 are satisfied by many different cost functions. In particular, cost functions that can be expressed as a sum of convex functions of the entries of a stochastic matrix are almost strictly convex.

Proposition 4. *Let $\{c_{i,j}(\cdot, \cdot)\}$, $i \in \{1, \dots, n\}$, $j \in \{1, \dots, |M|\}$ be a collection of continuous functions on $\Delta(\Theta) \times [0, 1]$ such that for fixed π , each $c_{i,j}(\pi, \cdot)$ is a twice continuously differentiable function with \mathbb{R} with $\frac{\partial^2 c_{i,j}(\pi, \cdot)}{\partial q^2} > 0 \forall q \in (0, 1), \forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, |M|\}$. Then $C(\pi, Q) := \sum_{i,j} c_{i,j}(\pi, q_{i,j})$ satisfies Restrictions B through D.*

1.4 Uniform Guess Tasks

Consider a task where there is some unknown true state of the world that a decision-maker (DM) has to identify, and learning about the true state is costly. There are n possible states, each of which is *a priori* equally likely. The DM receives a reward r for correctly identifying the state and no reward for incorrectly identifying the state. Therefore, the

¹⁶The reader may have noticed that strictly speaking, observing a discontinuity is technically impossible without an infinite data set. We expound upon this point in the following section. For now, simply assume that the data strongly suggest that gross payoffs are discontinuous in incentives.

DM’s goal is to maximize her probability of correctly identifying the state, net of whatever costs she incurs in gathering information about the true state. We refer to tasks with this setup as *uniform guess tasks*.

An example of such a task is the type of task we implement in our experiment. In this type of task, which we refer to as the “dots” task, the DM is shown a screen with a random arrangement of dots. Her goal is to determine the number of dots on the screen, which is between 38 and 42, inclusive, with each possible number equally likely. She receives a reward r for correctly guessing the number of dots and no reward otherwise.

In our example, information costs could include the cost of effort exerted in counting dots, cognitive costs incurred in employing an estimation heuristic, or the opportunity cost of time spent trying to determine the number of dots.

In a uniform guess task, the DM’s goal is to choose an information structure to maximize:

$$\frac{r}{n} \sum_x \Pr(a = x | \theta = x) - C(\pi_{\text{unif}}, Q) \tag{1.6}$$

where θ is a state of the world, a is the subject’s guess of the state, and $C(\pi_{\text{unif}}, Q)$ is the cost associated with information structure Q and the uniform prior. Put differently, the DM rationally chooses an information structure to maximize her net payoff (1.6). Thus, uniform guess tasks can be considered problems of rational inattention. The cost function $C(\pi_{\text{unif}}, Q)$ is the primary object of interest in our experiment.

1.4.1 Applying the General Model and Performance Functions

The general model we presented in Section 1.3 can be applied to uniform guess tasks. In these decision problems, since the DM is trying to determine the true state, we can set $A = \Theta$. Moreover, $U = rI_n$ for some $r > 0$. Therefore, the DM's ex-ante gross payoff in this task can be written as $r\text{tr}(QD\Pi)$. Note that $P := \text{tr}(QD\Pi) = \frac{1}{n}\text{tr}(QD)$ is the ex-ante probability of correctly guessing the state, which we call their *performance*. For each reward r , the DM's optimal choice of Q and D induce an optimal performance $P^*(r)$, which we call the *performance function*. We can estimate $P^*(r)$ from stochastic choice data, and by studying its properties, we can infer the properties of the DM's cost function.

Given some choice of Q , the DM should optimally choose D so that given a signal, she chooses the most likely state; put differently, she should choose the matrix D that selects the maximal element in each column of Q . Thus, if the DM's choice of D is optimal, then:

$$P^*(r) = \frac{1}{n} \sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} q_{i,j} \quad (1.7)$$

1.4.2 Rational Inattention

In uniform guess tasks, NIAC can be tested by looking at the performance function $P^*(r)$.

Proposition 5. *In a set of uniform guess tasks, the DM's behavior is consistent with NIAC iff $P^*(r)$ is nondecreasing.*

In other words, the DM's behavior is consistent with NIAC if and only if she performs no worse when given higher incentives; she allocates her attention such that she pays

more attention to more valuable tasks.

For testing NIAS in uniform guess tasks, we require a more detailed summary of the data than the performance function; simply looking at which questions were answered correctly or incorrectly is insufficient.

Proposition 6. *In a set of uniform guess tasks, the DM's behavior is consistent with NIAS iff*

$$\forall x \in A, \forall y \in \Theta, \Pr(\theta = x|a = x) \geq \Pr(\theta = y|a = x).$$

In other words, the DM's behavior is consistent with NIAS if and only if the mode of each posterior distribution of states given an action is equal to that action; if an outside observer were to see the DM's actions without observing the true states, then his best guess of the true state for any task would be the answer given by the DM.

1.4.3 Responsiveness

In a set of uniform guess tasks, any pair of tasks with different reward levels satisfies the conditions required for the definition of responsiveness. This is easy to see: the utilities associated with payoff-maximizing actions are constant within each corresponding payoff matrix and higher in one payoff matrix than the other. Furthermore, the utilities associated with non-payoff-maximizing actions are all zero. Therefore, responsiveness in uniform guess tasks boils down to there being at least one pair of reward levels such that the DM performs strictly better at the higher reward level than at the lower one.

Thus, responsiveness implies that the performance function cannot be flat everywhere. In other words, it must have a region of strict increase.

1.4.4 Continuity and Convexity

In uniform guess tasks, Proposition 3 has a simple implication for behavior: if the conditions of the proposition are satisfied, then performance is continuous in reward.

Proposition 7. *If C satisfies Restrictions B, D, E, and F, then $P^*(r)$ is continuous.*

Proof. Since optimal gross payoffs are given by $rP^*(r)$, this follows immediately from Proposition 3. □

Using the same contrapositive reasoning as we did in Subsection 1.3.5, this means that if we observe a discontinuous performance function, then the DM's cost function either is discontinuous or violates almost strict convexity.

1.4.5 Perceptual Distance

In uniform guess tasks, where the action space is identified with the state space, we can establish a meaningful concept of perceptual distance. Perceptual distance refers to the notion that distant states are easier to distinguish from each other than nearby ones. For example, if $\Theta = \{1, 2, 3, 4, 5\}$, and the true state is $\theta = 2$, then the DM may be more likely to answer 1 (which is 1 away from 2) than she is to answer 5 (which is 3 away from 2). This is especially plausible if the states in Θ represent physical, measurable quantities. To give a more concrete example, when shopping for televisions, one is much more likely to misperceive a 27-inch screen as a 23-inch screen than as a 40-inch screen. We formalize this notion below.

Definition 2. Let ρ be a metric on Θ . Then in this task, the DM *evinces the perception of distance* iff $\forall x, y, z \in \Theta, \rho(x, y) > \rho(x, z) \implies \Pr(a = y | \theta = x) < \Pr(a = z | \theta = x)$.

In other words, the DM evinces the perception of distance if for each possible true state, she is more likely to give an answer close to the true state than one farther away from it.

Though one can define a metric on a given set in many different ways, it makes sense to take ρ to be a “natural” metric on Θ . For instance, if Θ is a subset of the real line as in the example above, then absolute value, $\rho(x, y) = |x - y|$, may be a sensible metric to use. Since the state space in our experiment is such a subset, absolute value is the metric we use in analyzing our experimental results.

1.5 Cost Functions

The space of admissible cost functions is vast. Indeed, any cost function $C : \Delta(\Theta) \times \mathcal{Q} \rightarrow \bar{\mathbb{R}}$ leads to behavior consistent with NIAS and NIAC. In this subsection, we introduce the classes of cost functions that are most relevant for our analysis and derive their behavioral implications.

1.5.1 Mutual Information

Mutual information is one of the most common information cost functions in the economic literature, having been used since at least Sims (2003). Mutual information is defined as the expected reduction in entropy from the prior to the posterior. If we denote

the mutual information cost function by I and entropy by H , then:

$$\begin{aligned}
I(\pi, Q) &:= \alpha(H(\pi) - \mathbb{E}[H(\pi|Q)]) \\
&= \alpha \left[- \sum_{i=1}^n \pi_i \ln \pi_i - \left(- \sum_{i=1}^n \pi_i \sum_{j=1}^{|M|} q_{i,j} \ln \left(\frac{\pi_i q_{i,j}}{\sum_{k=1}^n p_k q_{k,j}} \right) \right) \right] \quad (1.8)
\end{aligned}$$

where $\alpha > 0$ and we adopt the convention that $0 \ln 0 = 0$ and $0 \ln \frac{0}{0} = 0$.

Mutual information satisfies some of the restrictions of Section 1.3.

Proposition 8. *Mutual information satisfies Restrictions A, C, E, and F.*

In uniform guess tasks, we can derive closed-form expressions for the DM's optimal behavior.

Proposition 9. *Suppose that $\pi_i = \frac{1}{n} \forall i \in \{1, \dots, n\}$ so that the prior is uniform on Θ and that $C(\pi, Q) = I(\pi, Q) := \alpha(H(p) - \mathbb{E}[H(\pi|Q)])$, $\alpha > 0$, i.e. C is the mutual information cost function. Then, the probability of guessing the correct state is continuous in the reward.*

Moreover, this probability is given by $P^(r) = \frac{\exp(\frac{r}{\alpha})}{n-1+\exp(\frac{r}{\alpha})}$, and the probability of guessing any given incorrect state is $\frac{1}{n-1+\exp(\frac{r}{\alpha})}$.*

Proof. This follows directly from Proposition 1 of Matějka and McKay (2015). □

Note that mutual information is convex, but it can be shown that it is not almost strictly convex, i.e. it does not satisfy Restriction D.¹⁷ Therefore, Proposition 9 cannot be

¹⁷ For example, let $Q_1 = \begin{pmatrix} \frac{1}{4} & \frac{1}{16} & \frac{1}{8} & \frac{9}{16} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{16} & \frac{1}{4} & \frac{1}{16} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} \end{pmatrix}$ and $Q_2 = \begin{pmatrix} \frac{1}{8} & \frac{7}{96} & \frac{7}{48} & \frac{21}{32} \\ \frac{1}{8} & \frac{48}{7} & \frac{32}{7} & \frac{96}{49} \\ \frac{1}{8} & \frac{32}{7} & \frac{24}{96} & \frac{96}{49} \\ \frac{1}{8} & \frac{1}{24} & \frac{1}{96} & \frac{1}{96} \end{pmatrix}$. These two in-

formation structures induce different distributions of posteriors (albeit with the same support). However, it can be shown that any convex combination $Q_\lambda := \lambda Q_1 + (1 - \lambda) Q_2$, $\lambda \in [0, 1]$ of them is such that $I(\pi, Q_\lambda) = \lambda I(\pi, Q_1) + (1 - \lambda) I(\pi, Q_2)$, thereby violating almost strict convexity.

derived as a special case of Proposition 3.

There are two important things to note about Proposition 9. The first is that mutual information implies a logistic performance function, which is defined by a single parameter α . The second is that for any given true state θ , the probability of answering each incorrect state is equally likely, i.e. a DM with a mutual information cost function does not evince the perception of distance.

1.5.2 Fixed Costs

Another common model of information costs in the literature is “all-or-nothing” costs, where the DM begins with no information but can become completely informed about the state of the world if she pays a cost (e.g. Grossman and Stiglitz, 1980; Hellwig et al., 2012). Here, we generalize this form of costs by allowing for the DM to receive some information for free and pay a fixed cost to receive more information; we do not stipulate that she must become fully informed.

We can represent this situation as follows. Let there exist \underline{Q}, \bar{Q} such that $\underline{Q} = \bar{Q}R$ for some garbling matrix R (so that \bar{Q} is more informative than \underline{Q} in the Blackwell sense) and:

$$C(Q) = \begin{cases} 0, & Q = \underline{Q} \\ \kappa, & Q = \bar{Q} \\ \infty, & \text{otherwise} \end{cases} \quad (1.9)$$

According to this cost function, the DM can receive the information contained in \underline{Q} for

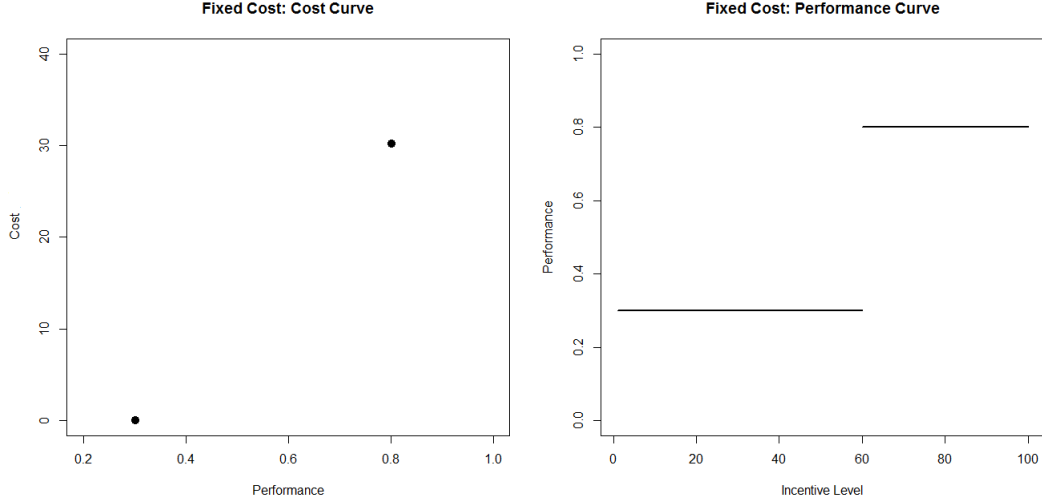


Figure 1.1: Fixed cost for information acquisition.

The left panel shows the cost function, and the right panel shows the resulting performance curve. Parameters are $\kappa = 30$, $\underline{q} = 0.3$, and $\bar{q} = 0.8$.

free, but she must pay a fixed cost κ to acquire the information in \bar{Q} .

Cost functions with fixed costs such as these can be seen as representing dual-system cognitive processes (cf. Stanovich and West, 2000; Kahneman, 2003). In such processes, a small amount of information may be acquired at a very low cost, but there is a fixed cost to acquiring more information. This implies a discontinuity in the cost function between information structures with “low” informativeness and those with “high” informativeness.

In uniform guess tasks, each of the two admissible information structures \underline{Q} and \bar{Q} induces a corresponding performance level \underline{q} and \bar{q} , respectively, the former of which is achievable for free and the latter of which costs κ . She is incapable of achieving a higher performance than \bar{q} . Therefore, the DM will pay the cost κ of acquiring information only when $r\bar{q} - \kappa \geq r\underline{q}$. This implies a binary performance function: for $r \leq \frac{\kappa}{\bar{q} - \underline{q}}$, the DM acquires no information and achieves \underline{q} , and for $r > \frac{\kappa}{\bar{q} - \underline{q}}$, the DM acquires enough information to achieve \bar{q} .

Cost functions of this subclass are easily recoverable from data by estimating the relationship depicted in the right panel of Figure 1.1 and finding the incentive level threshold at which the DM's performance level jumps.¹⁸

1.5.3 Normal Signals

Some authors, such as Verrecchia (1982) and Van Nieuwerburgh and Veldkamp (2010), have assumed that the DM receives normally distributed signals about the underlying state of the world, and she pays a higher cost for a more precise signal. In this subsection, we present such a situation. We then present a discretized version of it that is compatible with our model and generates the same predictions.

Let $\Theta \subset \mathbb{R}$, so that we can order its elements from smallest to largest as $\theta_1 < \theta_2 < \dots < \theta_n$,¹⁹ and suppose that the DM receives signals $\hat{m} \sim N(\theta, \sigma^2)$ about the state of the world θ . The DM can choose the precision $\zeta^2 := \sigma^{-2}$ of these signals, and she pays a cost $K(\zeta)$ accordingly, where K is increasing, convex, and differentiable.²⁰

Now suppose that the DM has been given a uniform guess task and has received a signal \hat{m} . Her belief that the state of the world is θ is, by Bayes' rule:

$$\Pr(\theta|\hat{m}) = \frac{\frac{1}{n} \frac{1}{\sigma} \phi\left(\frac{\hat{m}-\theta}{\sigma}\right)}{\sum_{i=1}^n \frac{1}{n} \frac{1}{\sigma} \phi\left(\frac{\hat{m}-\theta_i}{\sigma}\right)} = \frac{\frac{1}{\sigma} \phi\left(\frac{\hat{m}-\theta}{\sigma}\right)}{\sum_{i=1}^n \frac{1}{\sigma} \phi\left(\frac{\hat{m}-\theta_i}{\sigma}\right)} \quad (1.10)$$

where $\phi(\cdot)$ is the standard normal density. Notice that the denominator in (1.10) depends

¹⁸Technically speaking, what we recover is the performance levels \underline{q} and \bar{q} , not the specific \underline{Q} and \bar{Q} that induced them.

¹⁹Recall that Θ is finite, so it has a minimal element θ_1 and a maximal element θ_n .

²⁰Note that K is defined as a function of the positive square root of the precision. However, for the sake of parsimony, we will refer to it as the “cost of precision.”

only on \hat{m} ; it is the same for all θ . Therefore, if the DM is trying to determine the most likely state given her signal, she only needs to compare the numerators of (1.10) for each possible θ ; in other words, she only needs to find the state that maximizes the conditional probability density of her signal.

Since the normal probability density function is symmetric around its mean, the conditional probability density of her signal is maximized at θ_1 if $\hat{m} \leq \frac{1}{2}(\theta_1 + \theta_2)$, at θ_i if $\hat{m} \in [\frac{1}{2}(\theta_{i-1} + \theta_i), \frac{1}{2}(\theta_i + \theta_{i+1})]$ for $i \in \{2, 3, \dots, n-1\}$, and at θ_n if $\hat{m} \geq \frac{1}{2}(\theta_{n-1} + \theta_n)$.

This implies that if the DM guesses optimally given her signal, then her probabilities of guessing state i given true state j are:

$$\Pr(a = \theta_i | \theta = \theta_j) = \begin{cases} \Pr(\hat{m} \leq \frac{1}{2}(\theta_1 + \theta_2) | \theta = \theta_j), & i = 1 \\ \Pr(\hat{m} \in [\frac{1}{2}(\theta_{i-1} + \theta_i), \frac{1}{2}(\theta_i + \theta_{i+1})] | \theta = \theta_j), & i \in \{2, 3, \dots, n-1\} \\ \Pr(\hat{m} \geq \frac{1}{2}(\theta_{n-1} + \theta_n) | \theta = \theta_j), & i = n \end{cases} \quad (1.11)$$

Since signals are normally distributed, the probabilities in (1.11) can be rewritten as:

$$\Pr(a = \theta_i | \theta = \theta_j) = \begin{cases} \Phi(\zeta(\frac{1}{2}(\theta_1 + \theta_2) - \theta_j)), & i = 1 \\ \Phi(\zeta(\frac{1}{2}(\theta_{i+1} + \theta_i) - \theta_j)) - \Phi(\zeta(\frac{1}{2}(\theta_i + \theta_{i-1}) - \theta_j)), & i \in \{2, 3, \dots, n-1\} \\ \Phi(\zeta(\frac{1}{2}(\theta_{n-1} + \theta_n) - \theta_j)), & i = n \end{cases} \quad (1.12)$$

where Φ is the cumulative distribution function of the standard normal distribution.

Therefore, since the DM only receives a reward for a correct guess, her problem is:

$$\begin{aligned} \max_{\zeta \in [0, \infty)} \frac{r}{n} & \left[\Phi \left(\frac{1}{2} \zeta (\theta_1 - \theta_2) \right) + \sum_{i=2}^{n-1} \left(\Phi \left(\frac{1}{2} \zeta (\theta_{i+1} - \theta_i) \right) - \Phi \left(\frac{1}{2} \zeta (\theta_{i-1} - \theta_i) \right) \right) \right. \\ & \left. + 1 - \Phi \left(\frac{1}{2} \zeta (\theta_{n-1} - \theta_n) \right) \right] - K(\zeta) \end{aligned} \quad (1.13)$$

If the distance between consecutive states is constant so that $\exists \delta$ such that $\theta_i - \theta_{i-1} = 2\delta$ for $i \geq 2$, then (1.13) can be written as:

$$\max_{\zeta \in [0, \infty)} \frac{r}{n} [2\Phi(\zeta\delta) + (n-2)(2\Phi(\zeta\delta) - 1)] - K(\zeta) \quad (1.14)$$

This assumption of equidistant states allows us to draw some conclusions about whether a DM who receives normal signals necessarily evinces the perception of distance. The answer, in general, is no. This is because the lowest possible state θ_1 is guessed for any signal $\hat{m} \leq \frac{1}{2}(\theta_1 + \theta_2)$.²¹ If the costs of precision are very high, so that the DM selects a very low signal precision, then her distribution of signals may have fat enough tails that for some true state, guessing the lowest state is likelier than guessing the next outermost state, i.e. $\Pr(\hat{m} \leq \frac{1}{2}(\theta_1 + \theta_2) | \theta = \theta_j) > \Pr(\hat{m} \in [\frac{1}{2}(\theta_1 + \theta_2), \frac{1}{2}(\theta_2 + \theta_3)] | \theta = \theta_j)$ for some $j \geq 2$.

However, while we cannot conclude that a DM with normal signals necessarily evinces the perception of distance over the entire state space, we can say that she does if we restrict our focus to guesses of inner states (i.e. states θ_2 to θ_{n-1}).

Proposition 10. *In a uniform guess task with equidistant states, a DM with normal signals*

²¹A symmetric argument applies to the highest possible state.

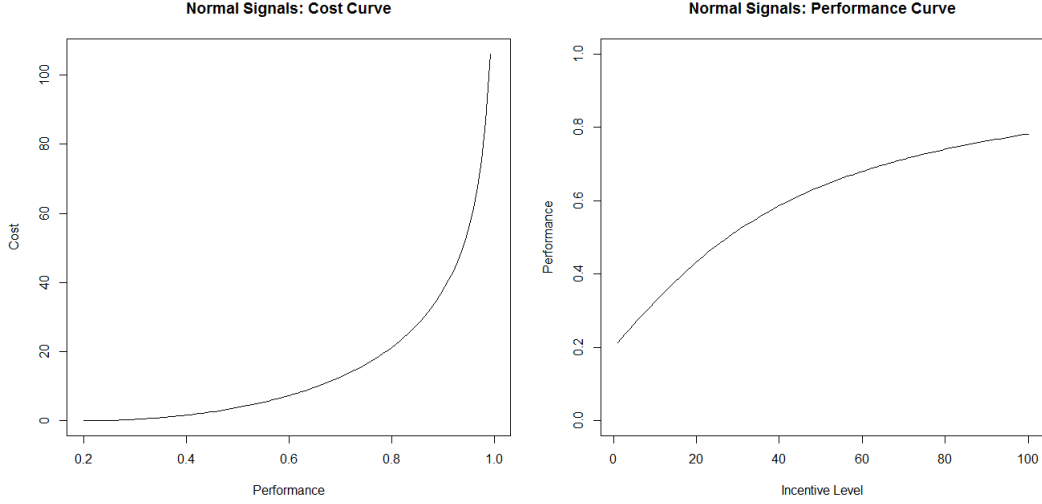


Figure 1.2: Normal signals with cost of precision given by $K(\zeta) = 4\zeta^2$.

The left panel shows the cost function, and the right panel shows the resulting performance curve.

evinces the perception of distance for guesses of inner states; that is to say, $\forall x \in \Theta, y, z \in \Theta \setminus \{\theta_1, \theta_n\}, |x - y| > |x - z| \implies \Pr(a = y | \theta = x) < \Pr(a = z | \theta = x)$.

The assumption of equidistant states also allows us to determine the shape of the performance function.

Proposition 11. *A DM with normal signals and a convex, increasing cost of precision $K(\cdot)$ with non-negative third derivative²² has a strictly concave performance function.*

This type of performance function is depicted in the right-hand panel of Figure 1.2.

Thus far in this subsection, we have implicitly assumed an uncountable signal space. However, the model of Section 1.3 assumes a finite signal space. We can reduce this setup to one with a finite signal space by assuming a set of signals $\{m_1, m_2, \dots, m_n\}$ that are

²²This assumption on the third derivative is a technical assumption. It holds if, for instance, K is linear in precision (i.e. quadratic in the square root of precision), as we assume later in the paper.

generated with the probabilities given in (1.12). In other words, the information structure

Q is such that:

$$\begin{aligned}
 q_{i,j} &= \Pr(m = m_i | \theta = \theta_j) \\
 &= \begin{cases} \Phi\left(\zeta\left(\frac{1}{2}(\theta_1 + \theta_2) - \theta_j\right)\right), & i = 1 \\ \Phi\left(\zeta\left(\frac{1}{2}(\theta_{i+1} + \theta_i) - \theta_j\right)\right) - \Phi\left(\zeta\left(\frac{1}{2}(\theta_i + \theta_{i-1}) - \theta_j\right)\right), & i \in \{2, 3, \dots, n-1\} \\ \Phi\left(\zeta\left(\frac{1}{2}(\theta_{n-1} + \theta_n) - \theta_j\right)\right), & i = n \end{cases}
 \end{aligned} \tag{1.15}$$

The cost associated with information structures in this setup is then $C(\pi, Q) = K(\zeta)$ if Q has the form of (1.15) and ∞ otherwise.

1.5.4 Performance-Dependent Cost Functions

As we showed in Subsection 1.4.1, in uniform guess tasks the DM's ex-ante gross payoff is given by $r\text{tr}(QD\Pi)$. Therefore, their gross payoff depends directly on their performance $\text{tr}(QD\Pi)$. It is possible that the DM selects a desired performance level and pays a cost that depends only on this performance level.

Proposition 12. *Let $K : [0, 1] \rightarrow \mathbb{R}$. Then if $C(\pi, Q) = K\left(\sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}\right)$, the optimal Q^* can be chosen such that $q_{i,i}^* = q \forall i \in \{1, \dots, n\}$ for some $q \in [\frac{1}{n}, 1]$, i.e. q is the DM's performance. Moreover, if K is continuously differentiable and strictly convex with $K'(\frac{1}{n}) < r < K'(1)$, then q solves $r = K'(q)$, and $q_{i,j} = 0 \forall j > n$, and moreover, $q = P^*(r)$ is continuous.*

If Q^* is chosen as in Proposition 12, then the optimal decision matrix D^* is such that

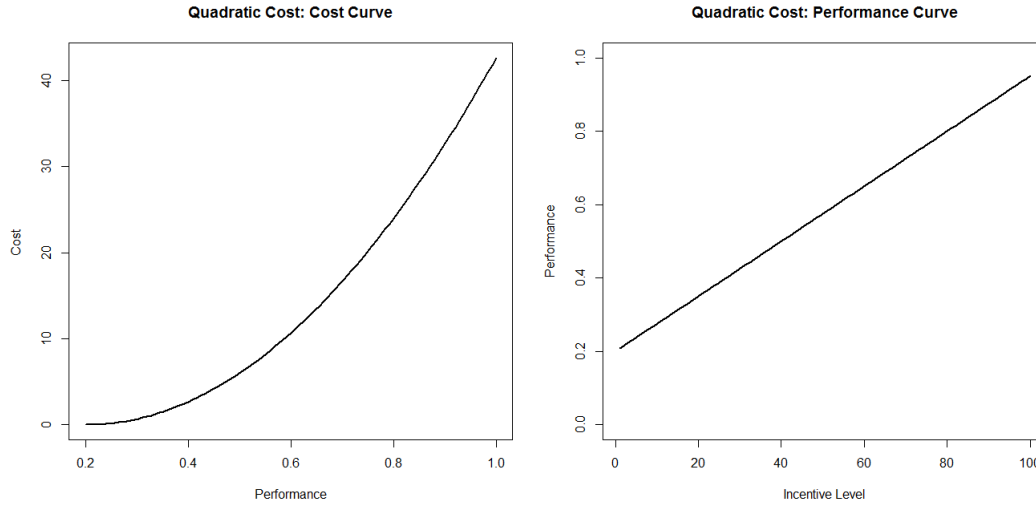


Figure 1.3: Quadratic costs.

The left panel shows the cost function, and the right panel shows the resulting performance curve.

its first $|A|$ rows form the identity matrix. This ensures that q is the subject's performance. Then, the cost function can be written as a function of q , i.e. $K(q)$. It is for this reason that we consider the class of cost functions described in Proposition 12 to be *performance-dependent*.

A subclass of cost functions that is easily recoverable from data includes the cost functions from the second part of Proposition 12, i.e. continuously differentiable and strictly convex K . In that case, as the proposition states, $r = K'(q)$. Therefore, the performance function is $P^*(r) = (K')^{-1}(r)$, and the cost function can be recovered by inverting and integrating the performance function. For example, if the performance function is affine, then its inverse is also affine, and the inverse's integral is quadratic; i.e. the cost function is quadratic in performance.

Table 1.1: Properties of cost functions

Cost Function	Continuity	Convexity	Perceptual Distance	Performance Function
Mutual information	Yes	(Weakly) convex	No	Logistic
Fixed costs	No	No	Can accommodate	Binary
Normal signals	In precision	In precision	On inner states	Concave
Strictly convex in performance	Yes	Strictly in performance	Can accommodate	Inverse of derivative

Note: Perceptual-distance and performance-function properties of normal-signal costs are for state spaces with equidistant spacing.

1.5.5 Summary

Table 1.1 summarizes the properties of the cost functions discussed in this section. It should be noted that each of these cost functions implies a different performance function in uniform guess tasks. Therefore, which of these cost functions best reflects the DM’s behavior can be determined by seeing which of the corresponding performance functions is closest to the DM’s observed performance function.

1.6 Experimental Design

The experiment we implemented involved a series of perceptual tasks, each for a potential reward. In each of these tasks, subjects were shown a screen with a random arrangement of dots and were asked to determine the number of dots on the screen. The number of dots was between 38 and 42, inclusive, and each number was equally likely. Subjects were informed of these facts; there was no deception or withholding of information about the structure of the tasks. Subjects also completed tasks involving the identification of angles. We refer to the first type of task as the “dots” task and to the second as the “angle” task. Subjects generally did not exhibit responsiveness in the “angle” tasks, and so we relegate their description and results to Appendix B.

Each task had a potential reward in a currency called “points.” At the start of each

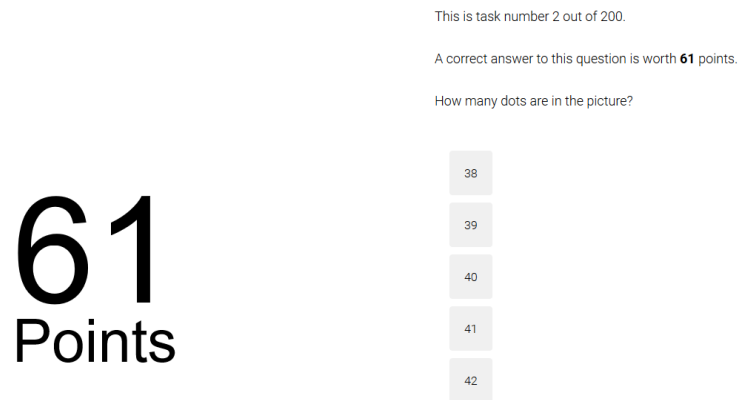


Figure 1.4: Incentive display for a task

task, subjects were shown this reward, which we refer to as the *incentive level*, in large characters for three seconds (e.g. Figure 1.4), before it was replaced with the random dot arrangement (e.g. Figure 1.5). (The incentive level continued to be displayed to the right of the arrangement.) Subjects then had as much time as they desired to determine the number of dots on the screen before proceeding to the next task. If they answered correctly, then they earned the potential reward; if not, then they earned no points for that task. Feedback was not given until the end of the experiment. After completing the tasks (but before receiving feedback) subjects completed a brief demographic questionnaire asking about age, gender, and education. The questionnaire also asked subjects about the strategies they used for determining the number of dots on each screen, as well as whether they varied their strategies depending on the level of reward.

Subjects completed 200 tasks, each at an integer incentive level between 1 and 100, inclusive. Blocks of tasks were balanced by incentive level to ensure roughly the same level of variation in incentive throughout the experiment. Subjects were first shown each



Figure 1.5: Arrangement of dots for a task

of the 50 odd incentive levels between 1 and 100 in a random order, and were then shown each of the 50 even incentive levels between 1 and 100 in a random order. This was repeated (in a different random order) for the next 100 tasks.

Experimental earnings were determined as follows. One task from the first half the experiment and one task from the second half of the experiment were randomly selected for payment. The incentive level of each selected task determined the probability of winning one of two monetary prizes. For example, if the first selected task had an incentive level of 84 and was answered correctly, and the second selected task had an incentive level of 33 and was answered incorrectly, then this would give the subject an 84% probability of winning the first prize and a 0% probability of winning the second prize. Determining earnings in this manner ensured that expected earnings were linear in the incentive level, which obviated the need to elicit risk preferences. In other words, this ensured that under the assumption of expected utility theory, the subjects' utilities (ignoring information costs) were known to us (up to a multiplicative constant). Thus, the estimated relation-

ship between performance and incentive level for each subject could be considered a valid estimate of their performance function, without the need to apply any additional transformation.

As mentioned above, subjects completed 200 tasks in total: 100 “dots” tasks and 100 “angle” tasks. They either completed all the “dots” tasks or all the “angle” tasks first, and this order was randomly determined. For 41 subjects, the prizes were \$10 US, and for 40 subjects, the prizes were \$20 US. In addition, subjects were paid a \$10 participation fee.

All sessions were conducted at the Columbia Experimental Laboratory in the Social Sciences (CELSS) at Columbia University, using the Qualtrics platform. We ran 8 sessions with a total of 81 subjects, who were recruited via the Online Recruitment System for Economics Experiments (ORSEE) (Greiner, 2015).

1.7 Basic Results and Categorization

In this section, we present the main results of our laboratory experiments. First, we consider choice data in the aggregate. Then we perform an individual-level analysis to classify subjects according to whether they are rationally inattentive, are responsive to incentives, have violations of convexity and/or continuity in their cost functions, and evince the perception of distance.

Table 1.2: Laboratory Demographics

Number of subjects	$n = 81$
Gender ($n = 80$)	41.3% male; 58.8% female
Age ($n = 80$)	Average: 23.00; St. dev.: 4.17
Highest level of education achieved ($n = 81$)	
Some post-secondary	44.4%
Completed bachelor’s degree	29.6%
Completed graduate or professional degree	25.9%
Area of study ($n = 80$)	
Economics, psychology, or neuroscience	24.7%

1.7.1 Demographic Data

Table 1.2 lists basic demographic data for the laboratory subjects. The pool is fairly gender-balanced;²³ the null of perfect gender balance cannot be rejected (two-sided test of proportions, $p = 0.146$). The pool is also highly educated; over 55% of laboratory subjects have completed a post-secondary degree.

1.7.2 Choice Data

The data we are most interested in for each task t are the incentive level r_t , the true state of nature θ_t , and the subject’s response a_t . For each task, define the subject’s *correctness* as $y_t := \mathbf{1}_{\{\theta_t\}}(a_t)$. That is, y_t takes the value 1 if the subject correctly determined the state of nature in task t and 0 otherwise.

We are primarily interested in the relationship between correctness and incentive level. We can think of the pattern of successes and failures that we observe as being generated by some underlying data-generating process that for every possible reward level tells us the probability of answering correctly. We denote this probability by

²³Subjects were given the option to list their gender as “other/non-binary.” No subjects used this option, though one subject declined to disclose their gender.

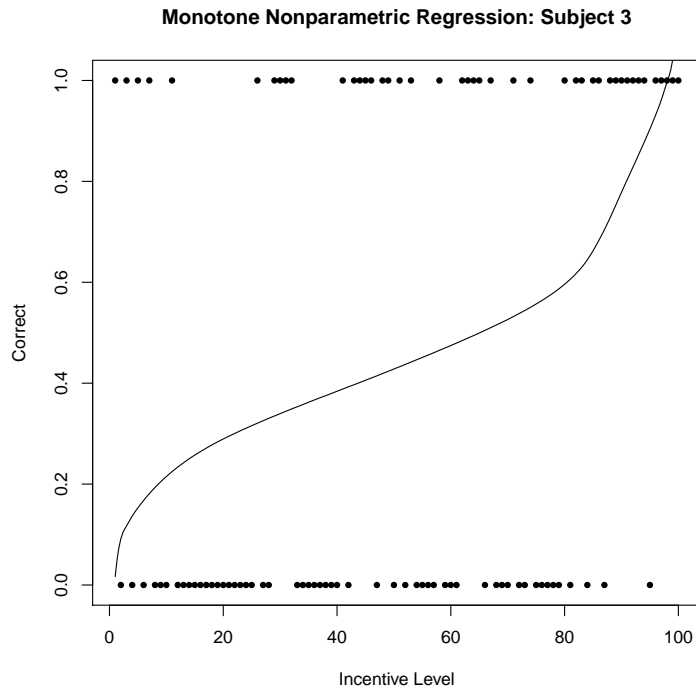


Figure 1.6: Isotone nonparametric regression of correctness on incentive level

Data for Laboratory Subject 3. Method of Dette et al. (2006).

$P_t := \Pr(y_t = 1|r_t) = \Pr(a_t = \theta_t|r_t)$ for each task t ; in other words, the underlying data-generating process is the performance function. Using the correctness data allows us to infer properties of the performance function, which in turn allows us to infer properties of the cost function that generated it. (Figure 1.6 provides an example of what such a performance function might look like.)

In particular, we can answer some of the questions raised in the introduction:

- Do subjects behave in a manner consistent with the predictions of rational inattention?
- Do they respond to incentives?
- Do they exhibit evidence of non-convexities or discontinuities in their cost func-

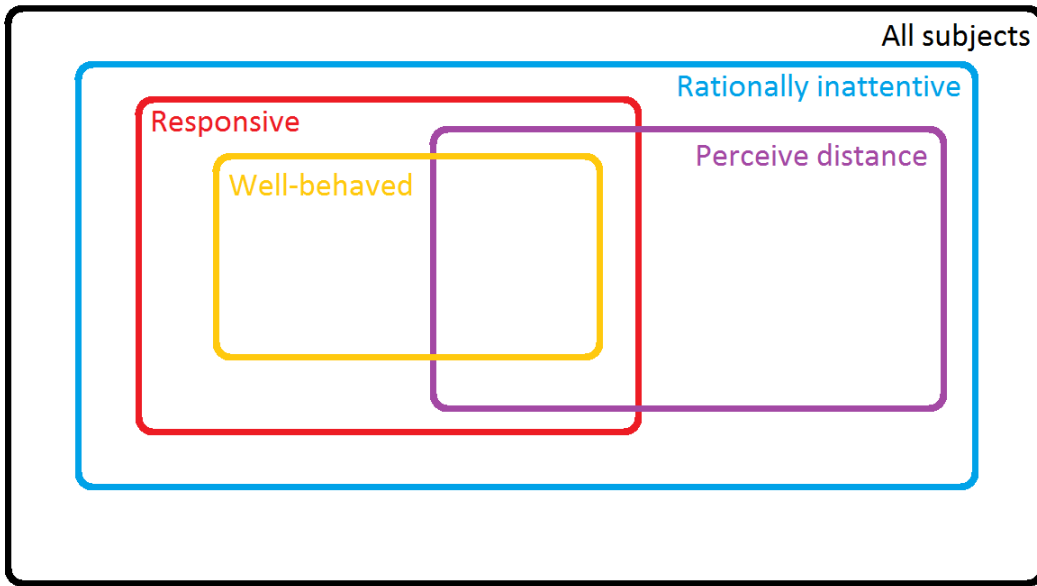


Figure 1.7: Categorization of subjects

tions? (Put differently, are their cost functions not “well-behaved”?)

- Do they evince the perception of distance?

We are able to categorize subjects according to their answers to these questions. First, we classify them by whether or not they are rationally inattentive. Then, we classify rationally inattentive subjects by whether or not they are responsive to incentives. This subset of subjects is the subset of greatest interest to us; these are the subjects for whom we can estimate performance functions and back out corresponding information cost functions. We classify responsive subjects according to whether or not their behavior is consistent with “well-behaved” (i.e. continuous, convex) cost functions. Finally, we classify all rationally inattentive subjects according to whether or not they evince the perception of distance. This categorization scheme is illustrated in Figure 1.7.

However, before proceeding with this categorization exercise, we analyze the subjects’

data in the aggregate.

1.7.3 Aggregate Analysis

Table 1.3: Regressions of correctness on incentive level and demographic covariates

	(1)	(2)
Incentive Level	0.003*** (0.0004)	0.003*** (0.0004)
Age		-0.0001 (0.006)
Male		0.004 (0.056)
Bachelor's		-0.062 (0.058)
Econ/Psych/Neuro		-0.097* (0.054)
\$20 Prize		0.023 (0.049)
Dots First		0.049 (0.052)
Task Number		-0.001*** (0.0003)
Constant	0.425*** (0.032)	0.498*** (0.140)
Observations	7900	7900
R ²	0.03799	0.05635

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered on subject.

Table 1.3 displays a regression of correctness on incentive level. The regression in

column 2 includes demographic covariates, including age (in years) and dummies for maleness, holding at least a bachelor's degree, studying economics, psychology, or neuroscience, participating in the \$20 prize treatment, and being shown the "dots" tasks before the "angle" tasks. It also controls for the order in which tasks were completed.

It is apparent that in the aggregate, performance is higher at higher incentive levels. In particular, on average each increase of 1 point in incentive level results in a 0.3% increase in the probability of answering correctly.

For the most part, demographic covariates have no significant effect on performance. Moreover, there is no significant effect of doing the "dots" tasks before the "angle" tasks. However, performance does decline slightly over time, indicating that subjects may experience some fatigue.²⁴

1.7.4 Rational Inattentiveness

We now proceed with the individual-level categorization exercise.

Before testing the properties of the subjects' cost functions, it is necessary to determine whether there exists a cost function that rationalizes their data in the first place. To that end, we test the necessary and sufficient "no improving attention cycles" and "no improving action switches" conditions by testing the equivalent conditions established in

²⁴The effect of task number on performance vanishes if we only consider the second half of the data, i.e. the last 50 tasks for each subject. (Recall that the first fifty tasks contained the odd-numbered incentives, and the last fifty tasks contained the even-numbered incentives, so each half of the data contains the same range variation in incentives as the whole data set.) This is consistent with some portion of the subjects choosing to exert effort early in the experiment before succumbing to fatigue. As further evidence of this explanation, we find fewer subjects who are responsive to incentives when considering only the last fifty tasks as compared to when considering all tasks (33 (40.7%) as compared to 45 (55.6%)). In that case, the subsequent individual-level analysis can be thought of as estimating behavior in the first half of the data, with random noise coming from the second half of the data for subjects who stop exerting effort.

Subsection 1.3.3.

1.7.4.1 No Improving Attention Cycles

As demonstrated in Proposition 5, a subject satisfies NIAC in our experiment if and only if their probability of correctly guessing the state is non-decreasing in the reward. This implies that rationally inattentive subjects have non-decreasing performance functions.

At this point, a clarification is in order. As we showed in Proposition 5, NIAC holds in a set of uniform guess tasks iff for any pair of decision problems (r_1, r_2) with $r_1 > r_2$, we have that $P^*(r_1) \geq P^*(r_2)$. Observationally, this means that the subject had more correct answers under incentive level r_1 than incentive level r_2 . However, in our experiment each subject is given each decision problem only once. Therefore, the empirically-observed probabilities of answering each decision problem correctly are either 0 or 1. If we were to apply the NIAC condition directly to our data, this would mean that the only subjects whose behavior is consistent with NIAC would be those who always answer incorrectly up to some incentive threshold after which they always answer correctly. Given the stochasticity of choice under limited attention, this scenario is implausible.

Therefore, rather than strictly interpreting our data as stochastic choice data and making direct pairwise comparisons of decision problems to test NIAC, we adopt an estimation-based approach. We estimate the performance function given correctness data and see if this estimate is significantly different from a non-decreasing function, in which case we reject NIAC. In theory, unless there is some reward threshold below which the subject is never correct and above which the subject is always correct, the fit of a monotone performance function can be improved by adding peaks and troughs. The question,

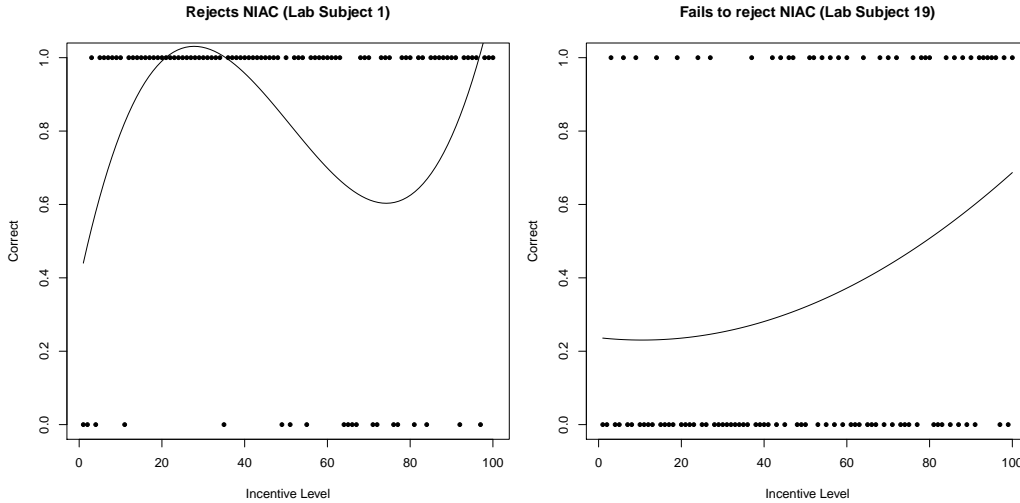


Figure 1.8: Unrestricted cubic polynomial regression of correctness on incentive level for Subjects 1 and 19. The former rejects NIAC (and therefore rejects rational inattentiveness), and the latter fails to reject NIAC.

then, that we wish to pose is not whether a non-monotone or decreasing function can fit the data, but whether we can reject the hypothesis that a non-decreasing function explains the data.

To test for weak positive monotonicity, we employ a method developed by Doyeh et al. (2002) and compare the estimation of an unrestricted cubic polynomial regression of correctness on incentive level for each subject to one with a positive derivative restriction. The null hypothesis for this test is that the response function is monotonic. At the 5% level, we fail to reject positive monotonicity for 77 out of 81 lab subjects (95.1%).²⁵ Examples of polynomial regressions and correctness data for two subjects, one who rejects NIAC and one who fails to reject NIAC, are depicted in Figure 1.8.

²⁵The optimization in the computation of the restricted regression for lab subject 35 failed to converge, and so we did not perform the test for them. For that subject, a one-tailed t-test of the coefficient on incentive level in a linear regression of correctness on incentive level failed to reject the null of the coefficient being non-negative at the 5% level, and so we classify them as having a non-decreasing performance function.

1.7.4.2 No Improving Action Switches

To test for the second necessary and sufficient condition for rational inattentiveness, NIAS, we cannot simply simply examine the estimated performance function; we must look at the posterior probabilities of each state given each response. We employ a bootstrap procedure. For each subject and response, we calculate the empirically observed distribution of true states, i.e. we calculate $\Pr(\theta|a)$. We then simulate 499 bootstrap samples for each distribution.²⁶ If the most common true state is the one corresponding to the response in at least 5% of samples for each response for a given subject, then that subject fails to reject NIAS. Overall, we find that 74 out of 81 (91.3%) laboratory subjects fail to reject NIAS.

Overall, 70 out of 81 (86.4%) laboratory subjects fail to reject both NIAC and NIAS. We refer to these subjects as “rationally inattentive,” or simply “rational,” subjects.

1.7.5 Responsiveness

Of the subjects who fail to reject rational inattentiveness, some of them may have flat response functions, i.e. while they could be rationally inattentive, they do not actually respond to incentives (within the range of incentives presented to them).

To determine which subjects are responsive to incentives, for each subject who failed to reject rational inattentiveness, we run a linear weighted least squares regression of correctness on incentive level and run a one-sided t-test of the coefficient on incentive level with the null of non-positivity, i.e. non-responsiveness to incentives. However, this

²⁶Simulating 1 fewer than 500 bootstrap samples ensures that Type I error probabilities are exact (cf. Section 4.6 of Davidson and MacKinnon, 2004).

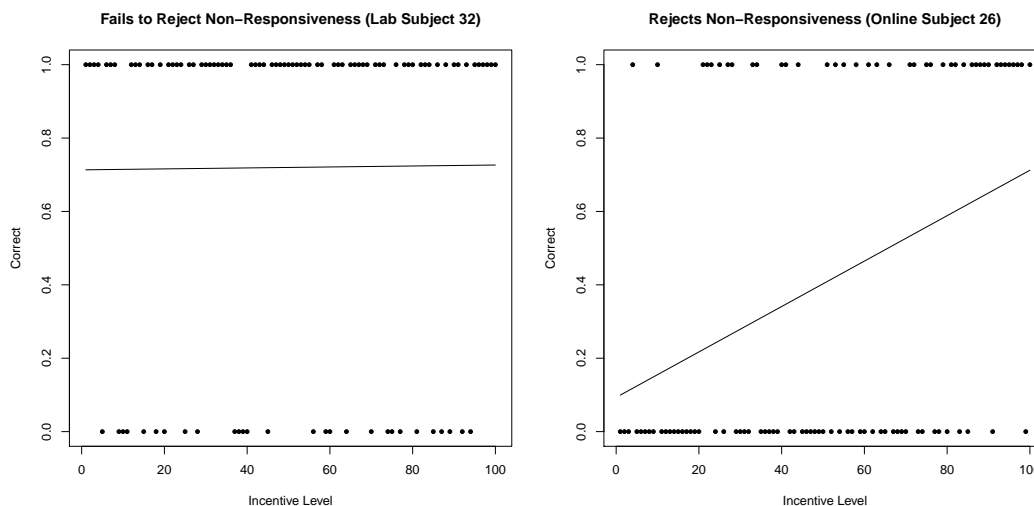


Figure 1.9: Linear regressions of correctness on incentive level for two subjects. The left panel shows an unresponsive subject, and the right panel shows a responsive one.

is insufficient to capture all responsive subjects; a subject may be responsive only within a small range of incentives. To address this issue, for each subject, we repeat this procedure on incentive levels 1 through 50 and on incentive levels 51 through 100.²⁷ If a subject has a significantly positive coefficient on incentive level in any of these three regressions, then we classify them as responsive.²⁸

At the 5% significance level, 42 out of 70 lab subjects (60.0%) who fail to reject rational inattentiveness are responsive to incentives. Examples of full-sample linear regressions and correctness data for two subjects, one who fails to reject non-responsiveness and one who rejects non-responsiveness, are depicted in Figure 1.9.

²⁷Further sample splitting leads to the spurious detection of responsiveness; it leads to some subjects with >95% success being classified as responsive.

²⁸We must consider the full-sample regressions in tandem with the split-sample regressions. If we considered only the split-sample regressions, then we would classify subjects who have binary-response performance functions with thresholds around 50 as non-responsive.

1.7.6 Continuity and Convexity

If the assumptions of Propositions 7, 9, 11, or 12 are satisfied (the cost function is almost strictly convex, the cost function is a strictly convex function of accuracy, the cost function implies convexity in the square root of the precision of a normal signal, or the cost function is mutual information, respectively), then the performance function should be continuous in r . Therefore, observing a discontinuity in the performance function is an indication that the cost function is not almost strictly convex in the information structure, strictly convex in the probability of correctly determining the state, strictly convex in the square root of the precision of a normally-distributed signal, or mutual information, i.e. there is some violation of convexity.

Strictly speaking, one cannot definitively observe a discontinuity without an infinite data set; a continuous function with a sufficiently steep slope at points of potential discontinuity can always be used to fit finite data. Therefore, for each subject, the question we wish to answer is whether it is more plausible that a discontinuous performance function or a continuous performance function generated their correctness data. This implies a statistical test where the null hypothesis is that the performance function belongs to some class of discontinuous functions, and the alternative is that the performance function belongs to some class of continuous functions.

We test for the presence of a discontinuity by applying a likelihood ratio test. We

estimate a regression of the form:²⁹

$$P_t = \beta_0 + \beta_1 \mathbf{1}_{\geq \delta}(r_t) \quad (1.16)$$

where β_0 , β_1 , and δ are the parameters to be estimated and compare its likelihood to an estimation of the following logistic relationship:

$$P_t = \frac{\beta_1}{1 + \exp(-\lambda(r_t - \delta))} + \beta_0 \quad (1.17)$$

It can be shown that (1.16) is the pointwise limit of (1.17) as λ goes to infinity. Therefore, (1.16) can be seen as the restricted null model, and a likelihood ratio test comparing these models is effectively a test of the null hypothesis that $\lambda = \infty$, i.e. it is a test against the null hypothesis that there is a jump discontinuity. Since we are performing this test only on responsive subjects, our estimates of β_1 for each subject should be positive, and therefore this procedure should not detect spurious downward jump discontinuities for those subjects.³⁰

Using this test, at the 5% level we cannot reject that 29 out of 42 responsive lab subjects (69.0%) have discontinuities in their response functions.

²⁹We use the procedure of Bai and Perron (1998) for this estimation.

³⁰Several procedures for detecting discontinuities have been proposed in the econometric literature. See, for example, Andrews (1993), Andrews and Ploberger (1994), Bai and Perron (1998), and Porter and Yu (2015). All of these procedures are designed to detect both positive and negative jump discontinuities, and so they are vulnerable to the detection of spurious negative jumps in our setting. A clarification is in order here. Bai and Perron (1998) propose both an estimation procedure and a testing procedure for models with structural breaks with unknown discontinuity points. We use their estimation procedure to estimate (1.16), but we do not use their testing procedure.

1.7.7 Perceptual Distance

A mutual-information cost function, among others, does not embed a concept of perpetual distance. What that means in our experiment is that mutual information predicts that given a true state of nature, each incorrect response is equally likely. For example, if the true number of dots is 39, reporting 42 should be just as likely as reporting 38, even though the difference between 42 and 39 is 3, whereas the difference between 38 and 39 is 1. Taking absolute value as the natural metric on $\{38, \dots, 42\}$, this means that if a subject evinces perceptual distance, then given a true state of 39, she should be more likely to report 38 than 42.

For each subject and trial t , we compute the error distance $\rho(a_t, \theta_t) = |a_t - \theta_t|$. In our experiment this distance is an integer in $\{1, 2, 3, 4\}$. In order to test this symmetry prediction of the mutual information cost function, for each responsive subject we compute the distribution of error distances that mutual information would predict, given the empirically observed distribution of true states and the subject's overall accuracy rate. We then compare the empirically observed distribution of error distances to this distribution using a chi-square test.

At the 5% level, we find that 26 out of 42 responsive lab subjects (61.9%) have a distribution of mistakes that evinces the perception of distance. Of course, the notion that perceptual distance matters for error distance distributions is not limited to responsive subjects; mutual information implies responsiveness (i.e. a strictly increasing performance function), so subjects who are not responsive have already rejected mutual information for other reasons. But as a test of the general notion that each possible mistake is equally

Table 1.4: Categorization of subjects

Category	Of All Subjects	Of R.I. Subjects	Of Resp. Subjects	Of P.D. Subjects
All subjects	81 (100%)	—	—	—
R.I. subjects	70 (86.4%)	70 (100%)	—	—
Resp. subjects	///	42 (60.0%)	42 (100%)	26 (57.8%)
W.B. subjects	///	///	13 (31.0%)	8 (17.8%)
P.D. subjects	///	45 (64.3%)	26 (61.9%)	45 (100%)

Note: “R.I.” = rationally inattentive; “Resp.” = responsive; “W.B.” = well-behaved, i.e. subjects whose behavior is consistent with continuous, convex cost functions; “P.D. Subjects” = subjects who evince the perception of distance. — denotes that the column category is a subset of the row category, and /// denotes that the row category is defined only on a subset of the column category.

likely given an true state of nature, it is worth running these tests on the entire pool of rationally inattentive subjects. At the 5% level, we find that 45 out of 70 rationally inattentive lab subjects (64.3%) reject this hypothesis.

1.7.8 Summary of Categorization

Table 1.4 summarizes the results of preceding subsections. Each cell indicates the number and percentage of row category subjects in the column category. It should be noted that the vast majority (86.4%) of subjects are rationally inattentive, and moreover, most rationally inattentive subjects are responsive (60.0%). Also of note is the fact that most rationally inattentive subjects evince the perception of distance (64.3%).

1.8 Model Selection

In this section, for each responsive subject we fit several possible parametric functional forms for performance functions, each of which can be generated by some cost function.

Table 1.5: Performance functions estimated and their corresponding cost functions

	Cost Function	Performance Function	Estimation
1	Very high or low marginal or absolute costs	Constant	OLS
2	Dual-process or concave	Binary	BP98
3	Quadratic in accuracy	Affine	WLS
4	Discontinuous or has non-convexity	Affine with break	BP98
5	Integral of inverse of 2nd degree polynomial in accuracy	2nd degree polynomial	WLS
6	Integral of inverse of 3rd degree polynomial in accuracy	3rd degree polynomial	WLS
7	Mutual information or logit cost	Logistic	MLE
8	Normal signals with linear cost of precision	Concave	MLE

Note: BP98 = Bai and Perron (1998)

These models are listed in Table 1.5.³¹

1.8.1 Description of Models

We now describe each of these models.

Model 1. Since we perform this model selection exercise only on responsive subjects, we fit a constant performance function as a robustness check. A flat performance function obtains when there is a single performance level such that a subject’s cost for higher performance is too high to warrant improvement for any incentive level and their savings in terms of information costs are too low to warrant lowering their performance at any incentive level.

Models 2 and 4. These models have discontinuous performance functions, implying by Proposition 7, 9, 11, or 12 that the corresponding cost functions have some violation of convexity or continuity. We estimate Model 2 by:

$$P_t = \beta_0 + \beta_1 \mathbf{1}_{\geq \delta}(r_t) \quad (1.18)$$

³¹The reason that we do not consider the channel capacity cost function is because since the prior distribution in our task is uniform, channel capacity would be consistent with the same behavior as mutual information (cf. Section 1.2.3 of Woodford, 2012)

and Model 4 by:

$$P_t = \beta_0 + \beta_1 \mathbf{1}_{\geq \delta}(r_t) + \beta_2 r_t + \beta_3 r_t \cdot \mathbf{1}_{\geq \delta}(r_t) \quad (1.19)$$

where δ is the location of the discontinuity, using the linear structural change estimation technique of Bai and Perron (1998).

Given the estimates of (1.18), we can recover the parameters of the corresponding fixed-cost information cost function (1.9). It is clear that $\underline{q} = \hat{\beta}_0$ and $\bar{q} = \hat{\beta}_0 + \hat{\beta}_1$. As shown in Subsection 1.5.4, the location δ of the discontinuity in the performance function is given by $\frac{\kappa}{\bar{q} - \underline{q}}$. Therefore, $\kappa = \hat{\delta}(\hat{\beta}_1 - \hat{\beta}_0)$.

A performance function of the form of (1.18) could also be obtained from a concave information cost function. To illustrate, consider for simplicity a subject who has a cost function K that is concave in performance, as depicted in Figure 1.10.

Net payoffs are maximized when the positive distance between gross payoffs and costs is largest. For low reward levels (such as r_1), this happens at the no-information performance level, 0.2. For high reward levels (such as r_2), this happens at the full-information performance level, 1. In this manner, a binary performance function obtains, with the subject acquiring no information if the incentive is low and acquiring full information if the incentive level is high.

Model 4 nests Model 2, and so it can correspond to a wider class of cost functions, including those of similar form to (1.9), but with two convex components as opposed to only two admissible information structures.

Models 3, 5, and 6. These models estimate continuous performance functions, and

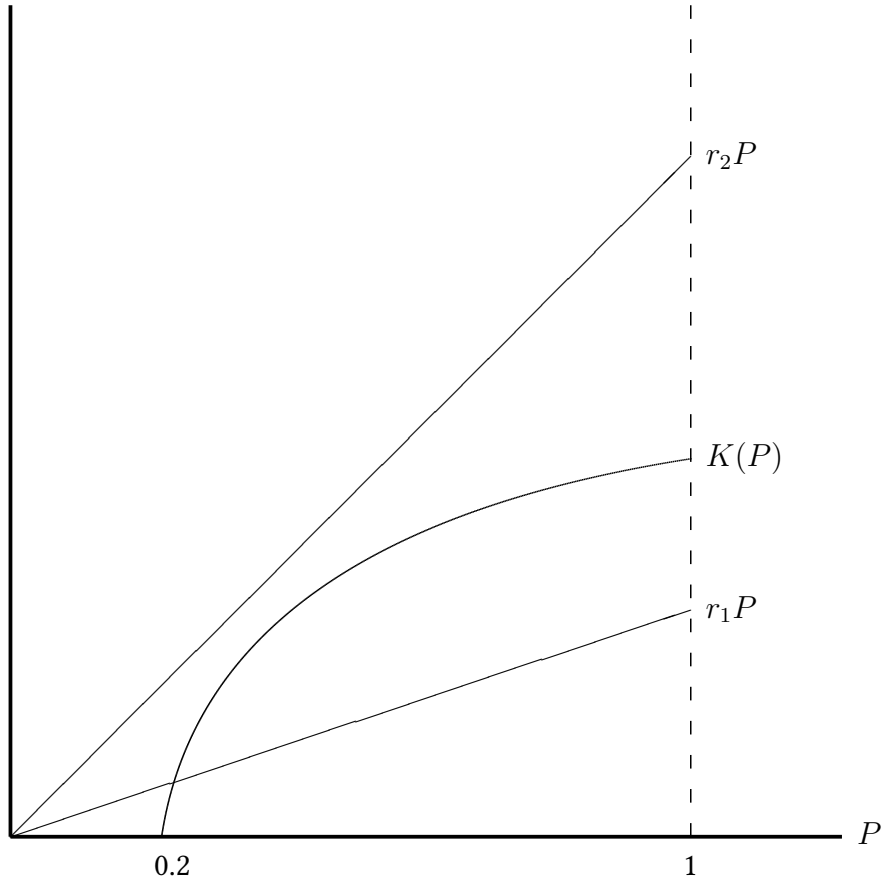


Figure 1.10: Concave costs

since these models are estimated only for responsive subjects, these estimated performance functions should be increasing. As explained in Subsection 1.5.4, this means we can recover the corresponding information cost functions by inverting and taking the antiderivative of the estimated performance functions. While the quadratic and cubic performance functions of Models 5 and 6 do not have very simple closed forms for their corresponding cost functions, Model 3's corresponding cost function can be easily recovered from its estimate:

$$P_t = \beta_0 + \beta_1 r_t \tag{1.20}$$

In this case, inverting and integrating the performance function gives us the performance-dependent cost function $K(P)$, where:

$$K(P) = \frac{1}{2\hat{\beta}_1}(P^2 - 0.04) - \frac{\hat{\beta}_0}{\hat{\beta}_1}(P - 0.2) \quad (1.21)$$

assuming that $K\left(\frac{1}{n}\right) = 0$, i.e. that acquiring no information entails no cost.

Model 7. As we showed in Proposition 9, a mutual information cost function implies a logistic performance function, which we can estimate with:

$$P_t = \frac{1}{4 \exp\left(-\frac{r_t}{\alpha}\right) + 1} \quad (1.22)$$

However, Proposition 9 also implies that if a subject evinces the perception distance, then they cannot have a mutual information cost function. To reconcile logistic performance with perceptual distance, one can assume that the subject actually has a performance-dependent cost function and apply the inversion-and-integration procedure of Subsection 1.5.4. This gives the following “logit” cost function (assuming that acquiring no information entails no cost):

$$K(P) = \hat{\alpha} \left(P \ln \left(\frac{4P}{1-P} \right) + \ln(1-P) + \ln(1.25) \right) \quad (1.23)$$

Model 8. For this model, we assume that the subject receives normal signals with precision ζ^2 and the cost of a signal is linear in precision, as in the numerical example of Section 5 of Verrecchia (1982). Therefore, the cost of a signal is $K(\zeta) = \alpha\zeta^2$. α is the parameter we wish to estimate for each subject.

Table 1.6: Model Selection for Responsive Subjects

Model	Binary (2)	Logistic (7)	Concave (8)
Number of Subjects	10 (23.8%)	26 (61.9%)	6 (14.3%)

The performance function for this cost function in our experiment is (from (1.14)):

$$P_t = \frac{8}{5} \Phi \left(\frac{\zeta^*(r_t)}{2} \right) - \frac{3}{5} \quad (1.24)$$

where ζ^* solves the first-order condition of (1.14) for our experiment:

$$\alpha \zeta^* = \frac{2}{5} r_t \phi \left(\frac{\zeta^*}{2} \right) \quad (1.25)$$

1.8.2 Comparison of Models

We now run a “horse race” to determine which model is the best fit for each subject. Since the models are non-nested and are estimated using different estimation techniques, we cannot use a traditional auxiliary regression method for model selection. To determine which model is the best fit for each responsive subject, we estimate each model for each such subject and then compare their Akaike Information Criteria (AIC) (cf. Section 8.5 of Cameron and Trivedi, 2005), selecting the model that yields the lowest AIC. The results of this selection are given in Table 1.6.

All responsive subjects are best fit by binary (fixed costs), logistic (mutual information or “logit” performance dependence), or concave performance (normal signals with linear precision cost). The first implies some sort of non-convexity or discontinuity in the cost function, whereas the latter two are consistent with convex cost functions. Figures 1.11,

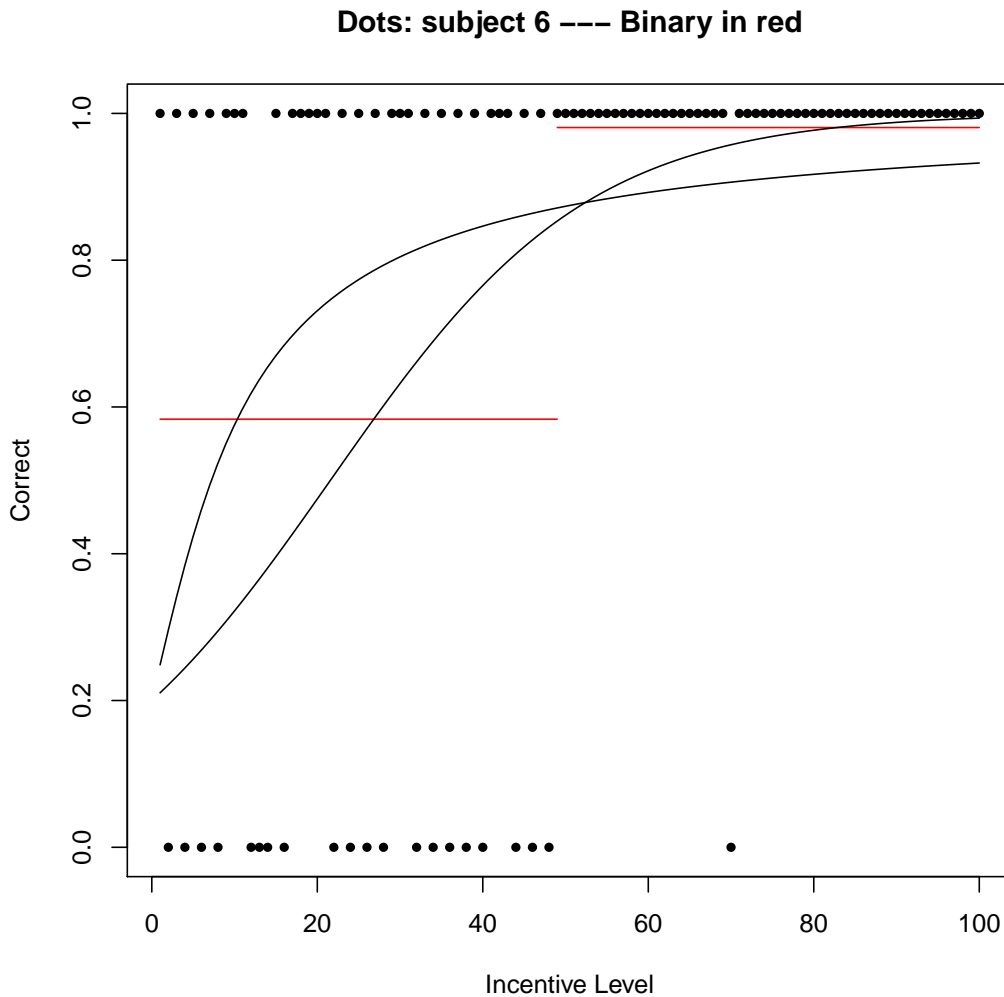


Figure 1.11: Best-fit binary performance

This figure shows the fits of binary, logistic, and concave (normal) performance for Subject 6, with the best-fitting binary model in red.

1.12, and 1.13 show what these performance functions look like for three subjects.

Table 1.7 shows the average estimated AIC and rank of each model in the “horse race.”³² Models 2 (binary), 3 (affine), and 7 (logistic) have the lowest ranks on average. Flexible polynomial fits do quite poorly; the average rank of a cubic performance function (Model 6) is higher than that of the constant performance model (Model 1).

³²A lower rank implies a lower AIC and therefore a better fit.

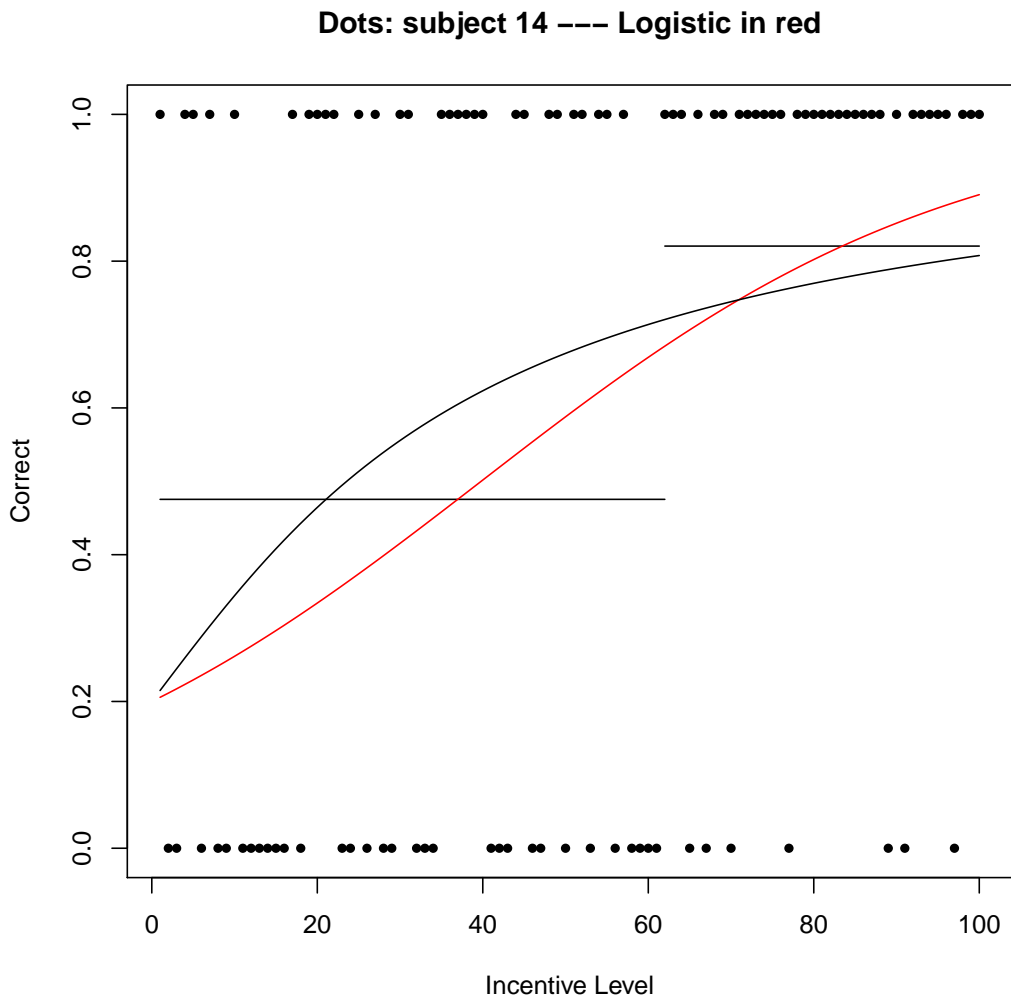


Figure 1.12: Best-fit logistic performance

This figure shows the fits of binary, logistic, and concave (normal) performance for Subject 14, with the best-fitting logistic model in red.

Table 1.7: Average AIC and Rank for Estimated Models

Model	AIC	Rank
1	131.165	6.548
2	114.060	2.143
3	117.832	3.524
4	119.982	4.905
5	131.123	5.595
6	132.929	6.643
7	116.008	2.308
8	121.967	4.333

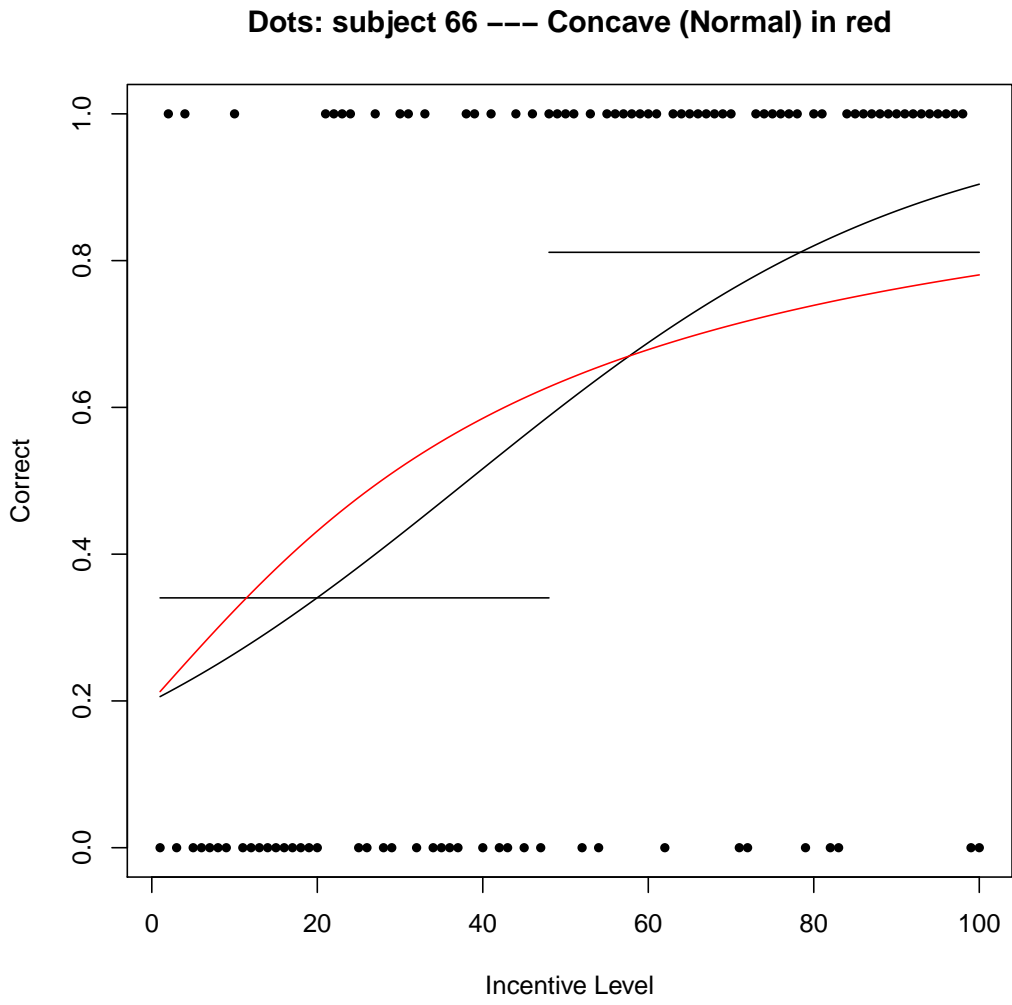


Figure 1.13: Best-fit concave (normal) performance

This figure shows the fits of binary, logistic, and concave (normal) performance for Subject 66, with the best-fitting concave (normal) model in red.

Note that the average AIC and rank for Model 7, the logistic performance function, is lower than that of Model 2, the binary performance function, despite the fact that significantly more subjects are best fit by Model 7 than by Model 2. This indicates that the binary model is a decent fit when the best-fitting model is logistic, but the logistic model is a poor fit when the best-fitting model is binary. When the logistic model is the best fit, the average rank of the binary model is 2.577; however, when the binary model is the best fit, the average rank of the logistic model is 4.100. Note also that the average rank of Model 8 (normal signals with linear precision cost) is fairly high at 4.333. This indicates that when Model 8 is not the best fit for a subject, it is a poor fit.

1.9 Additional Results

1.9.1 Categorization and Demographics

In this subsection, we determine the extent to which demographic covariates predict the categorization of subjects as rationally inattentive and responsive, as well as whether their best-fitting performance function is binary or logistic.

1.9.1.1 Rational Inattentiveness

To determine the extent to which demographics predict a failure to reject NIAC, we run a logit regression of an indicator for rational inattentiveness on demographic covariates. These covariates are age, an indicator for being male, an indicator for having attained at least a bachelor's degree, an indicator for studying economics, psychology, or neuroscience, an indicator for participating in the \$20 prize treatment, and an indicator for

Table 1.8: Demographics and Categorization: Logit Regressions

	Rational Inattentiveness	Responsiveness
	(1)	(2)
Age	-0.0002 (0.084)	-0.015 (0.064)
Male	-0.475 (0.718)	-0.691 (0.588)
Bachelor's	0.963 (0.804)	0.054 (0.623)
Econ/Psych/Neuro	0.749 (0.886)	1.261* (0.704)
\$20 Prize	0.128 (0.714)	0.263 (0.557)
Dots First	-0.602 (0.752)	-0.482 (0.562)
Constant	1.726 (1.939)	0.911 (1.488)
AIC	74.110	99.128

Note: *p<0.1; **p<0.05; ***p<0.01

having done the dots tasks first. We display the results of this regression in column 1 of Table 1.8.

Demographic covariates do not seem to be predictive of rational inattentiveness in this particular subject pool. Neither do experimental variables, such as the higher prize and completing the dots tasks first. This suggests that for a given set of tasks, rational inattentiveness is an innate characteristic that is not well captured by demographics, and moreover, it may be difficult to manipulate experimentally.

1.9.1.2 Responsiveness

To determine the extent to which demographics predict responsiveness, we run a logit regression of an indicator for responsiveness on demographic covariates for the subjects who fail to reject rational inattentiveness. We display the results of this regression in column 2 of Table 1.8.

As is the case with rational inattentiveness, demographic covariates are not significant predictors of responsiveness.

Table 1.9: Model Selection and Demographics

	Binary (2)	Concave (8)
Age	-0.201 (0.211)	0.201 (0.141)
Male	-0.417 (0.916)	-0.922 (1.243)
Bachelor's	-0.338 (1.124)	-0.609 (1.282)
Econ/Psych/Neuro	0.168 (0.897)	-0.690 (1.400)
\$20 Prize	0.505 (0.825)	-0.365 (1.255)
Dots First	-0.075 (0.888)	2.211* (1.342)
Constant	3.482 (4.435)	-6.688* (3.488)
AIC	94.635	94.635

Note: *p<0.1; **p<0.05; ***p<0.01

1.9.1.3 Cost Functions

To determine the extent to which demographics predict model selection, we run a multinomial logit regression of the best-fitting model on the same set of demographic covariates as in previous subsections, with logistic performance (Model 7, mutual-information costs) as the baseline. This regression shows us the extent to which these demographic factors affect the likelihood of selecting a model that implies a non-convexity or discontinuity in the cost function over one that is consistent with convexity. We display the results of this regression in Table 1.9.

As is the case with previous demographic regressions, demographic factors are not significant predictors. This seems to indicate that not only is rational inattentiveness not well captured by demographics, so is the nature of one's cost function for information in a given task.

1.9.2 Reaction Times

In addition to data on subject responses, we also collected data on how much time subjects spent on each task. We call this the *reaction time*.

1.9.2.1 Time and Attention

Thus far, in this paper, we have remained agnostic about the exact nature of what attention comprises, and by corollary, we have remained agnostic about the exact source of information costs. One possibility is that attention can be decomposed into a quantity component — time spent on a task — and a quality component — how much effort is ex-

erted during that time. Here, we provide some suggestive evidence that attention indeed has a quantity component.

Table 1.10: Linear regression of reaction time on incentive level

	Reaction Time
Incentive Level	0.178*** (0.017)
Constant	14.159*** (1.378)
Observations	8100
R ²	0.059

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered on subject.

Table 1.11: Linear regression of correctness on reaction time

	Correctness
Reaction Time	0.007*** (0.001)
Constant	0.427*** (0.036)
Observations	8100
R ²	0.096

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered on subject.

Tables 1.10 and 1.11 display linear regressions of reaction time on incentive level and correctness on incentive level, respectively, aggregating over the subject pool. The coefficients on the dependent variables in both regressions are positive and significant. In the case of the first regression, this indicates that subjects respond to higher incentives

by increasing the quantity of attention paid to the task at hand. In the case of the second regression, this indicates that increasing the quantity of attention results in higher performance; this is the speed-accuracy trade-off commonly noted in the literature on perceptual psychology (e.g. Schouten and Bekker, 1967).

1.9.2.2 Dual-Process Mechanisms

As we showed in Subsection 1.8.2, choice data for approximately one-third of responsive subjects are best fit by binary performance functions. This suggests that these subjects employ two different strategies for determining the number of dots on the screen — one for low incentives, and one for high incentives. In this subsection, we provide further suggestive evidence for this hypothesis.

Figure 1.14 shows the histogram of reaction time on every task for the subject population. The distribution of reaction times is clearly bimodal. There are at least two possible, non-mutually exclusive explanations for this. One is that some portion of the subjects simply do not exert any effort on the task and make a response at the earliest opportunity, while others exert effort in acquiring information. Another is that subjects have binary performance functions, choosing not to spend time acquiring information for some incentive levels but choosing to do so for others.

The fact that a significant portion of subjects are best fit by binary performance functions provides an explanation for the pattern observed in Figure 1.14. Some subjects make snap decisions when confronted with low incentives but take the time to acquire information at higher incentive levels. This can be seen more clearly in Figure 1.15, which shows the histogram of reaction time on every task for responsive subjects only. Observe that

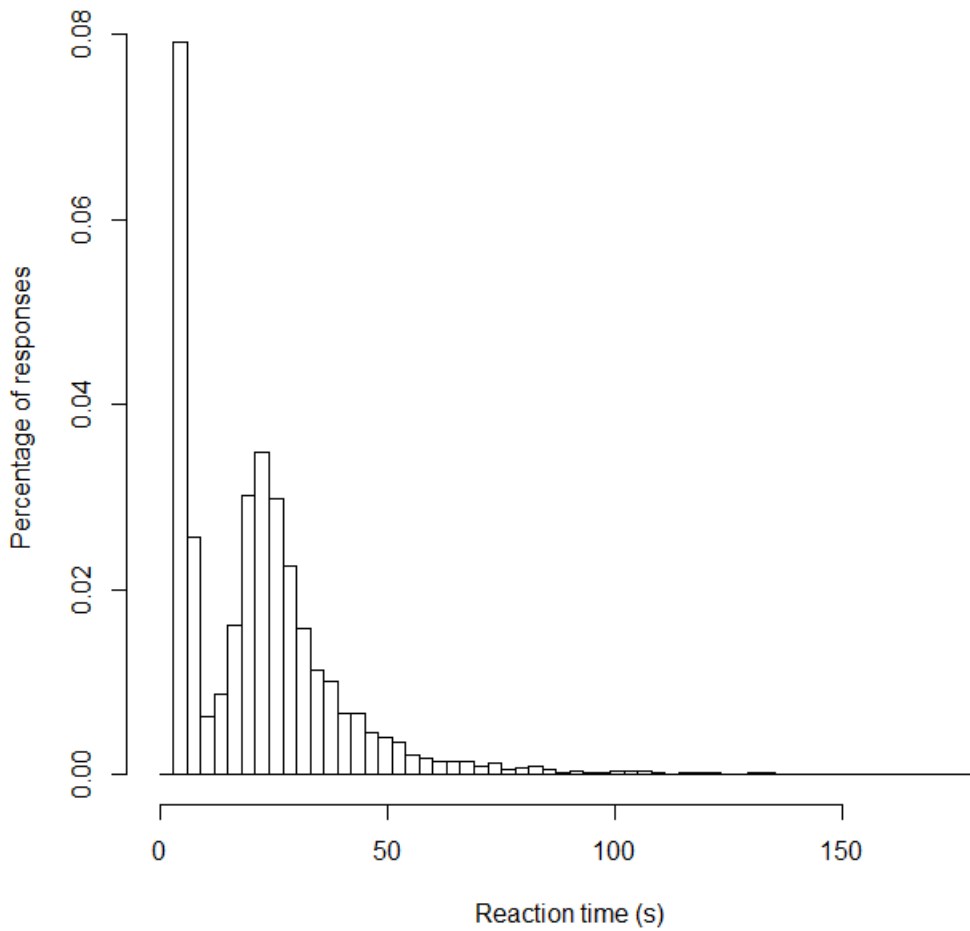


Figure 1.14: Histogram of reaction times for all subjects

this histogram is also clearly bimodal.

To interrogate this question further, we run the dip test of Hartigan and Hartigan (1985) on each subject's reaction times to determine which ones have multimodal reaction time distributions. We can reject the null of unimodality at the 5% level for 26 out of 42 responsive subjects (61.9%). This is more than the number of responsive subjects whose data are best fit by binary performance functions, meaning that some subjects with logistic or concave performance functions do not have unimodal reaction time distributions.

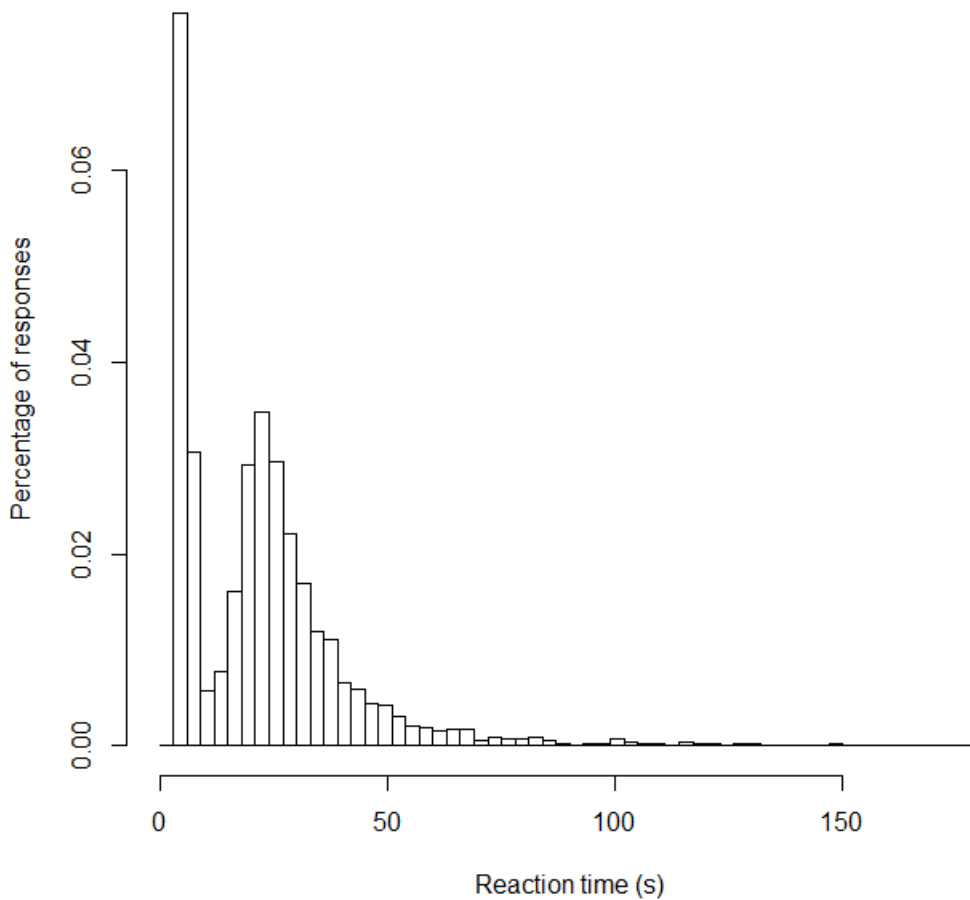


Figure 1.15: Histogram of reaction times for responsive subjects

This suggests that rather than continuously adjusting their quantity of attention as incentive levels increase, some subjects randomize between paying a high quantity and a low quantity of attention, and the probability of paying a high quantity of attention increases as incentive levels increase.

This evidence implies that for a large portion of the subject pool (61.9%), there are two information-acquisition processes that they can employ in this task. Still, there is a significant portion of the pool (38.1%) that is apparently able to adjust their quantity of

attention continuously. As was the case with previous categorizations of subjects, there is significant heterogeneity.

1.10 Application to the Delegation of Investment

The characteristics of the decision-maker's cost function can obviously have effects on her own decisions. But as we show in this section, these characteristics can also have effects on economically-relevant outcomes when there is strategic interaction.

In order to demonstrate this notion, let us consider a situation in which an investor is deciding which of n options to invest in, and he cannot split his investment across options. Suppose that only one of these options can be a winner, in which case an investment in it will pay a net return of x . Losing opportunities pay a net return of zero. This setup has the relevant features of a situation where the success of an investment depends on the outcome of a contest. Many economic situations, such as competing to be granted development rights by the government for a plot of land, take the form of contests. Another salient example is a patent race, where various firms compete to be the first to patent an invention, such as a drug or a piece of technology.

Suppose that the investor wishes to delegate researching these options to an expert. This is a common occurrence in reality; people frequently solicit the services of financial advisors, presumably because it is prohibitively difficult or costly for laypeople to research investment opportunities themselves, while financial advisors who are trained to seek and interpret financial information can research these opportunities at a much lower cost.

We can analyze this situation in a simple principal-agent framework, where the in-

vestor is the principal and the expert is the agent.³³ The agent acquires information about the available investment opportunities at a cost and selects one of the options on the principal's behalf. Suppose that the principal employs the agent with a contract that pays r if the agent correctly selects the winner and zero otherwise.³⁴ Furthermore, suppose that *a priori*, each option is equally likely to be the winner. Then, the agent's problem can be represented as a uniform guess task, with the reward for a correct answer being r . Consequently, the principal's problem is

$$\max_{r \in [0, x]} (x - r)P^*(r) \tag{1.26}$$

where $P^*(r)$ is the agent's performance function.

As we established in Proposition 5, if the agent is rationally inattentive, then her performance function is (weakly) increasing. Thus, the principal faces a trade-off between incentivizing the agent to acquire better information and giving up a larger portion of his net return upon success. The exact nature of this trade-off depends on the potential net return x and the agent's information cost function. In the following subsections, we analyze the properties of the principal's optimal payment strategy r^* under the three cost function models fit by our data: fixed costs; mutual information; and normally-distributed signals.³⁵

³³We use male pronouns for the principal and female pronouns for the agent.

³⁴This type of contract is optimal for the principal if we assume that (a) there is a limited-liability constraint so that the agent cannot earn a negative payoff in any state of the world, which implies that the principal cannot "sell the firm" to the agent; and (b) the agent's cost of an uninformative information structure is zero. As CD15 demonstrate, the latter assumption is without loss of generality; it is not a testable restriction on information cost functions.

³⁵Some caution is required in applying the assumption of normally-distributed signals, because it implies that the options have some existing ranking, and it is not clear what it means for the options to be

1.10.1 Fixed Costs

Suppose the agent has a fixed cost κ for acquiring information. If she pays the cost, then she learns the winner with certainty. If not, then she learns nothing about the identity of the winner. Thus, she chooses to acquire information if $r - \kappa \geq \frac{r}{n}$, i.e. when $r \geq \frac{\kappa n}{n-1}$.

Therefore, if $x < \frac{\kappa n}{n-1}$, then the reward required to incentivize the agent to acquire information is higher than the potential net return, so the principal is better off not hiring the agent at all and simply picking an option at random. If instead $x \geq \frac{\kappa n}{n-1}$, then the principal could incentivize information acquisition by paying as little as $r = \frac{\kappa n}{n-1}$. To ensure that this payment is not so high than the principal could do better on his own, he requires that $\frac{x}{n} \leq x - \frac{\kappa n}{n-1}$, which holds if and only if $x \geq \frac{\kappa n^2}{(n-1)^2}$. But since $\frac{\kappa n}{n-1} < \frac{\kappa n^2}{(n-1)^2}$, the principal will not hire the agent unless $x \geq \frac{\kappa n^2}{(n-1)^2}$.

To summarize: if $x < \frac{\kappa n^2}{(n-1)^2}$, then the principal does not hire the agent and selects an option at random. If $x \geq \frac{\kappa n^2}{(n-1)^2}$, then the principal hires the agent and gives her a payment of $\frac{\kappa n}{n-1}$, and the agent picks the winner with certainty. This implies a discontinuity in the principal's payment as a function of the potential net return x . Figure 1.16 shows what this payment scheme looks like for $\kappa = 40$.

1.10.2 Mutual Information

Suppose the agent has a mutual-information cost function with cost parameter α . Then, since her performance function is logistic (see Proposition 9), the principal chooses r to

“equidistant” from each other. In any case, if the normal-signals model is excluded from consideration, then in our data, the best-fitting model for each subject is either binary (fixed costs) or logistic (mutual information). (Results available from the authors on request.)

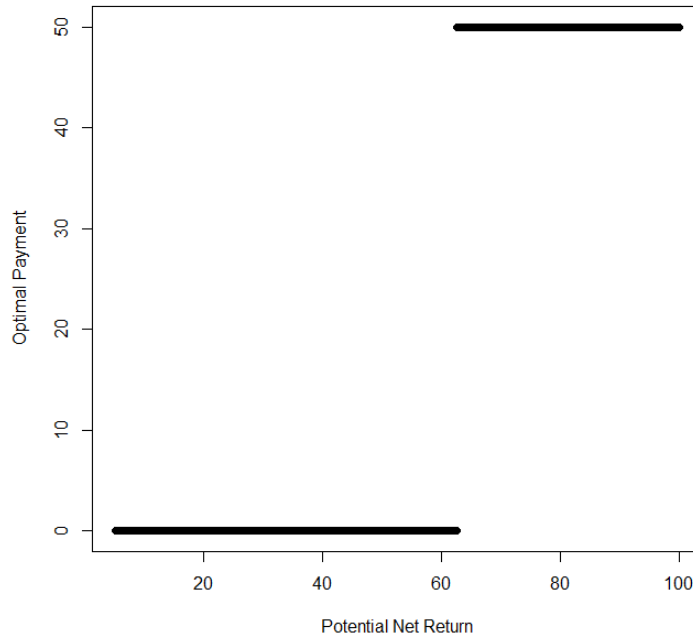


Figure 1.16: Optimal payment as a function of potential net return, fixed costs

maximize:

$$\frac{x - r}{(n - 1) \exp\left(-\frac{r}{\alpha}\right) + 1} \quad (1.27)$$

If this maximand is strictly quasiconcave, then this problem has a unique solution for each x , and the maximum theorem guarantees that the principal's optimal choice of r^* is continuous in x . This turns out to be the case.

Proposition 13. *If the agent has a mutual information cost function, then the principal's optimal payment strategy $r^*(x)$ is continuous.*

To provide an example, suppose $n = 5$, $\alpha = 10$, and $x \in [5, 100]$. As shown by Proposition 13, the maximand (1.27) is strictly quasiconcave in r (see Figure 1.17 for an

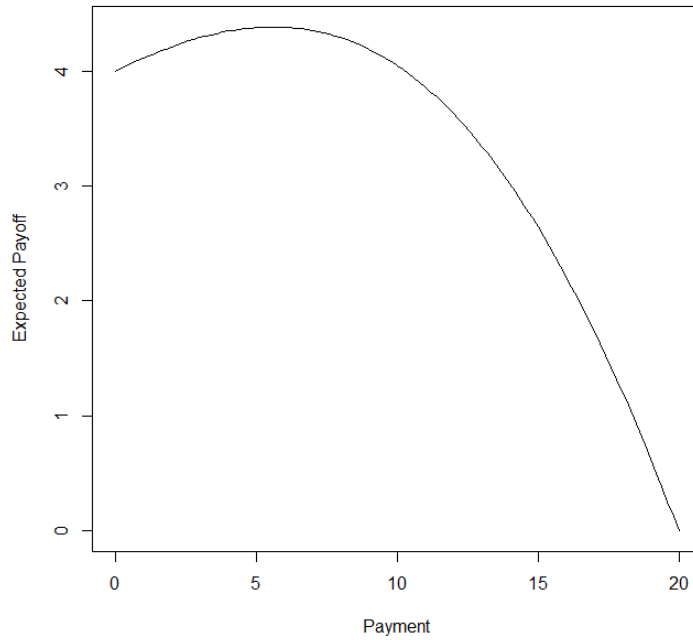


Figure 1.17: Principal's expected payoff as a function of payment for $x = 20$, mutual information costs

example). For these parameters, $r^*(x)$ is continuous and increasing, as shown in Figure 1.18.

1.10.3 Normally-Distributed Signals

Suppose that the options are ranked and equidistant on some scale. For example, in the case of bidding for development rights, the projects could be ranked by the estimated length of time until project completion.³⁶ In this case, if the agent's cost function satisfies the conditions of Proposition 11, then it can be shown that the principal's optimal choice of r^* is continuous in x .

³⁶Shorter completion times mean that the development will be more quickly available for public use, but may also signal poor craftsmanship.

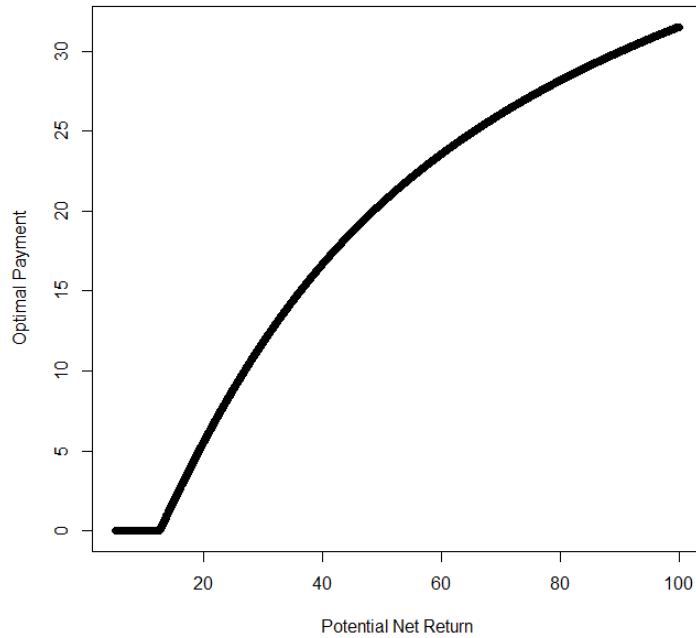


Figure 1.18: Optimal payment as a function of potential net return, mutual information costs

Proposition 14. *If the options are ranked and equidistant, and the agent has a convex, increasing cost of precision of normal signals with non-negative third derivative, then the principal's optimal payment strategy $r^*(x)$ is continuous.*

Figure 1.19 shows what this payment scheme looks like if costs are linear in the precision of normally-distributed signals, with a marginal cost of precision of 7.5.

1.10.4 Hybrid Models

The previous subsections have studied the investment delegation model under the assumption of a single type of cost function. However, the results of Section 1.8 indicate that there is heterogeneity in individuals' cost function types in the population.

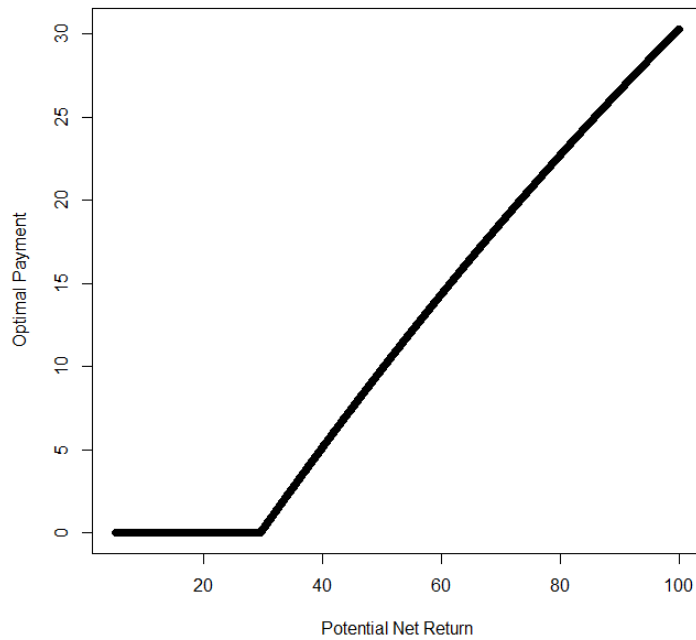


Figure 1.19: Optimal payment as a function of potential net return, normally-distributed signals

To address this heterogeneity, we simulate the investment delegation model with a population of agents, each of which could have a fixed costs ($\kappa = 40$), mutual information ($\alpha = 10$), or linear precision in normally-distributed signals (marginal cost of precision 7.5) for their cost function.³⁷

Figures 1.20 and 1.21 show the principal’s optimal payment scheme in this hybrid model. Figure 1.20 uses the proportion of each cost function type found among responsive subjects in our population, while Figure 1.21 uses a population with 50% fixed-cost agents, 40% mutual-information agents, and 10% normal-signal agents. Notice that the optimal payment scheme of the hybrid model resembles a mixture of the optimal payment schemes for each of the individual cost function types, with, going from left to right, a flat region,

³⁷We assume that agents are rationally inattentive and responsive.

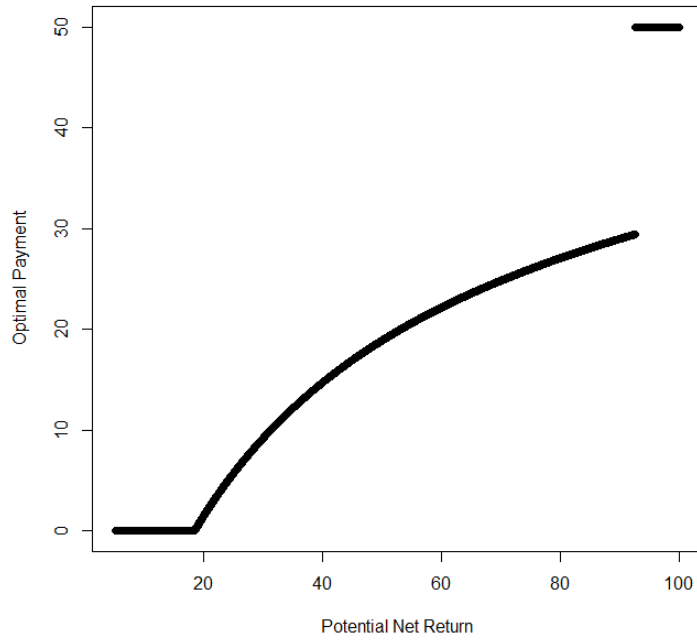


Figure 1.20: Optimal payment as a function of potential net return, hybrid model with empirical proportions of types

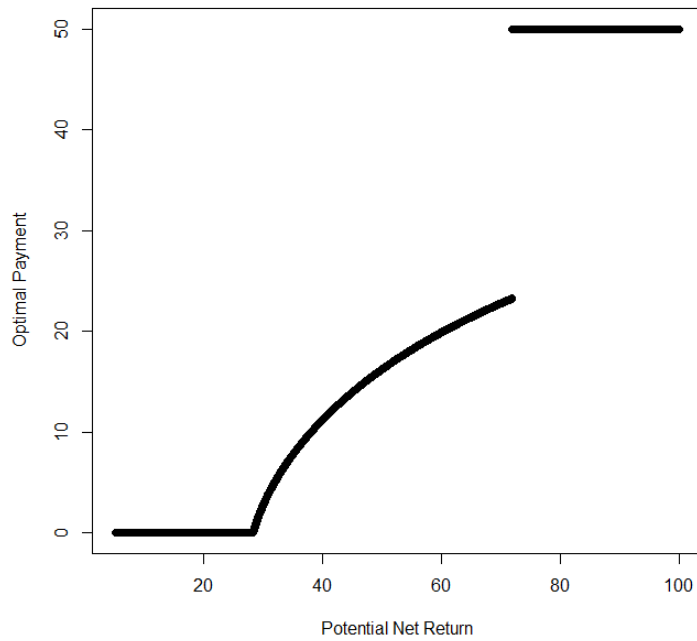


Figure 1.21: Optimal payment as a function of potential net return, hybrid model with 50% fixed-cost types

followed by a region of strict increase, followed by another flat region. These flat regions increase in size when the proportion of fixed-cost agents is increased.

The examples of this section demonstrate that the characteristics of an agent's information cost function can affect how a principal's decisions change with the parameters of his environment. Some cost functions, like the fixed-cost function described above, induce the principal to pay for information acquisition only when the potential net return is sufficiently high. Others, like the mutual information cost function, cause the principal to vary his payment continuously with changing potential net returns. In hybrid models combining these cost functions, the optimal payment scheme resembles a mixture of the optimal payments of the component cost function types.

1.11 Conclusion

This paper has provided a framework for testing properties of information cost functions in models of rational inattention. To the extent that the presence or absence of characteristics such as continuity, convexity, and perceptual distance can have an impact on people's decisions, it is worth knowing whether their cost functions satisfy such conditions. Decision-makers' cost functions are not directly observable, so instead we must infer their characteristics from observed behavior. We conducted a set of experiments that allowed us to implement these tests.

These experiments reveal substantial heterogeneity in behavior. Most subjects are rationally inattentive, but only about half are actually responsive to incentives. Many subjects have behavior that is consistent with continuous, convex cost functions, and

many subjects have behavior that evince the perception of distance as well. Moreover, there is considerable heterogeneity in whether or not subjects finely adjust their quantity of attention, though this heterogeneity is limited to three classes of cost functions: fixed costs, mutual information, and normal signals are the only best-fitting cost functions for responsive subjects in terms of performance.

This has implications for modeling. In models where agents acquire information, it is perhaps a better reflection of reality to allow for some heterogeneity in their fundamentals. In particular, one may want to model some portion of the population as non-responsive to incentives for information acquisition, or among responsive incentives, one could designate some as having fixed costs for information acquisition and others as having mutual-information or normal-signal cost functions.

Two possible avenues for future experimental research present themselves. The first is to obtain more detailed data on what subjects are actually paying attention to. Eye-tracking has already been used in several economics experiments (e.g. Wang et al., 2010; Krajbich et al., 2010; Arieli et al., 2011) to track subjects' gaze, which allows researchers to find out what visual information the subjects are acquiring. Tracking subjects' mouse movements in computer-based tasks (e.g. Gabaix et al., 2006) is another potential approach. The second is to use choice data in tandem with reaction time data to fit models of dynamic information acquisition. This would also allow researchers to determine to what extent subjects trade off speed and accuracy in their decision-making.

This page intentionally left blank.

Chapter 2

Promises and Pronouncements

This chapter is coauthored with Charles Maurin.

2.1 Introduction

Researchers have long been interested in what motivates dishonest behavior. Mazar et al. (2008) theorize that it is the result of a trade-off between positive self-concept maintenance and rational self-interest; when the latter dominates the former, then an individual lies.¹ This is consistent with previous findings by Kashy and DePaulo (1996), who note that people who are more concerned with impression management tell more lies.

Another notable finding of Kashy and DePaulo (1996) is that people whose personalities indicate higher levels of social responsibility are less likely to tell selfish lies. This raises the question: if the level of responsibility that one felt over a situation could be exogenously manipulated, would that have an influence on lying behavior?

For example, consider the case of an investment manager who can either report potential returns to a client truthfully or guarantee a higher rate of return that he knows will not be achieved. Suppose that the investment manager lies and guarantees the higher

¹In this paper, we purposely avoid the use of the words "deceive" and "deception," so as not to invite confusion with the concept of "experimental deception," where experimenters mislead or withhold crucial information from subjects. Our experiment involves no such deception and was in fact conducted in a lab with a strict no-deception policy.

return. Then, when the low return is realized, he may feel more culpable if this was due to his own intended lack of effort than if it was due to unfavorable market conditions (of which he was aware). That is, the level of control that the investment manager has over the final outcome may influence how responsible he feels for a bad outcome. In turn, this may influence how likely he is to lie about the possibility of a bad outcome in the first place.

Our experiment is designed to test if those who communicate information are more or less likely to lie if the final outcome of an interaction is under their control. Put differently, we aim to study if there is a different propensity to lie when making a *promise* — a statement of intent to commit actions in the future — versus when making a *pronouncement* — a statement about a variable outside of the communicator’s control. This research question is relevant in management and economics, as illustrated by the example of the investment manager, but not just to questions of investment. If our hypothesis that statements of intent are less likely to be false than statements about uncontrolled variables, then this provides an argument against delegation; a supervisor would be less likely to lie about a project under her direct control than one delegated to her subordinates.²

The paper is organized as follows. Section 2 presents a survey of related literature. Section 3 introduces the game we study and derives predictions for how the game is played under both standard cheap talk and lying costs. Section 4 presents our experimental design and explain how the game studied in Section 3 is implemented in a laboratory setting. Section 5 presents experimental results. Section 6 discusses the results and concludes.

²Of course, delegation may provide benefits in terms of efficiency. To the extent that lies are socially harmful, there may be a trade-off between the efficiency of delegation and the resulting dishonesty.

2.2 Related Literature

Since Gneezy (2005), there has been a large experimental literature on lying in economic interactions.³ Most of this literature studies games with variations on a simple sender-receiver paradigm and finds evidence of lying aversion. However, the literature does not usually explicitly distinguish between experiments involving promises and experiments involving pronouncements; i.e. it does not study these as separate phenomena. Here, we separate the two strands of literature.

2.2.1 Promises

Charness and Dufwenberg (2006) study a trust game where the trustee can send a free-form pre-play message to the truster. They find that pre-play communication increases trust, especially when the message is promissory, and that it also increases cooperation. In a related paper (2010), the authors perform the same experiment but with a pre-defined message, and relative to their earlier paper, they find that the increases in trust and cooperation are reduced but still significant.

Vanberg (2008) studies a similar game to Charness and Dufwenberg (2006), but randomly rematches a portion of subjects after pre-play communication, while letting re-matched subjects look at the exchange of communication between the previous pair. He finds that breaking the promissory link between truster and trustee reduces the level of cooperation. This evidence strongly suggests that people's propensity to keep promises not only depends on the potential outcomes of others but also on a feeling of who should

³An inexhaustive survey of this literature can be found in Rosenbaum et al. (2014)

be responsible for engendering those outcomes.

More recent experiments have also used a design similar to Charness and Dufwenberg's (2006) design. Tadelis (2011) studies the same game as Charness and Dufwenberg (2006), but varies the degree to which the trustee's decision to break a promise is revealed with certainty ex post across treatments. He finds that the potential for such revelation increases cooperation. Ismayilov and Potters (2015) run an experiment where it is possible for a message not to be delivered to the recipient and find that more trustworthy senders make promises, rather than promises inducing trustworthiness. In another variation of the design, Ederer and Stremitzer (2015) reduce senders' ability to keep their promises and find that the resulting changes in players' expectations have a negative impact on cooperation.

2.2.2 Pronouncements

There are many experiments involving sender-receiver games where the sender takes no additional action after communicating information to the receiver; payoffs are determined solely by a randomly chosen state of the world and the actions chosen by the receiver. Games such as these are very much in the spirit – and indeed are often special cases – of the “cheap talk” games analyzed by Crawford and Sobel (1982) and Green and Stokey (2007).

At least two papers experimentally implement a version of Crawford and Sobel's (1982) game. Cai and Wang (2006) confirm Crawford and Sobel's (1982) experimental findings, but also discover evidence that senders “overcommunicate” in that their mes-

sages are more informative about the true state of the world than would be predicted by the most informative equilibrium. They illustrate that these findings can be explained by boundedly rational senders who are level- k thinkers (e.g. Nagel, 1995), implying that people may demonstrate lying aversion because it is cognitively costly. Wang et al. (2010) confirm these findings with eyetracking data.

Other studies restrict receivers to simpler, typically binary, action spaces. Gneezy (2005) studies a game where senders tell receivers which of two options will give them more money, and then receivers select one of the options. He finds evidence of lying aversion and suggests that this aversion may be increasing in the potential loss for the receiver. Hurkens and Kartik (2009) reinterpret Gneezy's (2005) evidence and show that it is consistent with a model in which one portion of the population is willing to lie to get their preferred outcome, and the remainder never lies.

Many authors have used variations of this simple cheap-talk setup to study various issues in communication. Sánchez-Pagés and Vorsatz (2007) add a punishment phase to the game to test if people have an aversion to lying *per se* or if they simply have preferences over material outcomes, and they find evidence for both types of preferences. Lundquist et al. (2009) use pre-specified messages of varying strength and find that the aversion to lying is increasing in the strength and severity of a lie. Erat and Gneezy (2012) study situations where lies could actually benefit receivers and find evidence of aversion to telling even these "white lies." Sheremeta and Shields (2013) elicit the second-order beliefs of senders and find that liars tend to believe that their lies will be believed. Greenberg et al. (2015) study treatments where receivers can detect *ex post* if senders have lied to them and find that such treatments result in less lying, implying that senders are shame-averse.

2.2.3 Both Promises and Pronouncements

To our knowledge, the only other paper in experimental economics that studies both promises and pronouncements is Charness and Dufwenberg (2011). They implement a sender-receiver game where different types of senders have different levels of ability to complete a task; thus, senders can send messages both about their ability levels and future intentions, i.e. a message may contain both a promise and a pronouncement. Their research question differs substantially from ours; they are primarily concerned with whether giving all types of senders the ability to implement Pareto-improving outcomes results in communication that can sustain cooperative outcomes — indeed, they find that it does — whereas we are concerned with whether the propensity to tell a false pronouncement and the propensity to break a promise differ.

2.2.4 Other Experiments

Other studies find evidence of lying aversion even in the absence of explicit strategic considerations.⁴ Of particular note are the well-known die roll experiment of Fischbacher and Föllmi-Heusi (2013), which finds evidence of “partial lying,” where individuals lie but not necessarily maximally, and an experiment by Gibson et al. (2013), which finds significant between- and within-person heterogeneity in lying aversion. The latter study is particularly relevant to this paper, because it provides support for the hypothesis that

⁴It is possible that experimental subjects perceive these experiments as “games” played against the experimenters, but it is also possible that social preference effects are attenuated when subjects are not explicitly playing against other players. For example, Fischbacher and Föllmi-Heusi (2013) notice a shift towards less severe lies when these lies affect another player, but this shift is not significant. However, it should be noted that each session of their experiment was conducted immediately following another experiment in which the subjects participated, and they did not control for this previous experiment. As such, the insignificance of the shift should be interpreted with caution.

lying aversion is highly context-dependent.

2.2.5 Theoretical Approaches

Various approaches have been devised to explain lying aversion. Kartik et al. (2007) and Kartik (2009) study communication games where senders have explicit lying costs that increase in the severity of the lie. Both papers find that if there is a non-zero possibility for the receiver to be credulous, then the sender will engage in “language inflation” and bias his report in the direction that he favors.

Battigalli and Dufwenberg (2007) study psychological games (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009) where people feel guilt for causing others not to meet their payoff expectations (not specifically for lying). The amount of guilt felt is proportional to the amount by which others fall short of their expectations. Battigalli et al. (2013) apply this guilt aversion model to Gneezy (2005) and find that their theory can explain his results.

None of these approaches explicitly varies the costs of dishonesty with the degree of control that the sender has over the outcome, though differences in sender behavior between promises and pronouncements may arise in Battigalli and Dufwenberg’s (2007) approach because of strategic considerations, as illustrated in the next section.

2.3 The Game

In this section, we outline the game on which our experimental design is based. The game is similar in structure to the one studied by Sánchez-Pagés and Vorsatz (2007). For the sake

of readability, we assume in this section that gross utility (before applying costs of lying, guilt, etc.) is linear and strictly increasing in money, or material payoffs, so we use the terms “gross utility” and “material payoffs” interchangeably. The qualitative results would be essentially unchanged if we only assumed that gross utility were strictly increasing in material payoffs, as we demonstrate in the appendix.

There are two players: a Sender (S, “he”), and a Receiver, (R, “she”), and they could play one of two possible scenarios, each with equal probability, a Promises scenario (μ), and a Pronouncements scenario (ν). The Sender is informed of what scenario is being played, but the Receiver is not.

In the Promises scenario (see Figure 2.1), the Sender chooses whether to play Left (λ) or Right (ρ). We refer to this selection as the “state.” Without observing the Sender’s selection, the Receiver also chooses whether to play Left (ℓ) or Right (r). If both players go in the same direction, then they both obtain a payoff of 10. If the players go in opposite directions, then the Sender obtains a payoff of 15, and the Receiver obtains a payoff of 0.

The Pronouncements scenario (see Figure 2.2) is identical in structure to the Promises scenario, except the Sender’s Left/Right decision is made by Nature (N). Nature selects either Left (λ) or Right (ρ) as the state, each with equal probability, and then the Receiver selects whether to go In or Out. The payoffs are as in the Promises scenario.

After the Sender learns what scenario will occur, he can choose to send a message that encourages the Receiver to go Left (m_λ) or to go Right (m_ρ). This message is pre-defined and cannot be used to identify the scenario being played. In the Pronouncements scenario, the Sender makes his messaging decision after having learned Nature’s selection of Left or Right, and in the Promises scenario, he makes his messaging decision after having chosen

Left or Right. In either scenario, we say that the sender “told the truth” if he sent the message m_s when the state was s and that he “lied” if he sent the message m_s when the state was y , with $s \neq y$. The structure of the entire game is depicted in Figure 2.3.

The advantage of using this game to study promises and pronouncements is that the Receiver’s beliefs about the probability of Left or Right being selected are constant across scenarios and depend only on which message was sent; i.e., she cannot infer different meanings for a message in the different scenarios. As a result, this also fixes the Sender’s second-order beliefs (i.e. his beliefs about the Receiver’s beliefs about the probability of Left or Right being selected), and so the observed message-sending behavior of the Sender cannot be affected by strategic responses to belief shifts between a Promises scenario and Pronouncements scenario.

2.3.1 Perfect Bayesian Equilibrium

In the absence of social preferences or constraints on rationality, the theoretical prediction for the outcome of the game is that the Receiver randomizes between Left and Right, and the Sender’s communication has no “bite” in determining the Receiver’s behavior. This is because the two players have opposing interests; the Receiver gains when the Sender loses, and vice versa. The Receiver randomizes evenly between Left and Right so that the Sender cannot take advantage of her actions, and the Sender communicates uninformatively to the Receiver so that she cannot take advantage of any information extracted from his messages. The following proposition formalizes this:

Proposition 15. *In any perfect Bayesian equilibrium (PBE) of the game, the Sender is equally*

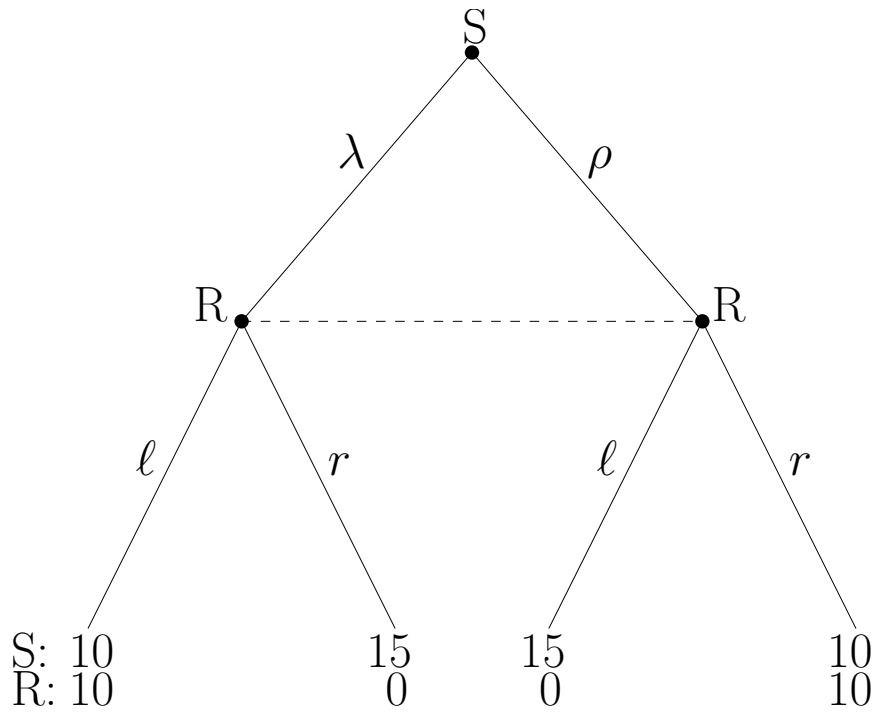


Figure 2.1: The Promises scenario

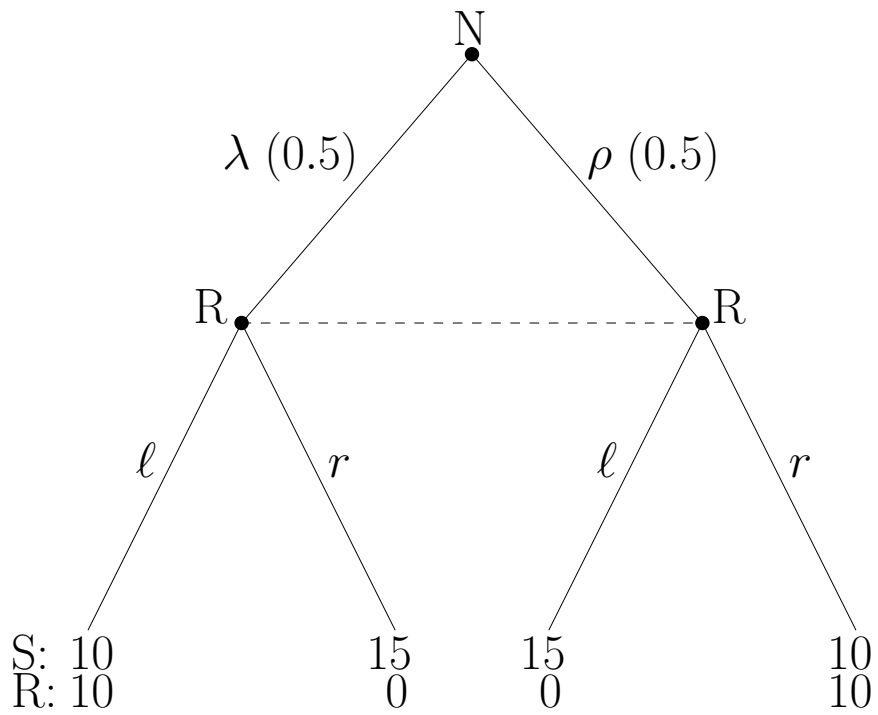


Figure 2.2: The Pronouncements scenario

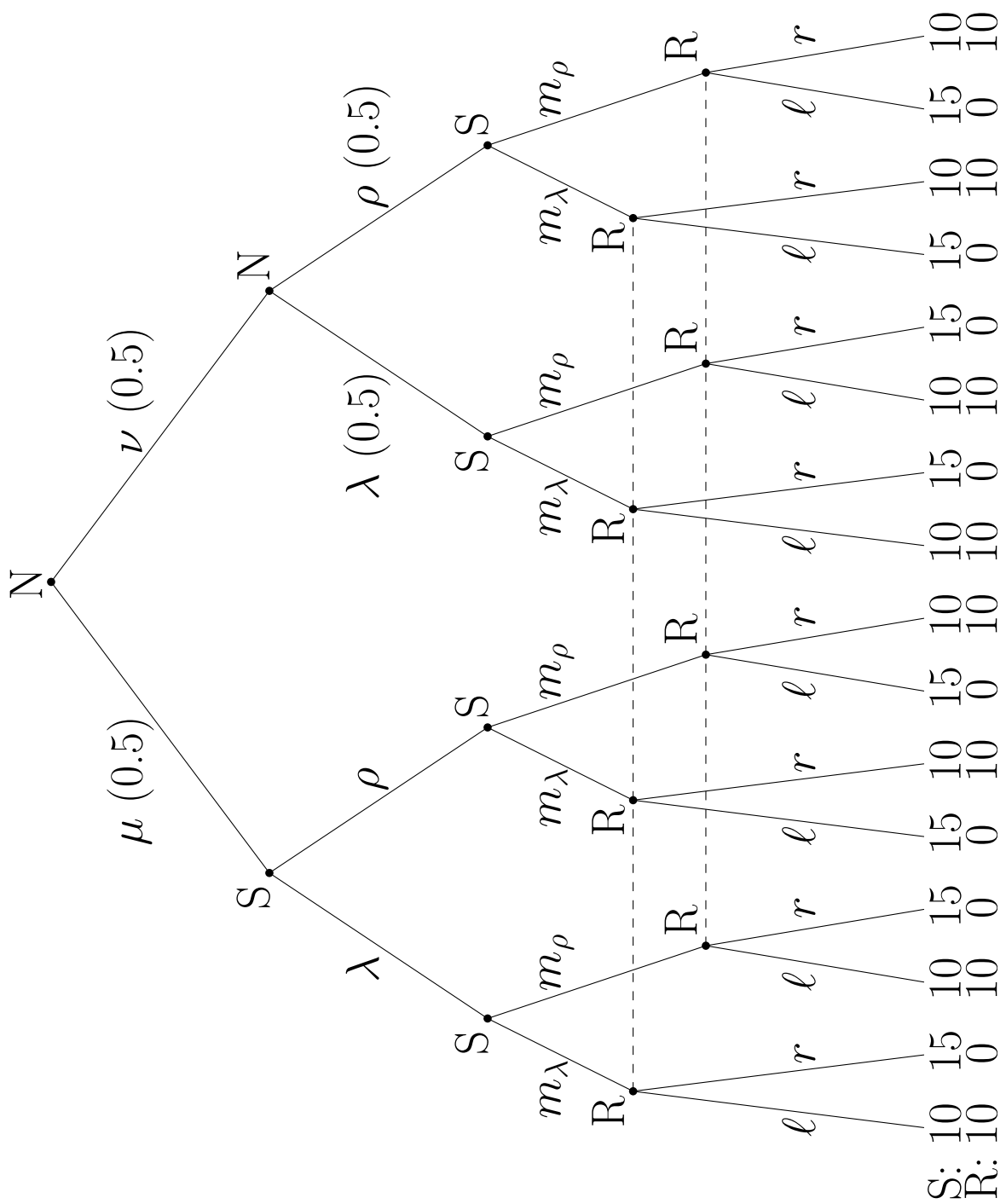


Figure 2.3: The full game

likely to send a given message in one state as in the other, and the receiver is equally likely to go Left as to go Right upon receiving either message. Moreover, the Sender is equally likely to select either direction in the Promise scenario.

Proof. Consider first the Receiver's strategy. Suppose the Receiver were, WLOG, strictly more likely to select ℓ than r upon receiving m_λ , i.e. $\Pr(\ell|m_\lambda) > \Pr(r|m_\lambda)$. Suppose also, WLOG, that $\Pr(\ell|m_\rho) \in [\Pr(r|m_\lambda), \Pr(\ell|m_\lambda)]$. Then, the Sender's best response would be to always choose ρ and send m_λ in the Promise scenario, and to send m_λ when the state is ρ and send m_ρ when the state is λ in the Pronouncement scenario. But the Receiver's best response to that strategy would be to always play r upon receiving m_λ and ℓ upon receiving m_ρ . Therefore, the strategy originally proposed for the Receiver cannot be part of a PBE, and so she must be equally likely to select either action after receiving either message in equilibrium.

Now consider the Sender's choice of action. Suppose, WLOG, that the Sender plays λ strictly more often than ρ in the Promises scenario, i.e. $\Pr(\lambda|\mu) > \Pr(\rho|\mu)$, which further implies that $\Pr(\lambda) > \Pr(\rho)$. Then for at least one of the messages m_x , it must be the case that $\Pr(\lambda|m_x) > \Pr(\rho|m_x)$. WLOG, this message is m_λ . But then, the Receiver's best response upon receiving m_λ is to always play ℓ . There are now two cases. 1) If $\Pr(\lambda|m_\rho) \geq \Pr(\rho|m_\rho)$, then the Receiver's best response upon receiving m_ρ is to play ℓ . In that case, the Sender's optimal actions would be to always play ρ in the Promise scenario and send, WLOG, m_λ . 2) If $\Pr(\lambda|m_\rho) < \Pr(\rho|m_\rho)$, then the Receiver's best response upon receiving m_ρ is to play r . In that case, the Sender could improve his expected payoff by always playing ρ when he would have played ℓ in the original strategy and sending m_λ

in those situations. In either case, the strategy originally proposed for the Sender cannot be part of a PBE, and so he must be equally likely to select either direction in the Promise scenario, i.e. $\Pr(\lambda|\mu) = \Pr(\rho|\mu) = \frac{1}{2}$, which implies that $\Pr(\lambda) = \Pr(\rho) = \frac{1}{2}$

Finally, consider the sender's communication choices. Since the Sender benefits from the Receiver mismatching the state, the Sender wishes to minimize the probability that the Receiver can match the state. This probability is minimized when $\Pr(\lambda|m_\lambda) = \Pr(\lambda|m_\rho) = \frac{1}{2}$.⁵ Therefore, since $\Pr(\lambda) = \Pr(\rho) = \frac{1}{2}$, we have by Bayes' rule:

$$\begin{aligned} \Pr(\lambda|m_\lambda) &= \frac{\Pr(m_\lambda|\lambda) \Pr(\lambda)}{\Pr(m_\lambda|\lambda) \Pr(\lambda) + \Pr(m_\lambda|\rho) \Pr(\rho)} \\ \iff \frac{1}{2} &= \frac{\Pr(m_\lambda|\lambda)}{\Pr(m_\lambda|\lambda) + \Pr(m_\lambda|\rho)} \\ \iff \Pr(m_\lambda|\lambda) &= \Pr(m_\lambda|\rho) \end{aligned}$$

Similarly, $\Pr(m_\rho|\lambda) = \Pr(m_\rho|\rho)$. □

An easy corollary of this proposition is that in any PBE, the Sender is equally likely to lie as to tell the truth. To see this, note that the *ex ante* probability of truth-telling is $\Pr(\lambda) \Pr(m_\lambda|\lambda) + \Pr(\rho) \Pr(m_\rho|\rho)$, and the *ex ante* probability of lying is $\Pr(\lambda) \Pr(m_\rho|\lambda) + \Pr(\rho) \Pr(m_\lambda|\rho)$. But since $\Pr(m_\lambda|\lambda) = \Pr(m_\lambda|\rho)$ and $\Pr(m_\rho|\lambda) = \Pr(m_\rho|\rho)$, these probabilities are the same.

⁵It is possible that one of the messages, say WLOG, m_ρ , is unused in equilibrium. But the Receiver's off-path beliefs must still be such that $\Pr(\lambda|m_\rho) = \frac{1}{2}$, or else the Sender would benefit from sending m_ρ when the state is ρ if the belief is greater than $\frac{1}{2}$ and sending m_ρ when the state is λ if the belief is less than $\frac{1}{2}$ (provided the Receiver is sequentially rational in their choice of action).

2.3.2 Theories of Lying Aversion

As outlined in Section 2.2, in practice it is observed that people lie less often than equilibrium would predict. There are many possible explanations for this, including simple lying costs, guilt aversion (Battigalli and Dufwenberg, 2007), altruistic preferences, and inequity aversion (Fehr and Schmidt, 1999). In this subsection, we show that under the assumption that Receivers are more likely to follow a message's recommendation than not to, none of these theories could account for differences in lying behavior across scenarios.

2.3.2.1 Lying Costs

In this subsection, we assume that Senders simply have a positive cost of lying c , of which they are aware. For now, we proceed under the assumption that the cost of lying is not necessarily known to the Receiver to derive the central experimental hypothesis that our experiment is designed to test. We can distinguish between two types of lying costs: *intentional*, which the Sender feels whenever he lies, and *consequentialist*, which the Sender feels only when his lie harms the Receiver, i.e. if she actually follows the recommendation of his lie. Put differently, if the Sender has intentional lying costs, he feels them if he sends the message m_x when the state is y , with $x \neq y$, and if he has consequentialist lying costs, he feels them if in addition the Receiver takes the action $f(x)$, where $f(\lambda) := \ell$ and $f(\rho) := r$.

Let $\text{Pr}_R(y|m_x)$ denote the Receiver's belief that the state is y after receiving message m_x . Suppose that the Receiver is just as likely to believe a recommendation to go Left as she is a recommendation to go Right, so that $\text{Pr}_R(\lambda|m_\lambda) = \text{Pr}_R(\rho|m_\rho) =: \alpha$.⁶ In order for

⁶This is an easily justifiable assumption, given the symmetry of the game.

a Receiver to follow the Sender's recommendation, she requires:

$$10\alpha \geq 10(1 - \alpha) \iff \alpha \geq \frac{1}{2} \quad (2.1)$$

Now let $\beta := \Pr_S(\alpha \geq \frac{1}{2})$, i.e. the Sender's belief that the Receiver is more likely to follow his recommendation than not. α is the Receiver's first-order belief that the Sender is being truthful, and β is the receiver's second-order belief about the Receiver's belief about the Sender's truthfulness. This second-order belief is crucial for analyzing the Sender's incentive to lie.

Proposition 16. *If lying costs are intentional, the Sender is willing to lie iff $c \leq 10\beta - 5$, regardless of the scenario. If lying costs are consequentialist, the Sender is willing to lie iff $c \leq \frac{10\beta-5}{\beta}$, regardless of the scenario.*

Proof. Under intentional lying costs, the Sender's expected payoff from lying is $15\beta + 10(1 - \beta) - c$, and his expected payoff from telling the truth is $10\beta + 15(1 - \beta)$. The former is greater than the latter iff $c \leq 10\beta - 5$.

Under consequentialist lying costs, the Sender's expected payoff from lying is $(15 - c)\beta + 10(1 - \beta)$, and his expected payoff from telling the truth is $10\beta + 15(1 - \beta)$. The former is greater than the latter iff $c \leq \frac{10\beta-5}{\beta}$. \square

There are several important implications to note here. Note that as long as lying costs are sufficiently low, the Sender is willing to lie, regardless of scenario; neither type of lying costs implies a different propensity of breaking promises than of telling false pronouncements. The upper bound below which lying is desirable increases in the second-order

belief (for both intentional and consequentialist lying costs), meaning that Senders are more likely to lie if they believe that their lies will be believed. Finally, note that the upper bound for consequentialist lying costs is higher than that for intentional ones; if the Sender only feels a lying cost when it harms the Receiver, then he is less likely to incur that cost and therefore more inclined to lie.

2.3.2.2 Social Preferences

It is possible that instead of being averse to lying *per se*, people simply have preferences over final monetary payoffs such that they feel good about good outcomes for other players or bad about bad outcomes for other players. We call such preferences *social preferences*. Formally, a social preference is a function $u_i, i \in \{R, S\}$ mapping monetary payoff vectors (π_S, π_R) to real numbers. Examples include altruistic preferences, where a player's psychological payoff is increasing in other's monetary payoffs, e.g. $u_S(\pi_S, \pi_R) = \pi_S + \gamma\pi_R, \gamma > 0$, and inequity aversion (Fehr and Schmidt, 1999).

Because social preferences only depend on monetary distributions of payoffs, in our game, they would be unable to explain any potential observed discrepancies between the propensity to lie across scenarios. Moreover, if the propensity to lie does not depend continuously on beliefs. The following proposition formalizes this:

Proposition 17. *Let $u_A := u(15, 0)$ and $u_B := u(10, 10)$. The Sender is willing to lie iff $u_A \geq u_B$ when $\beta \geq \frac{1}{2}$ and is willing to lie iff $u_A \leq u_B$ when $\beta \leq \frac{1}{2}$, regardless of scenario.*

Proof. The Sender's expected payoff from lying is $u_A\beta + u_B(1 - \beta)$, and his expected payoff from telling the truth is $u_A(1 - \beta) + u_B\beta$. The former is weakly greater than the

latter when $u_A \geq u_B$ and $\beta \geq \frac{1}{2}$, or when $u_A \leq u_B$ and $\beta \leq \frac{1}{2}$. □

In other words, the Sender is willing to lie when his message is more likely to be believed than disbelieved, provided the unequal outcome gives him a greater psychological payoff than the equal one, and he is willing to when his message is less likely to be believed than disbelieved, provided the equal outcome gives him a greater psychological payoff than the unequal one.⁷

[No social preference lie threshold (a) depends on beliefs (b) can explain discrepancy b/w prom/pron]

2.3.2.3 Guilt Aversion

Another theory that has been used to explain lying aversion is Battigalli and Dufwenberg's (2007) theory of guilt aversion (e.g. Charness and Dufwenberg, 2006; Battigalli et al., 2013). Under this theory, Senders feel a guilt cost that is proportional to the amount by which they cause Receivers not to meet their expected payoff. Formally:

$$u_S = \pi_S - \theta \max\{0, \mathbb{E}_\alpha[\pi_R] - \pi_R\} \quad (2.2)$$

where u_S is the Sender's utility, π_S is his monetary payoff, π_R is the Receiver's monetary payoff, and α is her first-order belief.

Similarly to social preferences, guilt aversion produces a lying condition that does not depend on second-order beliefs or the scenario being played.

⁷The latter case can be considered an example of "white lie," as defined by Erat and Gneezy (2012). As stated in Subsection 2.2.2, they find evidence of aversion to telling even these lies.

Proposition 18. *If the Sender is guilt-averse, then he is willing to lie iff $\theta \leq \frac{1}{2}$, regardless of scenario.*

Proof. The Sender's expected payoff from lying is $10(1 - \beta) + (15 - 10\theta\beta)\beta$, and his expected payoff from telling the truth is $10\beta + (15 - 10\theta(1 - \beta))(1 - \beta)$. Simple algebra shows that the former is greater than or equal to the latter iff $\theta \leq \frac{1}{2}$. \square

2.3.3 Equilibrium with Intentional Lying Costs

In the preceding section, we analyzed lying incentives in our game under various theories of lying aversion. In this section, we assume intentional lying costs and examine the equilibria that arise.

2.3.3.1 Common Knowledge of Lying Costs

We begin by assuming common knowledge of lying costs.⁸ This is a very strong epistemic assumption, since there is no reason why a Receiver should know an anonymous Sender's lying costs, but it provides a stark, useful benchmark.

In this case, as long as $c < 5$, the Sender prefers outcomes where the Receiver mismatches the states to ones where she matches it. Therefore this game has the same equilibria as the game without lying costs, as outlined in Proposition 15. However, if $c > 5$, then the Sender would rather tell the truth and have the Receiver match the state than tell a lie and have the Receiver mismatch the state; therefore, the Sender always tells the truth.

⁸In other words, we assume that the distribution of lying costs is degenerate. We examine a specific non-degenerate distribution of lying costs in the following section.

2.3.3.2 Ethical and Economic Types

Degenerate lying cost distributions provide very stark equilibrium results; lying costs are either too small to be relevant or so large that they cause Senders to always be honest. While small lying costs allow for some heterogeneity in observed behavior – Senders randomize between lying and not lying – they do not, for instance, allow for Senders who always lie.

To that end, in this sub-subsection we derive equilibrium results for a model of the kind suggested by Hurkens and Kartik (2009), where there are two types of Senders: “ethical” types, with infinite lying costs, and “economic” types, with zero lying costs. We show that the ethical types, who never lie, can in a sense provide cover for the economic types to lie.

Since our game is symmetric, we are interested in equilibria with a symmetric structure. We define a *symmetric PBE* of this game to have the following characteristics:

- $\zeta := \Pr(\ell|m_\lambda) = \Pr(r|m_\rho)$ (The Receiver is equally likely to follow a Right message as a Left message.)
- $\Pr(\lambda|\mu) = \Pr(\rho|\mu) = \frac{1}{2}$ (The Sender is equally likely to choose Left or Right in the Promises scenario.)
- $\phi := \Pr(m_\lambda|\lambda, c = 0) = \Pr(m_\rho|\rho, c = 0)$ (Economic types are equally likely to tell truth if the state is Left or if it is Right.)

Proposition 19. *Let $\xi = \Pr(c = \infty)$, i.e. the proportion of ethical types. If $\xi \geq \frac{1}{2}$, then there is a symmetric PBE where $\zeta = 1$ and $\phi = 0$. If $\xi \geq \frac{1}{2}$, then there is a symmetric PBE where*

$$\zeta = \frac{1}{2} \text{ and } \phi = \frac{1-2\xi}{2-2\xi}.$$

Proof. First, note that an ethical type is never willing to lie. An economic type, on the other hand, is willing to lie as long as his expected payoff from doing so is at least as large as his expected payoff from telling the truth. Put differently, they are willing to lie as long as $15\zeta + 10(1 - \zeta) \geq 10\zeta + 15(1 - \zeta)$, which holds iff $\zeta \geq \frac{1}{2}$.

The Receiver is willing to follow the Sender's message as long as their expected payoff from doing so is at least as great as their expected payoff from doing the opposite of what the message suggests. Put differently, she is willing to follow the message as long as $10\xi + 10(1 - \xi)\phi \geq 10(1 - \xi)(1 - \phi)$ (since ethical types never lie), which holds iff:

$$\xi - \xi\phi + \phi \geq \frac{1}{2} \tag{2.3}$$

(2.3) holds $\forall \phi \in [0, 1]$, as long as $\xi \geq \frac{1}{2}$. In other words, a Receiver always follows the Sender's message if a Sender is at least as likely to be ethical as economic. Knowing that his message will be followed, an economic type can maximize his expected payoff with $\phi = 0$.

If $\xi \leq \frac{1}{2}$, then (2.3) does not hold for all ϕ and in particular not for $\phi = 0$ (except when $\xi = \frac{1}{2}$), so if economic types always lie, then the Receiver will never follow messages. Therefore, economic types must tell the truth at least some of the time. In order for economic types to be indifferent between lying and not lying, we must have $\zeta = \frac{1}{2}$. And in order for the Receiver to be indifferent between following and not following a message, we need (2.3) to hold with equality, i.e. $\phi = \frac{1-2\xi}{2-2\xi}$. \square

If there are enough ethical types in the population, then their existence lowers the probability of a Receiver being lied to enough so that economic types can lie with the knowledge that a Receiver will have enough confidence in their messages to follow them. In this manner, ethical types provide cover for economic types to lie.

It is also worth noting that since Receivers do not observe the scenario being played, this theory does not predict how lying behavior depends on the scenario being played when $\xi < \frac{1}{2}$. While lying equally often under either scenario can be part of an equilibrium, it is also possible, for instance, for the Sender to always lie in Pronouncements scenario and tell the truth with probability $\frac{1-2\xi}{1-\xi}$ in the Promises scenario. Therefore, observed differences in lying behavior between the two scenarios are neither ruled in nor ruled out by a theory of intentional lying costs.

2.3.4 An Alternative Game

As written, in the Promises scenario (μ) of our game, the Sender chooses the state before sending a message. This facilitates comparison of behavior with the Pronouncements scenario (ν). However, one typically thinks of promises are being statements of commitment to a certain action, not statements about a previously chosen action.

For that reason, we also consider a version of our game where the order of message and state choice is flipped in the Promises scenario. (The Receiver continues not to observe the choice of state.) It is obvious that this does not affect equilibrium behavior, nor does it affect the behavior predicted by the aforementioned theories of lying aversion. This is because both in our original game and our alternative game, the Sender's decisions in the

Promises scenario could be collapsed into a single decision; in either case, the Receiver observes nothing until both decisions have been made.

2.4 Experimental Design

The experimental sessions were conducted at the Columbia Experimental Laboratory for the Social Sciences (CELSS), using the popular z-Tree software (Fischbacher, 2007). Subjects were placed at individual carrels with computer terminals whose monitors were covered by tinted glass and a shade so that subjects could not easily see each other's screens. Participants were recruited through CELSS's common subject pool, which consists of Columbia University students and alumni,⁹ via ORSEE (Greiner, 2015). We ran 4 sessions of the baseline game and 2 sessions of the alternative game.

Our experiment used a within-subjects design. Subjects played the game described in the preceding sections over the course of 4 stages.¹⁰ That is to say, each subject played the game four times. Each subject was randomly matched with another at the start of the experiment and was then randomly rematched between each stage. At the start of the first stage, one subject in each pair was assigned the role of Sender, and the other was assigned the role of Receiver. Each subject retained the same role for the first two stages ("Round 1") and then switched to the other role for the final two stages ("Round 2"). If a subject played Scenario 1 (Promises) in the first (third) stage, then she would play Scenario

⁹The subject pool is primarily undergraduate, but also has a significant number of graduate students. The number of alumni in the pool is negligible.

¹⁰The use of the term "stage" instead of the usual "round" here is deliberate. In our experimental instructions, the first two and last two "stages" each composed a "round," for a total of 4 stages and 2 rounds. We used the term "round" in this manner to make the rematching procedure clearer to the subjects.

2 (Pronouncements) in the second (fourth) stage, and vice versa. This ensured that each subject played both scenarios in both roles over the course of the experiment. When in the Sender role, subjects were informed of the order in which the scenarios would be played; when in the Receiver role, subjects were not informed of this order and were told that both orders were equally probable.

The messages that Senders could transmit to Receivers were fixed. They were "Going Left is in your best interests." and "Going Right is in your best interests." These messages were chosen so that they could not be used to identify which of the two scenarios was being played.

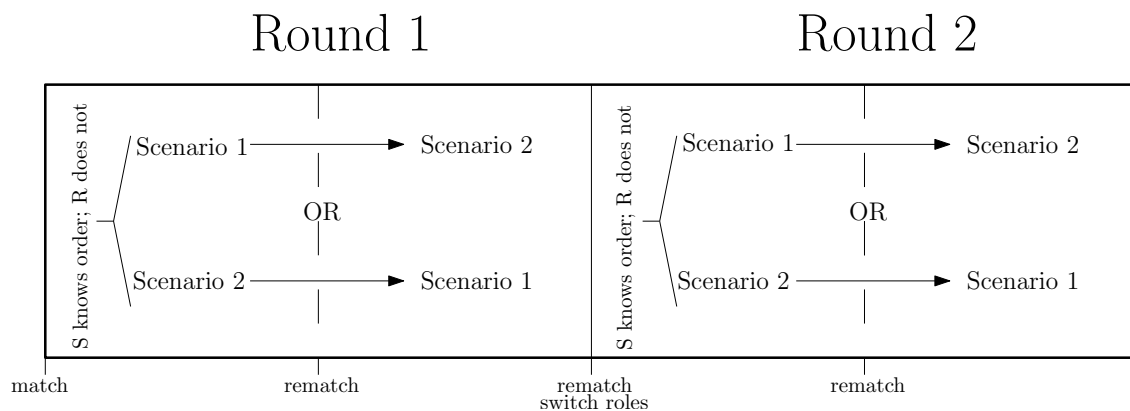


Figure 2.4: The rematching procedure

Prior to the four iterations of the game, subjects were given a brief questionnaire to test their understanding of the rules. After the four iterations of the game, the subjects' first-order and second-order beliefs about whether following the Sender's message would be in the Receiver's best interests when a message was sent or was not sent were elicited.¹¹

¹¹Specifically, when eliciting first-order beliefs, subjects were told how many interactions there were between subjects in the session. They were then asked to guess in how many of those interactions listening to the Sender's message was in the Receiver's best interests. When eliciting second-order beliefs, subjects were asked in how many interactions the Receiver thought that following the Sender's message was in her best interests.

Subjects were also asked a question to test if they truly understood that these beliefs could not shift between the two scenarios. Finally, subjects were given a brief demographic questionnaire asking about age, gender, and education.

One of the first two stages and one of the last two stages were randomly selected for payment, each at a rate of 2 points to \$1 US. So as not to blunt incentives, the same stages were selected for every subject. To incentivize belief elicitation, one of the two belief elicitation questions was also randomly selected for payment. If a subject's stated beliefs were within 20% of the correct value for that question,¹² then she could earn an additional \$2. Subjects also earned a show-up fee of \$5. Total earnings for each subject were rounded up to the nearest dollar to facilitate payment.

2.5 Results

Subsection 2.5.1 presents basic demographic data from the experiment. Subsection 2.5.2 analyzes the behavior of Receivers (subjects in the role of Individual B), and subsection 2.5.3 analyzes the behavior of Senders (subjects in the role of Individual A).

2.5.1 Summary Statistics

Table 2.1 lists basic demographic data for the subjects. The pool is fairly gender-balanced;¹³ the null of perfect gender balance cannot be rejected (two-sided test of pro-

¹²For the first-order belief question, answers were compared to the proportion of interactions in which the state matched the Sender's message. For the second-order belief question, answers were compared to the proportion of interactions in which the Receiver's choice matched the Sender's message.

¹³Subjects were given the option to list their gender as "Other / Non-binary"; no subjects used this option.

Table 2.1: Laboratory Demographics

Number of subjects	$n = 108$
Gender	Male: 52.8%; Female: 47.2%
Age (years)	Average: 23.3; St. Dev.: 4.64
Highest level of education achieved	
Some post-secondary	52.8%
Completed bachelor's degree	13.0%
Completed graduate or professional degree	34.3%
Primary area of study is economics	14.8%

portions, $p = 0.630$). The pool is also highly-educated; nearly half of the subjects have a post-secondary degree. Also note that most subjects are not economics majors.

Since there were 108 subjects, each subject played the game four times, and the game is played in pairs, there were 216 interactions between subjects in this experiment.

2.5.2 Receiver Behavior

In 75.0% (162/216) of interactions between subjects, Receivers followed Senders' messages. In other words, in 75.0% of interactions, the Receiver's action matched the Sender's message.

This is driven primarily by subjects who follow the Sender's message in both their interactions as Receivers: 64.8% (70/108) of Receivers follow the Sender's message twice. 20.4% (22/108) of Receivers follow the Sender's message only once, and 14.8% (16/108) of Receivers never follow the Sender's message. As shown by Proposition 15, the PBE of the game without lying costs or social preferences predicts that Receivers follow Senders' messages half the time, which means that 25% of Receivers should follow both messages, 50% should follow one message, and 25% should follow no message. A chi-squared test

rejects this hypothesis ($\chi^2 = 91.9, p < 0.001$).

The mean first-order belief (i.e. the Receiver's belief that the Sender is telling the truth) is 55.3%, and the standard deviation is 19.7%. The PBE would predict a first-order belief of 50%, and the data reject this (two-sided Mann-Whitney U test, $p = 0.0053$). Figure 2.5 shows a histogram of first-order beliefs. Note that the distribution appears bimodal, with one mode around 0.5, and another between 0.7 and 0.8; indeed, a dip test (Hartigan and Hartigan, 1985) rejects the null of unimodality ($D = 0.052, p = 0.029$).

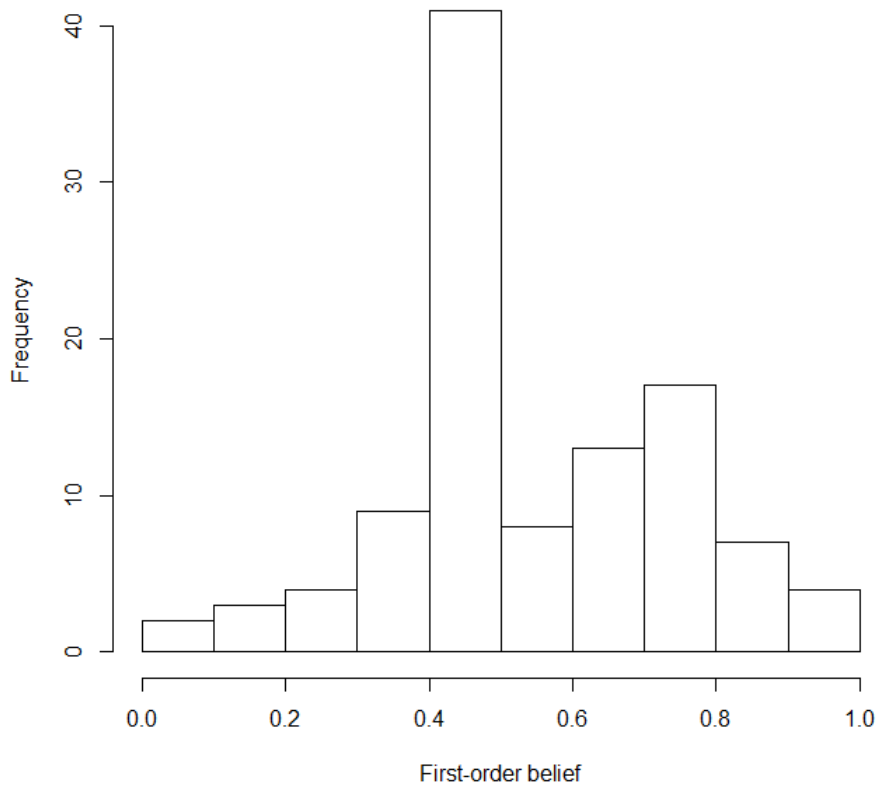


Figure 2.5: Histogram of first-order beliefs

Table 2.2 shows the results of a multinomial regression of belief data and demographic covariates on the categorization of Receivers by often they followed the Sender's message,

with no trust as the baseline category. Demographic variables have no significant effect on trusting behavior. Curiously, Receivers with higher second-order beliefs are more likely to trust both Senders' messages; those who believe others are credulous are more trusting. However, this effect is not highly significant.

Table 2.2: Categorization of Receivers

	Trust Once	Trust Twice
1st-order belief	2.326 (2.723)	3.455 (2.468)
2nd-order belief	0.301 (2.510)	3.942* (2.197)
Female	0.387 (0.735)	0.707 (0.659)
Economics major	0.014 (0.932)	-0.117 (0.862)
Age	-0.053 (0.088)	0.028 (0.071)
Sender first	0.559 (0.734)	0.088 (0.662)
Alternative treatment	1.557* (0.919)	1.084 (0.883)
Constant	-0.504 (2.696)	-3.494 (2.314)
Akaike Inf. Crit.	194.066	194.066
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	
<i>Baseline:</i>	<i>Trust Never</i>	

2.5.3 Sender Behavior

We now turn our attention to Sender behavior. Senders lie in 31.0% (67/216) of interactions, which means that they lie significantly less than Receivers believe.

This honesty is driven by the 60.2% (65/108) of Senders who never lie. 17.6% (19/108) lie once, and 22.2% (24/108) lie twice. As shown by Proposition 15, the PBE of the game without lying costs or social preferences predicts that Senders lie half the time, which means that 25% of them should never lie, 50% of them should lie once, and 25% of them should lie twice. A chi-squared test rejects this hypothesis ($\chi^2 = 78.5, p < 0.001$).

To further subcategorize the Senders who lie once, 47.4% (9/19) lie only in the Promises scenario, and 52.6% (10/19) lie only in the Pronouncements scenario. This is statistically indistinguishable from the Senders who lie once being equally likely to lie in either scenario (one-sided test of proportions, $p = 1.000$).

The mean second-order belief (i.e. the Sender's belief about the Receiver's belief that the Sender is telling the truth) is 55.4%, and the standard deviation is 21.3%. The PBE would predict that this belief would be 50%, but the data reject this (two-sided Mann-Whitney U test, $p = 0.0048$). Figure 2.6 shows a histogram of first-order beliefs. Note that the distribution appears bimodal, with one mode around 0.5, and another between 0.7 and 0.8; indeed, a dip test (Hartigan and Hartigan, 1985) rejects the null of unimodality ($D = 0.067, p < 0.001$). The correlation between first- and second-order beliefs is 0.710, which indicates that the credulous tend to believe others are credulous.

Tables 2.3 and 2.4 display multinomial regressions showing how the categorization of Senders by how often they lie depends on beliefs, experimental variables, and demo-

Table 2.3: Categorization of Senders (1)

	Lie Once	Lie Twice
1st-order belief	0.263 (1.997)	-7.650*** (2.293)
2nd-order belief	0.492 (1.826)	3.945** (1.897)
Female	0.245 (0.552)	-0.077 (0.558)
Economics major	-0.202 (0.893)	0.984 (0.792)
Age	0.053 (0.056)	-0.074 (0.071)
Sender first	-0.058 (0.560)	-1.142* (0.598)
Promises first as sender	-0.513 (0.542)	-0.262 (0.532)
Alternative treatment	0.707 (0.590)	-0.784 (0.703)
Constant	-2.959 (1.841)	3.205 (2.135)
Akaike Inf. Crit.	217.509	217.509
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	
<i>Baseline:</i>	<i>Lie Never</i>	

Table 2.4: Categorization of Senders (2)

	Lie Only Pronouncements	Lie Only Promises	Lie Twice
1st-order belief	2.513 (2.608)	-2.715 (3.055)	-7.853*** (2.325)
2nd-order belief	-1.897 (2.506)	3.143 (2.766)	4.086** (1.920)
Female	-0.217 (0.748)	0.922 (0.849)	-0.055 (0.558)
Economics major	-14.581*** (0.00001)	1.092 (1.048)	1.035 (0.794)
Age	0.133* (0.072)	-0.054 (0.103)	-0.074 (0.072)
Sender first	-0.551 (0.774)	0.352 (0.817)	-1.157* (0.599)
Promises first as sender	-0.886 (0.754)	-0.231 (0.742)	-0.248 (0.533)
Alternative treatment	0.452 (0.837)	0.882 (0.800)	-0.785 (0.703)
Constant	-4.756** (2.315)	-2.144 (3.118)	3.235 (2.124)
Akaike Inf. Crit.	250.189	250.189	250.189

Note:

Baseline:

*p<0.1; **p<0.05; ***p<0.01

Lie Never

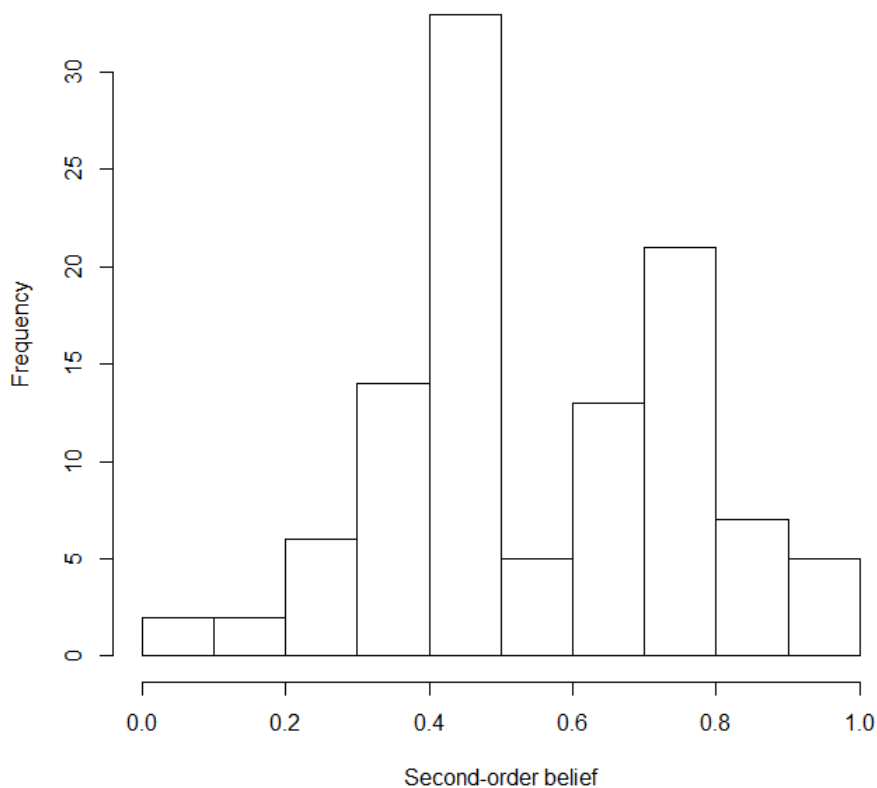


Figure 2.6: Histogram of first-order beliefs

graphic covariates, with the latter table splitting those who lie once by the scenario in which they lie. Experimental variables — namely playing in the Sender role before the Receiver role, playing the Promises scenario first as the Sender, and playing the alternative game — do not have a significant effect on this categorization. For the most part, demographic covariates do not either, but being an economics major reduces the chance of lying only in the pronouncements scenario significantly relative to the baseline of never lying. Most interesting, however, is the dependence of lying categorization on beliefs. All else being equal, there is no significant difference in beliefs between the Senders who never lie and those who lie once. However, those with lower first-order beliefs and higher

second-order beliefs are more likely to lie twice. In other words, the less credulous are more likely to lie twice, and those who believe others are more credulous are also more likely to lie twice.

Given this evidence on beliefs, we classify subjects according to whether their first-order belief was strictly less than their second-order belief or not (i.e. whether or not they were strictly less credulous than they believed others to be). Using a one-sided Fisher’s exact test, we find that people with higher second- than first-order beliefs are more likely to lie twice than those who do not hold such beliefs ($p < 0.001$). (The relevant two-way contingency table is Table 2.5.)

Table 2.5: Two-way contingency table for lying and beliefs

	Lie Once or Less	Lie Both
1st-order Belief Weakly Greater	67	10
1st-order Belief Strictly Less	17	14

Table 2.6 displays a linear regression of Sender behavior on beliefs, demographic covariates, and experimental variables. Each interaction is an observation, and the dependent variable is whether the Sender lied in that interaction. The main variable of interest, an indicator for the Pronouncements scenario, does not have a significant effect on the Sender’s behavior. It also does not have a significant effect when interacted with an indicator for playing the alternative game. As in the previous regressions in this subsection, the primary driver of lying is the first-order belief; those who are less credulous are more likely to lie.

Table 2.6: Regression of Lying on Scenario

	Lie
Pronouncement scenario	0.025 (0.044)
1st-order belief	-0.860*** (0.309)
2nd-order belief	0.468 (0.303)
Female	-0.014 (0.084)
Economics major	0.121 (0.119)
Age	-0.004 (0.008)
Sender first	-0.137 (0.083)
Promises first as sender	-0.052 (0.078)
Alternative treatment	0.007 (0.100)
Alternative treatment * Pronouncement scenario	-0.061 (0.107)
Constant	0.699*** (0.263)
R^2	0.091
Adjusted R^2	0.046

Note:

Standard errors clustered at the subject level.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

2.6 Discussion and Conclusion

From the results, it is evident that the degree of control over the final outcome of an interaction is not a major determinant of lying behavior; broken promises are no more or less likely than false pronouncements. These results stand in contrast with Vanberg's (2008) results, which find that the propensity to follow through on a commitment depends heavily on being responsible for having made said commitment. This suggests that some nuance is required when discussing lying aversion: control over final outcomes does not affect the aversion to lying; however, control over communication affects the aversion to breaking commitments.

In contrast to Sheremeta and Shields (2013), we find beliefs to be a significant factor in determining lying behavior. We find that higher first-order beliefs imply a lower probability of lying, and higher second-order beliefs imply a higher probability of lying. In other words, the less credulous people are and the more they believe others are credulous, the more likely they are to lie. Arguably, this implies that cynics – people who believe others are gullible but not trustworthy – are most likely to lie. Future research should examine the extent to which beliefs and lying behavior are correlated with the “Big Five” personality factors (cf. Goldberg and Saucier, 1996), such as conscientiousness and agreeableness, in order to study this hypothesis further.¹⁴

Previous studies in lying aversion have presented evidence consistent with a two-type model in which there are some people who will lie when it is in their monetary self-

¹⁴Some personality research has shown evidence consistent with a six-factor model, in which honesty/humility is one factor (Ashton et al., 2004). Presumably, lying behavior would be highly correlated with such a factor.

interest and others who will never lie (e.g. Hurkens and Kartik, 2009). The small number of Senders who lie only once (as compared to equilibrium predictions) provides some evidence consistent with this. However, subjects' beliefs paint a slightly more nuanced picture. First-order and second-order beliefs are strongly positively correlated; however, of the subjects who lie twice, more have higher second-order beliefs than first-order beliefs than do not. This indicates that there is heterogeneity of beliefs within classifications of lying and non-lying types.

Overall, we find responsibility for a final outcome not to be a major driver of lying behavior. This suggests that putting spokespeople in supervisory roles or putting supervisors in communicational roles is of limited usefulness; they are no more or less likely to lie about a project under their control. What may reduce lying in organizations, however, is putting less cynical people in charge of communication.

This page intentionally left blank.

Costly Information and Multiattribute Choice

3.1 Introduction

Many of the goods that people purchase have multiple attributes that consumers may care about. For instance, when purchasing a laptop computer, one may be interested in its hard drive space, its processor clock speed, its maximum screen resolution, and its battery life, among other features. Some features of the laptop may draw a consumer's attention more than others. This may be because some features will have more impact on the usefulness of the laptop to the consumer than others. For example, if the consumer travels a lot, then she may care a great deal about the computer's battery life.

However, the consumer's attention may be drawn to specific features not because of the direct impact on the utility she draws from them, but because of aspects of her decision environment. Here are two examples: 1) If the laptops available for sale vary greatly in their maximum screen resolution, then she may want to pay special attention to this attribute, because a laptop with the wrong maximum screen resolution may have a very different maximum screen resolution from what she desires. 2) If one of the laptops for sale has three times more hard drive space than any of the others, then her attention may be drawn to that attribute of that particular laptop, since it "stands out" relative to

the others.

The first example is an example of a *focusing illusion*, modeled by Kőszegi and Szeidl (2013) (henceforth KS), and the second is an example of *salience*, modeled by Bordalo et al. (2013) (henceforth BGS). Both of these papers assume context-dependent distortions on the evaluation of multiattribute options and derive implications therefrom.¹ However, while these distortions have compelling and justifiable properties, a question remains as to whether these distortions reflect an efficient use of cognitive resources, i.e. whether they are “optimal” relative to some maximization problem.

There is reason to believe that people make rational trade-offs when acquiring information about their decision environments. More information can help people make better decisions, but it can also be costly to obtain. Caplin and Dean (2014) document laboratory evidence consistent with such trade-offs in a ball-counting task. There is also empirical evidence that people engage in such trade-offs when making real-life purchases of the kind relevant to the present paper. For instance, De los Santos et al. (2012) document that people only visit a small number of online retailers before making book purchases, despite the fact that there are many online book sellers. This finding could reflect the fact that while searching through a large number of stores could help consumers get better deals, doing so is time consuming and takes a great deal of effort. To cite another example, Chetty et al. (2009) find that consumers do not fully account for sales taxes when making purchasing decisions.

Given that trade-offs between the costs and benefits of acquiring information may be

¹There is a growing body of literature on the evaluation of multiattribute options. See also, for example, Bordalo et al. (2012), Soltani et al. (2012), Cunningham (2013), and Bushong et al. (2016).

a feature of many decision environments, including consumer-relevant ones, it is worth exploring what these models of “rational inattention” imply in a multiattribute choice context. Some work has already been done in this direction. Woodford (2012) has provided an information-theoretic explanation for focusing illusions using a variant of Sims’s (2003) theory of rational inattention, and though it is not technically an optimizing foundation, Landry and Webb (2017) have provided a neuroscientific foundation for the salience model based on how the brain processes visual information. To the extent that the brain processes information efficiently, this could provide a basis for an optimizing framework.

The present paper provides a unified, optimizing, information-theoretic framework under which either focusing illusions or salience-like distortions can present themselves, depending on what information the decision-maker observes. Section 3.2 presents the model. Section 3.3 explains focusing illusions in greater detail and demonstrates how the framework can account for them. Section 3.4 does the same for salience distortions. Section 3.5 concludes. Proofs are appendicized.

3.2 The Model

Suppose the decision-maker (DM) must evaluate I options, or items, ($i = 1, \dots, I$), each with J attributes ($j = 1, \dots, J$). Let $a_{i,j}$ be the value of the j -th attribute of the i -th item, let A be the $I \times J$ matrix of item-attribute pair values, and let $\vec{a}_i := (a_{i,j})_{j=1}^J$ be the vector

of attribute values for item i . The DM's utility from selecting item i is:²

$$u_i := u(\vec{a}_i) := \sum_{j=1}^J a_{i,j} \quad (3.1)$$

The DM has a continuous, non-atomic prior over each item-attribute pair with cumulative distribution function (CDF) $F_{i,j}$ and associated density $f_{i,j}$, with full support on the real line. Assume that each $F_{i,j}$ has a finite second moment. Each item-attribute pair is independently distributed from any other item-attribute pair, and $\forall j, f_{i,j} = f_{k,j}$. Put differently, the same attribute is independently and identically distributed across items, and within an item, each attribute is independently (but not necessarily identically) distributed. Let f denote the joint density of all item-attribute pairs, and let f_i denote the joint density of all attributes within an item.

The DM wishes to learn the value of each u_i by minimizing the the mean squared error (MSE) $\mathbb{E}[(u_i - \hat{u}_i)^2]$, where \hat{u}_i denotes her prediction of u_i .³ Since item-attribute pairs are independent, this is equivalent to minimizing the sum of mean squared errors:

$$\sum_{j=1}^J \mathbb{E}[(a_{i,j} - \hat{a}_{i,j})^2] \quad (3.2)$$

where $\hat{a}_{i,j}$ denotes her prediction of $a_{i,j}$. She does so by choosing probabilistic subjective

²It is a fairly strong assumption that utility is additively separable in attribute values. However, linearity is not a strong assumption, since as long as utility is monotonically increasing in each attribute and additively separable, then one could simply apply a transformation to the underlying attribute values to make utility linear in the transformed values. Moreover, the lack of attribute weights is also not a restrictive assumption, since attribute values can simply be rescaled multiplicatively. For a further discussion of these points, see Subsection 2.1 of Woodford (2012).

³The importance of evaluating each item separately will become clear in Section 3.4 when discussing salience.

representations $r_{i,j}$ of these values, distributed according to the joint density $g_i(\vec{r}_i|\vec{a}_i)$, where $\vec{r} = ((r_{i,j})_{j=1}^J)$. This incurs a cost $C(f_i, g_i)$ given by a multiplicative scaling of the mutual information (cf. Cover and Thomas, 2006) between f and g :

$$C(f_i, g_i) = \kappa \left[- \int_{\vec{a}_i} f_i(\vec{a}_i) \ln(f_i(\vec{a}_i)) d\vec{a}_i + \int_{\vec{r}_i} \int_{\vec{a}_i} g_i(\vec{r}_i|\vec{a}_i) f_i(\vec{a}_i) \ln \left(\frac{g_i(\vec{r}_i|\vec{a}_i) f_i(\vec{a}_i)}{\int_{\vec{b}_i} g_i(\vec{r}_i|\vec{b}_i) f_i(\vec{b}_i) d\vec{b}_i} \right) d\vec{r}_i d\vec{a}_i \right] \quad (3.3)$$

where $\kappa > 0$. Since the item-attribute pairs are independent, this can be rewritten as:

$$C(f_i, g_i) = \kappa \sum_{j=1}^J \left[- \int_{a_{i,j}} f_{i,j}(a_{i,j}) \ln(f_{i,j}(a_{i,j})) da_{i,j} + \int_{r_{i,j}} \int_{a_{i,j}} g_{i,j}(r_{i,j}|a_{i,j}) f_{i,j}(a_{i,j}) \ln \left(\frac{g_{i,j}(r_{i,j}|a_{i,j}) f_{i,j}(a_{i,j})}{\int_{b_{i,j}} g_{i,j}(r_{i,j}|b_{i,j}) f_{i,j}(b_{i,j}) db_{i,j}} \right) dr_{i,j} da_{i,j} \right] \quad (3.4)$$

where $g_{i,j}$ is the density distribution of the probabilistic subjective representation $r_{i,j}$. Mutual information measures the expected reduction in entropy going from the prior distribution to the posterior, and it can be thought of as how much of the uncertainty in the prior is explained by the posterior. It is intuitive that the more one wishes to learn about an attribute value, the more effort is required, and so the greater the cost incurred.

Assume that the cost enters the DM's objective function additively, so that she must

choose g_i to minimize:

$$\mathbb{E}[(u_i - \hat{u}_i)^2] + C(f_i, g_i) \quad (3.5)$$

The MSE is minimized by the selecting the posterior mean for each $\hat{a}_{i,j}$. In other words, define:

$$h_{i,j}(a_{i,j}, r_{i,j}) := \frac{g_{i,j}(r_{i,j}|a_{i,j})f_{i,j}(a_{i,j})}{\int_{b_{i,j}} g_{i,j}(r_{i,j}|b_{i,j})f_{i,j}(b_{i,j})db_{i,j}} \quad (3.6)$$

i.e. the posterior density of the value of $a_{i,j}$ given $r_{i,j}$. The MSE is minimized by selecting the mean value of this distribution as $\hat{a}_{i,j}$. The MSE of u_i is then minimized by selecting:

$$\hat{u}_i = \sum_{j=1}^J \hat{a}_{i,j} \quad (3.7)$$

Since the posterior mean minimizes the MSE, the optimal MSE is actually the posterior variance of \hat{u}_i , which by independence across attributes, is the sum of the posterior variances of $\hat{a}_{i,j}$.

Before proceeding further, it is necessary to have a quantitative definition for the amount of attention paid to an attribute of an item. Define:

$$\alpha_{i,j} := \frac{\sigma_{f_{i,j}}^2 - \sigma_{h_{i,j}}^2}{\sigma_{f_{i,j}}^2} \quad (3.8)$$

to be the *amount of attention* paid to attribute j of item i , where σ_q^2 is the variance of the distribution with density q . In other words, I define the amount of attention paid to an

item-attribute pair to be proportional reduction in variance from observing its subjective representation. Since $\sigma_{h_{i,j}}^2 \in [0, \sigma_{f_{i,j}}^2]$, attention takes values between 0 and 1, where 0 implies no attention was paid to an attribute – the posterior variance is the same as the prior variance – and 1 implies full attention was paid – the posterior variance is 0, implying a 1-to-1 deterministic mapping between the subjective representation and the true attribute value. It is easy to see that $\alpha_{i,j}$ decreases in the posterior variance (i.e. increases in the posterior precision); the more the DM decreases the spread of her belief, the more attention she pays.

3.3 Focusing

3.3.1 The KS Model

I begin this section by presenting a brief overview of KS’s model of focusing effects. They assume, as this paper does, that the DM’s utility for option i is given by:

$$u(\vec{a}_i) = \sum_{j=1}^J a_{i,j} \quad (3.9)$$

However, she evaluates options according to the following criterion:

$$u(\vec{a}_i, A) = \sum_{j=1}^J \gamma_j(\Delta_j(A)) a_{i,j} \quad (3.10)$$

where γ_j is increasing in its argument, and $\Delta_j(A) := \max_j a_{i,j} - \min_j a_{i,j}$ is the range of attribute j . Therefore, when evaluating options, the DM puts greater emphasis on

attributes with a wider range, i.e. those that vary more.

3.3.2 Focusing in a Probabilistic Model

Since my model is probabilistic, instead of using the range of an attribute within the choice set as a measure of how much an attribute varies, I use the variance.⁴ In this subsection, I show that the attention paid to an attribute of an item is weakly increasing in the prior variance.

Proposition 20. *Let Φ and ϕ denote the CDF and density of the standard normal distribution. Then either $h_{i,j} = f_{i,j}$ or $h_{i,j}(a_{i,j}|r_{i,j}) = \frac{1}{\sigma} \phi\left(\frac{a - \mu_{h_{i,j}}}{\sigma}\right)$, where $\mu_{h_{i,j}}$ is the posterior mean and $\sigma = \sqrt{\frac{\kappa}{2}}$, whichever has lower variance.*

A version of this result was stated but not proven by Sims (2003). It may seem remarkable that the optimal posterior is normally distributed regardless of the prior (provided the prior has full support on the real line), but this is not surprising in light of Caplin and Dean’s (2013) “locally invariant posteriors” result, which shows in a discrete setting that if a set of posteriors is optimal for some prior, if it is feasible for another prior, then it is optimal for that prior. Proposition 20 can be seen as a continuous version of that result for a specific utility function.

Corollary 1. *Attention paid to an attribute increases in the prior variance.*

This is easy to see from the fact that $\sigma_{h_{i,j}} = \min\left\{\sqrt{\frac{\kappa}{2}}, \sigma_{f_{i,j}}\right\}$. This implies that either posterior variance is fixed, given prior variance, or no attention is paid if the prior variance

⁴The range of the attribute distribution’s support is not an appropriate measure, because it is infinite.

is small enough. Therefore, if attention is non-zero, then $\alpha_{i,j} = 1 - \frac{\sqrt{\kappa}}{\sigma_{f_{i,j}}^2 \sqrt{2}}$, which is clearly increasing in the prior variance.

If the priors are normally distributed, then there is a closed form for the posterior mean of h .

Proposition 21. *Suppose $a_{i,j} \sim N(\mu_j, \sigma_j^2)$ and $\sigma_j > \sqrt{\frac{\kappa}{2}}$.⁵ Then the optimal prediction for each attribute of the item i is $\hat{a}_{i,j} = \frac{2\sigma_j^2 - \kappa}{2\sigma_j^2} r_{i,j} + \frac{\kappa}{2\sigma_j^2} \mu_j$.*

When priors are normally distributed, the optimal prediction is a weighted average of the subjective representation $r_{i,j}$ and prior mean μ_j . The weight on the former is increasing in the prior variance σ_j^2 , and the weight on the latter is decreasing in the prior variance.

Suppose, for the sake of illustration, that the prior means are all zero, so that by (3.7):

$$\hat{u}_i = \sum_{j=1}^J \tilde{\gamma}(\sigma_j^2) r_{i,j} \quad (3.11)$$

where $\tilde{\gamma}(x) := \frac{2x - \kappa}{2x}$. The weights $\tilde{\gamma}(\sigma_j^2)$ in (3.11) have a similar property to the weights $\gamma_j(\Delta_j(A))$ in (3.10). Those in the former are increasing in the prior variance of each attribute, while those in the latter are increasing in the range of each attribute. In this instance, in my model, not only is the attention paid to each attribute increasing in the prior variance, but a focusing effect similar to the one presented by KS also obtains.

⁵This inequality ensures that the optimal posterior is not simply the prior.

3.4 Saliency

3.4.1 The BGS Model

I begin this section by presenting a brief overview of BGS's model of saliency distortions.

Suppose that the DM's utility for option i is given by:⁶

$$u(\vec{a}_i) = \sum_{j=1}^J a_{i,j} \quad (3.12)$$

However, she evaluates options according to the following criterion:

$$u(\vec{a}_i, A) = \sum_{j=1}^J \omega_{i,j} a_{i,j} \quad (3.13)$$

The weights $\omega_{i,j}$ reflect how much $a_{i,j}$ differs from its mean value in the choice set \bar{a}_j .

They are computed as follows. Consider a symmetric, continuous *saliency function* $s(\cdot, \cdot)$

with the following two properties:

1. *Ordering.* Let $\epsilon, \eta \geq 0$ with at least one strictly positive. If $a_{i,j} > \bar{a}_j$, then $s(a_{i,j} + \epsilon, \bar{a}_j - \eta) > s(a_{i,j}, \bar{a}_j)$. If $a_{i,j} < \bar{a}_j$, then $s(a_{i,j} - \epsilon, \bar{a}_j + \eta) > s(a_{i,j}, \bar{a}_j)$.
2. *Diminishing Sensitivity.* Let $a_{i,j}, \bar{a}_j \geq 0$ and $\epsilon > 0$.⁷ Then, $s(a_{i,j} + \epsilon, \bar{a}_j + \epsilon) > s(a_{i,j}, \bar{a}_j)$.

⁶Given that BGS's model is applied specifically to consumer choice, it also includes a price component that is subtracted away from the sum of attribute values. For the sake of compatibility with the present paper, I ignore it here. Bordalo et al. (2012) presents a version of their model applied to choice under risk that does not include a price component, but does include utility weights.

⁷BGS only consider positive attribute values. They allow for negative attribute values in Bordalo et al. (2012) when analyzing choice under risk and impose an additional requirement on the saliency function called *reflection*. My model will allow for negative attribute values.

The salience function measures how much an attribute of an item “stands out” relative to its mean in the choice set. The ordering property implies that the further away an attribute value is from its mean, the more salient it is. The diminishing sensitivity property implies that increasing attribute values uniformly in the choice set diminishes their salience.

After computing the salience $s(a_{i,j}, \bar{a}_j)$ for each item-attribute pair, the attributes of each item are ranked by salience and given a rank $\rho_{i,j}$, with the highest receiving a rank of 1 and the lowest receiving a rank of J . The weights $\omega_{i,j}$ are then computed as:

$$\omega_{i,j} = \frac{\delta^{r_{i,j}}}{\sum_{\ell=1}^J \delta^{r_{i,\ell}}} \quad (3.14)$$

where $\delta \in (0, 1)$.

In this model, the attributes that “stand out” compared to other attributes for a given item are given more weight when evaluating that item. This reflects the idea that attributes of an item that are irregular or atypical in some way may draw a consumer’s attention more than others.

3.4.2 Salience in a Probabilistic Model

My model can account for some features of the BGS model, but will have a few important differences. It will allow for multiple attributes to be “salient” for a given item. In fact, in the example I present, salience will be a binary property; an item-attribute pair is either salient or not.

Before proceeding further, a discussion of the crucial assumption that will lead to

salience effects is in order. It has long been observed in perceptual psychology that comparative judgements are easier to make than absolute ones (e.g. Fernberger, 1931; Miller, 1956). For example, in an experiment by Fernberger (1931), subjects did worse at classifying a series of weights as “light,” “intermediate,” or “heavy” than they did at classifying those weights as “lighter,” “equal to,” or “heavier” than a reference weight.

In this section, we make the extreme assumption that comparative judgements can be made for free, but absolute judgements are costly. Suppose that the DM can observe the ranking of each item-attribute pair across items at zero cost. In other words, before making her information-acquisition decision, the DM observes, for each j , $a_{d_j^{-1}(1),j} < \dots < a_{d_j^{-1}(I),j}$, where d_j is a permutation that ranks items from smallest to largest, from 1 to I , on attribute j .⁸ Let $\tilde{f}_{i,j}(a_{i,j}) := f_{i,j} \left(a_{i,j} \mid a_{d_j^{-1}(1),j} < \dots < a_{d_j^{-1}(I),j} \right)$ be the prior density of attribute value (i, j) conditional on the ranking.

It can be shown that attribute values are not necessarily independent across items conditional on this ranking information. This is where the assumption that the evaluation of each item is a separate decision problem becomes crucial. This means that the DM cannot use the information from the evaluation of one item in the evaluation of another. Moreover, the attribute values of a given item are still independent across attributes, conditional on rankings. These facts allows me to apply Proposition 20 and Corollary 1 in this context.

For the sake of illustration, I present an example. Suppose that $I = 3$, so that there are only 3 items and that the attribute values are independently distributed according to

⁸Since $f_{i,j}$ are independent and atomless, two items sharing the same value for a given attribute is a probability-zero event, and so d_j is almost surely well-defined.

$N(\mu_j, \tau_j^2)$, with $\tau_j > \frac{\sqrt{\kappa}}{2}$, before the ranking is observed.

Proposition 22. *In this context, with 3 items, more attention is paid to the highest- and lowest-ranked attribute values for a given attribute than to the middle-ranked one. If it is further assumed that $\tau_j = \tau$ for all j , then for a given item, more attention is paid to an attribute value ranked highest or lowest than to an attribute value that is middle-ranked.*

Let $V_{E,j}$ be the variance of an extreme-ranked attribute value for attribute j ,⁹ and let $V_{M,j}$ be the variance of a middle-ranked attribute value for attribute j . It can be shown that $V_{E,j} = \tau^2 \left(1 - \frac{9-2\sqrt{3}}{4\pi}\right)$, $V_{M,j} = \tau^2 \left(1 - \frac{\sqrt{3}}{\pi}\right)$, and $V_{M,j} < V_{E,j}$. By Corollary 1, this implies that more attention should be paid to extreme attribute values than middle-ranked ones. Furthermore, if the attributes are *a priori* identically distributed, then $V_{E,j}$ and $V_{M,j}$ are fixed across j , and the second part of Proposition 22 obtains.

As mentioned earlier, in contrast to the BGS model, this framework allows for multiple attributes to be equally salient for each item. In fact, in the three-item case, attributes are either salient or not. Attribute values that are highest- or lowest-ranked for a given item are salient, and those that are middle-ranked are not. This presents an advantage over the BGS model in that it allows for different items in a choice set to draw different amounts of attention from the DM. Items with a lot of extreme attribute values will draw the DM's attention more than items with mostly moderate attribute values.

To return to the laptop example of the introduction, a budget laptop that has the lowest value on many of its attributes or an enthusiast laptop that has the highest value on many of its attributes may draw a consumer's attention more than a mid-range laptop

⁹The variances of the highest- and lowest-ranked values for a given attribute are the same, by the symmetry of the normal distribution.

that is not particularly remarkable on most of its dimensions. However, if, for instance, the mid-range laptop has the best battery life, then more attention will be drawn that attribute of that laptop as compared to that attribute of other laptops.

3.5 Conclusion

This paper contributes to the growing body of literature on context effects and multi-attribute choice by providing an information-theoretic foundation for some of the phenomena the literature documents, namely focusing illusions (KS) and salience distortions (BGS).

Several potential avenues for future research present themselves. Most immediate is determining whether the salience effects obtained in this information-theoretic framework generalize to larger choice sets and other priors; the documented effects may be in part due to the assumption of normality.

Secondly, it is worth noting that Bushong et al. (2016) have analyzed a model with the opposite assumption of KS. In their “relative thinking” model, the weighting function γ_j is a *decreasing* function of an attribute’s range. It would be worth exploring whether such a distortion can be rationalized in this information-theoretic framework, using perhaps a different cost function or different assumptions on free information.

Finally, and relatedly, there is evidence to suggest that mutual information does not accurately reflect the costs of information acquisition (Caplin and Dean, 2013). Different functional forms for the cost function may lead to different types of distortions in multi-attribute choice and could potentially explain “relative thinking,” among other phenomena.

Bibliography

- Hunt Allcott and Dmitry Taubinsky. Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–38, 2015.
- Donald W. K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–856, 1993.
- Donald W. K. Andrews and Werner Ploberger. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62(6):1383–1414, 1994.
- Amos Arieli, Yaniv Ben-Ami, and Ariel Rubinstein. Tracking decision makers under uncertainty. *American Economic Journal: Microeconomics*, 3(4):68–76, 2011.
- Michael C. Ashton, Kibeom Lee, and Lewis R. Goldberg. A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*, 87(5):707–721, 2004.
- Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, 1998.
- Gadi Barlevy and Pietro Veronesi. Information acquisition in financial markets. *Review of Economic Studies*, 67(1):79–90, 2000.
- Pierpaolo Battigalli and Martin Dufwenberg. Guilt in games. *American Economic Review: AEA Papers & Proceedings*, 97(2):170–176, May 2007.
- Pierpaolo Battigalli and Martin Dufwenberg. Dynamic psychological games. *Journal of Economic Theory*, 144(1):1–35, January 2009.
- Pierpaolo Battigalli, Gary Charness, and Martin Dufwenberg. Deception: The role of guilt. *Journal of Economic Behavior & Organization*, 93:227–232, September 2013.
- Patrick Billingsley. *Probability and Measure*. Wiley, 1995.
- David Blackwell. Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24(2):265–272, 1953.
- Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285, 2012.

- Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, 2013.
- Benjamin Bushong, Matthew Rabin, and Joshua Schwartzstein. A model of relative thinking. *Working paper*, 2016.
- Hongbin Cai and Joseph Tao-Yi Wang. Overcommunication in strategic information transmission games. *Games and Economic Behavior*, 56(1):7–36, July 2006.
- A. Colin Cameron and Pravin K. Trivedi. *Microeconometrics: Methods and Applications*. Cambridge University Press, 2005.
- Andrew Caplin and Mark Dean. The behavioral implications of rational inattention with Shannon entropy. *Working paper*, 2013.
- Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *Working paper*, 2014.
- Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.
- Andrew Caplin and Daniel Martin. Defaults and attention: The drop-out effect. *Working paper*, 2013.
- Andrew Caplin and Daniel Martin. A testable theory of imperfect perception. *The Economic Journal*, 125(582):184–202, 2015.
- Gary Charness and Martin Dufwenberg. Promises and partnership. *Econometrica*, 74(6):1579–1601, November 2006.
- Gary Charness and Martin Dufwenberg. Bare promises: An experiment. *Economics Letters*, 107(2):281–283, May 2010.
- Gary Charness and Martin Dufwenberg. Participation. *American Economic Review*, 101(4):1211–1237, June 2011.
- Anton Cheremukhin, Anna Popova, and Antonella Tutino. A theory of discrete choice with information costs. *Journal of Economic Behavior & Organization*, 113:34–50, 2015.
- Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177, 2009.
- Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2006.
- Vincent P. Crawford and Joel Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, November 1982.
- Tom Cunningham. Comparisons and choice. *Working paper*, 2013.
- R. Davidson and J.G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2004.

- Babur De los Santos, Ali Hortaçsu, and Matthijs R. Wildenbeest. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6):2955–2980, 2012.
- Stefano DellaVigna. Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2):315–372, 2009.
- Holger Dette, Natalie Neumeyer, and Kay F. Pilz. A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, 12(3):469–490, 2006.
- Adele Diederich. Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, 41(3):260–274, 1997.
- E. Doveh, A. Shapiro, and P.D. Feigin. Testing of monotonicity in parametric regression models. *Journal of Statistical Planning and Inference*, 107(1&2):289–306, 2002.
- Florian Ederer and Alexander Stremitzer. Promises and expectations. *Working paper*, 2015.
- Sanjiv Erat and Uri Gneezy. White lies. *Management Science*, 58(4):723–733, April 2012.
- Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, August 1999.
- Samuel W. Fernberger. On absolute and relative judgments in lifted weight experiments. *The American Journal of Psychology*, 43(4):560–578, 1931.
- Urs Fischbacher. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, June 2007.
- Urs Fischbacher and Franziska Föllmi-Heusi. Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547, June 2013.
- John P. Frisby and James V. Stone. *Seeing: The Computational Approach to Biological Vision*. MIT Press, 2010.
- Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Stochastic choice and optimal sequential sampling. *Working paper*, 2015.
- Xavier Gabaix, David Laibson, Guillermo Moloche, and Stephen Weinberg. Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4):1043–1068, 2006.
- John Geanakoplos, David Pearce, and Ennio Stacchetti. Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1):60–79, March 1989.
- Rajna Gibson, Carmen Tanner, and Alexander F. Wagner. Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*, 103(1):532–548, February 2013.

- Uri Gneezy. Deception: The role of consequences. *American Economic Review*, 95(1): 384–394, March 2005.
- Lewis R. Goldberg and Gerard Saucier. Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, 10(1):61–77, 1996.
- Jerry R. Green and Nancy L. Stokey. A two-person game of information transmission. *Journal of Economic Theory*, 135(1):90–104, July 2007.
- Adam Eric Greenberg, Paul Smeets, and Lilia Zhurakhovska. Lying and shame aversion. *Working paper*, March 2015.
- Ben Greiner. Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125, 2015.
- Sanford J. Grossman and Joseph E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.
- J.A. Hartigan and P.M. Hartigan. The dip test of unimodality. *The Annals of Statistics*, 13(1):70–84, 1985.
- Christian Hellwig, Sebastian Kohls, and Laura Veldkamp. Information choice technologies. *American Economic Review*, 102(3):35–40, 2012.
- Sjaak Hurkens and Navin Kartik. Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, 12(2):180–192, June 2009.
- Huseyn Ismayilov and Jan Potters. Why do promises affect trustworthiness, or do they? *Experimental Economics*, pages 1–12, 2015.
- Daniel Kahneman. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5):1449–1475, December 2003.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion: Web appendix. *American Economic Review*, 2010. URL https://assets.aeaweb.org/assets/production/articles-attachments/aer/data/oct2011/20090933_app.pdf.
- Navin Kartik. Strategic communication with lying costs. *Review of Economic Studies*, 76(4):1359–1395, October 2009.
- Navin Kartik, Marco Ottaviani, and Francesco Squintani. Credulity, lies, and costly talk. *Journal of Economic Theory*, 134(1):93–116, May 2007.
- Deborah A. Kashy and Bella M. DePaulo. Who lies? *Journal of Personality and Social Psychology*, 70(5):1037–1051, May 1996.
- E.L. Kaufman, M.W. Lord, T.W. Reese, and J. Volkman. The discrimination of visual number. *American Journal of Psychology*, 62(4):498–525, 1949.
- Botond Köszegi and Adam Szeidl. A model of focusing in economic choice. *The Quarterly Journal of Economics*, 128(1):53–104, 2013.

- Ian Krajbich, Carrie Armel, and Antonio Rangel. Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10):1292–1298, 2010.
- Peter Landry and Ryan Webb. Pairwise attribute normalization. *Working paper*, 2017.
- Moshe Leshno and Yishay Spector. An elementary proof of Blackwell’s theorem. *Mathematical Social Sciences*, 25(1):95–98, 1992.
- Tobias Lundquist, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1–2):81–92, May 2009.
- Charles R. MacCluer. *Calculus of Variations: Mechanics, Control and Other Applications*. Dover Publications, 2013.
- Filip Matějka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.
- Nina Mazar, On Amir, and Dan Ariely. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6):633–644, December 2008.
- George A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97, 1956.
- Stephen Morris and Ming Yang. Coordination and the relative cost of distinguishing nearby states. *Working paper*, 2016.
- Giuseppe Moscarini and Lones Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.
- Rosemarie Nagel. Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326, December 1995.
- Maxim L. Pinkovskiy. Rational inattention and choice under risk: Explaining violations of expected utility through a Shannon entropy formulation of the costs of rationality. *Atlantic Economic Journal*, 37(1):99–112, 2009.
- Jack Porter and Ping Yu. Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics*, 189(1):132–147, 2015.
- Roger Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978.
- Roger Ratcliff and Philip L. Smith. A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2):333–367, 2004.
- Stephen Mark Rosenbaum, Stephan Billinger, and Nils Stieglitz. Let’s be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45:181–196, 2014.
- Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, 1974.

- I.J. Saltzman and W.R. Garner. Reaction time as a measure of span of attention. *Journal of Psychology*, 25(2):227–241, 1948.
- Santiago Sánchez-Pagés and Marc Vorsatz. An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior*, 61(1):86–112, October 2007.
- J.F. Schouten and J.A.M. Bekker. Reaction time and accuracy. *Acta Psychologica*, 27:143–153, 1967.
- Marilyn L. Shaw and Peter Shaw. Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, 3(2): 201–211, 1977.
- Roman M. Sheremeta and Timothy W. Shields. Do liars believe? Beliefs and other-regarding preferences in sender-receiver games. *Journal of Economic Behavior & Organization*, 94:268–277, October 2013.
- Christopher A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.
- Christopher A Sims. Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163, 2006.
- Philip L. Smith. Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, 44(3):408–463, 2000.
- Alireza Soltani, Benedetto De Martino, and Colin Camerer. A range-normalization model of context-dependent choice: A new model and evidence. *PLoS Computational Biology*, 8(7):1–15, 2012.
- Keith E. Stanovich and Richard F. West. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23(5):645–665, 2000.
- Knut Sydsæter, Peter Hammond, Atle Seierstad, and Arne Strøm. *Further Mathematics for Economic Analysis*. Financial Times Prentice Hall, 2008.
- Steven Tadelis. The power of shame and the rationality of trust. *Working paper*, March 2011.
- Satohiro Tajima, Jan Drugowitsch, and Alexandre Pouget. Optimal policy for value-based decision-making. *Nature Communications*, 7:1–12, 2016.
- Stijn Van Nieuwerburgh and Laura Veldkamp. Information acquisition and under-diversification. *Review of Economic Studies*, 77(2):779–805, 2010.
- Christoph Vanberg. Why do people keep their promises? An experimental test of two explanations. *Econometrica*, 76(6):1467–1480, November 2008.
- Robert E. Verrecchia. Information acquisition in a noisy rational expectations economy. *Econometrica*, pages 1415–1430, 1982.

Joseph Tao-Yi Wang, Michael Spezio, and Colin F. Camerer. Pinocchio's pupil: Using eye-tracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review*, 100(3):984–1007, June 2010.

Michael Woodford. Inattentive valuation and reference-dependent choice. *Working paper*, 2012.

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

Ming Yang. Coordination with flexible information acquisition. *Journal of Economic Theory*, 158, Part B:721–738, 2015.

This page intentionally left blank.

Appendices

Appendix A: Chapter 1 Proofs

Proof of Proposition 1

Proof. The ‘if’ direction is obvious, since our model generalizes CD15 with finite action sets.

The ‘only if’ direction can be seen as follows. Suppose $\{(U_i, \theta_i, a_i)\}$ can be rationalized by (1.2) with some cost function $C(\pi, Q)$. Since there are finitely many decision problems, C is pinned down for a finite set of points (i.e. a closed set), and so by the Tietze extension theorem (cf. Rudin, 1974, pg. 422), it may be assumed continuous. Define \mathcal{Q}_{γ_Q} to be set of information structures that induce posterior γ_Q , and define $\tilde{C}(\pi, Q) := \min_{R \in \mathcal{Q}_{\gamma_Q}} C(\pi, R)$,

assuming it is well-defined. γ_Q always induces the same maximum gross payoff, no matter which information structure in \mathcal{Q}_{γ_Q} generated it. Therefore, since the DM is a payoff maximizer, for each distribution of posteriors she generates, she will always select the lowest-cost method of doing so. This implies that behavior that can be rationalized by C can also be rationalized by \tilde{C} , which obviously satisfies cost equivalence.

Now we must verify that \tilde{C} is actually well-defined. Let $b : \mathcal{Q} \rightarrow \Delta(\Delta(\Theta))$ be the function that maps an information structure to the distribution of posteriors it induces. First, we must show that b is continuous when $\Delta(\Delta(\Theta))$ is equipped with the weak-* topology, i.e. the topology of weak convergence of measure.

By Bayes’ rule, $\text{Supp}(b(Q)) = \left\{ \left(\frac{\pi_s q_{s,k}}{\sum_{l=1}^n \pi_l q_{l,k}} \right)_{s=1}^n \mid k \in \{1, \dots, |M|\}, \sum_{l=1}^n \pi_l q_{l,k} > 0 \right\}$, and each element $\zeta \in \text{Supp}(b(Q))$ is induced with probability $\sum_{k \in Q^\zeta} \sum_{l=1}^n \pi_l q_{l,k}$, where Q^ζ is the set of columns of Q that generate the posterior ζ .

Consider a sequence of information structures $Q_1, Q_2, \dots \in \mathcal{Q}$ converging to Q . We must show that $\lim_{j \rightarrow \infty} b(Q_j) = b(Q)$ (in the sense of weak convergence of measure). By Theorem 25.8 of Billingsley (1995), this is equivalent to showing that $\lim_{j \rightarrow \infty} b(Q_j)(X) = b(Q)(X)$ for all continuity sets X in the Borel σ -algebra of $\Delta(\Theta)$.¹⁰

Since X is a continuity set, $\partial X \cap \text{Supp}(b(Q)) = \emptyset$. There are two cases. Either $X \cap \text{Supp}(b(Q)) = \emptyset$ or $\text{int}(X) \cap \text{Supp}(b(Q)) \neq \emptyset$.

Case 1: $X \cap \text{Supp}(b(Q)) = \emptyset$. If $\exists J \in \mathbb{N}$ such that $b(Q_j)(X) = 0 \forall j > J$, then clearly $\lim_{j \rightarrow \infty} b(Q_j)(X) = b(Q)(X) = 0$. If not, then $\forall J \in \mathbb{N}$, $\exists j > J$ such that $X \cap \text{Supp}(b(Q_j)) \neq \emptyset$. Suppose, for a contradiction, that $\lim_{j \rightarrow \infty} b(Q_j)(X) \neq 0$. Then

¹⁰A continuity set is a set X whose boundary ∂X has measure zero.

$\exists \varepsilon > 0$ such that $\forall J \in \mathbb{N}$, $\exists j > J$ such that $b(Q_j)(X) > \varepsilon$. Therefore, there must exist a subsequence Q_{j_h} such that $\left(\left(\frac{\pi_s q_{s,k}}{\sum_{l=1}^n \pi_l q_{l,k}} \right)_{s=1}^n \right)_{j_h}$ converges in $\text{cl}(X)$ for some k .¹¹ If it converges to a point in $\text{int}(X)$, then this contradicts the fact that $b(Q)(X) = 0$. If it converges to a point in ∂X , then $b(Q)(\partial X) > 0$, contradicting the fact that X is a continuity set. Thus, $\lim_{j \rightarrow \infty} b(Q_j)(X) = b(Q)(X)$.

Case 2: $\text{int}(X) \cap \text{Supp}(b(Q)) \neq \emptyset$. Note that since (Q_j) is a convergent sequence, each entry of the matrices in (Q_j) also defines a convergent sequence. Then each $((z_k)_j) := ((\sum_{l=1}^n \pi_l q_{l,k})_j)$ is a convergent sequence with limit z_k , and each $((y_k)_j) := \left(\left(\left(\frac{\pi_s q_{s,k}}{\sum_{l=1}^n \pi_l q_{l,k}} \right)_{s=1}^n \right)_{j_h} \right)$ either converges to some limit y_k (for $z_k > 0$) or else has an undefined limit (when $z_k = 0$).¹² Since they are continuous functions of the entries of π and (Q_j) , and because (Q_j) is convergent, $y_k = \left(\frac{\pi_s q_{s,k}}{\sum_{l=1}^n \pi_l q_{l,k}} \right)$ (when it exists) and $z_k = \sum_{l=1}^n \pi_l q_{l,k}$, where the entries $q_{l,k}$ are taken from Q . Consider the set $K \subseteq \{1, \dots, M\}$ such that $\{((y_k)_j) | k \in K\}$ is the collection of sequences that converge to points in $\text{int}(X)$. Then, because $\text{int}(X)$ is open, $\forall \varepsilon > 0$ and $\forall k \in K$, $\exists N_k$ such that $\forall j > N_k$, $(y_k)_j \in \text{int}(X)$. Let $\bar{N} = \max_{k \in K} N_k$. Then $\forall j > \bar{N}$, $b(Q_j)(X) \geq (\sum_{k \in K} (\sum_{l=1}^n \pi_l q_{l,k}))_j$.

We now show that $b(Q_j)(X) - \sum_{k \in K} (\sum_{l=1}^n \pi_l q_{l,k})_j$ goes to zero as j grows large. Suppose there does not exist $J \in \mathbb{N}$ such that this sequence has the value 0 $\forall j > J$. Then there must exist a subsequence (Q_{j_h}) such that $b(Q_{j_h})(X) - \sum_{k \in K} (\sum_{l=1}^n \pi_l q_{l,k})_{j_h} > 0$ for all j_h . Then for each j_h , there is some $k' \notin K$ such that $(y_{k'})_{j_h} \in \text{Supp}(b(Q_{j_h}))$. Because $|M|$ is finite, we may assume that this k' is fixed. If $((y_{k'})_{j_h})$ is convergent, it must converge in $\text{cl}(X)$. If $y_{k'} \in \text{int}(X)$, then this contradicts the fact that $k' \notin K$. If $y_{k'} \in \partial X$, then this contradicts the fact that X is a continuity set. If $((y_{k'})_{j_h})$ has no defined limit, then $z_{k'} = 0$. Therefore, $\lim_{h \rightarrow \infty} [b(Q_{j_h})(X) - \sum_{k \in K} (\sum_{l=1}^n \pi_l q_{l,k})_{j_h}] = 0$.

This establishes the continuity of b . Therefore, for a given γ with finite support in $\Delta(\Delta(\Theta))$, $b^{-1}(\{\gamma\})$ is closed (since singletons are closed). Because $b^{-1}(\{\gamma\}) \subseteq \mathcal{Q}$ and \mathcal{Q} is a bounded subset of $\mathbb{R}_n \times \mathbb{R}_{|M|}$, then by the Heine-Borel theorem, $b^{-1}(\{\gamma\})$ is compact.

In particular \mathcal{Q}_{γ_Q} is compact, and since C is continuous (fixing π), by the Weierstrass theorem, it attains its minimum on \mathcal{Q}_{γ_Q} . Therefore, \tilde{C} is well-defined. This concludes the proof. \square

¹¹We can take k fixed here because even if we construct a subsequence where the sequence of posteriors is constructed by different columns of Q_{j_h} for different sequence elements, we can merely take a subsequence of that subsequence, but with k fixed.

¹²It is possible that $(y_k)_{j'}$ maybe be undefined for some k and j' . This occurs when $(z_k)_{j'} = 0$. If there are finitely many such j' , then we can simply consider a sequence (Q_j) with these j' removed. If there are infinitely many such j' , then $(z_k)_j$ must converge to zero. Therefore, WLOG, either (Q_j) is such that $(z_k)_j \neq 0 \forall j, k$ and possibly converges to zero, or $(z_k)_j$ definitely converges to zero.

Proof of Proposition 3

Proof. First we show that we may assume a fixed decision matrix \bar{D} such that:

$$\bar{D} := \begin{bmatrix} I_{|A|} \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \quad (3.15)$$

where $I_{|A|}$ is the identity matrix of dimension $|A| \times |A|$ and the lower block of \bar{D} has dimension $(|M| - |A|) \times |A|$.

Given $Q \in \mathcal{Q}$, select a decision matrix D such that the gross payoff $\text{tr}(QDU\Pi)$ is maximized. Since the gross payoff is linear in the entries of D , its entries may be chosen so that all of its nonzero entries are 1, WLOG. Consider the mapping $\delta : \{1, \dots, |M|\} \rightarrow \{1, \dots, |A|\}$ so that $\delta(j)$ is the 1 entry in the j -th row of D . The signal (and associated posterior) induced by the j -th column of Q result in action $a_{\delta(j)}$ being taken.

Now consider Q' constructed from Q and δ in the following manner. For $k \in \{1, \dots, |A|\}$, the k -th column of Q' is $\sum_{j \in \delta^{-1}(k)} q_{\bullet, j}$, where $q_{\bullet, j}$ is the j -th column of Q and empty sums are taken to be columns of zeroes. For $k > |A|$, the columns of Q' are all zeroes.

(Q, D) and (Q', \bar{D}) imply the same $\Pr(a|\theta)$ for each $a \in A, \theta \in \Theta$, since the (i, j) -th entry of $Q'\bar{D}$ is $q'_{i, j} = \sum_{k \in \delta^{-1}(j)} q_{i, k}$. For $j \leq |A|$, $q'_{i, j} = \Pr(a_j|\theta_i)$, since under \bar{D} , action j is taken only when signal j is received. Furthermore, for $j \leq |A|$, $\sum_{k \in \delta^{-1}(j)} q_{i, k} = \Pr(a_j|\theta_i)$, since under D , action j is taken if and only if a signal in $\delta^{-1}(j)$ is received. Thus, $\text{tr}(QDU\Pi) = \text{tr}(Q'\bar{D}U\Pi)$, i.e. the gross payoffs are the same under either (Q, D) or (Q', \bar{D}) .

Consider the $|M| \times |M|$ matrix P with entries $p_{i, j}$ such that $p_{i, \delta(i)} = 1$ for each i and all other entries are 0. P takes the i -th column of a matrix and shuffles it to the $\delta(i)$ -th column. Thus, $Q' = QP$. P is right-stochastic since each of its rows has a single 1 entry and zeroes for the rest of its entries. Therefore, by monotonicity of information, $C(\pi, Q') \leq C(\pi, Q)$. Note that Q' uses as many signals as there are posteriors in its support, with each posterior inducing a different action. Therefore, given π , there is a one-to-one correspondence between a column of Q' and the posterior it induces along with its associated probability. Therefore, the only other matrices that induce the same distribution of posteriors as Q' with the same number of signals are permutations of the columns of Q' . By cost symmetry, each of these information structures has the same cost as Q' . Furthermore, as we showed above, given \bar{D} , Q' represents the probability of taking each action in each state. Therefore, the only information structures that induce the same distribution of actions given states at the same cost as Q' are permutations of Q' , and there are no information structures that do so at lower cost.

Thus, WLOG, we can fix \bar{D} as the decision matrix and simply consider the problem:

$$\max_{Q \in \mathcal{Q}} \text{tr}(Q\bar{D}U\Pi) - C(\pi, Q) \quad (3.16)$$

Denote the maximand in (3.16) by $F(Q)$. Since $F(Q)$ is continuous in Q and U , and \mathcal{Q} is compact, by the maximum theorem, the optimal choice of information structure for each payoff matrix, $Q^*(U)$, is upper hemicontinuous in U .

Since the first term of $F(Q)$ is linear and the second is almost strictly convex, it inherits its convexity properties from the second term. In other words, $F(Q)$ is almost strictly concave, with almost strict concavity defined analogously to almost strict convexity. For each U , either $Q^*(U)$ is unique or it is multivalued. Suppose it is multivalued, and $Q_1^*, Q_2^* \in Q^*(U)$. Then $F(Q_1^*) = F(Q_2^*)$. If Q_1^* and Q_2^* induce different distributions of posteriors, then $\forall \lambda \in (0, 1)$, $F(\lambda Q_1^* + (1 - \lambda)Q_2^*) \geq \lambda F(Q_1^*) + (1 - \lambda)F(Q_2^*) = F(Q_1^*)$, contradicting the optimality of Q_1^* and Q_2^* .

Now suppose that Q_1^* and Q_2^* induce the same distribution of posteriors and therefore induce the same gross payoffs. Then either $\nexists \lambda \in (0, 1)$ such that $F(\lambda Q_1^* + (1 - \lambda)Q_2^*) = \lambda F(Q_1^*) + (1 - \lambda)F(Q_2^*)$, in which case the argument of the preceding paragraph applies, or else there does exist such λ , in which case $Q_\lambda \in Q^*(U)$ as well, and we denote the corresponding information structure by Q_λ . Then, by the linearity of the trace function and the fact that Q_1^* and Q_2^* induce the same distribution of posteriors, $\text{tr}(Q_\lambda \bar{D}U\Pi) = \text{tr}(Q_1^* \bar{D}U\Pi) = \text{tr}(Q_2^* \bar{D}U\Pi)$.

This implies that $\text{tr}(Q^*(U) \bar{D}U\Pi)$ is single-valued, and since it is the composition of a continuous function (which can be viewed as an upper hemicontinuous correspondence) with an upper hemicontinuous correspondence, it is itself upper hemicontinuous (cf. Theorem 14.1.5 of Sydsæter et al., 2008). Together, its upper hemicontinuity and single-valuedness imply that it is a continuous function of U , thereby completing the proof. \square

Proof of Proposition 4

Proof. The continuity of C follows directly from the continuity of each of the functions $c_{i,j}$.

The Hessian matrix of $C(\pi, Q)$ is an $n|M| \times n|M|$ matrix with entries $\frac{\partial^2 c_{i,j}(\pi, \cdot)}{\partial q_{i,j}^2}$ on the diagonal and zeroes elsewhere. Since these diagonal entries are all strictly positive, the Hessian is clearly positive-definite, and we conclude that C is strictly convex for each π . \square

Proof of Proposition 5

Proof. We begin by proving the “only if” direction. Let $r_1 \geq r_2$ be two possible rewards. Let Q_i be the information structure optimally chosen under reward r_i , $i = 1, 2$. Let D_i^j be the decision matrix chosen under information structure Q_i and reward r_j , $i, j = 1, 2$. Since decisions can be thought of as being made optimally given signals from the information structure, WLOG, we can take $D_i := D_i^i = D_i^{-i}$, $i = 1, 2$.

The NIAC condition gives us:

$$\begin{aligned}
& r_1 \text{tr}(Q_1 D_1 \Pi) + r_2 \text{tr}(Q_2 D_2 \Pi) \geq r_2 \text{tr}(Q_1 D_1 \Pi) + r_1 \text{tr}(Q_2 D_2 \Pi) \\
\implies & r_1 P^*(r_1) + r_2 P^*(r_2) \geq r_2 P^*(r_1) + r_1 P^*(r_2) \\
\implies & (r_1 - r_2)[P^*(r_1) - P^*(r_2)] \geq 0
\end{aligned} \tag{3.17}$$

Since $r_1 \geq r_2$, in order for (3.17) to hold, we require that $P^*(r_1) \geq P^*(r_2)$. This proves the “only if” direction.

For the “if” direction, consider a set of reward levels $r_1 \geq r_2 \geq \dots \geq r_N$ and associated performances $P_1 \geq P_2 \geq \dots \geq P_N$, where $P_i := P^*(r_i)$. (We can order the performances in this manner since P^* is nondecreasing.)

Consider an assignment of performances to rewards $(r_i, P_{\sigma_1(i)})_{i=1}^N$, where σ_1 is a cyclic permutation. Let σ_2 be defined as follows:

$$\sigma_2(i) := \begin{cases} 1, & i = 1 \\ \sigma_1(1), & i = \sigma_1^{-1}(1) \\ \sigma_1(i), & \text{otherwise} \end{cases}$$

Now we compute the difference in total gross payoffs between the assignments defined by σ_2 and σ_1 .

$$\begin{aligned}
& \sum_{j=1}^N r_j P_{\sigma_2(i)} - \sum_{j=1}^N r_j P_{\sigma_1(i)} \\
&= r_1 P_1 + r_{\sigma_1^{-1}(i)} P_{\sigma_1(1)} - (r_1 P_{\sigma_1(1)} + r_{\sigma_1^{-1}(i)} P_1) \\
&= (r_1 - r_{\sigma_1^{-1}(i)}) (P_1 - P_{\sigma_1(1)}) \\
&\geq 0, \quad \text{since } r_1 \geq r_{\sigma_1^{-1}(i)} \text{ and } P_1 \geq P_{\sigma_1(1)}
\end{aligned}$$

Now we repeat this process for $j \geq 2$, at each step constructing the permutation σ_{j+1} as follows:

$$\sigma_{j+1}(i) := \begin{cases} j, & i = j \\ \sigma_j(j), & i = \sigma_j^{-1}(j) \\ \sigma_j(i), & \text{otherwise} \end{cases}$$

By the preceding argument, the total gross payoffs to the assignment increase (weakly) at each step. Since there are N rewards, this process must finish in $N - 1$ steps, ending with $\sigma_N(i) = i$ and the highest possible gross payoff. Since the initial assignment $(r_i, P_{\sigma_1(i)})_{i=1}^N$ was arbitrary, this implies the NIAC condition for our data. \square

Proof of Proposition 6

Proof. Fix some $x \in A$ and $y \in \Theta$. Then:

$$\begin{aligned}
& \Pr(\theta = x|a = x) \geq \Pr(\theta = y|a = x) \\
\iff & r \Pr(\theta = x|a = x) + 0 \cdot \sum_{z \neq x} \Pr(\theta = z|a = x) \\
& \geq r \Pr(\theta = y|a = x) + 0 \cdot \sum_{z \neq y} \Pr(\theta = z|a = x) \\
\iff & \sum_{z \in \Theta} u(x, z) \Pr(\theta = z|a = x) \geq \sum_{z \in \Theta} u(y, z) \Pr(\theta = z|a = x) \\
\iff & \sum_{z \in \Theta} u(x, z) \frac{\Pr(a = x|\theta = z) \Pr(\theta = z)}{\Pr(a = x)} \geq \sum_{z \in \Theta} u(y, z) \frac{\Pr(a = x|\theta = z) \Pr(\theta = z)}{\Pr(a = x)} \\
\iff & u_{k, \bullet} \Pi Q^* d_{\bullet, k}^* \geq u_{l, \bullet} \Pi Q^* d_{\bullet, k}^*,
\end{aligned}$$

where x and y are the k -th and l -th elements of Θ , respectively

The last implication holds because the (i, j) -th entry of $Q^* D^*$ is $\Pr(a_j|\theta_i)$. Since all these implications are bidirectional, and x and y were chosen arbitrarily, this completes the proof. \square

Proof of Proposition 8

Proof. Note that the expected posterior entropy is:

$$\mathbb{E}[H(\pi|Q)] = \sum_{\hat{\pi} \in \text{Supp}(\gamma_Q)} \gamma_Q(\hat{\pi}) H(\hat{\pi}) \tag{3.18}$$

since γ_Q has finite support. It is clear from (3.18) that the expected posterior entropy depends only on γ_Q and not the specific Q that induced it. Therefore, $I(\pi, Q) = \alpha(H(\pi) - \mathbb{E}[H(\pi|Q)])$ must satisfy cost equivalence.

Consider an information structure Q and a garbling P . Since P is itself a stochastic matrix, $\pi \rightarrow Q \rightarrow QP$ form a Markov chain in that order (cf. Cover and Thomas, 2006, Section 2.8). Therefore, by the data processing inequality (cf. Cover and Thomas, 2006, Theorem 2.8.1), $I(\pi, Q) \geq I(\pi, QP)$.

I is a composition of continuous functions and so it is clearly continuous except possibly when an entry in one of its arguments is zero. However, $\lim_{(x,y) \downarrow (0,0)} x \ln y = 0$ and

$\lim_{(x,y,z) \downarrow (0,0,0)} x \ln \frac{y}{z} = 0$, so continuity does not fail there either.

Convexity follows from Theorem 2.7.4 of Cover and Thomas (2006).

Finally, symmetry follows from cost equivalence. \square

Proof of Proposition 12

Proof. Since payoffs are symmetric across states, if the DM is optimizing, then she selects the most likely state given her posterior beliefs. Therefore, given an information structure Q^* , the optimal decision matrix D^* can be chosen such that $\forall j \in \{1, \dots, |M|\}$, the $\left(\min_{i \in \{1, \dots, n\}} \operatorname{argmax}_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^* \right)$ -th entry of $d_{\bullet,j}^*$ is 1 and all other entries are 0. Then, the probability

of correctly guessing the state is $\operatorname{tr}(Q^* D^* \Pi) = \sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^*$.

Now we show that this optimal probability can be achieved using no more than n signals. Suppose that for $j, j' \in \{1, \dots, |M|\}$, $G_{j,j'} := \operatorname{argmax}_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^* \cap \operatorname{argmax}_{i \in \{1, \dots, n\}} \pi_i q_{i,j'}^* \neq \emptyset$, and the corresponding maxima are strictly positive. Then, construct Q^{**} from Q^* where the columns of Q^{**} are the columns of Q^* , except for the j -th and j' -th columns. $q_{\bullet,j}^{**} = q_{\bullet,j}^* + q_{\bullet,j'}^*$, and $q_{j'}^{**}$ is a column of zeroes. Then, $\operatorname{argmax}_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^{**} = G_{j,j'}$, and $\max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^{**} = \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^* + \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j'}^*$, and since the rest of the columns of Q^{**} are the same

as those of Q^* , it must be that $\sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^{**} = \sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^*$. Furthermore, this quantity is unaffected by rearrangements of the columns of Q^{**} .

Using this logic, we construct Q' as follows. For $j > n$, let $q'_{\bullet,j}$ be a column of zeroes. For $j \leq n$, $q'_{\bullet,j} = \sum \left\{ k \mid \min_{i \in \{1, \dots, n\}} \operatorname{argmax}_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^* = j \right\} q_{\bullet,k}^*$, with an empty sum taken to be a column of zeroes. Then, $\sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q'_{i,j} = \sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q_{i,j}^*$. Therefore, if we construct D' analogously to D^* , then $\operatorname{tr}(Q' D' \Pi) = \operatorname{tr}(Q^* D^* \Pi)$. In particular, $\operatorname{tr}(Q' D' \Pi) = \sum_{j=1}^{|M|} \max_{i \in \{1, \dots, n\}} \pi_i q'_{i,j} = \sum_{i=1}^n \pi_i q'_{i,i}$.

Therefore, $C(\pi, Q') = K \left(\sum_{i=1}^n \pi_i q'_{i,i} \right)$, and the subject's problem can be rewritten as:

$$\max_{Q' \in \mathcal{Q}} r \sum_{i=1}^n \pi_i q'_{i,i} - K \left(\sum_{i=1}^n \pi_i q'_{i,i} \right) \quad \text{subject to } \pi_i q'_{i,i} = \max_k \pi_k q'_{k,i} \quad \forall i \leq n \quad (3.19)$$

Since both the ex-ante gross payoff and the cost depend only on the probability of answering correctly. Let this probability be q . Then, the maximand in (3.19) can be rewritten as $rq - K(q)$. In particular, $q'_{i,i}$ can be chosen such it is q for all $i \leq n$. Obviously $q \leq 1$, since it is a probability, and $q \geq \frac{1}{n}$, because the DM can do no worse than her prior.

If K is strictly convex and continuously differentiable, then the first-order condition (FOC) of this problem with respect to each $q'_{i,i}$ is:

$$r = K' \left(\sum_{i=1}^n \pi_i q'_{i,i} \right) \quad (3.20)$$

Because of the strict convexity of K and the restriction on its derivative, the FOC, together with the constraints in (3.19), is both necessary and sufficient for a solution to this problem. Therefore, any Q' such that $(q'_{i,i})_{i=1}^n$ satisfies (3.19) and the constraints in (3.19) is optimal. In particular, we can select Q' such that $q'_{i,i} = q \forall i \leq n$, where $q := (K')^{-1}(r)$, and $q'_{i,j} = \frac{1-q}{n-1} \forall j \neq i, i \leq n$.

Finally, note that $P^*(r) = \sum_{i=1}^n \pi_i q'_{i,i} = q \sum_{i=1}^n \pi_i = q = (K')^{-1}(r)$, which is continuous in r , since K is continuously differentiable. \square

Proof of Proposition 10

Proof. To proceed, we need a lemma:

Lemma 1. *Let β , ζ , and δ be strictly positive. Then $\Phi((\xi + \beta)\zeta\delta) - \Phi(\xi\zeta\delta)$ is strictly decreasing in ξ for positive ξ and strictly increasing for in ξ for negative ξ .*

This lemma is easily proven by differentiating to obtain $\zeta\delta[\phi((\xi + \beta)\zeta\delta) - \phi(\xi\zeta\delta)]$. Since the normal density is decreasing on the positive real line and increasing on the negative real line, this derivative is negative for positive ξ and positive for negative ξ .

For guesses of inner states that are not the true state, the result follows from setting $\xi = 2k + 1$ and $\beta = 2$ for $k \neq -1$ and comparing it to the expression in Lemma 1 when $\xi = 2k + 3$. This shows that guessing an inner state that is not the true state is likelier than guessing the inner state that is immediately farther from it. Applying this logic iteratively and exploiting the symmetry of the normal distribution to compare guesses of inner states on opposite sides of the true state gives the result.

In order to show that guessing the true state is likelier than guessing any other inner state, assume that the true state is not θ_{n-1} or θ_n , so that state immediately above the true state is also an inner state. (An obvious symmetric argument applies in case the true state is θ_{n-1} or θ_n .) Lemma 1 implies that:

$$\begin{aligned} \Phi(\zeta\delta) - \Phi(0) &> \Phi(2\zeta\delta) - \Phi(\zeta\delta) \quad \text{and} \quad \Phi(\zeta\delta) - \Phi(0) > \Phi(3\zeta\delta) - \Phi(2\zeta\delta) \\ \implies 2[\Phi(\zeta\delta) - \Phi(0)] &> \Phi(3\zeta\delta) - \Phi(\zeta\delta) \\ \implies \Phi(\zeta\delta) - \Phi(-\zeta\delta) &> \Phi(3\zeta\delta) - \Phi(\zeta\delta) \end{aligned}$$

Since the probability of guessing the true state is at least $\Phi(\zeta\delta) - \Phi(-\zeta\delta)$ (the true state could be the lowest state), combining this implication with the result for inner states that are not the true state proves the result. \square

Proof of Proposition 11

Proof. From (1.14), the DM's problem is:

$$\max_{\zeta \in [0, \infty)} \frac{r}{n} [2\Phi(\zeta\delta) + (n-2)(2\Phi(\zeta\delta) - 1)] - K(\zeta)$$

We can rewrite this as:

$$\max_{\zeta \in [0, \infty)} \frac{r}{n} [(2n-2)\Phi(\zeta\delta) - (n-2)] - K(\zeta) \quad (3.21)$$

The first-order condition is:

$$F(r, \zeta) \equiv \frac{(2n-2)r\delta}{n} \phi(\zeta\delta) - K'(\zeta) = 0 \quad (3.22)$$

In order to ensure that the first-order condition is sufficient, we verify the second-order condition:

$$-\frac{(2n-2)r\delta^3}{n} \zeta \phi(\zeta\delta) - K''(\zeta) < 0, \quad \text{since } \zeta \text{ is positive}$$

The DM's performance function is:

$$P^*(r) = \frac{1}{n} [(2n-2)\Phi(\zeta(r)\delta) - (n-2)] \quad (3.23)$$

In order to show that $P^*(r)$ is strictly concave, we compute:

$$\begin{aligned} \frac{d^2 P^*}{dr^2} &= \frac{dP^*}{dr} \left[\frac{(2n-2)\delta}{n} \phi(\zeta\delta) \frac{d\zeta}{dr} \right] \\ &= -\frac{(2n-2)\delta^3}{n} \zeta \phi(\zeta\delta) \frac{d\zeta}{dr} + \frac{(2n-2)\delta}{n} \phi(\zeta\delta) \frac{d^2 \zeta}{dr^2} \end{aligned} \quad (3.24)$$

In order to determine the sign of (3.24), we must compute $\frac{d\zeta}{dr}$ and $\frac{d^2 \zeta}{dr^2}$. By the implicit function theorem:

$$\begin{aligned} \frac{d\zeta}{dr} &= \frac{-\frac{\partial F}{\partial r}}{\frac{\partial F}{\partial \zeta}} \\ &= \frac{\frac{(2n-2)\delta}{n} \phi(\zeta\delta)}{\frac{(2n-2)r\delta^3}{n} \zeta \phi(\zeta\delta) + K''(\zeta)} \\ &> 0 \end{aligned} \quad (3.25)$$

Differentiating (3.25) with respect to r gives:

$$\begin{aligned}
\frac{d^2\zeta}{dr^2} &= \left[\frac{(2n-2)r\delta^3}{n} \zeta\phi(\zeta\delta) + K''(\zeta) \right]^{-2} \\
&\quad \times \left\{ -\frac{(2n-2)\delta^3}{n} \zeta\phi(\zeta\delta) \frac{d\zeta}{dr} \left(\frac{(2n-2)r\delta^3}{n} \zeta\phi(\zeta\delta) + K''(\zeta) \right) \right. \\
&\quad - \left[\left(\frac{(2n-2)\delta^3}{n} \zeta\phi(\zeta\delta) + \frac{(2n-2)r\delta^3}{n} \frac{d\zeta}{dr} \phi(\zeta\delta) \right. \right. \\
&\quad \quad \left. \left. - \frac{(2n-2)r\delta^5}{n} \zeta^2 \frac{d\zeta}{dr} \phi(\zeta\delta) + K'''(\zeta) \right) \right. \\
&\quad \left. \left. \times \left(\frac{(2n-2)\delta}{n} \phi(\zeta\delta) \right) \right] \right\} \\
&= \left[\frac{(2n-2)r\delta^3}{n} \zeta\phi(\zeta\delta) + K''(\zeta) \right]^{-2} \left\{ -\frac{(2n-2)\delta^3}{n} \zeta\phi(\zeta\delta) \frac{d\zeta}{dr} K''(\zeta) \right. \\
&\quad - \left[\left(\frac{(2n-2)\delta^3}{n} \zeta\phi(\zeta\delta) + \frac{(2n-2)r\delta^3}{n} \frac{d\zeta}{dr} \phi(\zeta\delta) + K'''(\zeta) \right) \right. \\
&\quad \left. \left. \times \left(\frac{(2n-2)\delta}{n} \phi(\zeta\delta) \right) \right] \right\} \tag{3.26} \\
&< 0
\end{aligned}$$

Substituting (3.25) and (3.26) back into (3.24) gives us that $\frac{d^2 P^*}{dr^2} < 0$, since $\frac{d\zeta}{dr} > 0$ and $\frac{d^2\zeta}{dr^2} < 0$. This concludes the proof. \square

Proof of Proposition 13

Proof. The principal's maximand is:

$$\frac{x-r}{(n-1)\exp\left(-\frac{r}{\alpha}\right)+1} \tag{3.27}$$

As argued in the main text of the paper, if this maximand is strictly quasiconcave in r , then this problem has a unique solution for each x , and since it is continuous in both x and r , the maximum theorem guarantees that the principal's optimal payment strategy $r^*(x)$ is continuous. Therefore, it simply remains to be shown that the maximand is strictly quasiconcave. We begin by differentiating it with respect to r :

$$\frac{\left(\frac{x-r-\alpha}{\alpha}\right)(n-1)\exp\left(-\frac{r}{\alpha}\right)-1}{\left((n-1)\exp\left(-\frac{r}{\alpha}\right)+1\right)^2} \tag{3.28}$$

Since the denominator in (3.28) is always strictly positive, the sign of (3.28) depends only on the sign of the numerator. The numerator is strictly positive (negative) when:

$$\begin{aligned} & \left(\frac{x - r - \alpha}{\alpha} \right) (n - 1) \exp \left(-\frac{r}{\alpha} \right) > (<) 1 \\ \iff & (n - 1) \left(\frac{x - r - \alpha}{\alpha} \right) > (<) \exp \left(\frac{r}{\alpha} \right) \end{aligned} \quad (3.29)$$

The LHS of (3.29) is strictly decreasing, and diverges to positive infinity as r is taken to negative infinity and to negative infinity as r is taken to positive infinity. The RHS of (3.29) is strictly increasing, and it approaches zero as r is taken to negative infinity and diverges to positive infinity as r is taken to positive infinity. Therefore, by the intermediate value theorem, the LHS and RHS must intersect, and they do so only at a single r .

Therefore, (3.27) exhibits a region of strict increase up until the point where $(n - 1) \left(\frac{x - r - \alpha}{\alpha} \right) = \exp \left(\frac{r}{\alpha} \right)$, after which it is strictly decreasing. Thus, (3.27) is strictly quasi-concave. \square

Proof of Proposition 14

Proof.

$$\begin{aligned} & \frac{d^2}{dr^2} [(x - r)P^*(r)] \\ &= \frac{d}{dr} \left[-P(r) + (x - r) \frac{d}{dr} P^*(r) \right] \\ &= -2 \frac{d}{dr} P^*(r) + (x - r) \frac{d^2}{dr^2} P^*(r) \end{aligned} \quad (3.30)$$

(3.30) is negative, since $P^*(r)$ is strictly increasing and strictly concave, and $x > r$, so the principal's ex-ante expected payoff is strictly concave in r . Therefore, there is a unique r^* for each x , and by the maximum theorem, $r^*(x)$ is continuous. \square

Appendix B: Angle Task

In addition to the “dots” tasks discussed in the main body of the paper, laboratory subjects also completed 100 “angle” tasks. For each of these tasks, subjects were shown a pair of intersecting line segments of random length¹³ and orientation and were told to identify the angle between them. This angle could have been 35°, 40°, 45°, 50°, or 55°, with each being equally likely. Subjects were rewarded for a correct answer and received no reward for an incorrect answer. Therefore, the “angle” tasks were uniform guess tasks of the same format as the “dots” tasks. Figure B.1 shows what this screen looks like to the subjects.

¹³Giving the arms of the angle random length ensured that subjects could not simply measure the distance between the endpoints of the arms to estimate the size of the angle.

Table 1: Linear regression of correctness on incentive level and demographic covariates in the “angle” tasks

	(1)	(2)
Incentive Level	0.0001 (0.0002)	0.0001 (0.0002)
Age		−0.001 (0.002)
Male		−0.007 (0.014)
Bachelor’s		−0.001 (0.017)
Econ/Psych/Neuro		0.028* (0.017)
\$20 Prize		0.017 (0.013)
Dots First		−0.015 (0.014)
Task Number		−0.00001 (0.0002)
Constant	0.444*** (0.014)	0.461*** (0.036)
Observations	7900	7900
R ²	1.16×10^{-5}	0.0009818

Note: *p<0.1; ** p<0.05; *** p<0.01
Standard errors clustered on subject.

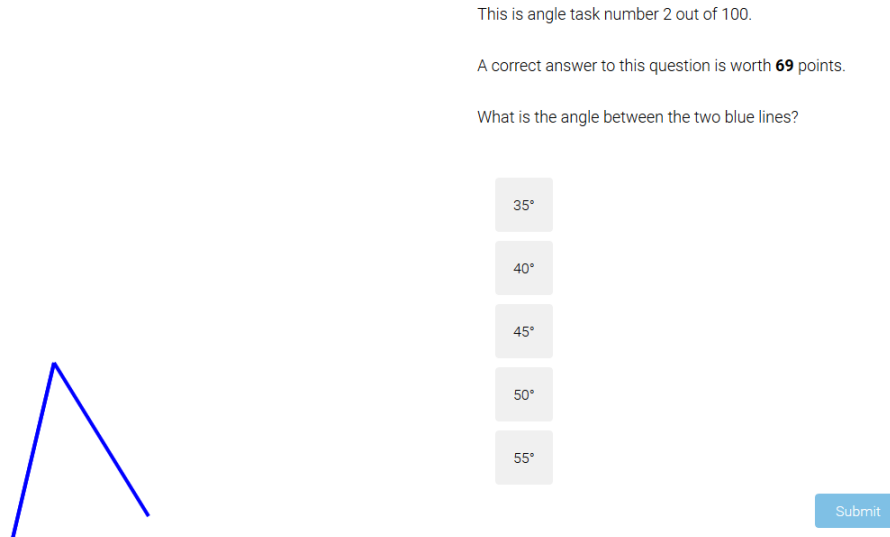


Figure 1: Angle display for a task

Table B.1 presents linear regressions of correctness on incentive level and demographic covariates for the entire laboratory subject pool. As was the case with the “dots” task, demographics are not significant predictors of correctness. However, neither is incentive level. This evidence indicates that this is not a task in which subjects generally respond to incentives.

Appendix C: Online Rational Inattention Experiment

In this appendix, we describe and present results from the online experiments mentioned in the main body of the paper.

Subjects were recruited using the Amazon Mechanical Turk platform and participated in the experiment on the Qualtrics platform. A total of 118 subjects completed the experiment. Subjects completed 200 tasks, each of the “dots” type. Roughly half the subjects (57 subjects) were given a participation fee of \$3 US and potential monetary prizes of \$3, while the other half (61 subjects) were given a participation fee of \$5 US and potential monetary prizes of \$5 US.

C1: Demographics

Table C.1 lists basic demographic data for the online subjects. The pool is fairly gender-balanced,¹⁴ though it is slightly more male than female, and highly educated; over 55% of the pool has a post-secondary degree.

The online pool is significantly different from the laboratory pool in some ways. In particular, the online pool is significantly older (one-tailed t-test of unpaired samples,

¹⁴One online subject declined to disclose their gender.

Table 2: Online Demographics

Number of subjects	$n = 118$
Gender ($n = 117$)	52.5% male; 47.5% female
Age ($n = 118$)	Average: 32.48; St. dev.: 8.88
Highest level of education achieved ($n = 118$)	
Some post-secondary	43.2%
Completed bachelor's degree	50.0%
Completed graduate or professional degree	6.8%

$p < 0.001$) and has a significantly greater proportion of subjects with bachelor's degrees but no advanced degrees (one-sided test of equality of proportions, $p = 0.003$).

C2: Rational Inattentiveness

No Improving Attention Cycles

We test against weak positive monotonicity using the method of (Doveh et al., 2002). At the 5% level, we fail to reject positive monotonicity for 103 out of 118 online subjects (87.3%).¹⁵

No Improving Action Switches

We test for NIAS using the bootstrap procedure outlined in Section 1.7. 82 out of 118 online subjects (69.5%) fail to reject NIAS.

Overall, this gives us 72 out of 118 online subjects (61.0%) whom we classify as rationally inattentive. This is a significantly smaller portion than in the laboratory pool (one-sided test of proportions, $p < 0.001$).

C3: Responsiveness to Incentives

We test for responsiveness using the full-sample and split-sample tests outlined in Section 1.7. At the 5% significance level 28 out of 72 online subjects (38.8%) who fail to reject rationality are responsive to incentives. This is a significantly smaller portion than in the laboratory pool (one-sided test of proportions, $p = 0.009$).

C4: Model Selection

We follow the same model selection procedures as in Section 1.7. As with the laboratory subjects, the only models that best fit the subjects are binary response and logistic response. 2 out of 28 responsive subjects (7.1%) are best fit by constant performance 8

¹⁵The optimization in the computation of the restricted regression for online subject 93 failed to converge, and so we did not perform the test for them. That subject has a success rate in the tasks of 99% (i.e. they identify the true state of nature correctly in 198 out of 200 tasks), and so we include them in the 103 online subjects who fail to reject positive monotonicity.

out of 28 responsive subjects (28.6%) are best fit by binary performance, 17 out of 28 responsive subjects (60.7%) are best fit by logistic performance, and 1 out of 28 responsive subjects (3.6%) are best fit by the concave performance function implied by normal signals. Ignoring the subjects who are best fit by constant response, these are similar to the proportions found in the laboratory. This seems to indicate that once the subset of responsive subjects is identified, the incidence of different types of cost functions within it is stable across contexts.

Appendix D: Instructions for Promises and Pronouncements Experiment

Included below are the instructions read out to participants for the baseline treatment. “Individual A” refers to the Sender role, and “Individual B” refers to the Receiver role.

Welcome to the Columbia Experimental Laboratory for the Social Sciences (CELSS)! Your participation in this experiment is much appreciated. During this session, we require your complete, undivided attention. As such, we ask that you remain quiet for the duration of the session, refrain from opening other applications on your computer, refrain from talking or passing notes to other participants, and put away all of your possessions, including your cell phones, which must be turned off. Do not touch the computer terminals until the session begins.

Before we begin, please read and sign both copies of the consent form located at your terminal.

[COLLECT CONSENT FORMS AND ENSURE THAT THEY ARE ALL SIGNED AND DATED]

You will be paid in cash for your participation in this experiment. Payment will occur once the session is over. This payment will be based partially on your own decisions and partially on the decisions of others; different participants may earn different amounts.

The entire session takes place at your computer terminals, including interaction with other participants. During the session, you may not communicate with other subjects, except through the computer terminals, as described later in the instructions.

The session will take place in four phases. The first phase will consist of a short quiz to test your understanding of the instructions. Your answers on this quiz will not determine your payment. The second phase is the actual experiment, which we will describe shortly. The third phase consists of three short questions about the outcomes of the second phase. Your responses in the second and third phases will determine your payment. The fourth phase consists of a brief questionnaire. Your answers to this questionnaire will not determine your payment. The first phase will begin once I have finished reading out the instructions. If you have any questions while I am reading out the instructions, please raise your hand, and your question will be answered out loud for the other participants to hear. If you have questions during the experiment, please raise your hand and remain silent; we will come to your terminal to assist you.

The second phase will consist of 2 rounds. At the start of each round, you will be randomly matched with another participant. Each pair of participants will follow the same

rules, and the actions of no pair will have any effect on any other pair. The matching is anonymous; at no point during or after the experiment will you be told with whom you were paired in any round. Each participant in a pair will be assigned one of two roles, A or B. If you were assigned role A in one round, you will be assigned role B in the next, and vice versa. You will not be informed of the outcomes of any round until after the third phase.

Each round consists of two stages, scenario 1 and scenario 2. Scenario 1 may happen before scenario 2, or scenario 2 may happen before scenario 1; both orders are equally probable. Individual A will be told which scenario occurs first; individual B will not. Between the two stages, you will be randomly rematched with another participant in the role that is opposite yours.

Scenario 1 looks like this. [SHOW DIAGRAM OF 'PROMISES' GAME] First, Individual A chooses to go Left or to go Right. Then, without seeing which direction A chose, Individual B also chooses to go Left or to go Right. If both individuals go in the same direction, then they both earn 10 points. If they choose to go in different directions, then A earns 15 points, and B earns 0 points. After selecting a direction but before B makes their choice, A chooses a message to send to B. A has a choice between sending: "Going Left is in your best interests" or "Going Right is in your best interests."

To recap the order of actions: First, Individual A chooses Left or Right from a screen that looks like this [SHOW SCREENSHOT OF A'S PROMISES ACTION SCREEN]. Then A chooses a message to send to Individual B from a screen that looks like this [SHOW SCREENSHOT OF A'S PROMISES MESSAGE SCREEN]. Finally, B receives the message and chooses Left or Right from a screen that looks like this [SHOW SCREENSHOT OF B'S PROMISES CHOICE].

Scenario 2 looks like this. [SHOW DIAGRAM OF 'PRONOUNCEMENTS' GAME] It is exactly like Scenario 1, except that A's choice of direction is instead randomly made by the computer; the computer chooses Left or Right, each with 50% probability. Just as in Scenario 1, before B makes their choice, A selects a message to send to them: "Going Left is in your best interests"; or "Going Right is in your best interests."

To recap the order of actions: First, the computer randomly selects Left or Right. Then, A learns of the computer's choice and selects a message to send to B from a screen that looks like this. [SHOW SCREENSHOT OF A'S PRONOUNCEMENTS MESSAGE SCREEN] Finally, B receives the message and chooses Left or Right from a screen that looks like this [SHOW SCREENSHOT OF B'S PROMISES CHOICE].

After both rounds are over, we will move on to the third phase of the experiment, where you will be asked three questions about the outcomes of the second phase. A correct answer to one of these questions is worth 4 points. Following these three questions, we will move on to the questionnaire phase. The questionnaire will have no impact on your payment.

To reiterate: you will not be informed about the outcome or the decisions made by the people with whom you were paired in any round until after the third phase. After the third phase, one stage from each of the 2 rounds of Phase 2 and one of the first two of the three questions of Phase 3 will be randomly selected by the computer, and you will be paid your earnings for those stages and that question at a rate of 2 points to 1 dollar. Additionally, you will be paid a show-up fee of 10 points, or 5 dollars. Earnings will be

rounded up to the nearest dollar. Payment will occur in private, and you are under no obligation to reveal to other participants how much you were paid.

We will now begin the quiz to test your understanding of the instructions. Please answer the questions on your computer screen. When you are done, you will receive immediate feedback about your performance.

[QUIZ]

Now that you have all completed the quiz, are there any questions? We will now begin the actual experiment. From now on, if you have any questions, please remain silent and raise your hand, and we will come to your terminal to assist you.

[EXPERIMENT AND QUESTIONNAIRE]

The session is now over. Please wait quietly while we deliver further instructions.

You will be paid in the adjacent room. We will call you one by one, in order of your lab ID numbers, located on the side of your carrel. Please remain silent and seated until we call your ID. When we call your number, please take all your belongings with you, as you will be asked to leave the lab once you have received your payment.

Thank you very much for your participation.

[CALL OUT ID'S, GIVE PAYMENT]

Appendix E: Proofs for Chapter 3

Proof of Proposition 20

Proof. By (3.2) and (3.4), the problem of minimizing (3.5) is additively separable across attributes. Therefore, I can consider each attribute value separately.

Consider attribute j of item i . The DM's problem is to minimize:

$$\mathbb{E}[(a_{i,j} - \hat{a}_{i,j})^2] + C(f_{i,j}, g_{i,j}) \quad (3.31)$$

subject to $\int_r g_{i,j}(r_{i,j}) dr_{i,j} = 1 \forall a \in \mathbb{R}$.

In what follows, we drop the subscripts i and j to save on notation. The Lagrangian of this problem can be written as:

$$\begin{aligned} L = & \int_r \left[\int_a a^2 \frac{g(r|a)f(a)}{\int_b g(r|b)f(b)db} da - \left(a \frac{g(r|a)f(a)}{\int_b g(r|b)f(b)db} da \right)^2 \right] \int_b g(r|b)f(b)db dr \\ & + \int_r \int_a \kappa g(r|a)f(a) \ln \left(\frac{g(r|a)f(a)}{\int_b g(r|b)f(b)db} \right) da dr + \int_r \int_a \lambda(a)g(r|a)f(a)da dr \end{aligned} \quad (3.32)$$

where $\lambda(a) = \nu(a)/f(a)$ and $\nu(a)$ are Lagrange multipliers.¹⁶

I now apply the variational method (cf. MacCluer, 2013). Consider $\eta(r, a)$ smooth such that $\lim_{r \rightarrow -\infty} \eta(r, a) = \lim_{r \rightarrow \infty} \eta(r, a) = \lim_{a \rightarrow -\infty} \eta(r, a) = \lim_{a \rightarrow \infty} \eta(r, a) = 0$.

¹⁶The values $\lambda(a)$ can be thought of as rescaled Lagrange multipliers.

Then, define:

$$\begin{aligned}
J(\varepsilon) := & \int_r \left[\int_a a^2 (g(r|a) + \varepsilon \eta(r, a)) f(a) da - \frac{(\int_a a (g(r|a) + \varepsilon \eta(r, a)) f(a) da)^2}{\int_a g(r|a) f(a) da} \right. \\
& + \int_r \int_a \kappa (g(r|a) + \varepsilon \eta(r, a)) f(a) \ln \left(\frac{(g(r|a) + \varepsilon \eta(r, a)) f(a)}{\int_b (g(r|b) + \varepsilon \eta(r, b)) f(b) db} \right) da \\
& \left. + \int_r \int_a \lambda(a) (g(r|a) + \varepsilon \eta(r, a)) f(a) da \right] dr \tag{3.33}
\end{aligned}$$

Computing $\frac{\partial J}{\partial \varepsilon}|_{\varepsilon=0}$ and setting it equal to zero gives:

$$\begin{aligned}
0 = & \int_r \left[\int_a \eta(r, a) f(a) \left(a^2 + \kappa \ln \left(\frac{g(r|a) f(a)}{\int_b g(r|b) f(b) db} \right) + \kappa + \lambda(a) \right) da \right. \\
& + \int_a \eta(r, a) f(a) da \left(\frac{\int_a (a g(r|a) f(a) da)^2}{\int_a (g(r|a) f(a) da)^2} \right) \\
& - 2 \int_a a \eta(r, a) f(a) da \left(\frac{\int_a a g(r|a) f(a) da}{\int_a g(r|a) f(a) da} \right) \\
& \left. + \int_a \int_b \kappa \eta(r, b) f(b) db \frac{g(r|a) f(a)}{\int_b g(r|b) f(b) db} da \right] dr \tag{3.34}
\end{aligned}$$

Given that η is arbitrary, this must hold for any η satisfying the criteria. Therefore, after rearranging, I have:

$$\ln h(a|r) = -\frac{\lambda(a)}{\kappa} - \frac{(a - \mu_h)^2}{\kappa} \tag{3.35}$$

where h is the posterior density and μ_h is its mean. A further rearrangement gives:

$$h(a|r) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{a - \mu_h}{\sigma} \right)^2 \right) \tag{3.36}$$

where $\sigma = \sqrt{\frac{\kappa}{2}}$ and $\lambda(a) = \kappa \ln(\sqrt{\kappa\pi})$.

(3.36) gives the density of a normal distribution with mean μ_h and variance $\frac{\kappa}{2}$.

Since the DM wishes to minimize posterior variance, she selects h as her posterior, provided $\frac{\kappa}{2} < \sigma_f^2$, where σ_f^2 denotes the prior variance. Otherwise, she opts not to collect information about this particular attribute and simply uses her prior. \square

Proof of Proposition 21

Proof. It is known that if $a_{i,j}$ is normally distributed and $r_{i,j}$ is normally distributed centred around $a_{i,j}$, then the posterior belief of the value of $a_{i,j}$ upon observing $r_{i,j}$ is also normally distributed. In what follows, we drop subscripts for convenience's sake.

Suppose $a \sim N(\mu, \sigma_f^2)$ and $r \sim N(a, \sigma_g^2)$. Then, tedious calculations show that the

posterior is such that $a|r \sim N\left(\frac{\sigma_f^2 r + \sigma_g^2 \mu}{\sigma_f^2 + \sigma_g^2}, \frac{\sigma_f^2 \sigma_g^2}{\sigma_f^2 + \sigma_g^2}\right)$.

From Proposition 20, it is known that the posterior variance σ_h^2 is equal to $\frac{\kappa}{2}$. Therefore, $\frac{\kappa}{2} = \frac{\sigma_f^2 \sigma_g^2}{\sigma_f^2 + \sigma_g^2}$. Rearranging this equation gives:

$$\sigma_g^2 = \frac{\kappa \sigma_f^2}{2\sigma_f^2 - \kappa} \quad (3.37)$$

Substituting this back into the expression for the posterior mean, $\mu_h := \frac{\sigma_f^2 r + \sigma_g^2 \mu}{\sigma_f^2 + \sigma_g^2}$ gives:

$$\frac{2\sigma_f^2 - \kappa}{2\sigma_f^2} r + \frac{\kappa}{2\sigma_f^2} \mu \quad (3.38)$$

□

Proof of Proposition 22

Proof. First, I must compute the relevant conditional densities. For the sake of convenience, I drop subscripts and assume that the DM is trying to estimate the attribute values x , y , and z , with the knowledge that $x > y > z$. Assume that without knowledge of the ranking, $x, y, z \sim N(\mu, \sigma^2)$.

Let ϕ and Φ denote the standard normal density and CDF, respectively. I begin by computing the conditional density of x :

$$\begin{aligned} & f(x|x > y > z) \\ &= \frac{f(x \cap x > y > z)}{\Pr(x > y > z)} \\ &= \frac{\int_{-\infty}^x \int_{-\infty}^y f(x, y, z) dz dy}{\frac{1}{6}}, \quad \text{by the fact that } x, y, \text{ and } z \text{ are unconditionally i.i.d.} \\ &= 6 \int_{-\infty}^x \int_{-\infty}^y \frac{1}{\sigma^3} \phi\left(\frac{x-\mu}{\sigma}\right) \phi\left(\frac{y-\mu}{\sigma}\right) \phi\left(\frac{z-\mu}{\sigma}\right) dz dy \\ &= 6 \int_{-\infty}^x \frac{1}{\sigma^2} \phi\left(\frac{x-\mu}{\sigma}\right) \phi\left(\frac{y-\mu}{\sigma}\right) \Phi\left(\frac{y-\mu}{\sigma}\right) dy \\ &= 6 \frac{1}{\sigma^2} \phi\left(\frac{x-\mu}{\sigma}\right) \int_{-\infty}^{\Phi\left(\frac{x-\mu}{\sigma}\right)} t dt, \quad \text{by substitution} \\ &= \frac{3}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi^2\left(\frac{x-\mu}{\sigma}\right) \end{aligned} \quad (3.39)$$

Similarly, for the middle-ranked attribute value y :

$$\begin{aligned}
& f(x|x > y > z) \\
&= 6 \int_{-\infty}^y \int_y^{\infty} \frac{1}{\sigma^3} \phi\left(\frac{x-\mu}{\sigma}\right) \phi\left(\frac{y-\mu}{\sigma}\right) \phi\left(\frac{z-\mu}{\sigma}\right) dz dx \\
&= 6 \frac{1}{\sigma} \phi\left(\frac{y-\mu}{\sigma}\right) \Phi\left(\frac{y-\mu}{\sigma}\right) \left[1 - \Phi\left(\frac{y-\mu}{\sigma}\right)\right] \\
&= \frac{6}{\sigma} \phi\left(\frac{y-\mu}{\sigma}\right) \left[\Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi^2\left(\frac{y-\mu}{\sigma}\right)\right] \tag{3.40}
\end{aligned}$$

And finally, for the lowest-ranked attribute value z :

$$\begin{aligned}
& f(z|x > y > z) \\
&= 6 \int_x^{\infty} \int_y^{\infty} \frac{1}{\sigma^3} \phi\left(\frac{x-\mu}{\sigma}\right) \phi\left(\frac{y-\mu}{\sigma}\right) \phi\left(\frac{z-\mu}{\sigma}\right) dy dx \\
&= \frac{3}{\sigma} \phi\left(\frac{z-\mu}{\sigma}\right) \left[1 - \Phi\left(\frac{z-\mu}{\sigma}\right)\right]^2 \tag{3.41}
\end{aligned}$$

Now, to save on notation, I assume that $\mu = 0$ and compute the expectations of each of these densities. I can do this because the variance of a normal distribution is invariant to translation, and ultimately the expectation is only needed to compute the variance.

The expectation of x conditional on the ranking is:

$$\begin{aligned}
& 3 \int_{-\infty}^{\infty} \frac{x}{\sigma} \phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) dx \\
&= 3\sigma \int_{-\infty}^{\infty} \frac{x}{\sigma^2} \phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) dx \\
&= 3\sigma \left[-\phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) \Big|_{-\infty}^{\infty} + 2 \int_{-\infty}^{\infty} \frac{1}{\sigma} \phi^2\left(\frac{x}{\sigma}\right) \Phi\left(\frac{x}{\sigma}\right) dx \right] \\
&= \frac{3\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\sqrt{2}}{\sigma} \phi\left(\frac{\sqrt{2}x}{\sigma}\right) \Phi\left(\frac{x}{\sigma}\right) dx \\
&= \frac{3\sigma}{2\sqrt{\pi}}, \quad \text{since the standard normal distribution is symmetric about 0} \tag{3.42}
\end{aligned}$$

By symmetry, the conditional expectation of z is $-\frac{3\sigma}{2\sqrt{\pi}}$, and the conditional expectation of y is 0, since its distribution is symmetric and centred around 0.

Now I compute the second moment of x :

$$\begin{aligned}
& 3 \int_{-\infty}^{\infty} \frac{x^2}{\sigma} \phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) dx \\
&= 3\sigma^4 \int_{-\infty}^{\infty} \left(\frac{x^2}{\sigma^5} \phi\left(\frac{x}{\sigma}\right) - \frac{1}{\sigma^3} \phi\left(\frac{x}{\sigma}\right) \right) \Phi^2\left(\frac{x}{\sigma}\right) dx + 3\sigma^4 \int_{-\infty}^{\infty} \frac{1}{\sigma^3} \phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) dx \\
&= 3\sigma^4 \left[-\frac{x}{\sigma^3} \phi\left(\frac{x}{\sigma}\right) \Phi^2\left(\frac{x}{\sigma}\right) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{2x}{\sigma^4} \phi^2\left(\frac{x}{\sigma}\right) \Phi\left(\frac{x}{\sigma}\right) dx \right] + \sigma^2 \\
&= \frac{3\sigma^2}{\sqrt{2\pi}} \left[-\phi\left(\frac{\sqrt{2}x}{\sigma}\right) \Phi\left(\frac{x}{\sigma}\right) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sigma} \phi\left(\frac{\sqrt{2}x}{\sigma}\right) \phi\left(\frac{x}{\sigma}\right) dx \right] + \sigma^2 \\
&= \sigma^2 + \frac{\sqrt{3}\sigma^2}{2\pi} \int_{-\infty}^{\infty} \frac{\sqrt{3}}{\sigma} \phi\left(\frac{\sqrt{2}x}{\sigma}\right) dx \\
&= \sigma^2 \left(1 + \frac{\sqrt{3}}{2\pi} \right) \tag{3.43}
\end{aligned}$$

By symmetry, this is also the second moment of z . Therefore, the variances of x and z are:

$$V(x) = V(z) = \sigma^2 \left(1 + \frac{\sqrt{3}}{2\pi} - \frac{9}{4\pi} \right) \tag{3.44}$$

Now I compute the second moment of y , which is also its variance, since its mean is zero.

$$\begin{aligned}
& V(y) \\
&= 6 \int_{-\infty}^{\infty} \frac{y^2}{\sigma} \phi\left(\frac{y}{\sigma}\right) \left[\Phi\left(\frac{y}{\sigma}\right) - \Phi^2\left(\frac{y}{\sigma}\right) \right] dy \\
&= 6\sigma^4 \int_{-\infty}^{\infty} \left(\frac{y^2}{\sigma^5} \phi\left(\frac{y}{\sigma}\right) - \frac{1}{\sigma^3} \phi\left(\frac{y}{\sigma}\right) \right) \Phi\left(\frac{y}{\sigma}\right) dy + 6\sigma^4 \int_{-\infty}^{\infty} \frac{1}{\sigma^3} \phi\left(\frac{y}{\sigma}\right) \Phi\left(\frac{y}{\sigma}\right) dy \\
&\quad - 2\sigma^2 \left(1 + \frac{\sqrt{3}}{2\pi} \right), \quad \text{by (3.43)} \\
&= 6\sigma^4 \left[-\frac{y}{\sigma^3} \phi\left(\frac{y}{\sigma}\right) \Phi\left(\frac{y}{\sigma}\right) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{y}{\sigma^4} \phi^2\left(\frac{y}{\sigma}\right) dy \right] + 3\sigma^2 - 2\sigma^2 \left(1 + \frac{\sqrt{3}}{2\pi} \right) \\
&= \frac{3\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} y \frac{\sqrt{2}}{\sigma} \phi\left(\frac{\sqrt{2}y}{\sigma}\right) dy + 3\sigma^2 - 2\sigma^2 \left(1 + \frac{\sqrt{3}}{2\pi} \right) \\
&= \sigma^2 \left(1 - \frac{\sqrt{3}}{\pi} \right) \tag{3.45}
\end{aligned}$$

$V(y) < V(x) = V(z)$, so by Corollary 1, this concludes the proof of the first part of the

proposition. The second follows easily from the fact that if the attributes are identically distributed, then $V(x)$, $V(y)$, and $V(z)$ are constant across attributes. \square