# Misperceiving Mechanisms: Imperfect Perception and the Failure to Recognize Dominant Strategies* 

Daniel Martin ${ }^{\dagger}$ and Edwin Muñoz-Rodriguez ${ }^{\ddagger}$

May 5, 2019


#### Abstract

Because of cognitive constraints, individuals may not perfectly perceive a mechanism's payoff structure, and thus may not realize that the mechanism they face has a dominant strategy. We study the link between misperception of mechanisms and the use of dominant strategies by applying standard techniques from the literatures on imperfect perception, rational attention, and contingent thinking to a widely-used single agent mechanism. For this mechanism, we provide a novel Bayesian model of limited game form recognition and contingent thinking, and by taking this model to experimental data, we find that misperception of this mechanism appears to respond in line with the incentives to perceive the mechanism better.


JEL Codes: I30, C91, D12
Keywords: Imperfect perception, rational inattention, contingent thinking, experiments

[^0]Strategic dominance is a widespread concept in the theory of incentives and its applications. Incentivized single-agent tasks such as proper scoring rules and the Becker-DeGrootMarschak (BDM) mechanism -broadly used to elicit beliefs and preferences- hinge on having a dominant strategy. In the theory of mechanism design, implementation in dominant strategies plays a major role, in part because "we can feel fairly confident that a rational agent who has a dominant strategy will indeed play it" (Mas-Colell, Whinston, Green, et al. 1995). ${ }^{1}$ In applications, dominant strategy incentive compatible mechanisms are viewed as providing a safe environment for participants to reveal their information (Niederle, Roth, and Sönmez 2008), a level playing field when participants are heterogeneous in their level of sophistication (Pathak and Sönmez 2008), and a simplified choice problem that spares "participants the need for elaborate strategic calculations" (Milgrom 2004).

However, recognizing that one strategy dominates all others requires understanding the payoffs to all actions in all contingencies. When a payoff rule is complicated, so that agents must determine the consequences for multiple contingencies based on the payoff rule and/or jointly consider numerous contingencies and actions, it may be unrealistic to assume that they have a complete or correct mental representation of the mechanism's extensive form when choosing their strategy, so they may not realize there is a dominant course of action.

In line with this, there is increasing evidence from the lab and the field that individuals can have trouble identifying the dominant strategies of mechanisms. For example, Hassidim, Marciano, Romm, and Shorrer (2017) indicate that one barrier to truthful revelation is that agents may fail to identify which strategies are dominant, and point to evidence that relates "low cognitive abilities with higher rates of misrepresentation". ${ }^{2}$ Relatedly, Rees-Jones and Skowronek (2018) conduct a large-scale experiment with a group of medical students in a version of the algorithm used by the National Residency Match Program, and find evidence that many students do not realize telling the truth about their preferences is optimal.

Experimental evidence of a link between perception of a mechanism's payoffs and truthful revelation of preferences is provided by Cason and Plott (2014) (CP hereafter). They show that subjects can misperceive the incentives of a widely-used single agent mechanism, the BDM mechanism. This mechanism is designed so that valuations of goods and services are elicited in a truthful way. ${ }^{3}$ CP implement this mechanism experimentally and find that subjects do not act in line with the dominant strategy of the mechanism and that they misperceive the payoff rule. Because of this, CP conclude that the choices of subjects do not

[^1]reveal their true preferences.
In this paper, we use a novel model and a new experimental protocol to show that misperception of this mechanism appears to respond systematically to the costs and benefits of perceiving the mechanism better. While our results pertain just to the BDM mechanism, it is our hope is that these results can offer guidance to market designers and policymakers about how to respond to misperception of mechanisms more generally. One trivial takeaway from our results is that higher incentives are likely to decrease misperception. However, this does not imply that high incentives are sufficient to eliminate misperception entirely. Completely eliminating misperception may be unattainable for mechanisms in which the payoff rule is inherently complex. In this regard, the BDM mechanism is ideally suited for studying misperception of mechanisms because while it has a dominant strategy, its payoff rule involves a number of different contingencies for each action, which creates an opening for misperception to have a sizable impact on behavior.

The fact that agents appear to weigh the costs and benefits of perceiving the mechanism accurately also means that it is important which mechanism or mechanisms that the mechanism in question can be confused with. If a certain alternative mechanism has a very distorting effect on behavior, market designers may want to emphasize that agents are not facing that particular extensive form. However, this does not mean that more information about the mechanism is necessarily better. Our results also suggests that how hard or costly it is for the agent to understand the payoff rule is also an important consideration.

In addition, our finding that agents behave in line with the costs and benefits of perceiving the mechanism better preserves, at least in one setting, ex-ante rationality in the face of seemly dominated actions. Further, these actions can be explained without the need to impose non-rational assumptions that may not be falsifiable. However, because there are many other potential explanations for mistakes in BDM experiments besides misperception, ${ }^{4}$ we run a variant on the CP experiment where payoffs are described contingency-by-contingency. Because we do not change the extensive form of the BDM as we vary the protocol, most alternative explanations for choice mistakes in the BDM do not provide an answer for why our new protocol substantially reduces mistakes. So while other behavioral biases may still impact behavior with our new protocol, our experimental variation helps to identify a portion of mistakes that have fewer alternative explanations besides misperception.

Having identified a portion of mistakes that appear to be driven by misperception, we try to identify whether misperception responds systematically to variation in the mechanism. Existing models of mechanism misperception cannot answer this question because they do not leave room for uncertainty or learning. For example, CP provide a non-Bayesian model

[^2]of game form recognition in which agents are certain (correctly or incorrectly) of the game form they face, and Li (2017) proposes a model of misperceiving a mechanism's game form based on a complete failure of contingent thinking.

Instead, we introduce a Bayesian model of limited game form recognition and contingent thinking that explicitly allows for uncertainty and learning. We model misperception of the BDM mechanism using standard techniques from the psychology, neuroscience, and economics literatures on perception and attention. A ubiquitous assumption in these literatures is that agents receive a noisy mental signal about an uncertain state of the world. ${ }^{5}$ Our modeling innovation is to assume that this uncertain state of the world is the mechanism's extensive form and that agents receive noisy mental signals about the mechanism's extensive form. We follow these literatures in assuming that agents correctly form posterior beliefs about the state based on these signals, and in the context of our model, these beliefs summarize an agent's uncertainty about the contingent payoffs to each action.

Relative to existing models of misperceiving mechanisms, our approach has two main advantages for modeling behavior in the BDM. First, unlike the model of Cason and Plott (2014), our approach is Bayesian, so it allows for partial recognition of the BDM's game form. In practice, this matters because partial game form recognition allows us to rationalize offers in the BDM that cannot be explained by a belief in a single game form. In addition, a Bayesian framework is needed to accommodate standard models of perception and attention, such as rational inattention theory, which predict that decision makers will not internalize all available information. ${ }^{6}$

Second, unlike the model of Li (2017), our approach allows for agents to be incorrect about the set of possible payoffs to taking an action. The set of payoffs the agent thinks are possible is determined by the set of mechanisms that the agent might confuse with the BDM. Our approach requires the modeler to specify the game forms that an agent thinks are possible, which presents a modeling challenge and provides the modeler an additional degree of freedom. However, for the BDM, CP provide external justification for this modeling choice: they find that subjects appear to confuse the BDM with a first-price sealed-bid procurement auction (FPA hereafter). Collecting further non-choice data, such as beliefs about outcomes, can be useful in providing external justification. For example, CP asked subjects to report their payment after the posted price was realized, and many subjects incorrectly reported monetary payments in line with the FPA payoff rule. ${ }^{7}$

[^3]Based on the findings of CP, we assume that the set of game forms an agent thinks are possible when facing the BDM are the BDM and FPA. With this assumption, our model generates a one-to-one map between an agent's beliefs about the game form and their offers, which can be used to rationalize many dominated offers. This map can also be used to infer the beliefs agents hold about the likelihood of the BDM, and with these implied beliefs we measure the extent of misperception subjects possess in BDM experiments.

To provide comparative static predictions for beliefs and offers, we further specialize our model by assuming that mental signals are chosen optimally and signals have costs proportional to their informativeness, as in rational inattention theory. We examine the comparative static predictions for two model parameter values: one that represents the benefits of perceiving the mechanism better and another that represents the costs of perceiving the mechanism better. First, we investigate the comparative static prediction for a model parameter that represents the benefits of perceiving the mechanism better: the maximum posted price. As the maximum posted price in the BDM increases, subjects benefit more from understanding the mechanism, so our model predicts that their misperception should on average decrease. Thus, their implied beliefs (of the likelihood that they face the BDM) should be higher on average. We can test this prediction using data from the CP experiment because they varied the maximum posted price between and across subjects. Using their data, we find that implied beliefs are higher when the maximum posted price increases, which is consistent with the theoretical prediction.

It may be surprising that offers suggest misperception has decreased given that offers are further from the dominant strategy as the maximum posted price increases. However, our theoretical results show that such a divergence is possible as the maximum posted price increases: subjects can make larger mistakes even though their misperception has decreased. ${ }^{8}$

Second, we investigate the comparative static prediction for a model parameter that represents the costs of perceiving the mechanism better: the cost of information. As the cost of information falls, our model predicts both a decrease in misperception (higher average beliefs of the correct game form) and a reduction in the size of mistakes (lower average offers). To test this prediction, we vary the cost of information by comparing a replication of the CP experiment to a new experiment in which the payoffs to each action are specified contingency-by-contingency, as in the experimental approach of Esponda and Vespa (2014). ${ }^{9}$ Seen through the lens of Martinez-Marquina, Niederle, and Vespa (2017), our contingent protocol is easier computationally than the CP protocol because our contingent protocol

[^4]focuses a subject's attention on a subset of random variable realizations, instead of the whole set of potential random variable realizations.

When we implement this novel "contingent thinking" protocol for the BDM, we find that the fraction of subjects playing in line with the dominant strategy doubles and that there is a general shift in offers towards the dominant strategy. Using our approach to measure misperception, we find that subjects are $45.6 \%$ more correct about the game form on average, which corresponds to an average increase of 20.8 percentage points in the likelihood of the correct game form. Given these results, our comparative static prediction for the cost of information appears to be consistent with the data.

In Section 2, we discuss the relationship of our paper to existing literatures. In Section 3, we provide our model of mechanism misperception and its comparative static predictions. Section 4 tests our comparative static prediction for benefits empirically using the data from CP. Section 5 presents our replication of the CP experiment, introduces our new protocol, and tests our comparative static prediction for costs. Section 6 concludes.

## 1 Literature Review

Our paper relates to several literatures, including ones that study mistakes in dominantstrategy mechanisms, rational inattention in games and experiments, and failures of contingent reasoning.

First, our paper relates to a literature on mistakes in dominant-strategy mechanisms. In the context of public good provision, Chen (2008) provides a systematic review of empirical evidence on the failure to truthfully report valuations in the pivotal mechanism. In this order of ideas, Cason, Saijo, Sjöström, and Yamato (2006) introduce the notion of secure implementation, and compare the rate of dominant strategy play in the pivotal and Groves mechanisms, finding that in the latter, which is securely implementable, individuals are more likely to play dominant strategies. Both notions coincide for the BDM, so a lack of secure implementability does not offer an explanation for choice mistakes in our setting.

Kagel, Harstad, and Levin (1987) find that English auctions induce truthful bidding behavior more often and in general closer to the truth-telling dominant strategy than its counterpart, the second-price sealed-bid auction, even though they are strategically equivalent. This fundamental finding has become part of the "folk wisdom" among auction theorists (Ausubel 2004) and is a motivating fact in Li (2017). Harstad (2000) studies experimentally the sealed-bid second-price auction and finds that even after allowing for learning effects with feedback ( 70 rounds), subjects still fail to identify the dominant strategy.

A concurrent paper by Bull, Courty, Doyon, and Rondeau (2019) provides a replication
of CP, and they also find failures to play a dominant strategy of the BDM mechanism. They conduct three variations on the CP experiment and conclude that game form misperception cannot explain differences in behavior across treatments. However, unlike our approach, they assume that the probability of misperceiving the game form is not impacted by treatment variation.

Second, our paper relates to the literature on rational inattention theory (Sims 2003) because we use Shannon entropy to model the costs of optimally-chosen mental signals. For instance, we draw on theoretical results for rational inattention in individual decision problems (Caplin and Dean 2013; Caplin, Dean, and Leahy 2018) and costly information in persuasion games (Kamenica and Gentzkow 2011).

When rational inattention theory has been applied to games, players are modeled as being inattentive to an exogenous payoff-relevant source of information (for instance, see Matějka 2015, Yang 2015, and Ravid 2014). Alternatively, Martin (2017) considers the possibility that players are inattentive to the strategic implications of actions. ${ }^{10}$ In this paper, we propose a new form of inattention in games: being rationally inattentive to the game form itself.

Several other studies have looked for evidence of rational inattention using lab experiments. These experiments have studied rational inattention to the objects of choice (see Cheremukhin, Popova, and Tutino 2011, Martin 2016, and Caplin and Martin 2017) and rational inattention in perceptual tasks (see Dean and Neligh 2017, Dewan and Neligh 2017, Khaw, Stevens, and Woodford 2017, and Ambuehl, Ockenfels, and Stewart 2018). ${ }^{11}$

Third, our paper relates to a growing literature on contingent reasoning, which is a kind of thinking "that entails reasoning about events without knowing whether or not these events are true or will occur" (Esponda and Vespa 2017). Papers in this literature have studied contingent reasoning in a number of settings, such as common value and private value auctions, as well as single decision maker problems that illustrate the Allais and Ellsberg paradoxes. ${ }^{12}$

Esponda and Vespa (2017) provide a formal notion of contingent thinking in a Savage setting with testable implications. One of the five settings they study is equivalent to the BDM. Our new protocol is related to their contingent treatment, as it is intended to help agents with contingent reasoning. However, a crucial difference in design is that we describe the consequences of taking different actions in different states, whereas they focus the attention

[^5]of subjects onto a particular state. Martinez-Marquina, Niederle, and Vespa (2017) decompose the difficulty of contingent reasoning into two parts: the number of states that need to be considered (computational complexity) and the difficulty of coping with uncertainty while having to think through the state space (lost of power of certainty).

As mentioned previously, the behavioral model of misperceiving mechanisms in Li (2017) is based on a total failure of contingent reasoning. With this assumption, he shows that extensive forms can be organized into equivalence classes based on the set of possible payoffs, which leads to "local misunderestandings" of a game form within a class. Our approach allows agents to instead possess something between perfect contingent reasoning and a total failure of contingent reasoning. For instance, an agent might think through the contingencies for one action, but not the other. This might happen if one action is more salient, for instance if it is the default or recommended action. Such partial contingent thinking might even be optimal if the payoffs to one action vary more across possible game forms and thinking requires costly effort. Along these lines, the rational inattention version of our model endogenously produces limits to contingent thinking based on the costs and benefits of understanding contingent payoffs within the context of the BDM.

## 2 Model of Misperceiving the BDM

In this section, we first introduce our approach to imperfect perception of the BDM, and then solve for optimal offers in the BDM with misperception. Finally, we consider how offers and beliefs change with model parameter values if the misperception of agents responds to the costs and benefits of correctly perceiving the BDM game form.

### 2.1 Misperception and the BDM

In the literatures on perception and attention, agents are assumed to receive noisy signals about a state of the world $\omega \in \Omega$. Our innovation is to assume that $\Omega$ is the set of possible extensive forms the agent could confuse the BDM with. ${ }^{13}$ While agents start off uncertain about which $\omega \in \Omega$ they face, they are able to reduce their uncertainty by thinking about the extensive form of the game they face. When there are cognitive limits in thinking about game form, this can produce game form misrecognition, which is "a failure of the decision maker to recognize the proper connections between the acts available for choice and the consequences of choice" (Cason and Plott 2014, p. 1237).

Following CP, we analyze the case where subjects could confuse this game form with

[^6]the FPA. Thus, the states of the world $\Omega=\left\{\omega_{0}, \omega_{1}\right\}$, where $\omega_{0}$ corresponds to the BDM and $\omega_{1}$ corresponds to the FPA. ${ }^{14}$ While other forms of misconceptions are possible, CP primarily focus on this one alternative game form because it can explain many offers and because it reflects the mistakes that some subjects make in indicating the payment they should receive. ${ }^{15}$ As in CP, the restriction to having just two possible game forms greatly simplifies our analysis. For similar reasons, papers in the perception and attention literature often analyze settings where there are just two possible states of the world. However, while expanding the size of the state space is likely to reduce analytic tractability, our model is still in principle solvable with more than two possible game forms.

The prior beliefs that an agent has about the likelihood of facing alternative extensive forms is $\mu \in \Delta(\Omega)$. This prior could be set exogenously by an experimenter, determined by an individual's experience in a particular laboratory, or formed from day-to-day interactions. Also, it could be correct given the likelihood of encountering a particular game in a particular setting, but it need not be correct. The prior probability an agent assigns to payoffs being determined by the BDM is given by $\mu\left(\omega_{0}\right)$. For convenience, we abuse notation and denote $\mu:=\mu\left(\omega_{0}\right)$ as the prior probability of state of world being $\omega_{0}$. In other words, $\mu$ is the agent's belief of the likelihood they could face the BDM. Determining the prior of an agent about game forms is an empirically challenging task, ${ }^{16}$ any of the theoretical results in this paper, such as our comparative static predictions, will be robust to changes in the prior.

Before choosing a strategy, the agent receives a mental signal about the game form the face, which represents their subjective perception of the game they are facing. This signal can be fully informative about the game form, but it can also be fuzzy and stochastic, such as the mental signal received by a subject who does not put much cognitive effort into thinking about the game they are facing. ${ }^{17}$ Thus, while our model allows for misperception, it also nests the standard model without misperception.

As in Kamenica and Gentzkow (2011), the agent's signal is represented by an information structure $\pi: \Omega \rightarrow \Delta(\Delta(\Omega))$, where $\Gamma(\pi) \in \Delta(\Omega)$ is the set of posterior beliefs supported by $\pi$. We restrict information structures to those that contain Bayes plausible posteriors

[^7]according to the prior $\mu$, which we denote by $\Pi(\mu)$. This posterior-based approach is used to model imperfect perception and limited attention by Caplin and Martin (2015), Caplin and Dean (2015), and Matějka and McKay (2015). Note that this can include any information structure, including those that are chosen optimally and those that are not. For convenience, we abuse notation and denote $\gamma:=\gamma\left(\omega_{0}\right)$ as the posterior probability of state of world being $\omega_{0}$. In other words, $\gamma$ is the agent's belief of the likelihood they are actually facing the BDM when choosing their offer strategy.

### 2.2 Offers with Misperception

With the BDM, a subject who values a good at $\theta$ is asked to choose an "offer" price $b$, which corresponds to the minimum price they are willing to sell the good. Next, a "posted" price $p$ is randomly drawn from a distribution $f$ on $[0, \bar{p}]$. We will follow the standard assumption that $f$ is the uniform distribution. If $p \geq b$ the good is sold and the agent is paid $p$. Otherwise, the agent keeps the good and obtains $\theta$. Thus, the agent's ex-ante payoff from offering a minimum price $b$ is given by

$$
\begin{equation*}
\int_{0}^{\bar{p}}[\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b]] d f(p) \tag{1}
\end{equation*}
$$

with $b_{B D M}^{*}=\theta$ being a weakly dominant offer strategy.
With the FPA, the timing is the same, but if $p \geq b$ the good is sold, the agent is paid $b$ instead. Thus, the agent's ex-ante payoff from offering a minimum price $b$ is given by

$$
\begin{equation*}
\int_{0}^{\bar{p}}[\theta \mathbb{1}[p<b]+b \mathbb{1}[p \geq b]] d f(p) . \tag{2}
\end{equation*}
$$

Since $f$ is uniform, the optimal offer strategy is $b_{F P A}^{*}=\frac{\theta+\bar{p}}{2}$.
With misperception of mechanisms, the agent may unsure of which payoff rule they face, so their optimal offer strategy will reflect this uncertainty. Based on their posterior belief $\gamma$, the agent chooses a minimum selling price $b$ to maximize

$$
\int_{0}^{b} \theta d f(p)+\int_{b}^{\bar{p}}[\gamma p+(1-\gamma) b] d f(p)
$$

Since $f$ is uniform, the above expression is strictly concave in $b$, which guarantees existence. In addition, the unique optimal offer strategy is given by

$$
\begin{equation*}
b^{*}(\gamma)=\frac{\theta+(1-\gamma) \bar{p}}{2-\gamma} \tag{3}
\end{equation*}
$$

In practice, offers are monetary amounts, so are limited to discrete increments, such as dollars or pennies. In the appendix, we solve for the optimal offer strategy when offers are


Figure 1: Map between an agent's belief of the likelihood they are facing the $\operatorname{BDM}(\gamma)$ and their optimal offer ( $b^{*}$ ).
restricted in this way. The solution for discrete increments converges to the solution above as the increments become infinitesimally small.

The map between optimal offers and beliefs is illustrated in Figure 1. As long as $\theta \leq \bar{p}$, the optimal offer is decreasing in $\gamma$ (moving to the right along the line), and for a given $\gamma$, the optimal offer is increasing in $\bar{p}$ (the line shifts out as the y-intercept moves up and the $x$-intercept remains fixed).

Note that the optimal offer strategy reduces to the standard optimal strategies when posteriors are degenerate. Specifically, $b^{*}(1)=b_{B D M}^{*}=\theta$ and $b^{*}(0)=b_{F P A}^{*}=\frac{\theta+\bar{p}}{2}$. As a result, all offers in the range $\left[\theta, \frac{\theta+\bar{p}}{2}\right]$ can be rationalized by our model. This expands the set of explainable offers relative to either the standard approach (without misperception) or an all-or-nothing model of misperception (as in CP). By extending our model to incorporate other behavioral forces (such as the endowment effect), it could also explain offers made outside of this range.

### 2.3 Offers with Optimal Misperception

Without any additional restrictions on the form of misperception, our model does not indicate how posterior beliefs and offers will change as model parameter values change. A natural restriction on misperception is to assume that information structures are chosen optimally. We assume further that information carries entropic costs, as in rational inattention theory.

The timing of this specialization is summarized in Figure 2. After observing the value of the good, the agent chooses an information structure. Once a posterior is observed, an offer


Figure 2: The timeline for optimal misperception of the BDM.
price is chosen, a posted price is realized, and payoffs are determined.
We solve this model backwards. First, fixing a posterior, the optimal offer strategy is given by (3), which has an expected payoff of

$$
\begin{equation*}
\frac{\bar{p}^{2}+2 \bar{p}(1-\gamma) \theta+\theta^{2}}{2 \bar{p}(2-\gamma)} \tag{4}
\end{equation*}
$$

Next, considering the cost it will entail, an information structure is chosen optimally, subject to Bayesian plausibility. Specifically, let the cost of any information structure $\pi$ be

$$
K(\pi, \kappa, \mu)=\kappa E_{\pi}[H(\mu)-H(\gamma)]
$$

with $H(\gamma)=-\sum_{\omega \in \Omega} \gamma(\omega) \ln (\gamma(\omega))$. The parameter $\kappa$ linearly scales the cost of an information structure, so can be interpreted as the marginal cost of information.

Given prior $\mu$, the value of choosing the optimal information structure can be written as

$$
\begin{aligned}
V(\mu)=\max _{\pi \in \Pi(\mu)} & E_{\pi}[\hat{v}(\gamma)] \\
\text { s.t. } & E_{\pi}[\gamma]=\mu
\end{aligned}
$$

where for all $\gamma \in \Gamma(\pi)$

$$
\begin{equation*}
\hat{v}(\gamma)=\frac{\bar{p}^{2}+2 \bar{p}(1-\gamma) \theta+\theta^{2}}{2 \bar{p}(2-\gamma)}-\kappa(H(\mu)-H(\gamma)) \tag{5}
\end{equation*}
$$

Since $\hat{v}$ is continuous, an optimal information strategy exists.
This problem can be mapped into the sender's problem in the "costly persuasion" framework of Gentzkow and Kamenica (2014) by choosing the reference belief to be the prior $\mu$ and the value of the induced belief (adjusted for the cost of inducing that belief) to be $\hat{v}(\gamma) .{ }^{18}$

[^8]Because $V(\mu)$, the value of the above program, is the concave closure of $\hat{v}(\mu)$, the concavity of $\hat{v}$ determines whether or not the agent benefits from choosing an informative information structure. We say an information structure is "informative" if it generates posterior beliefs that are different from the prior.

The function $\hat{v}$ is twice differentiable on $(0,1)$, and its concavity is fully characterized by

$$
\frac{\partial^{2} \hat{v}}{\partial^{2} \gamma}=\frac{(\bar{p}-\theta)^{2}}{\bar{p}(2-\gamma)^{3}}-\frac{\kappa}{\gamma(1-\gamma)}
$$

It is easily inferred from this expression that there are two cases: On the one hand, if $\frac{\kappa \bar{p}}{(\bar{p}-\theta)^{2}}<\frac{1}{6 \sqrt{3}}$ the domain of the function $\hat{v}$ can be partitioned in three regions (strictly concave, strictly convex, and then strictly concave) . For the set of priors between the strictly concave regions, $V(\mu)>\hat{v}(\mu)$, so it is optimal to choose an informative information structure. Outside this region, the agent regards the BDM game form to be either very unlikely or very likely, so it is not worth choosing an informative, but costly, information structure.

On the other hand, if $\frac{\kappa \bar{p}}{(\bar{p}-\theta)^{2}}>\frac{1}{6 \sqrt{3}}, \hat{v}$ is strictly concave for all $\mu$, so it is never optimal to choose an informative information structure. This establishes the following claim.

Claim 1 There exists a set of priors such that choosing an informative information structure is optimal if and only if $\frac{\kappa \bar{p}}{(\bar{p}-\theta)^{2}}<\frac{1}{6 \sqrt{3}}$.

While an optimal information structure exists for all model parameter values, this condition indicates when the optimal information structure will or will not contain informative mental signals (in other words, when it is degenerated for all priors). It follows that when this condition holds, the optimal information structure has two posteriors in its support, that we denote by $\gamma_{1}^{*}$ and $\gamma_{2}^{*}$. Also, these posteriors are unique for a fixed prior. Otherwise, a contradiction with strict concavity would be reached.

Since $\hat{v}$ is differentiable everywhere in $(0,1)$, it is possible to give a differential characterization of the optimal posteriors, which turns out to be useful for deriving comparative statics. In the spirit of the ILR conditions of Caplin and Dean (2013), these restrictions are:

1. Same slope of $\hat{v}$ at the optimal posteriors:

$$
\begin{equation*}
\left.\frac{\partial \hat{v}}{\partial \gamma}\right|_{\gamma_{1}^{*}}=\left.\frac{\partial \hat{v}}{\partial \gamma}\right|_{\gamma_{2}^{*}} \tag{6}
\end{equation*}
$$

2. Same tangent of $\hat{v}$ and $V$ at the optimal posteriors:

$$
\begin{equation*}
\hat{v}\left(\gamma_{1}^{*}\right)-\hat{v}\left(\gamma_{2}^{*}\right)=\left.\gamma_{1}^{*} \frac{\partial \hat{v}}{\partial \gamma}\right|_{\gamma_{1}^{*}}-\left.\gamma_{2}^{*} \frac{\partial \hat{v}}{\partial \gamma}\right|_{\gamma_{2}^{*}} \tag{7}
\end{equation*}
$$



Figure 3: An example of $\hat{v}$ (curved line in black) and the tangent of $\hat{v}$ (straight line in red) at the optimal posteriors.

This differential characterization is broadly discussed in Caplin and Dean (2013), and we refer the interested reader to their paper. A geometric intuition can be provided using Figure 3. The curved black line corresponds to $\hat{v}$, and the straight red line corresponds to the unique line tangent to $\hat{v}$ at the optimal posteriors. The concave closure of $\hat{v}$ coincides with $\hat{v}$ on the intervals $\left[0, \gamma_{1}^{*}\right)$ and $\left(\gamma_{2}^{*}, 1\right]$ and with the tangent line on $\left[\gamma_{1}^{*}, \gamma_{2}^{*}\right]$. The first restriction says that the slope of $\hat{v}$ at the optimal posteriors is the same. The second adds the restriction that the red line has the same slope as $\hat{v}$ at the optimal posteriors.

Finally, we consider the impact of this optimal information structure on offers. If the agent has a prior $\mu \in\left[0, \gamma_{1}^{*}\right) \cup\left(\gamma_{2}^{*}, 1\right]$, the prior and posterior will coincide, so the agent's optimal offer is $b^{*}(\mu)$. Otherwise, if the agent's prior belongs to $\left[\gamma_{1}^{*}, \gamma_{2}^{*}\right]$, the posterior is stochastically determined to be one of the two optimal posteriors characterized above, so the agent's offer will sometimes be $b^{*}\left(\gamma_{1}^{*}\right)$ and sometimes be $b^{*}\left(\gamma_{2}^{*}\right)$. To simply testing, we will establish our comparative static predictions for average offers.

For a given $\mu$, the average offer is a straight line if we do not condition on the state, which is illustrated by Figure 4. However, if we just look at average offers when agents face the BDM (when the state is $\omega_{0}$ ), the average offer bows towards the standard optimal strategy without misperception, as illustrated by Figure 5.


Figure 4: An example of the impact of optimal misperception on average offers (without conditioning on the state).


Figure 5: An example of the impact of optimal misperception on average offers when agents face the BDM (conditional on the state being $\omega_{0}$ ).

### 2.4 Comparative Statics of Optimal Misperception

In this section we establish comparative static predictions for two parameters of our specialized model: the maximum posted price and the cost of information. We first examine how a change in each parameter impacts the optimal posteriors, and then average optimal posteriors. Last, we determine the impact of a change in each parameter on average offers.

### 2.4.1 Comparative Statics: Optimal Posteriors

The maximum posted price $(\bar{p})$ and the cost of information $(\kappa)$ impact optimal posteriors in systematic ways. As $\bar{p}$ increases, an agent has more incentive to obtain informative mental signals of the game form, so the region of priors for which it is optimal to do so enlarges. On the other hand, as $\kappa$ increases, the region of priors for which it is optimal to obtain informative mental signals of the game form shrinks. These comparative statics are established with the following two claims.

Claim 2 Assume $\frac{\kappa \bar{p}}{(\bar{p}-\theta)^{2}}<\frac{1}{6 \sqrt{3}}$ and without loss of generality that $\gamma_{1}^{*}<\gamma_{2}^{*}$. Optimal posterior $\gamma_{1}^{*}$ decreases with $\bar{p}$, and optimal posterior $\gamma_{2}^{*}$ increases with $\bar{p}$.

Proof. Use the implicit function theorem to derive conditions (6) and (7) implicitly with respect to $\bar{p}>0$ and solve for $\hat{v}^{\prime \prime}\left(\gamma_{i}\right) \frac{\partial \gamma_{i}}{\partial \bar{p}}, i=1,2$ to obtain

$$
\hat{v}^{\prime \prime}\left(\gamma_{i}\right) \frac{\partial \gamma_{i}}{\partial \bar{p}}=\frac{1}{\gamma_{j}-\gamma_{i}} \frac{\left(\gamma_{j}-\gamma_{i}\right)^{2}\left(\bar{p}^{2}-\theta^{2}\right)}{2\left(2-\gamma_{j}\right)\left(2-\gamma_{i}\right)^{2} \bar{p}^{2}}
$$

$i \neq j, i, j=1,2$. Since $\gamma_{1}<\gamma_{2}$, the right-hand term is positive for $i=1$ and negative for $i=2$. Combined with the fact that $\hat{v}$ is strictly concave in the optimal posteriors, this implies $\frac{\partial \gamma_{1}}{\partial \bar{p}}<0$ and $\frac{\partial \gamma_{2}}{\partial \bar{p}}>0$, as desired.

Claim 3 Assume $\frac{\kappa \bar{p}}{(\bar{p}-\theta)^{2}}<\frac{1}{6 \sqrt{3}}$ and without loss of generality that $\gamma_{1}^{*}<\gamma_{2}^{*}$. Optimal posterior $\gamma_{1}^{*}$ increases with $\kappa$, and optimal posterior $\gamma_{2}^{*}$ decreases with $\kappa$.

Proof. Since by assumption $\gamma_{1}^{*}<\gamma_{2}^{*}$, use again the implicit function theorem to derive conditions (6) and (7) implicitly with respect to $\kappa$, and solving for $\hat{v}^{\prime \prime}\left(\gamma_{i}\right) \frac{\partial \gamma_{i}}{\partial \kappa}$ for $i=1$, 2, we obtain:

$$
v^{\prime \prime}\left(\gamma_{1}\right) \frac{\partial \gamma_{1}}{\partial \kappa}=\frac{H\left(\gamma_{1}\right)-H\left(\gamma_{2}\right)-\left(\gamma_{1}-\gamma_{2}\right) H^{\prime}\left(\gamma_{1}\right)}{\gamma_{1}-\gamma_{2}}
$$

and

$$
v^{\prime \prime}\left(\gamma_{2}\right) \frac{\partial \gamma_{2}}{\partial \kappa}=\frac{H\left(\gamma_{1}\right)-H\left(\gamma_{2}\right)-\left(\gamma_{1}-\gamma_{2}\right) H^{\prime}\left(\gamma_{2}\right)}{\gamma_{1}-\gamma_{2}}
$$

$i \neq j, i, j=1,2$. Since $\gamma_{1}<\gamma_{2}$ by assumption, the denominator in both expressions is negative. It follows from strict concavity and differentiability of the entropic cost function that $H\left(\gamma_{1}\right)-H\left(\gamma_{2}\right)+H^{\prime}\left(\gamma_{1}\right)\left(\gamma_{2}-\gamma_{1}\right) \geq 0$ and $H\left(\gamma_{1}\right)-H\left(\gamma_{2}\right)-H^{\prime}\left(\gamma_{2}\right)\left(\gamma_{1}-\gamma_{2}\right) \leq 0$. Therefore, the right-hand side in the first expression is negative and is positive in the second one. Since $\hat{v}$ is concave in the optimal posteriors, we have that $\frac{\partial \gamma_{1}}{\partial \kappa} \geq 0$ and $\frac{\partial \gamma_{2}}{\partial \kappa} \leq 0$, as desired.

### 2.4.2 Comparative Statics: Average Optimal Posteriors

Building on previous claims, we first show that when agents face the BDM, the average belief is increasing in the maximum posted price $(\bar{p}) .{ }^{19}$ Sensibly, as the benefits to perceiving the game form correctly increase, so do the average beliefs of the correct game form. Next, we show that when agents face the BDM, the average belief decreases as the cost of information $(\kappa)$ increases. As it gets harder and harder to disentangle states, the extent of misperception increases on average.

Proposition 4 Let $E\left[\gamma \mid \omega_{0}, \bar{p}\right]$ be the expected belief when facing the BDM and the maximum posted price is $\bar{p}$. If $\bar{p}^{\prime}>\bar{p}$, then $E\left[\gamma \mid \omega_{0}, \bar{p}^{\prime}\right] \geq E\left[\gamma \mid \omega_{0}, \bar{p}\right]$ for any prior $\mu$.

Proof. By definition,

$$
\begin{aligned}
E\left[\gamma \mid \omega_{0}\right] & =\pi\left(\gamma_{1}^{*} \mid \omega_{0}\right) \gamma_{1}^{*}+\pi\left(\gamma_{2}^{*} \mid \omega_{0}\right) \gamma_{2}^{*} \\
& =\frac{\pi \gamma_{1}^{*}}{\mu} \gamma_{1}^{*}+\frac{(1-\pi) \gamma_{2}^{*}}{\mu} \gamma_{2}^{*}
\end{aligned}
$$

with $\gamma_{1}^{*}, \gamma_{2}^{*}$, and $\pi$ implicit functions of $\bar{p}$. By definition $\pi$ is constrained to satisfy

$$
\gamma_{1}^{*} \pi+\gamma_{2}^{*}(1-\pi)=\mu
$$

so we can rewrite

$$
\begin{equation*}
E\left[\gamma \mid \omega_{0}\right]=\frac{1}{\mu}\left(\frac{\gamma_{2}^{*}-\mu}{\gamma_{2}^{*}-\gamma_{1}^{*}}\right)\left(\gamma_{1}^{*}\right)^{2}+\frac{1}{\mu}\left(1-\frac{\gamma_{2}^{*}-\mu}{\gamma_{2}^{*}-\gamma_{1}^{*}}\right)\left(\gamma_{2}^{*}\right)^{2}=\gamma_{2}^{*}+\gamma_{1}^{*}-\frac{\gamma_{2}^{*} \gamma_{1}^{*}}{\mu} \tag{8}
\end{equation*}
$$

By using the chain rule to derive $E\left[\gamma \mid \omega_{0}\right]$ w.r.t. $\bar{p}$, we obtain

$$
\frac{d E\left[\gamma \mid \omega_{0}\right]}{d \bar{p}}=\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{1}^{*}} \frac{\partial \gamma_{1}^{*}}{\partial \bar{p}}+\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{2}^{*}} \frac{\partial \gamma_{2}^{*}}{\partial \bar{p}}
$$

and because $\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{1}^{*}}<0$ and $\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{2}^{*}}>0$ (from 8), we conclude from Claim 2 that $\frac{d E\left[\gamma \mid \omega_{0}\right]}{d \bar{p}}>0$ as desired.

[^9]Proposition 5 Let $E\left[\gamma \mid \omega_{0}, \kappa\right]$ be the expected belief when facing the BDM and the cost of information is $\kappa$. If $\kappa^{\prime}>\kappa$, then $E\left[\gamma \mid \omega_{0}, \kappa^{\prime}\right] \leq E\left[\gamma \mid \omega_{0}, \kappa\right]$ for any prior $\mu$.

Proof. Using the chain rule to derive $E\left[\gamma \mid \omega_{0}\right]$ w.r.t. $\kappa$, we obtain

$$
\frac{d E\left[\gamma \mid \omega_{0}\right]}{d \kappa}=\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{1}^{*}} \frac{\partial \gamma_{1}^{*}}{\partial \kappa}+\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{2}^{*}} \frac{\partial \gamma_{2}^{*}}{\partial \kappa}
$$

and because $\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{1}^{*}}<0$ and $\frac{\partial E\left[\gamma \mid \omega_{0}\right]}{\partial \gamma_{2}^{*}}>0$ (see eq. 8), from Claim 3 we conclude that $\frac{d E\left[\gamma \mid \omega_{0}\right]}{d \kappa} \leq 0$, as desired.

### 2.4.3 Comparative Statics: Average Offers

Last, we examine how average offers change as the maximum posted price ( $\bar{p}$ ) and the cost of information ( $\kappa$ ) change. When agents are facing the BDM, the average offer can either increase or decrease with $\bar{p}$ depending on the prior, but the average offer increases with $\kappa$ for all priors.

Because the average optimal posterior increases with $\bar{p}$, we might expect the average offer to decrease with $\bar{p}$. However, this is not true in general, as the example in Figure 6 illustrates. Note that in general, the expression for the average offer, conditional on facing the BDM, is given by

$$
E\left[b(\gamma) \mid \omega_{0}, \mu\right]= \begin{cases}{\left[\gamma_{2}^{*} A-\gamma_{1}^{*} B\right] \frac{1}{\mu}-(A-B)} & \text { if } \mu \in\left[\gamma_{1}^{*}, \gamma_{2}^{*}\right] \\ b(\mu) & \text { otherwise }\end{cases}
$$

where $A=\frac{\gamma_{1}^{*}}{\gamma_{2}^{*}-\gamma_{1}^{*}} b\left(\gamma_{1}^{*}\right)$ and $B=\frac{\gamma_{2}^{*}}{\gamma_{2}^{*}-\gamma_{1}^{*}} b\left(\gamma_{2}^{*}\right)$. Observe that the expected offer is convex in the region of the domain where information is valuable.

When it is not valuable to get informative mental signals, $b^{*}(\gamma)$ shifts out as $\bar{p}$ increases, which means that there is an increase in the expected offer for all such $\mu$. At the same time, an increase in $\bar{p}$ widens the region of priors for which it is optimal to get informative mental signals. For such priors, when $\bar{p}$ increases, the rate at which the average offer moves closer to the standard one increases. The net result is that the comparative static is uncertain and any behavior can be rationalized by assuming a particular prior (or a particular distribution of priors across subjects).

On the other hand, when $\kappa$ increases, the average offer moves in a consistent direction for all priors. A higher information cost decreases the value of disentangling the two states, and as a result, the agent misperceives the game form more often, resulting in a higher average offer. This is illustrated in Figure 7 and formally established in the following proposition.


Figure 6: An example of $E\left[b^{*}(\gamma) \mid \omega_{0}\right]$ where $\bar{p}=4$ (blue), $\bar{p}=5$ (black) $\bar{p}=6$ (red), and $\kappa=0.1$.


Figure 7: An example of the impact of an increase in the cost of information on the average expected offer when agents are facing the BDM. The dashed line corresponds to a higher $\kappa$ than the one for the solid line.

Proposition 6 Let $E\left[b^{*}(\gamma) \mid \omega_{0}, \kappa\right]$ be the expected offer when facing the BDM and the cost of information is $\kappa$. For $\kappa^{\prime}>\kappa$, then $E\left[b^{*}(\gamma) \mid \omega_{0}, \kappa^{\prime}\right] \geq E\left[b^{*}(\gamma) \mid \omega_{0}, \kappa\right]$ for any prior $\mu$.

Proof. For convenience, let $g(\mu, \kappa)=E\left[b^{*}(\gamma) \mid \omega_{0}\right]$. First, note that by Claim 1, for $\kappa^{\prime} \geq \frac{\theta^{2}}{6 \sqrt{3} \bar{p}}$, $g\left(\mu, \kappa^{\prime}\right)=b^{*}(\mu)$ as choosing an uninformative information structure is optimal for every prior and the claim is trivially satisfied. For $\kappa<\frac{\theta^{2}}{6 \sqrt{3} \bar{p}}$, let $\gamma_{1}^{*}(\kappa) \leq \gamma_{2}^{*}(\kappa)$ be the unique posteriors in the optimal signal structure for any $\mu \in\left[\gamma_{1}^{*}(\kappa), \gamma_{2}^{*}(\kappa)\right]$. Claim 3 implies that for any $\kappa^{\prime}>\kappa,\left[\gamma_{1}^{*}\left(\kappa^{\prime}\right), \gamma_{2}^{*}\left(\kappa^{\prime}\right)\right] \subseteq\left[\gamma_{1}^{*}(\kappa), \gamma_{2}^{*}(\kappa)\right]$. Therefore, we have three cases: In the first case, $\mu \in\left[0, \gamma_{1}^{*}(\kappa)\right] \cup\left[\gamma_{2}^{*}(\kappa), 1\right]$, so we have $g(\mu, \kappa)=b^{*}(\mu)=g\left(\mu, \kappa^{\prime}\right)$.

In the second case, $\mu \in\left[\gamma_{1}^{*}(\kappa), \gamma_{1}^{*}\left(\kappa^{\prime}\right)\right] \cup\left[\gamma_{2}^{*}\left(\kappa^{\prime}\right), \gamma_{2}^{*}(\kappa)\right]$, we have $g\left(\mu, \kappa^{\prime}\right)=b^{*}(\mu)$, and since $b^{*}(\mu)$ is strictly concave in $\mu$ we have $g\left(\mu, \kappa^{\prime}\right)>\pi b^{*}\left(\gamma_{1}^{*}(\kappa)\right)+(1-\pi) b^{*}\left(\gamma_{2}^{*}(\kappa)\right)$ for $\pi$ s.t. $\pi \gamma_{1}^{*}(\kappa)+(1-\pi) \gamma_{2}^{*}(\kappa)=\mu$. Since $b^{*}(\cdot)$ is decreasing, and $\frac{\pi \gamma_{1}^{*}}{\mu}<\pi, \frac{(1-\pi) \gamma_{2}^{*}}{\mu}>1-\pi$, we have that $\pi b^{*}\left(\gamma_{1}^{*}(\kappa)\right)+(1-\pi) b^{*}\left(\gamma_{2}^{*}(\kappa)\right)>\frac{\pi \gamma_{1}}{\mu} b^{*}\left(\gamma_{1}^{*}(\kappa)\right)+\frac{(1-\pi) \gamma_{2}}{\mu} b^{*}\left(\gamma_{2}^{*}(\kappa)\right)=g(\mu, \kappa)$, so the conclusion follows.

In the third case, $\mu \in\left[\gamma_{1}^{*}\left(\kappa^{\prime}\right), \gamma_{2}^{*}\left(\kappa^{\prime}\right)\right]$, with $g\left(\mu, \kappa^{\prime}\right)=\frac{\pi^{\prime} \gamma_{1}}{\mu} b^{*}\left(\gamma_{1}^{*}\left(\kappa^{\prime}\right)\right)+\frac{(1-\pi) \gamma_{2}}{\mu} b^{*}\left(\gamma_{2}^{*}\left(\kappa^{\prime}\right)\right)$ for $\pi^{\prime}$ s.t. $\pi^{\prime} \gamma_{1}^{*}\left(\kappa^{\prime}\right)+\left(1-\pi^{\prime}\right) \gamma_{2}^{*}\left(\kappa^{\prime}\right)=\mu$. Since $g(\mu, \kappa)$ can be written as (see next paragraph) $\pi^{\prime} g\left(\gamma_{1}^{*}\left(\kappa^{\prime}\right), \kappa\right)+\left(1-\pi^{\prime}\right) g\left(\gamma_{2}^{*}\left(\kappa^{\prime}\right), \kappa\right)$, from strict concavity of $b$ and the definition of $g\left(\mu, \kappa^{\prime}\right)$ we conclude the desired inequality.

To see that $g(\mu, \kappa)=\pi^{\prime} g\left(\gamma_{1}^{*}\left(\kappa^{\prime}\right), \kappa\right)+\left(1-\pi^{\prime}\right) g\left(\gamma_{2}^{*}\left(\kappa^{\prime}\right), \kappa\right)$, note that $g(\mu, \kappa), g\left(\gamma_{1}^{*}\left(\kappa^{\prime}\right), \kappa\right)$, and $g\left(\gamma_{2}^{*}\left(\kappa^{\prime}\right), \kappa\right)$ are by definition convex combinations of $b^{*}\left(\gamma_{i}^{*}(\kappa)\right), i=1,2$ with the distribution over $\gamma_{i}^{*}(\kappa)$ such that they average average to the corresponding prior, i.e. $\pi \gamma_{1}^{*}(\kappa)+(1-$ $\pi) \gamma_{2}^{*}(\kappa)=\mu, \pi_{\gamma_{1}^{*}\left(\kappa^{\prime}\right)} \gamma_{1}^{*}(\kappa)+\left(1-\pi_{\gamma_{1}^{*}\left(\kappa^{\prime}\right)}\right) \gamma_{2}^{*}(\kappa)=\gamma_{1}^{*}\left(\kappa^{\prime}\right), \pi_{\gamma_{2}^{*}\left(\kappa^{\prime}\right)} \gamma_{1}^{*}(\kappa)+\left(1-\pi_{\gamma_{2}^{*}\left(\kappa^{\prime}\right)}\right) \gamma_{2}^{*}(\kappa)=\gamma_{2}^{*}\left(\kappa^{\prime}\right)$. These necessary conditions combined with the fact that $\pi^{\prime} \gamma_{1}^{*}\left(\kappa^{\prime}\right)+\left(1-\pi^{\prime}\right) \gamma_{2}^{*}\left(\kappa^{\prime}\right)=\mu$ leads to what is required for the equality of interest to hold.

## 3 Using Existing Data to Study Comparative Statics

In this section, we test our comparative static prediction for benefits by re-examining existing data. In the BDM experiment of CP , the benefits for accurate perception are varied because the maximum posted price varies, which allows us to examine if the level of misperception changes when the incentives to correctly perceive the game form change. Looking across maximum posted prices, we find evidence that misperception decreases when there are higher incentives to correctly perceive the game form, which is consistent with the theoretical prediction.

### 3.1 Offers in the CP Experiment

In the CP experiment, subjects were provided with a physical card that gave instructions for a seller version of a sealed-bid BDM mechanism. CP used the BDM to elicit the amount of money subjects would be willing to accept for the card itself, which could be exchanged later for $\$ 2$ if not sold. Thus, if the game form was perceived perfectly, the dominant strategy was for subjects to offer $\$ 2$ for their card.

After receiving the card, subjects provided an offer price for their card and then turned over the card to reveal a posted price and filled in their actual payments in light of this posted price. In a second round, subjects were then given a new card, also worth $\$ 2$, and completed the BDM again.

The minimum posted price was always $\$ 0$, but between subjects, the maximum draw for the posted price $\bar{p}$ varied between $\$ 4, \$ 5, \$ 6, \$ 7$, and $\$ 8$, and within subject, the maximum posted price $\bar{p}$ varied in the same way between cards.

As described in detail in the preceding section, our model predicts that offers will increase for some parameter values and decrease for others. CP report that the mean offer in their experiment is mostly increasing with the maximum draw $\bar{p}$, and for offers outside of $\$ 0.05$ of $\$ 2$, the mean offer is monotonically increasing in the value of the maximum draw $\bar{p}$. In a linear regression of offers onto the maximum posted price and round, the coefficient on maximum posted price is positive and statistically significant (coefficient $=0.216$, p-value $<0.001$ with robust standard errors clustered at the subject level). This means that a $\$ 1$ increase in the maximum posted price corresponds to an approximately $\$ 0.22$ increase in offers.

While average offers differ between maximum posted prices, the percent of subjects who offered within $\$ 0.05$ of $\$ 2$ does not appear to increase with the value of the maximum posted price $\bar{p}$. In a linear regression of a dummy variable indicating whether an offer is near $\$ 2$ onto the maximum posted price and round, the coefficient on maximum posted price is not statistically significant (coefficient $=0.004$, p-value $=0.783$ with robust standard errors clustered at the subject level). This suggests that changes in the benefits to accurate perception largely impact the intensive margin of perception, which helps to justify models of partial game form misperception.

### 3.2 Using Implied Beliefs to Measure Misperception

Assuming subjects misperceive game form, we can use their offers to determine their "implied" beliefs of how likely they are to be facing the BDM by inverting the optimal offer function (3). Specifically,

$$
\begin{equation*}
\gamma=\frac{\theta+\bar{p}-2 b}{\bar{p}-b} \tag{9}
\end{equation*}
$$



Figure 8: Offers and implied beliefs for the posted price ranges in CP.

Because $\bar{p}$ is in this equation, the same offer implies different beliefs depending on the range of posted prices. ${ }^{20}$ The map between implied beliefs and offers for the different posted price ranges used in the CP experiment is illustrated in Figure 8. The fact that the same offer corresponds to different implied beliefs for different maximum posted prices will play a central role in our subsequent analysis because when looking across posted price ranges, increases in offers (moving away from the dominant strategy) may not correspond to increases in misperception about the mechanism's extensive form.

Because beliefs are bounded between 0 and 1, this relationship places restrictions on the offers that are consistent with our model. ${ }^{21}$ Across posted price ranges, $75.9 \%$ of offers are consistent with our model for the first card and $79.5 \%$ are consistent with our model for the second card. ${ }^{22}$ These rates are not statistically different (two-sided p-value $=0.3402$ using a test of proportions), which means that we do not have evidence that experience increases consistency with our model.

[^10]
### 3.3 Comparative Statics for Benefits

As we have shown in our theoretical comparative static analysis, the average belief of the BDM should increase with $\bar{p}$. In other words, misperception should fall as benefits to accurate perception increase. We now use the implied beliefs of subjects in the CP experiment to test this prediction.

We find a clear trend in implied beliefs for consistent subjects. Looking just at consistent subjects, a linear regression of offers onto the maximum posted price and round, the coefficient on maximum posted price is once again both positive and statistically significant (coefficient $=0.032$, p-value $=0.022$ with cluster robust standard errors). This means that holding the round fixed, increasing the maximum posted price by $\$ 1$ corresponds to an approximately 3 percentage point increase in beliefs of the correct game form.

These results provide a key insight about mistakes in the BDM. Because choice mistakes are increasing in size with $\bar{p}$, it might seem that subjects are becoming less informed about the mechanism. However, our regression results suggest that misperception is actually decreasing with the maximum posted price.

As established previously, a higher maximum posted price should lead subjects to be more certain of game form, but the impact of changing the maximum posted price on offers is ambiguous. Using the map between beliefs and offers, we can determine whether both mistakes and misperception are increasing or whether mistakes are increasing despite a decrease in misperception.

## 4 Using New Data to Study Comparative Statics

In this section, we test our comparative static prediction for costs by running new experiments. To vary the cost of information, we use variation in the experimental protocol. Specifically, we implement both a replication of the CP experiment and a new protocol in which payoffs are explained contingency-by-contingency.

### 4.1 Our Replication of the CP Experiment

We first replace the CP experiment with a new subject pool and setting. Instead of undergraduate students completing the experiment in a classroom as in CP, our participants were all part of the Kellogg School of Management panel on Amazon's Mechanical Turk
(MTurk) and completed our experiment online. ${ }^{23}$ Subjects in this panel are required to be U.S. residents over the age of 18. Following Cavallo, Cruces, and Perez-Truglia (2017), we used many of the best practices identified in the literature for getting high quality responses when running studies on MTurk. ${ }^{24}$ Even with these methods, we received some offers well over the maximum posted price $\bar{p}$. These outliers could represent data entry errors and are not consistent with any of the explanations we considered, so we did not include offers over the maximum posted price in any of the analyses in this section. For our baseline replication, this reduced our sample by $3.6 \% .^{25}$ After removing these offers, our analysis sample for the baseline replication consisted of 190 subjects ( $57.4 \%$ Female; $M_{\text {age }}=38.9, S D_{\text {age }}=12.6$ ).

Because our subjects took our experiment online, we needed to make one substantial change to the design, which was to replace the card with a digital token. ${ }^{26}$ We made two other changes. First, we made the value of the token $\$ 1$ instead of $\$ 2$. Second, we told subjects the distribution from which posted prices were drawn. In our baseline replication, we chose to draw posted prices uniformly from $\$ 0$ to $\$ 3$ in increments of $\$ 0.50$. In practice, it is challenging to implement the BDM at finer increments because payments become cumbersome, but as a robustness check, we also run a version with increments of $\$ 0.01$.

We find that the baseline replication produces similar results to the first round of the CP experiment. To compare offers directly, we rescaled offers in the CP experiment by dividing them by 2 , as the value of the card is twice the value of the token. Given the differences in pools, settings, and design, it may be surprising that the results replicated so closely.

As shown in Table 1, the proportion of offers near $\$ 1$ is similar (17.6 for rescaled CP offers across posted prices ranges and 17.4 for our replication) and not statistically different using a two-tailed test of proportions ( p -value $=0.9603$ ). The mean of the rescaled CP offers across posted price ranges in the first round (1.70) is also similar to the mean of the offers in our baseline replication (1.63), and the distributions are not statistically different using a twotailed Wilcoxon rank-sum test ( p -value $=0.3764$ ). Figure 9 provides these two distributions side-by-side.

As mentioned previously, we also tested the robustness of our replication by having 92 subjects complete our replication but with posted prices drawn from $\$ 0.01$ increments. We find more offers between $\$ 0.50$ increments when posted prices are drawn from smaller increments, but as shown in Table 1, the percent of subjects offering near $\$ 1$ is similar ( $17.4 \%$

[^11]|  | Baseline replication | $\begin{gathered} \text { CP round 1 } \\ \mathrm{n} \\ \hline \end{gathered}$ | 1 Baseline=CP p-value |
| :---: | :---: | :---: | :---: |
| Observations | 190 | 245 |  |
| Mean offer | \$1.70 | \$1.63 | 0.3764 |
|  | (\$0.73) | (\$0.56) |  |
| Percent offer near \$1 | 17.4\% | 17.6\% | 0.9603 |
| Mean offer if not near \$1 | \$1 \$1.85 | \$1.76 | 0.0844* |
|  | (\$0.72) | (\$0.53) |  |
| (a) Baseline replication v. CP round 1. |  |  |  |
|  | Baseline replication | Robustness check | Baseline=Robustness |
|  |  |  | p-value |
| Observations <br> Mean offer | 190 | 92 | 0.1792 |
|  | \$1.70 | \$1.58 |  |
|  | (\$0.73) | (\$0.64) |  |
| Percent offer near \$1 | 17.4\% | 18.5\% | 0.8190 |
| Mean offer if not near \$1 | \$1.85 | \$1.71 | 0.0845* |
|  | (\$0.72) | (\$0.64) |  |

(b) Baseline replication v. robustness check.

Table 1: Summary statistics for our replications and rescaled offers from the CP experiment. Offers in the CP experiment are divided in half. An offer is "near $\$ 1$ " if it is within $\$ 0.05$ of $\$ 1$. Standard errors are in parentheses. p-values are from two-tailed tests of proportion for percents and from two-tailed Wilcoxon rank-sum tests otherwise. ${ }^{* * *} \mathrm{p}<0.01 ;{ }^{* *} \mathrm{p}<0.05$; * $\mathrm{p}<0.10$.


Figure 9: Offers in our baseline replication and scaled first round offers in CP pooled across posted price ranges.
for the baseline replication and $18.5 \%$ for the robustness check) and the distribution are not statistically different using a two-tailed Wilcoxon rank-sum test (p-value=0.1792). ${ }^{27}$

In practice, it is common to tell subjects the dominant strategy of a mechanism, with the hope of avoiding problems of misperception entirely. With this in mind, we ran an additional robustness check in which we informed subjects about the dominant strategy in the area just above the box where they entered their offers. Specifically, we stated: "The rule for selling the token is given below. It is a bit unusual, but its implications are straightforward. There is no way of gaming the rule, the BEST thing that you can do is to ask yourself how much you would be willing to exchange the token for, and then offer the number closest to that amount." For the 44 subjects who saw this statement, the average offer price (\$1.58) and the percent offering near $\$ 1(18.2 \%)$ were very similar to the corresponding figures for the robustness check without this statement. ${ }^{28}$ In a post-experiment questionnaire, $81.8 \%$ of subjects indicated that they trusted the statement, but $97.2 \%$ indicated that they decided to "read the rule anyway." If subjects are attempting to read the rule anyway, then it is possible that even in this treatment, subjects are misperceiving the mechanism and acting based on their fuzzy perception of the mechanism's game form.

[^12]
### 4.2 A New Contingent Protocol

To test the comparative static prediction, our goal is to make it "easier" for a subject to understand the BDM's extensive form. However, because our model does not provide an explicit map between experimental protocols and the cost of information, we draw inspiration from the literature, primarily the method of explaining payoffs contingency-by-contingency, as in Esponda and Vespa (2014).

In our "contingent" protocol, we frame each posted price as a separate computer bidder, one of which the subject will be paired with. Each bidder offers a different bid, and like the posted prices in the baseline replication, they are spaced in $\$ 0.50$ increments. The payoff rule is identical to the one in the baseline replication: if the offer (the "minimum amount you are willing to sell the token for") is at or below the computer's bid, the subject sells the token at the computer's bid. ${ }^{29}$

This protocol has similarities to methods for eliciting valuations that have been proposed in the experimental literature. It could be argued that the closest elicitation approach to ours is what Healy (2018) calls the "randomized binary choice" (RBC) elicitation mechanism. In this mechanism, subjects chose at each price whether they would like to sell at that price and a consistent switching point is enforced. For payment, a random choice is selected to be implemented. ${ }^{30}$ Mobius, Niederle, Niehaus, and Rosenblat (2011) and Coffman (2014) implement an RBC to elicit probabilities that has a very similar framing to ours. In their task, subjects chose which robots they would let choose for them (by selecting a threshold robot), where each robot had between a 1 and 100 chance of being correct (and a random robot was then selected). Like the RBC, our contingent protocol for the BDM highlights the payoffs for each contingency. While it could be argued that the RBC method for eliciting values is a simpler procedure, our contingent protocol is useful for studying misperception in the BDM , as it is closer to the CP protocol (because subjects set a price instead of choosing directly in which contingencies to sell).

### 4.3 Comparative Statics for Costs

Our analysis sample for the contingent protocol consisted of 192 subjects ( $59.4 \%$ Female; $M_{\text {age }}=39.3, S D_{\text {age }}=12.9$ ). As shown in Table 2, the average offer from these subjects was

[^13]

Figure 10: Offers in our baseline replication and new protocol.
$\$ 1.50$, which is $\$ 0.20$ lower than the average offer from subjects in the baseline replication. The distribution of offers for the contingent protocol and the baseline replication are provided in Figure 10, and these distributions are significantly different at a $5 \%$ level ( $\mathrm{p}=0.0016$ for a two-tailed Wilcoxon rank-sum test). Importantly, $34.9 \%$ of offers were near $\$ 1$ with the contingent protocol, which is roughly double the percentage in the baseline replication, and these proportions are significantly different at a $5 \%$ level ( $\mathrm{p}=0.0001$ for a two-tailed test of proportions). In addition, there is a much higher average implied belief for the contingent protocol. The average implied belief for consistent subjects increases by over 20 percentage points, and the average implied belief significantly different at a $5 \%$ level ( $\mathrm{p}<0.0001$ for a two-tailed Wilcoxon rank-sum test). Both of these findings - the decrease in average offers and the increase in average implied beliefs - are consistent with the theoretical comparative static prediction for the cost of information.

### 4.4 Other Explanations for Mistaken Offers

Because there are many possible explanations for mistakes in BDM experiments besides misperception, we use this section to investigate several of these possibilities. We examine not just their ability to explain the mistakes observed in the CP experiment and our experiments, but their ability to explain the difference in offers between our baseline replication and contingent protocol.

|  | Contingent <br> protocol | Baseline <br> replication | Contingent=Baseline <br> p-value |
| :--- | :---: | :---: | :---: |
| Observations | 192 | 190 |  |
| Mean offer | $\$ 1.50$ | $\$ 1.70$ | $0.0016^{* * *}$ |
| Percent offer near $\$ 1$ | $(\$ 0.70)$ | $(\$ 0.73)$ |  |
| Mean offer if not near $\$ 1$ | $34.9 \%$ | $17.4 \%$ | $0.0001^{* * *}$ |
|  | $\$ 1.76$ | $\$ 1.85$ | 0.1205 |
| Consistent with misperception | $(\$ 3.25)$ | $(\$ 2.97)$ |  |
| Mean belief of BDM (if consistent) | $78.6 \%$ | $70.0 \%$ | $0.0530^{*}$ |
|  | $(39.4 \%$ | $45.6 \%)$ | $(43.2 \%)$ |

Table 2: Summary statistics for our new protocol and baseline replication. An offer is "near $\$ 1$ " if it is within $\$ 0.05$ of $\$ 1$. Standard errors are in parentheses. p-values are from two-tailed tests of proportion for percents and from two-tailed Wilcoxon rank-sum tests otherwise. ${ }^{* * *} \mathrm{p}<0.01 ;{ }^{* *} \mathrm{p}<0.05 ;{ }^{*} \mathrm{p}<0.10$.

### 4.4.1 Other Behavioral Biases

A number of behavioral biases could impact the offers made in BDM experiments. ${ }^{31}$ For example, the endowment effect lead to higher offers in the BDM because ownership creates a reference point from which losses are experienced, and the direct benefits to ownership could lead to higher offers in the BDM because ownership creates positive feelings that make the good feel higher valued. With bad deal aversion, agents could set higher offers in the BDM as not to get a bad deal relative to a reference point when selling the good, and with the buy-low sell-high heuristic, they might follow an optimal rule from the world in which first offers are higher (in anticipation of future bargaining).

However, many of these behavioral theories do not predict a change in offers between our baseline replication of the CP BDM and our new protocol because we hold the extensive form, payoffs, and role fixed between the two protocols. For instance, behavioral biases based on ownership of the token, such as the endowment effect and the direct benefits of ownership such as positive feelings about the good, do not suggest a change in behavior, given that ownership is unaffected by our change in protocol. In addition, behavioral biases based on the maximum payment, such as anchoring on the maximum possible payoff and attraction to the maximum possible payoff, do not suggest a change, given that the distribution of posted prices (particularly the maximum posted price) does not change. Finally, behavioral biases based on being a seller, such as "bad deal" aversion and the "buy-low sell-high" heuristic,

[^14]do not suggest a change, given that the subject is still in the role of seller.
One behavioral bias that can explain the change in offers we see with our new protocol is framing effects. For instance, this protocol could induce a framing effect based on the repeated appearance of $\$ 1$ in the instructions, which could make offering $\$ 1$ a more salient action or even make $\$ 1$ the reference point.

### 4.4.2 Decision-Making Noise and All-Or-Nothing Misperception

In addition to these behavioral biases, mistakes could arise in the BDM if agents have decision-making noise unrelated to their perception of the game form. Further, the movement in offers towards the dominant strategy with our new contingent protocol could be explained with a reduction in such decision-making noise.

However, using maximum likelihood estimation (MLE), we find that a representative agent model of decision-making noise based on logit errors fits the data significantly better (in a statistical sense) if it also allows for partial game form recognition, as in our model. We also find that partial game form recognition fits the data significantly better than all-ornothing game form recognition. ${ }^{32}$

As in CP, we first examine the possibility that choice mistakes are driven purely by decision-making noise by using maximum likelihood to estimate the "noise" parameter $\lambda$ that best explains offers. This parameter is taken from the Quantal Response Equilibrium (QRE) approach in which the likelihood of taking an action takes the form of a multinomial logit. A feature of this approach is that the frequency of taking an action is increasing its relative payoff.

To nest several models, we will use the function $l_{i}(\gamma, \lambda)$, which is the likelihood of offer $b_{i}$ if the agent has belief of the $\operatorname{BDM} \gamma$ and has noise parameter $\lambda$. For the QRE specification,

$$
\begin{equation*}
\ln l_{i}(\gamma, \lambda)=\ln \frac{e^{\lambda E\left[\text { payoff for } \gamma \mid b_{i}\right]}}{\sum_{k \in K} e^{\lambda E\left[\text { payoff for } \gamma \mid b_{k}\right]}} \tag{10}
\end{equation*}
$$

where $E\left[\right.$ payoff for $\left.\gamma \mid b_{i}\right]=\frac{1}{\bar{p}}\left(\theta b_{i}+.5 \gamma \bar{p}^{2}+(1-\gamma) b_{i} \bar{p}-(1-.5 \gamma) b_{i}^{2}\right)$ and $K$ is the set of possible offers. ${ }^{33}$

Thus, to estimate $\lambda$ for a representative agent model with noise but no misperception, we set $\gamma=1$ and find

$$
\begin{equation*}
\arg \max _{\lambda} \sum_{i \in I} \ln l_{i}(1, \lambda)=\arg \max _{\lambda} \sum_{i \in I} \ln \frac{e^{\lambda E\left[\text { payoff for } \gamma=1 \mid b_{i}\right]}}{\sum_{k \in K} e^{\lambda E\left[\text { payoff for } \gamma=1 \mid b_{k}\right]}} \tag{11}
\end{equation*}
$$

[^15]where $b_{i}$ is the offer of subject $i$ and $I$ is set of subjects. We solve this problem using the Nelder-Mead method with 1,000 random started values, and standard errors were computed using 1,000 bootstrapping samples. As shown in Table 3, the parameter that best explains the data is 0.8040 .

CP also estimate an all-or-nothing model of game form misrecognition in which there is a probability $M$ that subjects believe they are facing the FPA. Specifically, they solve

$$
\begin{equation*}
\arg \max _{\lambda, M} \sum_{i \in I} \ln \left[(1-M) l_{i}(1, \lambda)+M l_{i}(0, \lambda)\right] \tag{12}
\end{equation*}
$$

We estimate their mixture model using the Nelder-Mead method with 1,000 random started values, and standard errors were computed using 1,000 bootstrapping samples. We estimate that $84.7 \%$ of subjects are playing as if they are facing a first price auction in the baseline replication and $36.9 \%$ in the contingent protocol. The estimate of $\lambda$ is higher than with just noise: rising from 0.8040 to 2.1864 , which means that less error is needed to explain the data ( $\lambda=0$ produces purely random choice).

To add partial game form recognition, we make one small change to the CP estimation. Instead of estimating a mixture between no game form misrecognition $\left(l_{i}(1, \lambda)\right)$ and full game form misrecognition $\left(l_{i}(0, \lambda)\right)$, we estimate a mixture between no game form misrecognition $\left(l_{i}(1, \lambda)\right)$ and partial game form misrecognition $\left(l_{i}(\gamma, \lambda)\right)$. By allowing for a representative posterior that captures uncertainty, we add another parameter to the model. Thus, we solve

$$
\begin{equation*}
\arg \max _{\lambda, M, \gamma} \sum_{i \in I} \ln \left[(1-M) l_{i}(1, \lambda)+M l_{i}(\gamma, \lambda)\right] \tag{13}
\end{equation*}
$$

With this, we can estimate a very simple representative agent version of our model, which can be interpreted as a "representative belief" model. In principle, we could consider richer models, such as a representative information structure with two $\gamma$ 's. In this case, $M$ could be interpreted as the probability of each belief.

We once again use the Nelder-Mead method with 1,000 random started values to perform our estimations, and compute standard errors using 1,000 bootstrapping samples. Using this mixture model, we find that $100 \%$ of subjects are classified as playing as if they are unsure of the mechanism in the baseline replication and the contingent protocol. The belief $\gamma$ that best explains the data is being $39.1 \%$ sure of the correct game form in the baseline replication and $66.2 \%$ in the contingent protocol. The estimate of $\lambda$ is once again higher than with just noise: rising from 0.8040 to 2.8786 , which means that less error is needed to explain the data. The levels of error needed for the CP mixture model and the $\gamma$ mixture model are not statistically different.

Because these models are nested (the noise-only model in the CP all-or-nothing mixture model and the CP all-or-nothing mixture model in the $\gamma$ mixture model), we can use a

|  | Noise-only model | CP mixture model | $\gamma$ mixture model |
| :--- | :---: | :---: | :---: |
| $\lambda$ | 0.8040 | 2.1864 | 2.8786 |
|  | $(0.4049)$ | $(0.9865)$ | $(0.5207)$ |
| $90 \%$ conf. interval | $[0.2197,1.5776]$ | $[1.5052,4.4470]$ | $[2.1203,3.8295]$ |
| $M$ |  | 0.8465 | 1.0000 |
|  |  | $(0.1550)$ | $(0.0000)$ |
| $90 \%$ conf. interval |  | $[0.5734,1.0000]$ | $[1.0000,1.0000]$ |
| $\gamma$ |  |  | 0.3914 |
|  |  |  | $(0.0972)$ |
| $90 \%$ conf. interval |  | -1.7929 | $[0.2194,0.5352]$ |
| Avg. log likelihood | -1.9322 | -340.6568 | -1.7693 |
| Log likelihood | -367.1215 | -336.1607 |  |

(a) Baseline replication

|  | Noise-only model | CP mixture model | $\gamma$ mixture model |
| :--- | :---: | :---: | :---: |
| $\lambda$ | 2.3937 | 3.8515 | 4.1224 |
|  | $(0.6101)$ | $(1.3467)$ | $(0.7113)$ |
| $90 \%$ conf. interval | $[1.6090,3.6029]$ | $[2.3677,6.7400]$ | $[3.1602,5.4883]$ |
| $M$ |  | 0.3693 | 1.0000 |
|  |  | $(0.0961)$ | $(0.0000)$ |
| $90 \%$ conf. interval |  | $[0.2521,0.5443]$ | $[1.0000,1.0000]$ |
| $\gamma$ |  |  | 0.6615 |
|  |  |  | $(0.0494)$ |
| $90 \%$ conf. interval |  | -1.8048 | $[0.5794,0.7445]$ |
| Avg. log likelihood | -1.8558 | -348.3308 | -1.7372 |
| Log likelihood | -358.1783 |  | -335.2868 |

(b) Contingent protocol

Table 3: Maximum likelihood estimates of different models. Standard errors are in parentheses.
likelihood ratio test to look for evidence of whether the fit of the model with more parameters is significantly better than the fit of the model with fewer parameters. With large samples, twice the difference in likelihoods should be distributed as a $\chi^{2}$ statistic with degrees of freedom equal to the difference in the number of parameters in the model. For one degree of freedom, the critical value for a significance level of $1 \%$ is 6.635 .

For the baseline replication, twice the difference in likelihoods between the CP mixture model and noise-only model is 52.9294 , which is well above 6.635 , so the CP mixture model has significantly better fit. For the CP mixture model and the $\gamma$ mixture model, twice the difference is 8.9922 , which is again above 6.635 , so the $\gamma$ mixture model has significantly better fit. For the contingent protocol, the twice the difference in likelihoods between the CP mixture model and noise-only model is 19.695 , so again the CP mixture model has significantly better fit. For the CP mixture model and the $\gamma$ mixture model, it is 26.088 , so here too the $\gamma$ mixture model has significantly better fit.

## 5 Conclusion and Discussion

In this paper, we take a standard model of imperfect perception - receiving a noisy mental signal of the environment - and use it to model an agent's misperception of the BDM mechanism. Our approach provides an as if representation for agents who may have trouble thinking through the mechanism's complex payoff rule.

The primary limitation of our approach is that the set of extensive forms that are confused with a given extensive form must be specified, which provides a significant modeling challenge and gives the modeler a substantial degree of freedom. However, for the BDM mechanism, the extensive form it might be confused with has been independently identified. Based on this external justification, we can explain most dominated offers in the CP experiment and measure the extent of misperception in their experiment.

We sharpen our model by assuming that mental signals are costly, and this generates comparative static predictions for the costs and benefits of more accurate perception. By reexamining the data from the CP experiment and generating data with new experiments, we are able to test these predictions, and we find that behavior is consistent with our theoretical predictions.

Because we leave the game form unchanged between our contingent protocol and baseline replication, the difference in mistakes between these experiments is plausibly related to a reduction in misperception. The remaining mistakes we observe with the contingent protocol could be due to lingering misperceptions or other behavioral biases. However, it should be noted that our model of misperception is compatible with other behavioral biases. For
example, our model could incorporate an endowment effect that increases the value of the good, which would generate an interaction between endowment effects and misperceptions.

## References

Ambuehl, Sandro, Axel Ockenfels, and Colin Stewart (2018). "Attention and Selection Effects". In: Working paper available at SSRN.
Ausubel, Lawrence M (2004). "An efficient ascending-bid auction for multiple objects". In: American Economic Review 94.5, pp. 1452-1475.
Avoyan, Ala and Andrew Schotter (2016). "Attention in games: An experimental study". In: Working paper.
Bartling, Björn, Florian Engl, and Roberto A Weber (2015). "Game form misconceptions are not necessary for a willingness-to-pay vs. willingness-to-accept gap". In: Journal of the Economic Science Association 1.1, pp. 72-85.
Brebner, Sarah and Joep Sonnemans (2018). "Does the elicitation method impact the WTA/WTP disparity?" In: Journal of behavioral and experimental economics 73, pp. 40-45.
Bull, Charles, Pascal Courty, Maurice Doyon, and Daniel Rondeau (2019). "Failure of the Becker-DeGroot-Marschak mechanism in inexperienced subjects: New tests of the game form misconception hypothesis". In: Journal of Economic Behavior \& Organization 159, pp. 235-253.
Caplin, Andrew and Mark Dean (2013). "Behavioral implications of rational inattention with shannon entropy". In: Working paper provided by NBER.

- (2015). "Revealed preference, rational inattention, and costly information acquisition". In: American Economic Review 105.7, pp. 2183-2203.
Caplin, Andrew, Mark Dean, and John Leahy (2018). "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice". In: The Review of Economic Studies, rdy037.
Caplin, Andrew and Daniel Martin (2015). "A testable theory of imperfect perception". In: The Economic Journal 125.582, pp. 184-202.
- (2017). "Defaults and attention: the drop out effect". In: Revue économique 5, pp. 747755.

Cason, Timothy N and Charles R Plott (2014). "Misconceptions and game form recognition: Challenges to theories of revealed preference and framing". In: Journal of Political Economy 122.6, pp. 1235-1270.
Cason, Timothy N, Tatsuyoshi Saijo, Tomas Sjöström, and Takehiko Yamato (2006). "Secure implementation experiments: Do strategy-proof mechanisms really work?" In: Games and Economic Behavior 57.2, pp. 206-235.
Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia (2017). "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments". In: American Economic Journal: Macroeconomics 9.3, pp. 1-35.
Charness, Gary and Dan Levin (2009). "The origin of the winner's curse: a laboratory study". In: American Economic Journal: Microeconomics 1.1, pp. 207-36.

Chen, Yan (2008). "Incentive-compatible mechanisms for pure public goods: A survey of experimental research". In: Handbook of experimental economics results 1, pp. 625-643.
Cheremukhin, Anton, Anna Popova, and Antonella Tutino (2011). "Experimental evidence on rational inattention". In: Working paper provided by Federal Reserve Bank of Dallas.
Coffman, Katherine Baldiga (2014). "Evidence on self-stereotyping and the contribution of ideas". In: The Quarterly Journal of Economics 129.4, pp. 1625-1660.
Compte, Olivier and Philippe Jehiel (2007). "Auctions and information acquisition: sealed bid or dynamic formats?" In: The Rand Journal of Economics 38.2, pp. 355-372.
Dean, Mark and Neligh (2017). "Experimental tests of rational inattention". In: Working paper.
Dewan, Ambuj and Neligh (2017). "Estimating information cost functions in models of rational inattention". In: Working paper.
Esponda, Ignacio and Emanuel Vespa (2014). "Hypothetical thinking and information extraction in the laboratory". In: American Economic Journal: Microeconomics 6.4, pp. 180202.

- (2017). "Contingent preferences and the sure-thing principle: Revisiting classic anomalies in the laboratory". In: Working paper.
Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2017). "Stochastic choice and optimal sequential sampling". In: Working paper.
Gentzkow, Matthew and Emir Kamenica (2014). "Costly persuasion". In: The American Economic Review 104.5, pp. 457-462.
Gershkov, Alex, Jacob K Goeree, Alexey Kushnir, Benny Moldovanu, and Xianwen Shi (2013). "On the equivalence of Bayesian and dominant strategy implementation". In: Econometrica 81.1, pp. 197-220.
Harstad, Ronald M (2000). "Dominant strategy adoption and bidders' experience with pricing rules". In: Experimental economics 3.3, pp. 261-280.
Hassidim, Avinatan, Déborah Marciano, Assaf Romm, and Ran I Shorrer (2017). "The Mechanism Is Truthful, Why Aren't You?" In: American Economic Review 107.5, pp. 22024.

Healy, Paul J (2018). "Explaining the BDM—Or any random binary choice elicitation mech-anism-To Subjects". In: Working paper.
Kagel, John H, Ronald M Harstad, and Dan Levin (1987). "Information impact and allocation rules in auctions with affiliated private values: A laboratory study". In: Econometrica: Journal of the Econometric Society, pp. 1275-1304.
Kalai, Ehud and Ehud Lehrer (1995). "Subjective games and equilibria". In: Games and Economic Behavior 8.1, pp. 123-163.
Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion". In: The American Economic Review 101.6, pp. 2590-2615.

Khaw, Mel Win, Luminita Stevens, and Michael Woodford (2017). "Discrete adjustment to a changing environment: Experimental evidence". In: Journal of Monetary Economics 91, pp. 88-103.
Li, Shengwu (2017). "Obviously strategy-proof mechanisms". In: American Economic Review 107.11, pp. 3257-87.

Martin, Daniel (2016). "Rational Inattention in Games: Experimental Evidence". In: Working paper available at SSRN.

- (2017). "Strategic Pricing with Rational Inattention to Quality". In: Games and Economic Behavior.
Martinez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa (2017). "Probabilistic States versus Multiple Certainties: The Obstacle of Uncertainty in Contingent Reasoning". In: Working paper provided by NBER.
Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green, et al. (1995). Microeconomic theory. Vol. 1. Oxford university press New York.
Matějka, Filip (2015). "Rigid pricing and rationally inattentive consumer". In: Journal of Economic Theory 158, pp. 656-678.
Matějka, Filip and Alisdair McKay (2015). "Rational inattention to discrete choices: A new foundation for the multinomial logit model". In: The American Economic Review 105.1, pp. 272-298.
Meier, Martin and Burkhard C Schipper (2014). "Bayesian games with unawareness and unawareness perfection". In: Economic Theory 56.2, pp. 219-249.
Milgrom, Paul (2004). Putting auction theory to work. Cambridge University Press.
Mobius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat (2011). "Managing self-confidence: Theory and experimental evidence". In:
Niederle, Muriel, Alvin E Roth, and Tayfun Sönmez (2008). "Matching". In: The New Palgrave Dictionary of Economics. Palgrave Macmillan 2.
Oechssler, Jörg and Burkhard Schipper (2003). "Can you guess the game you are playing?" In: Games and Economic Behavior 43.1, pp. 137-152.
Pathak, Parag A and Tayfun Sönmez (2008). "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism". In: American Economic Review 98.4, pp. 163652.

Ravid, Doron (2014). "Bargaining with rational inattention". In:
Rees-Jones, Alex and Samuel Skowronek (2018). "An experimental investigation of preference misrepresentation in the residency match". In: Proceedings of the National Academy of Sciences 115.45, pp. 11471-11476.
Sims, Christopher A (2003). "Implications of rational inattention". In: Journal of monetary Economics 50.3, pp. 665-690.
Woodford, Michael (2014). "Stochastic choice: An optimizing neuroeconomic model". In: American Economic Review 104.5, pp. 495-500.

Yang, Ming (2015). "Coordination with flexible information acquisition". In: Journal of Economic Theory 158, pp. 721-738.

## 6 Appendix

### 6.1 Discrete Offers

In this section we derive the optimal offer strategy when offers are discrete, which is often the case in practice. First we examine the simple case of unitary increments in offers (in the units of $\theta$ ) and then we generalize to increments of size $z$.

### 6.1.1 Unitary Increments

We first derive the optimal strategy for the BDM case when offers are discrete. As in the continuous model, we assume that agent sells as long as the posted price $p$ is at least equal to the offered price $b$. Offering the own valuation $\theta$ is a weakly dominant strategy. $\theta+1$ yields to the same payoff in expectation.

To see that $\theta$ is weakly dominant, consider a fixed offer price $b \leq \theta-1$, and compare it to the payoff of offering $\theta$. The agent gets the same payoff as offering $\theta$ if the realization of $p$ lies in $[0, b-1]$ (in which case gets $\theta$ ) or in $[\theta, \bar{p}]$ (in which case gets $p$ ). For the interval $[b, \theta-1]$, the payoff of offering $\theta$ dominates the payoff of offering $b$. On the other hand, offering $b \geq \theta+2$ there is a positive chance of making a negative payoff, so it is dominated by offering $\theta$. It is easy to see that $\theta+1$ yields the same payoff in expectation, by writing the payoff and manipulating the expression.

On the other hand, in a FPA with discrete offers, the agent's expected payoff from offering $b$ is

$$
\begin{array}{rlrl}
\sum_{p \in\{0,1, \ldots, \bar{p}\}} \theta \mathbb{1}[p<b]+b \mathbb{1}[p \geq b] & =\theta & & \operatorname{Prob}[p<b]+b \quad \operatorname{Prob}[p \geq b] \\
& =\theta & \operatorname{Prob}[p \leq b-1]+b \quad(1-\operatorname{Prob}[p \leq b-1])
\end{array}
$$

Because $p$ is uniformly distributed on the finite set $\{0,1, \ldots, \bar{p}\}$ the former expression becomes

$$
\theta \frac{b}{\bar{p}+1}+b\left(1-\frac{b}{\bar{p}+1}\right)
$$

Since $\bar{p}+1>0$, we are interested in finding $b$ s.t. $b(\theta+\bar{p}+1)-b^{2}$ is maximal. This can be rewritten as

$$
-\left(b-\frac{\theta+\bar{p}+1}{2}\right)^{2}+\frac{(\theta+\bar{p}+1)^{2}}{4}
$$

which is a quadratic function of $b$ that reaches the maximum at $b^{*}$, with

$$
b^{*}= \begin{cases}\frac{\theta+\bar{p}+1}{2}, & \text { if } \theta+\bar{p}+1 \text { is even } \\ \left\{\frac{\theta+\bar{p}}{2}, \frac{\theta+\bar{p}}{2}+1\right\}, & \text { if } \theta+\bar{p}+1 \text { is odd }\end{cases}
$$

With uncertainty about the game form, the expected payoff for offering $b$ when agent assigns probability $\gamma$ to the game form being the BDM is,

$$
\theta \operatorname{Prob}[p<b]+E_{p}[(\gamma p+(1-\gamma) b) \mathbb{1}[p \geq b]]
$$

which reduces to

$$
\theta \frac{b}{\bar{p}+1}+\gamma E_{p}[p \mathbb{1}[p \geq b]]+(1-\gamma) b\left(1-\frac{b}{\bar{p}+1}\right)
$$

Since $E_{p}[p \mathbb{1}[p \geq b]]=\sum_{i=b}^{\bar{p}} \frac{i}{\bar{p}+1}=\frac{1}{2(\bar{p}+1)}[\bar{p}(\bar{p}+1)-(b-1) b]$, we have

$$
\theta \frac{b}{\bar{p}+1}+\gamma \frac{1}{2(\bar{p}+1)}\left(\bar{p}^{2}-b^{2}\right)+\gamma \frac{1}{2(\bar{p}+1)}(b+\bar{p})-(1-\gamma) \frac{b^{2}}{\bar{p}+1}+(1-\gamma) b
$$

rearranging,

$$
-\left(\frac{1-\gamma}{\bar{p}+1}+\frac{\gamma}{2(\bar{p}+1)}\right) b^{2}+\left(\frac{\theta}{\bar{p}+1}+\frac{\gamma}{2(\bar{p}+1)}+1-\gamma\right) b+\frac{\gamma}{2(\bar{p}+1)}\left(\bar{p}^{2}+\bar{p}\right)
$$

since this is a concave function in $b$, we always have interior solution. By symmetry, it happens at either $K=\frac{1}{2} \frac{2 \theta+\gamma+2(1-\gamma)(\bar{p}+1)}{2-\gamma}=\frac{\theta+\bar{p}-\gamma \bar{p}}{2-\gamma}+\frac{1}{2}$ if $K$ is integer ${ }^{34}$; otherwise, at the closest integer. More explicitly,

$$
b^{*}= \begin{cases}\lfloor K\rceil, & \text { if } K-\lfloor K\rfloor \neq 0.5  \tag{14}\\ \{\lfloor K\rfloor,\lceil K\rceil\}, & \text { otherwise }\end{cases}
$$

Observe that if $\gamma=0,1$ we obtain the expected optimal offers for the BDM and FPA above derived.

### 6.1.2 Fractional Increments

We now consider the case where offers on a grid of mesh $z$ on $[0, \bar{p}]$ are allowed. In other words, agents can offer $0, z, 2 z, 3 z, \ldots, \bar{p}$.

This case is equivalent to solving the problem with unitary increments on $\left[0, \frac{\bar{p}}{z}\right]$ and realized type $\frac{\theta}{z}$ which yields the optimal offer $b^{*}$ as in (14). In the original game, the optimal offer corresponds then to the transformation $b_{z}^{*}=z b^{*}$.

More explicitly, for the problem with mesh $z$, define

$$
K^{\prime}=\frac{\frac{\theta}{z}+\frac{\bar{p}}{z}-\gamma \frac{\bar{p}}{z}}{2-\gamma}+\frac{1}{2}
$$

[^16]Hence, the optimal offer price is given by

$$
b_{z}^{*}= \begin{cases}z\left\lfloor K^{\prime}\right\rceil, & \text { if } K^{\prime}-\left\lfloor K^{\prime}\right\rfloor \neq 0.5, \\ \left\{z\left\lfloor K^{\prime}\right\rfloor, z\left\lceil K^{\prime}\right\rceil\right\}, & \text { otherwise } .\end{cases}
$$

Observe also that as $z$ goes to 0 , the solution converges to the continuous case.

### 6.2 Alternative Game Forms

In this section, we show how the optimal offer strategy changes as we change the game form agents confuse with the BDM. Some of these alternative game forms were proposed by CP after observing the distribution of offers, but were deemed to be unlikely.

As these optimal offers show, the testable implications of our approach depend on the alternative game form that is selected. For instance, in some cases only offers at or below $\theta$ are consistent, and in other cases only offers at or above $\theta$ are consistent. Some cases even allow offers to be above the maximum posted price $\bar{p}$.

| Alternative payment rule | Expected payoff | Map between offer and beliefs |
| :---: | :---: | :---: |
| Posted price $p$ is paid, regardless of winning | $E[\gamma(\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b])+(1-\gamma) p]$ | $b^{*}(\gamma)= \begin{cases}\theta, & \text { if } \gamma \in(0,1], \\ x \in \mathbb{R}, & \gamma=0 .\end{cases}$ |
| Offer $b$ is paid, regardless of winning | $E[\gamma(\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b])+(1-\gamma) b]$ | $b^{*}(\gamma)= \begin{cases}\theta+\frac{1-\gamma}{\gamma} \bar{p} & \text { if } \gamma \in(0,1] \\ \infty & \gamma=0 .\end{cases}$ |
| Paid theta only if offer is below posted price, nothing otherwise | $E[\gamma(\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b])+(1-\gamma) \theta \mathbb{1}[p \geq b]]$ | $b^{*}(\gamma)= \begin{cases}\frac{\theta}{\gamma} & \text { if } \gamma \in(0,1] \\ 0 & \text { if } \gamma=0\end{cases}$ |
| Paid offer only if offer is below posted price, nothing otherwise | $E[\gamma(\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b])+(1-\gamma) b \mathbb{1}[p \geq b]]$ | $b^{*}(\gamma)=\frac{\gamma \theta+(1-\gamma) \bar{p}}{2-\gamma}$ |
| Paid posted price if offer below posted price, nothing otherwise | $E[\gamma(\theta \mathbb{1}[p<b]+p \mathbb{1}[p \geq b])+(1-\gamma) p \mathbb{1}[p \geq b]]$ | $b^{*}(\gamma)=\gamma \theta$ |

### 6.3 Additional Robustness Check: Within-Subject Comparisons

Because our robustness check was based on variation between subjects, we ran an additional robustness check in which 137 subjects first completed the BDM with $\$ 0.01$ increments
and then either our baseline replication (with $\$ 0.50$ increments) or our contingent protocol, separated by 5 months. This design allows us to perform a within-subject analysis of the impact of changes in the protocol. However, there is the possibility that learning occurred between rounds, so we do not use this as our primary robustness check and we do not include these subjects in any of our other analyses.

For the 74 subjects who completed the BDM with $\$ 0.01$ increments and our baseline replication, there is an increase in offers of 0.081 on average, but this is not statistically different from 0 using a two-tailed t -test ( p -value $=0.3843$ ). On the other hand, because our baseline replication has more offer prices at increments of $\$ 0.50$, there is actually a $9.5 \%$ increase in the probability of offering near the value of the token, and this is statistically different from 0 ( p -value $=0.0073$ ).

On the other hand, for the 63 subjects who completed the BDM with $\$ 0.01$ increments and our contingent protocol, there is a decrease in offers on average (a drop of 0.075), but this is also not statistically different from zero using a two-tailed $t$-test ( $p$-value $=0.5588$ ). At the same time, there is a much larger increase in the probability of offering near the value of the token with this protocol (23.8\%), which is highly statistically different from a $0 \%$ increase ( p -value $<0.0001$ ).

### 6.4 MLE Results for the CP Experiment

Here we apply the MLE strategy introduced in the body of the text to the two rounds of the CP experiment, and the results are similar to our results for the baseline replication and contingent protocol.

First, we estimate the noise-only model and mixture model from CP using the NelderMead method with 1,000 random started values. ${ }^{35}$ Standard errors were computed using 1,000 bootstrapping samples. We find that $65.0 \%$ of subjects are classified as playing as if they are facing a first price auction on the first card and $41.1 \%$ on the second card.

Next, we estimate the representative agent version of our model, and we find that $84.4 \%$ of subjects are classified as playing as if they are unsure of the mechanism on the first card and $87.8 \%$ on the second card. While these rates are similar, the belief $\gamma$ that best explains the data is being $41.1 \%$ sure of the correct game form on the first card and $61.9 \%$ on the second card.

Because these models are nested, we can use a likelihood ratio test to look for evidence of whether fits are significantly better. With large samples, twice the difference in likelihoods should be distributed as a $\chi^{2}$ statistic with degrees of freedom equal to the difference in

[^17]the number of parameters in the model. For one degree of freedom, the critical value for a significance level of $1 \%$ is 6.635 , so for both rounds, the CP mixture model has significantly better fit than the noise-only model, and the $\gamma$ mixture model has significantly better fit than the CP mixture model.

| $\lambda$ | Noise-only model | CP mixture model | $\gamma$ mixture model |
| :---: | :---: | :---: | :---: |
|  | 0.9833 | 4.4932 | 3.9626 |
|  | (0.1609) | (0.9398) | (1.0408) |
| 90\% conf. interval | [0.7399,1.2649] | [3.2787,6.4129] | [3.1459,6.4437] |
| M |  | 0.6499 | 0.8443 |
|  |  | (0.0465) | (0.0798) |
| 90\% conf. interval |  | [0.5756,0.7302] | [0.7048,0.9704] |
| $\gamma$ |  |  | 0.4107 |
|  |  |  | (0.0870) |
| 90\% conf. interval |  |  | [0.2358,0.5244] |
| Avg. log likelihood | -4.0228 | -3.8055 | -3.7608 |
| Log likelihood | -985.5835 | -932.3379 | -921.4038 |

(a) Round 1 ( $\mathrm{N}=245$ )

|  | Noise-only model | CP mixture model | $\gamma$ mixture model |
| :--- | :---: | :---: | :---: |
| $\lambda$ | 1.2628 | 3.4536 | 2.6909 |
|  | $(0.2646)$ | $(1.1599)$ | $(0.5375)$ |
| $90 \%$ conf. interval | $[0.8928,1.7440]$ | $[2.1001,6.0094]$ | $[2.0811,3.6415]$ |
| $M$ |  | 0.4107 | 0.8781 |
|  |  | $(0.0561)$ | $(0.1052)$ |
| $90 \%$ conf. interval |  | $[0.3253,0.5099]$ | $[0.6968,1.0000]$ |
| $\gamma$ |  |  | 0.6188 |
|  |  |  | $(0.0711)$ |
| $90 \%$ conf. interval |  | -3.9198 | $[0.4947,0.7068]$ |
| Avg. log likelihood | -4.0013 | -948.5934 | -3.8824 |
| Log likelihood | -968.3104 | -939.5417 |  |

(b) Round $2(\mathrm{~N}=242)$

Table 4: Maximum likelihood estimates of different models. Standard errors are in parentheses.

### 6.5 Experimental Instructions

## Baseline Replication

You have one digital token that is worth $\$ 1$ to you. You can sell it. Name your offer price: \$ $\qquad$
A "posted" price will be drawn randomly between $\$ 0$ and $\$ 3$ (in increments of $\$ 0.50$ ). Every possible posted price ( $\$ 0, \$ 0.50, \$ 1, \$ 1.50, \$ 2, \$ 2.50$, and $\$ 3$ ) has an equal chance of being selected.

If your offer price is at or below the posted price, then you sell your token at the posted price.

If your offer price is above the posted price, then you do not sell your token, but you do collect the $\$ 1$ value of the token.

## Contingent Protocol

You have one digital token that can be exchanged for $\$ 1$ or sold to a computer bidder.
Here is how your bonus payment is determined:

1. You will name the minimum amount you are willing to sell your token (it might sell for more).
2. We will randomly select 1 of the 7 computer bidders listed below (each is equally likely to be selected).
3. You will keep your token if the minimum amount you are willing to sell it is above the computer's bid. Otherwise, you will sell your token at the computer's bid.

| Bidder | Bid | If your minimum amount is above their bid | Otherwise |
| :--- | :---: | :---: | :---: |
| A | $\$ 0$ | $\$ 1$ | $\$ 0$ |
| B | $\$ 0.50$ | $\$ 1$ | $\$ 0.50$ |
| C | $\$ 1$ | $\$ 1$ | $\$ 1$ |
| D | $\$ 1.50$ | $\$ 1$ | $\$ 1.50$ |
| E | $\$ 2$ | $\$ 1$ | $\$ 2$ |
| F | $\$ 2.50$ | $\$ 1$ | $\$ 2.50$ |
| G | $\$ 3$ | $\$ 1$ | $\$ 3$ |

Name the minimum amount you are willing to sell the token for: $\$$ $\qquad$

## Robustness Check

You have one digital token that is worth $\$ 1$ to you. You can sell it. Name your offer price: $\$$ $\qquad$
A "posted" price will be drawn randomly between $\$ 0$ and $\$ 3$ (in increments of $\$ 0.01)$. Every possible posted price has an equal chance of being selected.

If your offer price is at or below the posted price, then you sell your token at the posted price.

If your offer price is above the posted price, then you do not sell your token, but you do collect the $\$ 1$ value of the token.

## Robustness Check (with Statement)

You have one digital token that is worth $\$ 1$ to you. You can sell it.
The rule for selling the token is given below. It is a bit unusual, but its implications are straightforward. There is no way of gaming the rule, the BEST thing that you can do is to ask yourself how much you would be willing to exchange the token for, and then offer the number closest to that amount.

Name your offer price: $\$$ $\qquad$
A "posted" price will be drawn randomly between $\$ 0$ and $\$ 3$ (in increments of $\$ 0.01)$. Every possible posted price has an equal chance of being selected.

If your offer price is at or below the posted price, then you sell your token at the posted price.

If your offer price is above the posted price, then you do not sell your token, but you do collect the $\$ 1$ value of the token.


[^0]:    *We thank Ignacio Esponda, Antonio Rangel, and Emanuel Vespa for excellent suggestions about experimental design, Kellogg Research Support (particularly Mac Abruzzo and Kat Baker) for valuable assistance with our experiments, and Andrew Caplin, Tim Cason, Mark Dean, P.J. Healy, Philippe Jehiel, Peter Klibanoff, Shengwu Li, Sanket Patil, Nicola Persico, Charlie Plott, and Alvaro Sandroni for insightful feedback.
    ${ }^{\dagger}$ Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, dmartin@kellogg.northwestern.edu.
    ${ }^{\ddagger}$ Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, edwin.munoz@kellogg.northwestern.edu.

[^1]:    ${ }^{1}$ In addition, dominant strategy implementation of social choice functions is "robust to changes in agents' beliefs and does not rely on the assumptions of a common prior and equilibrium play" (Gershkov, Goeree, Kushnir, Moldovanu, and Shi 2013).
    ${ }^{2}$ The other barriers they identify are mistrust of the market maker (e.g., players doubt the recommendation to play a dominant strategy), self-selection into the mechanism, and non-standard utility functions.
    ${ }^{3}$ See section 3.1 for a description of the BDM mechanism.

[^2]:    ${ }^{4}$ For example, the endowment effect, positive feelings from ownership, anchoring or attraction to the maximum possible payoff, bad deal aversion, and the buy-low sell-high heuristic

[^3]:    ${ }^{5}$ For example, noisy mental signals are central to signal detection theory, the drift diffusion model, and rational inattention theory. In the economics literature, cognition is modeled with noisy mental signals in Woodford (2014), Caplin and Martin (2015), Caplin and Dean (2015), Matějka and McKay (2015), and Fudenberg, Strack, and Strzalecki (2017). As in these papers, our approach takes an as if perspective.
    ${ }^{6}$ Although we do not explore it here, a Bayesian framework also allows for learning about game form with repeated play, as in Subjective Games (see Kalai and Lehrer 1995 and Oechssler and B. Schipper 2003).
    ${ }^{7}$ In addition, the set of extensive forms an agent confuses with a given extensive form are potentially

[^4]:    testable. For instance, in the appendix we show how the set of optimal actions varies as the set of alternatives to the BDM varies, which produces testable content.
    ${ }^{8}$ We say subjects have made a "mistake" if a different offer would have increased their expected payoff.
    ${ }^{9}$ Mobius, Niederle, Niehaus, and Rosenblat (2011) and Coffman (2014) implement a BDM designed to elicit probabilities that also focuses attention on contingencies.

[^5]:    ${ }^{10}$ In the solution concept Rationally Inattentive Cursed Equilibrium (RICE) introduced in Martin (2017), players trade off the costs and benefits of improving on their "cursed" strategic beliefs.
    ${ }^{11}$ Relatedly, Avoyan and Schotter (2016) use experiments to study the allocation of attention across games when an agent faces more than one game at a time.
    ${ }^{12}$ See both the results and references presented in Esponda and Vespa (2014), Charness and Levin (2009), and Esponda and Vespa (2017), as well as footnote 1 in Martinez-Marquina, Niederle, and Vespa (2017).

[^6]:    ${ }^{13}$ We assume that agents are aware of the set of possible extensive forms, which rules out interesting possibilities for unawareness (such as those considered by Meier and B. C. Schipper 2014).

[^7]:    ${ }^{14}$ The action set is identical in both states of the world, as well as the information sets. The only difference is the way an agent believes payoffs are determined.
    ${ }^{15}$ In the appendix, we show how the optimal strategy changes if we assume instead that subjects confuse the BDM with other game forms besides the FPA. Some of these alternative game forms were proposed as possibilities in footnote 19 of CP .
    ${ }^{16}$ If we could elicit a subject's beliefs about the game form they think they are facing, this would produce strong testable predictions for our model. However, such beliefs are challenging to elicit directly because of their probabilistic nature. See Bartling, Engl, and Weber (2015) for an example where deterministic beliefs about the BDM are elicited.
    ${ }^{17}$ This differs from Compte and Jehiel (2007), who study the impact of an agent acquiring information about their valuations, not the game form itself.

[^8]:    ${ }^{18}$ In the framework of Gentzkow and Kamenica (2014), the reference belief does not need to coincide with the prior. Instead, it can be any fixed interior belief against which the cost of an information structure is assessed. In our case, the reference belief happens to coincide with the prior, which is inconsequential because the optimal information structure is independent of the fixed reference belief.

[^9]:    ${ }^{19}$ We condition the state on when agents face the BDM because our interest is in the comparative statics of optimal misperception when subjects are facing the BDM.

[^10]:    ${ }^{20}$ Because increments of the posted price are small (\$0.01), the discretized version of this equation is virtually indistinguishable. In practice, the difference in the continuous and discrete versions is less than half a percentage point, so we use the continuous version throughout our analysis.
    ${ }^{21}$ We allow a 5 percentage point margin, so implied beliefs between $-0.05 \%$ and $1.05 \%$ sure of the correct game form are considered consistent with our model.
    ${ }^{22}$ If an agent was to chose offers randomly between $\$ 0$ and $\bar{p}$, then approximately $25 \%$ of offers would be consistent with our model when $\bar{p}=4$, and approximately $38 \%$ would be consistent when $\bar{p}=8$.

[^11]:    ${ }^{23}$ Bull, Courty, Doyon, and Rondeau (2019) also replicate the CP experiment, but with university students and in a classroom as in CP.
    ${ }^{24}$ This included requiring subjects to complete a standard attentional check question at the start of the experiment, which was passed by $97.0 \%$ of subjects who started the baseline replication.
    ${ }^{25}$ For technical reasons, CP also removed offers over $\bar{p}$ in their maximum likelihood estimation procedure, but this was just $0.8 \%$ of second round offers.
    ${ }^{26}$ See the appendix for the instructions for this replication.

[^12]:    ${ }^{27}$ We also ran an robustness check in which 137 subjects first completed our robustness check and then either our baseline replication or our contingent protocol, and those results appear in the appendix.
    ${ }^{28}$ For this additional robustness check, we also used $\$ 0.01$ increments in the distribution of posted prices.

[^13]:    ${ }^{29}$ The exact instructions are provided in the appendix.
    ${ }^{30}$ Bartling, Engl, and Weber (2015) use a "multiple price list" RBC to elicit willingness-to-pay (WTP) and willingness-to-accept (WTA) for both money (as in CP) and goods. They also use a deterministic belief question to assess whether subjects understand the payoffs for money and find that subjects who appear to understand the payoffs for money still exhibit WTP-WTA disparities for the good. Brebner and Sonnemans (2018) compare choices between the BDM and the multiple price list RBC and find that the WTP-WTA gap is similar between for the protocols they use.

[^14]:    ${ }^{31} \mathrm{CP}$ and Brebner and Sonnemans (2018) provide a comprehensive and detailed discussion of the possible explanations for mistakes in the BDM.

[^15]:    ${ }^{32}$ In the body of the text, we present the results for the baseline replication and contingent protocol, and in the appendix, we show that the same holds for the data from the CP experiment.
    ${ }^{33}$ We follow CP by discretizing the space of offers, but because the increment of posted prices is $\$ 0.50$, we use $\$ 0.50$ instead of $\$ 0.10$.

[^16]:    ${ }^{34} K$ is decreasing in $\gamma$. Hence a tight upper bound for $K$ is $\bar{p}+\frac{1}{2}$, reached when $\theta=\bar{p}$. In this case, it is optimal to offer either $\bar{p}$ or $\bar{p}+1$, the latter interpreted as no selling, which was a priori obvious.

[^17]:    ${ }^{35}$ For comparability, we follow CP by removing all offers over $\bar{p}$ and rounding offers to $\$ 0.10$.

