Journal of Economic Behavior and Organization 159 (2019) 235–253

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Failure of the Becker–DeGroot–Marschak mechanism in inexperienced subjects: New tests of the game form misconception hypothesis☆

Charles Bull[a], Pascal Courty[b], Maurice Doyon[c], Daniel Rondeau[d,*]

[a] *Camosun College, Canada*
[b] *University of Victoria and CEPR, Canada*
[c] *Université Laval, Canada*
[d] *University of Victoria, Canada*

ARTICLE INFO

Article history:
Received 3 August 2017
Revised 29 December 2018
Accepted 6 January 2019

JEL classification:
C8
C9

Keywords:
Experiment
Game form recognition
Game form misconception
Mistake
Becker–DeGroot–Marschak
Preference elicitation

ABSTRACT

Cason and Plott (2014, hereafter CP) conclude that sub-optimal behavior in the (second price) Becker-DeGroot-Marschak mechanism (BDM) is consistent with the hypothesis that a significant proportion of subjects misconceive the BDM as a first price auction. We broadly replicate CP's results, formalize a game form recognition theory for the analysis of treatment effects, and explore the robustness of CP's conclusions across four treatments. We conclude that the pattern of misconception that explains the BDM data cannot simultaneously explain observed choices in closely related treatments.

1. Introduction

The Becker–DeGroot–Marschak mechanism (Becker et al., 1964, hereafter BDM) is a simple mechanism designed to elicit valuation. Suppose that an individual has a private value V for an object. In a procurement BDM, the individual first submits an offer price to sell the object. A random price is then drawn from a given distribution. If the offer price is less than or equal to the random price, the individual sells the item at the random price. Otherwise, the individual obtains the reservation value V.

☆ This research was sponsored by the Ministère des Ressources Naturelles et de la Faune du Québec. The analysis and opinions presented are exclusively those of the authors and errors remain our sole responsibility. We thank Tim Cason for helpful comments, and presentation audiences at IMEBESS and the CEA annual conference. We are responsible for mistakes and any misconceptions.

* Corresponding author.
E-mail addresses: bullc@camosun.bc.ca (C. Bull), pcourty@uvic.ca (P. Courty), maurice.doyon@eac.ulaval.ca (M. Doyon), rondeau@uvic.ca (D. Rondeau).

https://doi.org/10.1016/j.jebo.2019.01.003
0167-2681/

The BDM is incentive compatible under the assumptions of expected utility theory.[1] Yet, it has been broadly recognized that subjects who have not been trained in the use of the BDM do not typically bid optimally.[2] Attempts at explaining sub-optimal behavior in the BDM have led researchers to postulate a number of explanations that rely on non-standard preferences, irrationality, or theories of framing. After reviewing these possible explanations, Plott and Zeiler (2005) report that an anomalous gap between willingness to pay and willingness to accept (repeatedly observed since the paper by Kahneman et al., 1990) can be turned on and off by varying experimental procedures, indicating that the source of the anomalous behavior cannot be inherent to subjects' preferences. Following Chou et al. (2009), they propose that some subjects may misconceive the task in front of them. Without codifying a formal definition of misconception, Plott and Zeiler discuss five different channels through which it might arise.[3]

Cason and Plott (2014, hereafter CP) seek to develop a conceptual framework to define and analyze Game Form Recognition (GFR). The theory holds that participants in an economic transaction may confuse the game they are playing for other games. From the instructions they are given, some subjects may form an incorrect understanding of their choice set or how their choices translate into payoffs. They make systematic Game Form Misconception (GFM) errors. GFM offers a possible explanation for deviations from predicted behavior without the need to invoke non-standard preferences, irrationality, or theories of framing.

In a demonstration of their framework, CP run a procurement BDM similar to the one described above. Untrained and inexperienced subjects have an opportunity to sell a plain paper card with a redeemable value $V = \$2$. CP hypothesize that some subjects correctly recognize the second price game form of the BDM, while others mistakenly take the game to be a first price (FP) procurement auction (against one other random price). Subjects who properly recognize the BDM are hypothesized to optimally bid $V = \$2$. Subjects who suffer the FP misconception are expected to submit the optimal amount for a risk neutral participant in the equivalent FP auction. While CP note the possibility that some subjects could be differently misconceived or make other types of errors, CP conclude that the misconception hypothesis provides a good fit for their data.

This conclusion is the starting point of this paper. We pursue three objectives: 1) to test whether or not CP's result can be replicated; 2) to develop a formalization of the theory of GFR in order to provide explicit structure for its application to experiments with multiple treatments; and 3) to explore the robustness of CP's characterization of BDM players by presenting new evidence obtained from three additional treatments. Each of these objectives makes a distinct contribution to the literature.

Our deployment of CP's single treatment at a different University in a different country produces a broad replication of their published findings. The replication is, in itself, a test of robustness to a different set of players. It is also necessary for a meaningful extension of the analysis to additional treatments. It is also a positive result given repeated calls within the discipline to demonstrate that economic research is replicable (see Binmore and Shaked, 2010; Camerer et al., 2016; Camerer et al., 2018; Chang and Li, 2017; Duvendack et al., 2017; and Maniadis et al., 2014, among others).

Testing for treatment effects when subjects are suspected of harboring GFM can result in the presence of confounding effects. This makes researchers' tasks far more challenging and calls for a formalization of the GFM framework. We propose a definition of GFM as a mapping from a set of instructions to a probability distribution over player types. A player's type corresponds to the game that the player believes she is playing (correctly or not). Under our construct, the robustness of a GFM finding can be defined as an invariance of the mapping to treatment variations. The theory makes all assumptions explicit, providing avenues for separating incentive effects from variations in the distribution of subject types. This explicit theory is the methodological contribution of this paper. It should help researchers design more pointed experiments and readers better evaluate the results.

The empirical contribution of this paper is the finding that despite a broad validation of CP's characterization of participants in the single BDM replication treatment, the result is not robust to treatment variations. Several observations lead us to this conclusion.

i) Contrary to theoretical predictions, subjects suspected of first price misconception do not respond to an increase in the number of random prices they must "beat" in order to sell the card.
ii) We fail to statistically detect a clear group of optimally bidding BDM subjects in our samples. As a result, our finding that bids are lower in the BDM than in a parallel first price mechanism is inconsistent with the hypothesized mapping.
iii) We conduct a within-subject treatment in which subjects are asked to simultaneously complete the BDM and a parallel FP task. The resulting FP and BDM distributions are statistically indistinguishable. This could be interpreted as support for the FP misconception hypothesis. However, both distributions are different from the data obtained from separate implementations of the FP and BDM cards. This contradicts the prediction that, for risk neutral subjects, behavior in one

[1] Karni and Safra (1987) showed that incentive compatibility is not ensured if the good is a lottery. Similarly, Horowitz (2006) shows that incentive compatibility is not ensured for non-expected utility preferences.

[2] See for example, Bohm et al. (1997), Berry et al. (2018), Freeman et al. (2016), Georganas et al. (2017), Noussair et al. (2004), Rutström (1998) and Vecchio and Annunziata (2015).

[3] In a follow up, Isoni et al. (2011) confirm that misconception is a plausible explanation when mugs are traded but question this explanation when lotteries are bought and sold. A reply by Plott and Zeiler (2011) challenges this finding and sheds further light on the possible role of misconception when using lotteries.

task should not be affected by the parallel presentation of another, payoff-independent task. It is another indication that results from the single BDM treatment are not robust to small variations in the experimental setting.

An Appendix presents a slightly more involved analysis that extends these conclusions to risk averse subjects.

There are clearly substantial similarities between our BDM and FP data. Our overall finding is that while CP's characterization of subjects' behavior provides a good fit for individual treatments, no single type distribution can rationalize all treatments taken as a group. We conclude that different forms of misconceptions must be considered or alternative theories must be invoked in order to explain the data.

Theoretical developments are presented in Section 2, experimental procedures in Section 3 and results in Section 4. The paper concludes with a discussion and summary of our findings.

2. Game Form Recognition theory and misconception of the BDM as a first price mechanism

This section develops an explicit structure to guide the set up of null hypotheses and interpret statistical results from multi-treatment experiments when there is a suspicion that GFM is present. It fleshes out some of the ramifications of CP's theory of GFR for the purpose of making inferences from different treatments.

2.1. Game Form Misconception and the analysis of treatment effects

At the heart of the theory of GFM is the conjecture that some subjects' only departure from the rational optimizing agent model is that they misunderstand the instructions they are given, resulting in systematic and predictable deviations in their choices. Each subject may misconstrue the game being played for a different game. As such, GFM can be thought of as a source of player heterogeneity. The researcher does not know whether a subject properly recognizes the game or not. Instead, the researcher postulates a hypothesis or infers a subject's misconception from her behavior (perhaps from a within-subject treatment or from one or more distribution of choices made by one or more groups of subjects).

In order to highlight the key features of the GFM framework as it applies to treatment variations, we assume that subjects are risk neutral and expected utility maximizers. The Appendix shows that the main conclusions of the paper extend to risk averse subjects. More importantly we maintain that subjects' underlying preferences are fixed and constant across treatments.

A participant is presented with instructions describing a game g, and is asked to select a strategy $s \in S$. Let G_g denote the set of games that subjects may believe they are playing when facing g. This is the set of possible game forms recognized by players. If $G_g = \{g\}$, all players correctly recognize game g. However, if $g \notin G_g$ or if G_g is not a singleton, players may misconceive game g for at least one other game. We model GFM as a probabilistic mapping with the game set G_g as support. Let m generically denote the elements of G_g. This leads to the formal definition of a GFR mapping:

Definition 1. The Game Form Recognition Mapping of a game g is a pair $(G_g, \alpha(m|g))$ such that $\alpha(m|g)$ is a (typically multivalued) function from $G_g \to [0, 1]$ with $\sum_{m \in G_g} \alpha(m|g) = 1$.

The elements of G_g are the possible player types and the function $\alpha(m|g)$ is a probability distribution over these types. $\alpha(m|g)$ denotes the probability that a subject presented with game g recognizes it as game m. Modeling GFM through a probability mapping is sensible since it is generally not possible to observe a subject's type with certainty.[4]

Let $s^*(m)$ denote the optimal strategy in game m under standard economic theory. For simplicity, assume that the optimal strategy is unique. This construction leads to the prediction that in game g, subjects play the optimal strategy $s^*(m)$ with probability $\alpha(m|g)$. In practice, if subjects also make idiosyncratic errors around the optimal strategy for their type, the predicted distribution of play will be a conflation of the distributions of the optimal strategies for each game $m \in G_g$, with expected weights on each type given by $\alpha(m|g)$.

One of the challenges that arises from the fact that the mapping is not observable is that any empirical approach must rely on observing distributions of strategies chosen by subjects of different types. If a mapping is first hypothesized, the task consists of asking whether or not the observed strategy distributions are statistically consistent with the hypothesized mapping. More generally, the experimenter's task consists of finding a mapping that best fits empirical observations. However any test of CP's general theory is necessarily a joint test of i) the postulate that some subjects have misconceptions; and ii) the hypothesized mapping that precisely characterizes systematic mistakes made by subjects. Statistical rejection of a hypothesis test can then occur either because subjects errors are not of the GFM kind (e.g. non-standard preferences could be present); or perhaps they do, but an incorrect mapping was hypothesized by the researcher.

CP recognize that necessary restrictions on the set of possible mappings (as in Definition 1 above) pose challenges "since the nature of misconceptions is context dependent, falsifiability of any proposed general theory of the failure of game form

[4] One may think of $\alpha()$ as a general type distribution with independent draws for each individual. However, in the context of repeated games where individual decisions are correlated over time (e.g. as a result of learning), it may be more appropriate for an individual's type to be conditional on their own history. The researcher may then be able to make inference about the dynamics of the type distribution and also discern systematic patterns at the individual level (e.g. subjects who have discovered the correct game form and are no longer subject to misconception). This opens up the prospect of further developing our framework to integrate GFM into models with learning in repeated games. See also discussion in footnote 13.

recognition is problematic. There is a need to state a specific theory of the misconception in each case" (page 1266). These complications can be greatly exacerbated when comparing results from different treatments since, in principle, different mappings could arise from different treatments. Our approach to deal with treatment effects is to construct null hypotheses that assume that the mapping (G_g, $\alpha(m|g)$) is invariant to:

A.1 the value of game parameters (e.g. number of random prices in a BDM). The optimal strategy $s^*()$ may depend on parameters, but the mapping (G, $\alpha()$) remains constant;
A.2 whether or not the subject also plays other payoff-independent games simultaneously.

Assumption A.1 makes an explicit distinction between "game form" and "game parameter" and holds that a change in a game parameter alone does not modify the mapping. A.2 is akin to an assumption of "independence of irrelevant alternatives". These assumptions are quite strong but they are required in order to formally run tests across treatments. One way to think about A.1 and A.2 is that they are the maintained assumptions under the null hypothesis of a statistical test across treatments. As we will soon discover, one possible explanation for the rejection of a null hypothesis may very well be that one or both assumptions do not hold.

In principle, one could consider more general mappings than those allowed in Definition 1 and Assumptions A.1 and A.2 but doing so raises new problems. To illustrate, assume that a researcher wishes to compare the effect of changing a parameter from p to p'. The sample from the treatment with p is a joint distribution of the strategies chosen by each type $m \in G_g$. Each group is expected to be centered around its respective type optimum $s^*(m_p)$ with a type weight $\alpha(m|g)$. If the mapping is invariant to the change of parameter (as assumed in A.1), the expected weight of each type remains constant at $\alpha(m|g)$ in the treatment with p'. Any observed treatment effect can then be attributed to a shift in incentives resulting from the change of parameter (i.e. changes in $s^*(m)$). Such shifts will, in general, be different for different player types.

It is possible, however, that the mapping itself could be affected by treatment variations. A variation in a game parameter could affect the game sets $G_{g,p}$ and $G_{g,p'}$ and/or the functions $\alpha(m_p|g,p)$ and $\alpha(m_{p'}|g,p')$. This would violate A.1. Detecting and correctly identifying treatment effects would then require laying out hypotheses for how the mapping will change with the parameter change. Clearly, testing for treatment effects is then a joint test of: i) the treatment effect on the mapping (e.g. whether G and $\alpha()$ depend on (p, p')); and ii) the pure incentive effects of the parameter change (i.e. through a change in s^*). As in any joint test, identification strategies would then need to rely on clever experimental designs or econometric methods.

While Definition 1 is maintained in the core of this paper, Assumptions A.1 and A.2 constitute an explicit definition of the robustness of a GFR mapping to treatment variations. A mapping is said to be robust if it does not change with parameter variations or the presence of other games that leave the payoff function of the game of interest unchanged.

2.2. Testing for Game Form Misconception in the Becker–Degroot–Marschak mechanism

CP provide empirical evidence that a significant proportion of subjects misconceive the BDM as a FP mechanism. These subjects misunderstand the payoff function. They believe that they stand to receive their own offer price rather than the random posted price when they sell their card after a favorable draw. Thus, CP hypothesize that some unknown fraction of subjects, $\alpha(FP|BDM)$, choose actions consistent with the optimal risk-neutral offer in a first price version of the mechanism and that subjects make idiosyncratic errors in computing the optimal strategy. After considering alternative explanations for the data, they conclude that "The subjects consist of at least two groups. One group understands the game form as a second-price auction and behaves substantially as game theory predicts. Another group has a misconception of the game form as a first-price auction and under that model behaves substantially as game theory predicts" (page 1262-3).

Even though CP allow for the possibility of other types of players, such group(s) are not identified and are empirically unnecessary for a good fit of the data. Thus, we formally refer to a strict version of their conclusion as the FP-GFM mapping:

The FP-GFM Mapping: *The mapping for the BDM game is* (G_{BDM}, $\alpha(m|BDM)$) *with* $G_{BDM} = \{BDM, FP\}$. *With probability* $\alpha(FP|BDM)$ *a subject is a misconceived FP type who chooses a strategy centered at the risk neutral optimum* $s^*(FP)$. *With probability* $\alpha(BDM|BDM)$ *the subject correctly recognizes the game and chooses a strategy centered on the BDM optimum* $s^*(BDM)$.

We use the FP-GFM mapping as our working hypothesis throughout the analysis. We begin by attempting to replicate CP's results using their single BDM treatment before considering three new treatments closely related to this baseline. The FP-GFM mapping is therefore a natural and rigorous starting point to analyze subject behavior. It is also useful to illustrate how to test for misconceptions using treatment variations. As previously discussed, employing different treatments could alter the mapping. The joint analysis of treatments we propose is necessarily restricted to testing the joint hypotheses that i) subjects behave according to the theory of misconception; ii) FP-GFM is the correct mapping; and iii) the mapping is treatment-independent.

To sum up, our approach to codify CP's theory of misconception has several layers. At a high level of generality, we restrict ourselves to mappings that satisfy Definition 1 and Assumption A.1 and A.2. We also maintain throughout the analysis of our experimental data a null hypothesis that FP-GFM applies to all BDM treatments. This structure helps derive formal predictions for our treatments and more clearly identify what can and cannot be rejected. We revisit these assumptions when the evidence is inconsistent with predictions.

3. Experimental design

3.1. Treatments

The principal strength of the BDM experiment performed by CP is that the $2 card is worth exactly $2 to subjects under just about any reasonable conception of utility. This should greatly mitigate the possibility that unobserved preferences play a role in the data. The BDM is also a single person task, which eliminates the problem of strategic interdependence between players and off-equilibrium play that can arise for a variety of reasons in multi-player games. Thus, using induced values provides a focused context to study why so many subjects present an offer price different than the optimum of $2. The experiment performed in this study consists of four closely related treatments. Each provides a different perspective on the central question of whether the FP-GFM mapping is robust. They also illustrate some of the challenges posed by experimentation in an environment that admits GFM.

3.1.1. CP replication: the misconceived task (BDM1)

The BDM1 treatment uses the same BDM sale mechanism deployed by CP. Participants are instructed to read the card and write down an offer price. On the back side of the experiment card is a black piece of tape covering a randomly drawn 'posted' price taken from a uniform distribution on [$0,$5] (in one penny increments).[5] Subjects are instructed to read the entire card but to remove the tape only after writing down their offer price. If the subject's offer price is less than or equal to the random posted price, the subject sells their experiment card at the posted price. If the offer price is greater than the posted price, the card is redeemed for its nominal value of $2.

BDM1 replicates the experiment performed by CP with the exception of two small modifications. First, the wording on the card was slightly edited for clarity. Second, and more importantly, our subjects were told that all possible posted prices between $0 and $5 had an equal chance of being drawn as the hidden posted price.[6] This treatment allows us to verify that CP's results replicate and to perform additional comparisons with the next three treatments.

3.1.2. Two random posted prices (BDM2)

The BDM2 treatment differs from BDM1 in a single dimension. Rather than drawing a single random price, two are drawn. Subjects are instructed that if their offer price is less than or equal to the lowest of the two random prices, they sell the card at the lowest random price. If their offer price is greater than one or both of the random prices, they redeem the card for $2. Our motivation for including this treatment in the experiment is to test for the presence of a significant number of FP-misconceived subjects. This change in parameter does not modify the optimal strategy for BDM types. However, under the maintained assumption of risk-neutrality, the optimal offer price of a FP-misconceived subject decreases from $3.50 in BDM1 to $3 in BDM2.[7]

3.1.3. The task misconceived for: the first price auction (FP)

Instructions for the FP and BDM1 treatments are identical except for one critical detail. In FP, subjects are told that if the random posted price is greater than or equal to their offer price, they sell the card for their own offer price (rather than the random price in the BDM). If the posted price is less than their offer price, they redeem the card for $2 as in the BDM. The behavior of FP types in the BDM should match the behavior of subjects actually playing the FP game. The FP data can then help identify FP and BDM players in BDM1. Although the general theory of GFM does not impose this, finding that it is the case would seem to increase the predictive power of the theory. The FP treatment will also be used to interpret the results of our last treatment, a within-subject comparison of FP and BDM.

3.1.4. Simultaneous completion of both tasks (Within-Subject — WS)

In the within-subject (WS) treatment, subjects are presented with two cards simultaneously. One card is identical to BDM1 while the other is identical to FP. Subjects are alerted in the instructions that the cards implement different sales mechanisms. They are not told specifically what the difference is but they are asked to read both cards carefully before proceeding.[8] The cards are coded as WS-BDM and WS-FP respectively. This treatment serves as a test of procedural invariance when comparing the data to the single task treatments. It also provides a different perspective on individual idiosyncratic decision errors across the two tasks.

3.2. Procedures

Copies of instructions and decision cards are presented in the Appendix. Subjects in all except the WS treatment performed two successive rounds of the same task, though the second was unannounced. The core of the analysis focuses primarily on the results from the first round (the evidence on the second round is discussed mostly in the replication of

[5] We use a single upper bound ($5) for all subjects while CP used others as well.

[6] Examples of the material used during the experiment (instructions, consent forms, experiment cards) can be found in the Appendix.

[7] While CP do not consider the possibility that FP-misconceived subjects may be risk averse, there is evidence that participants in first price auctions may indeed have such preferences (Harrison, 1990; Kagel and Levin, 1993). With risk aversion, one would also expect a decrease in the optimal FP offer price. The Appendix generalizes the analysis and results to risk averse subjects.

[8] Subjects' instructions (item 2) read "Your envelope contains two decision cards. Carefully read the instructions on both decision cards before making any decision. **NOTE** that the way in which you can sell each card is **different**."

Table 1
Descriptive statistics.

	BDM1-1	BDM1-2	BDM2-1	BDM2-2	WS-BDM	WS-FP	FP-1	FP-2
mean ($)	2.93	2.82	3.27	2.93	3.19	3.24	3.30	3.20
median ($)	3.00	2.94	3.00	3.00	3.14	3.32	3.25	3.23
mode ($)	3.00	3.00	3.00	3.00	3.00	4.00	4.00	3.00
sd ($)	0.98	1.17	1.21	1.21	1.07	1.11	0.92	0.91
variance ($^{2})	0.97	1.36	1.47	1.47	1.14	1.24	0.85	0.83
max ($)	5.00	8.70	10.00	7.00	7.00	9.00	5.10	5.00
min ($)	0.00	0.00	1.00	0.00	0.01	0.49	1.00	1.00
offer $\leq$ 0.10 (count)	6	6	0	4	3	0	0	0
1.95 $\leq$ offer $\leq$ 2.05 (count)	13	22	14	21	10	12	7	3
3.45 $\leq$ offer $\leq$ 3.55 (count)	16	9	16	14	26	18	13	8
2.95 $\leq$ offer $\leq$ =3.05 (count)	43	25	32	21	35	24	11	16
offer $\geq$ 5.00 (count)	0	1	4	3	1	1	2	0
offer $\leq$ 0.10 (%)	3.7%	3.7%	0.0%	2.5%	1.8%	0.0%	0.0%	0.0%
1.95 $\leq$ offer $\leq$ 2.05 (%)	7.9%	13.5%	8.8%	13.3%	5.8%	7.1%	8.2%	3.6%
3.45 $\leq$ offer $\leq$ 3.55 (%)	9.8%	5.5%	10.1%	8.9%	15.2%	10.6%	15.3%	9.6%
2.95 $\leq$ offer $\leq$ 3.05 (%)	26.2%	15.3%	20.1%	13.3%	20.5%	14.1%	12.9%	19.3%
offer $\geq$ 5.00 (%)	0.0%	0.6%	2.5%	1.9%	0.6%	0.6%	2.4%	0.0%
First Price Misconception (per round, count)	10	15	2	1	17	–	–	–
First Price Misconception (in either round, count)	22		3		17	–	–	
Second Price Misconception (per round, count)	–	–	–	–	–	13	1	2
Second Price Misconception (in either round, count)	–		–		–	13	3	
Possible Misconception, but not shown (count)	51		105		67		42	
First Price Misconception (per round, %)	6.1%	9.2%	1.3%	0.6%	9.9%	–	–	–
First Price Misconception (in either round, %)	13.4%		1.9%		9.9%	–	–	
Second Price Misconception (per round, %)	–	–	–	–	–	7.6%	1.2%	2.4%
Second Price Misconception (in either round, %)	–		–		–	7.6%	3.5%	
Possible Misconception, but not shown (%)	31.1%		66.0%		39.2%		49.4%	
N	164	163	159	158	171	170	85	83

CP's analysis). We label first and second round data by adding "−1" and "−2" to treatment acronyms (e.g., BDM1-1 is the first round of BDM1; FP-2 is the second round of FP) where necessary.

In the BDM1, BDM2, and FP treatments each subject was handed a large envelope containing a research ethics consent form, detailed instructions on how to perform the experiment, an experiment card, and a smaller envelope. After subjects completed the experiment card, the instructions prompted them to open the smaller envelope containing a second experiment card and new instructions. The new instructions told subjects that they had an opportunity to sell a second card using the same procedure as before but with a new randomly drawn posted price. Subjects were instructed to complete this card and put all the material, except for a tag with an ID number on it, back in the large envelope. The tag was their ticket for payment. Payments to subjects were made in subsequent lab sections and lectures. The WS treatment followed a similar procedure except that both experiment cards were contained within the large envelope and subjects were asked to complete them simultaneously.

Experimental sessions lasted approximately twenty-five minutes and were run in first–year undergraduate classes at the University of Victoria (Canada). Subjects were not trained and were told that the purpose of the research project was to understand how participants take advantage of simple trading opportunities in different forms. All experimental material was prepackaged in envelopes and handed out in an alternating pattern, ensuring that each treatment's envelopes were homogeneously distributed across the entire classroom. For every envelope of FP handed out, two envelopes of BDM1, BDM2, and WS were distributed. Subjects were informed that their neighbors were completing a different task and instructed not to talk to others until the experiment was completed.

A total of 579 subjects participated in three separate sessions that took place in theatre-style classrooms over the period September 22–23, 2015. In total, 164 subjects participated in BDM1, 159 in BDM2, 171 in WS, and 85 in FP. In the first two sessions the subject to proctor ratio was roughly 30:1; in the third it was 18:1. Subjects were informed they could earn up to $10. The average payout was $4.91, the minimum was $2, and the maximum was $9.75. For convenience, payments were rounded up to the nearest quarter dollar (subjects were not aware of the rounding up at the time of the experiment). All amounts quoted in the paper are in Canadian dollars.

4. Results

Of the 164 subjects who participated in BDM1, 163 completed both experiment cards. The corresponding numbers are 158 of 159 subjects for the BDM2; 170 of the 171 subjects for the WS treatment; and 83 out of 85 subjects who participated in the FP auction. It is unknown why five subjects only completed one of the two cards presented to them. The data they provided is only used when possible.

Table 1 presents summary descriptive statistics for all four treatments. We refer to this table throughout the remainder of the paper. Statistics are presented separately for the first and second cards in treatments with repetition, and for the

Table 2
Results of Variance Ratio Tests for the equality of variances.

	BDM1-1	BDM1-2	BDM2-1	BDM2-2	WS-BDM	WS-FP	FP-1	FP-2
BDM1-1		0.7114** (0.0306)	0.6602*** (0.0088)	0.6586*** (0.0085)	0.8482 (0.2899)	0.7807 (0.1123)	1.138 (0.5140)	1.1645 (0.4439)
BDM1-2			0.9281 (0.6372)	0.9258 (0.6265)	1.1923 (0.2574)	1.0975 (0.5493)	1.5995** (0.0175).	1.637** (0.0136)
BDM2-1				0.9975 (0.9874)	1.2847 (0.1090)	1.1825 (0.2839)	1.7234*** (0.0063)	1.7638*** (0.0048)
BDM2-2					1.2879 (0.1060)	1.1855 (0.2774)	1.7277*** (0.0061)	1.7682*** (0.0047)
WS-BDM						1.0864 (0.5901)	1.3415 (0.1322)	1.373 (0.1077)
WS-FP							0.6862* (0.0544)	1.4916** (0.0432)
FP-1								1.0235 (0.9169)

Table shows *F*-statistics with *p*-values in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, $^{****}p < 0.001$.

Table 3
Results of Mann–Whitney Two-Sample *U*-tests or Wilcoxon matched-pairs signed-rank tests (whether two samples drawn from the same distribution).

	BDM1-1	BDM1-2	BDM2-1	BDM2-2	WS-BDM	WS-FP	FP-1	FP-2
BDM1-1		1.095 (0.2737)	1.890* (0.0588)	0.456 (0.6487)	2.163** (0.0306)	−2.758*** (0.0058)	−2.498** (0.0125)	−2.044** (0.0410)
BDM1-2			−3.106*** (0.0019)	−0.762 (0.4459)	−3.347**** (0.0008)	−3.601**** (0.0003)	−3.359**** (0.0008)	−2.948*** (0.0032)
BDM2-1				2.846*** (0.0044)	−0.104 (0.9173)	−0.568 (0.5703)	−0.763 (0.4455)	−0.274 (0.7840)
BDM2-2					−2.276** (0.0228)	−2.594*** (0.0095)	−2.498** (0.0125)	−2.060** (0.0394)
WS-BDM						−1.020 (0.3079)	−0.710 (0.4780)	−0.138 (0.8901)
WS-FP							−0.172 (0.8635)	0.410 (0.6820)
FP-1								0.809 (0.4183)

Table shows *U*-statistics with *p*-values in parentheses. Wilcoxon matched-pairs signed-rank test used when data are matched. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, $^{****}p < 0.001$.

Table 4
Results of Pearson's Chi-Squared Tests (Equality of sample medians).

	BDM1-1	BDM1-2	BDM2-1	BDM2-2	WS-BDM	WS-FP	FP-1	FP-2
BDM1-1		– –	1.5744 (0.210)	0.0267 (0.870)	3.8147* (0.051)	7.5763*** (0.006)	5.939** (0.015)	4.3484** (0.037)
BDM1-2			5.4765** (0.019)	1.5719 (0.210)	9.3036*** (0.002)	14.7374**** (0.000)	11.2209*** (0.001)	8.9643** (0.003)
BDM2-1				– –	0.4531 (0.501)	1.1102 (0.292)	2.1845 (0.139)	1.0833 (0.2980)
BDM2-2					3.1331* (0.077)	6.5691** (0.010)	5.2218** (0.022)	3.751* (0.053)
WS-BDM						– –	0.863 (0.353)	0.1611 (0.688)
WS-FP							0.1255 (0.723)	2.0719 (0.150)
FP-1								– –

Table shows χ^2 statistics with *p*-values in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, $^{****}p < 0.001$.

two simultaneous tasks in the WS treatment. The table first presents simple statistics on the distribution of offer prices and then computes the fraction of offers measured at key points in the distribution corresponding to optimal offer prices in the various treatments ($2 for all BDM treatments, $3.5 for FP with one random price and $3 for FP with two random prices). The table also reports the fraction of outliers (offers below $.10 and above $5) and data corresponding to CP's definitions of misconception discussed in Section 4.5.2.

We will also refer to Tables 2–4 for statistical test results. Table 2 provides results of variance ratio tests for the equality of variances across treatments; Table 3 shows the results of Mann-Whitney U-tests and Wilcoxon matched-pair tests, which

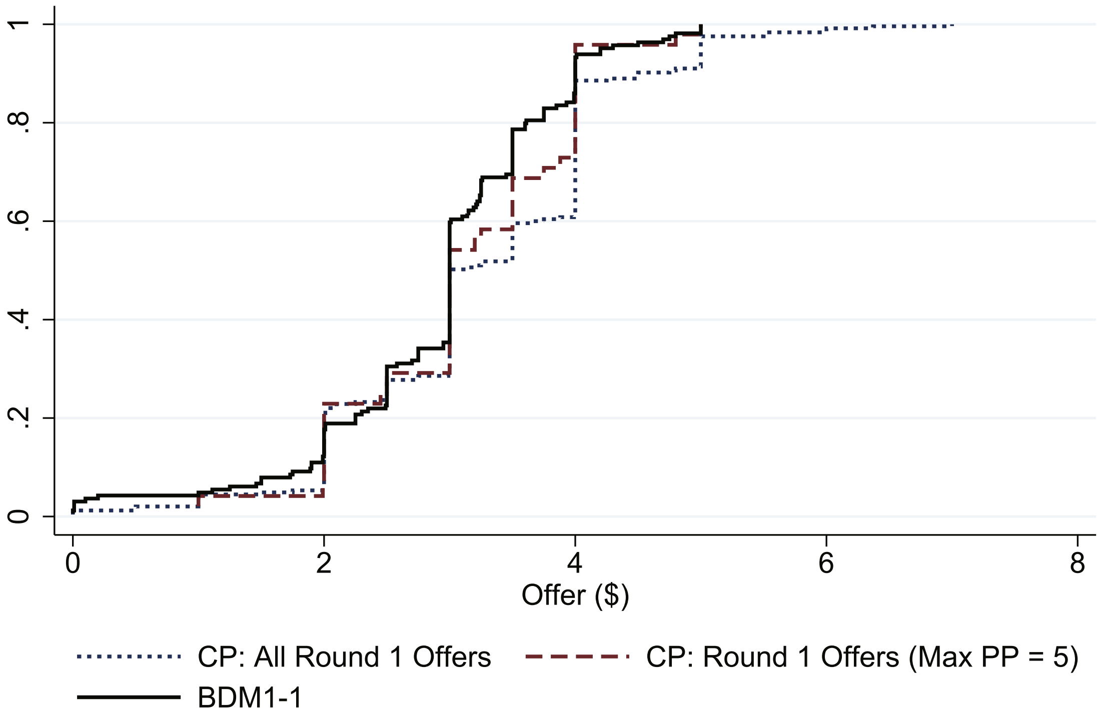

Fig. 1. Cumulative distribution functions for BDM1-1 and CP's Round 1.

test whether two samples come from populations with the same distribution. The Wilcoxon test is applied when the two distributions tested are matched pairs (i.e. from the same subject), while the *U*-Test is applied to compare samples across treatments. Table 4 reports the results of Pearson's chi-squared tests of whether two samples come from distributions with the same median.[9]

4.1. Replication of Cason and Plott's main results (BDM1)

First, we investigate whether our BDM1 treatment data is consistent with CP's findings.

Hypothesis 1. Offer prices in BDM1 replicate the results of CP.

Evidence from the BDM1 treatment is broadly consistent with CP's results (one exception is CP result 2 where the effect we observe is in the same direction as CP's but our measurement is not statistically significant). Fig. 1 provides a comparative overview of ours and CP's data.[10] It shows the cumulative density function of offer prices for our BDM1-1 treatment against i) all round one data from CP; and ii) CP's round 1 data for the subset of subjects who faced a maximum random price draw of $5.

Result 1. Results 1–3 and 5 from CP are broadly replicated.

Five main results on the BDM emanate from CP's paper. We can directly test and replicate four of them. They are referenced herein as 'CP Result 1' to 'CP Result 5'.[11]

CP Result 1. With simple instructions but no training or feedback, the BDM does not provide a reliable measure of preferences for the induced-value object.

In BDM1-1, only 7.9% of subjects make an offer price within 5 cents of the optimal $2 offer. The corresponding number in CP is higher at 16.7%. Therefore, our results show an even lower proportion of subjects making offers around the BDM

[9] Mann-Whitney U-tests and chi-squared tests are used rather than t-tests, as treatment data are not normally distributed. We also ran a full complement of KolmogorovSmirnov tests as an alternative to the *U*-Test. We did not produce a separate Table as they almost universally produce the same diagnostics as the Mann–Whitney *U*-test. Only one slightly different outcome is directly relevant and we present the test result in the text.

[10] CP's data is available at http://www.journals.uchicago.edu/doi/suppl/10.1086/677254. CP ran five treatments with different upper bounds of the random price (4–8).

[11] CP Result 4 analyzes how subjects respond to a change in the upper bound of the random posted price. We cannot investigate it since we kept this parameter constant in our study.

Table 5
Adjustment of Round 1 to Round 2 Offers for Subjects Choosing Incorrectly in BDM1.

	Exposed to Round 1 Error	Not Exposed to Round 1 Error
Total Subjects	41 (100%)	109 (100%)
Move onto optimum ($2)	3 (7.3%)	8 (7.3%)
Move Toward Optimum	25 (61.0%)	35 (32.1%)
Choose same offer ratio	7 (17.1%)	25 (22.9%)
Move away from optimum	6 (14.6%)	41 (37.6%)

optimum. A Wilcoxon signed-rank test strongly reject the null hypotheses that the median of the BDM1-1 distribution is equal to $2 ($p < 0.0001$). This result confirms what one might suspect from inspecting Fig. 1: untrained subjects do not generally behave as theory predicts for the BDM.

CP Result 2. A second round of decisions (after payoffs from the first round have been computed) increases the number of subjects making the optimal offer.

The fraction of subjects offering amounts within 5 cents of $2 increases to 13.5% in BDM1-2. Although the shift is in the predicted direction, a Fisher's exact test indicates that it is not statistically significant ($p = 0.111$).

CP Result 3. Subjects who chose the theoretically optimal offer price (near $2) on the first card also usually choose the theoretically optimal offer price on the second card. Subjects who did not choose optimally on the first card tend to choose a different offer price on the second card.

Table 5 provides detailed statistics for results 3 and 5. Of the 163 subjects who completed both BDM1 cards, 13 made offers within 5 cents of $2 on the first card. Of these 13 subjects, ten (76.9%) offered the same amount in BDM1-2. Of the 150 subjects who did not offer within 5 cents of $2 on the first card, 79.3% offered a different amount in BDM1-2. Similarly to CP, we reject the hypothesis that those who offer near $2 on the first card, and those who do not, exhibit the same degree of choice stability (Fisher's exact test p-value $<$ 0.001).

CP Result 5. Subjects who were "exposed" to their mistake were more likely to choose a correct offer in round 2.

A subject is said to be exposed to their mistake if, on the first BDM card, they submitted a sub-optimal offer and could have increased their payoff by offering a different amount (i.e. closer to or equal to the optimal offer). Most exposed subjects redeem their card for its nominal value of $2 but missed an opportunity to sell it for more through the BDM (i.e. their offer price is greater than $2 and the random price is between $2 and their offer price).[12]

Our result is not as clear as CP's. The strong version of the result says that subjects exposed to their error in BDM1-1 should move exactly to the optimal offer in BDM1-2. Table 5 shows that exposed subjects were just as likely to move to a $2 offer in round 2 as non-exposed subjects. However, and consistently with CP's results, exposed subjects are more likely to move in the direction of the optimum than non-exposed participants. They are also less likely to move away from the optimum. Both results are statistically significant (Fisher's exact test p-value $<$ 0.002) and replicate CP's data.[13]

To sum up, our baseline results substantially agree with CP's main findings and confirm Hypothesis 1. This general validation of prior results is valuable in itself, but our main interest in them is that they establish solid grounds for the interpretation of new treatment data. We do note a lower proportion of $2 offers in our study and return to its implication below.

4.2. Response to an increase in the number of random prices: BDM1-1 versus BDM2-1

The purpose of the BDM2 treatment is to inspect the response of subjects who exhibit FP-GFM in the BDM, to a change in the number of random prices they have to "beat" in order to sell the card (at the lowest of the two random prices). Technically, this is a parameter variation akin to CP's change in the upper bound of the distribution for random prices (see CP result 4). That is, if the hypothesized FP-GFM mapping is correct, FP-misconceived subjects acting optimally for a FP mechanism with two random prices will produce a distribution of offer prices centered around $s^*(FP_2|BDM_2) = 3.00$. This is lower than the FP-misconceived optimum of $3:50 in BDM1.[14] On the other hand, BDM types have no incentive to alter their strategy. For the entire treatment sample, therefore, the BDM2 distribution should be characterized by lower offer prices than BDM1. In addition, the data for BDM2 should show a higher proportion of offers around $s^*(FP_2|BDM_2) = \$3.00$ and a

[12] Very few subjects end up selling the card for less than $2 (they offer less than $2 and the random price is between their offer price and $2).

[13] CP acknowledge the presence of learning (see the last paragraph of section VI.A on p.1255), but do not formally analyze this feature of the data. In our theoretical framework, the observed shifts in the distribution of BDM bids between Rounds 1 and 2 indicate a change in the $\alpha()$ function. Formal modeling of this learning process could be accomplished by conceiving of the type distribution as conditional on past realizations at the individual level.

[14] This comparative static hypothesis is supported by results from laboratory experiments in standard FP auctions. Participants in these auctions respond to increases in the number of bidders by submitting more aggressive bids, as predicted by standard theory (Kagel and Levin, 1993).

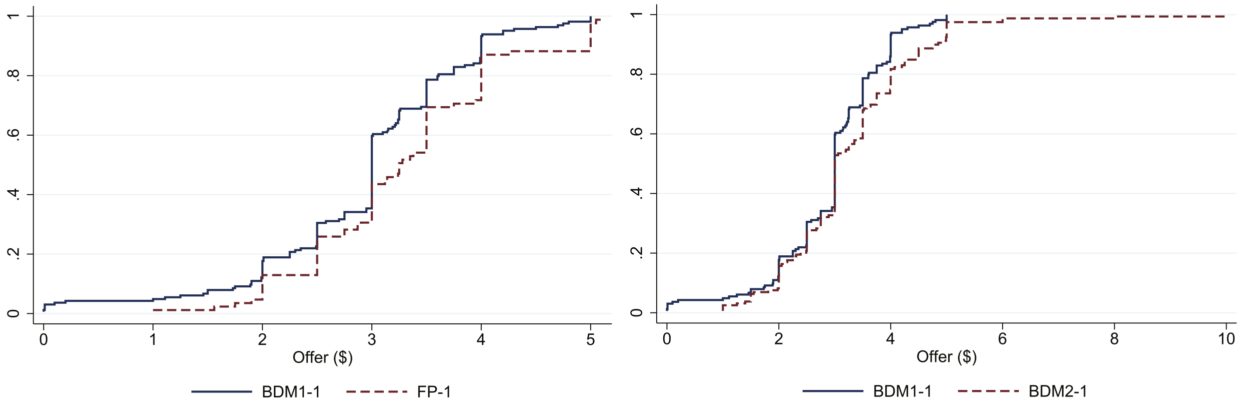

Fig. 2. Cumulative distribution function of BDM1-1 against FP-1 and BDM2-1.

lower proportion around $s^*(FP_1|BDM_1) = 3.50$ relative to BDM1. These comparisons provide potent tests of the robustness of the FP-GFM mapping hypothesis.

Hypothesis 2. Under the FP-GFM mapping, offer prices in BDM2 will be lower than in BDM1, with increased mass around $3.00 and decreased mass around $3.50.

Result 2. An increase in the number of random prices in the BDM does not decrease the subjects' offer prices.

Table 1 reports a mean offer price of $3.27 in BDM2-1. This is larger, not smaller, than the mean of $2.93 in BDM1-1. Notice also that the medians and modes of the two distributions are all equal to $3.00. Fig. 2 shows the CDFs for BDM1-1 and BDM2-1 graphed together. The Mann–Whitney's two sample *U*-test indicates a marginally significative difference between the two distributions (p=0.059). On the other hand, a Kolmogorov-Smirnov test fails to reject that the two distributions are equal (p=0.155). The Pearson's chi-squared test returns a p-value of 0.21, failing to reject equality of the two medians. Finally, the variance ratio test shows that the two samples have different variances ($p < 0.01$). There is therefore some evidence that the two samples come from different distributions, but no support for a systematic shift in the direction predicted by Hypothesis 2.

The proportions of offers within 5 cents of $3.00 are 26.2% in BDM1-1 and 20.1% in BDM2-1. These figures are not statistically different from one another (p-value = 0.1947) and in a direction opposite of expectations. Similarly, the proportions of offers within 5 cents of $3.50 in BDM1-1 and BDM2-1 are also statistically indistinguishable at all conventional levels of significance.

These results cast doubt on the claim that a significant proportion of subjects harbor FP-GFM and behave substantially as game theory predicts. However, it is possible that the FP-GFM mapping is indeed correct but that the shift in the distribution cannot be detected for statistical reasons. A first possibility is that the proportion of misconceived FP subjects in the samples is too small. Since the testable hypothesis relies entirely on the FP types responding to the treatment variable, a small probability $\alpha(FP|BDM)$ would leave too few observations in the full samples to allow detection of the shift. Empirically, this would be nothing other than a sample size issue. However, all of our data from BDM1 as well as CP's own results clearly reach the conclusion that a majority of participants in BDM1 appear to be FP misconceived and our sample sizes are quite large by experimental economics standards.

A second possibility is that the mapping is affected by the parameter change (a violation of Assumption A.1) or by the more detailed instructions required by the presence of two random prices. For instance, we might expect that the BDM2 instructions would help participants recognize the correct game form since the task requires greater attention than the single price BDM. Subjects have to inspect two random prices, determine the lowest of the two, and compare it to their offer price. If they sell the card, they have to make a calculation using the lowest of the two random prices. One might expect that this task is less likely to be misconceived as a first price mechanism than BDM1. Under a constant game set, fewer FP-misconceived subjects implies a higher proportion of BDM types who submit a bid around the $2 optimum and a lower proportion of misconceived FP types. This, however, would lead to an even greater leftward shift in the distribution and exacerbate the difference rather than mask the treatment effect. Without an observable shift and without a statistically significant increase in offers near $2 (7.9% for BDM1-1 vs. 8.8% for BDM2-1, $p < 0.78$), neither of these alternative explanations hold.

The absence of expected differences between BDM1-1 and BDM2-1 does not exclude the possibility that the mapping changed from FP-GFM in BDM1 to a new mapping in BDM2. Methodologically, it remains possible that the FP-GFM mapping is correct for BDM1 but that it is not for BDM2. For this to be true requires the conversion of a large proportion of subjects from FP-GFM types to one or more new game types, but given the empirical evidence, with very small if any modification in behavior. In the absence of a clear response to an increase in the number of random draws, we find it more likely that the FP-GFM is not the correct mapping at all. However, this is speculative since we can only conclude that the differences between BDM1-1 and BDM2-1 are not inconsistent with a robust FP-GFM mapping.

Our result 2 must be contrasted with CP result 4. This result showed that subjects in the BDM were sensitive to variations on the upper bound of the distribution for random prices, $\overline{p}$. The upper bound took the values 4,5,6,7, and 8. Optimal risk neutral FP bids are then given by $1 + 0.5\overline{p}$. Note that from a theoretical standpoint, increasing the number of random offers from one to two has an equivalent predicted impact on the optimal strategy as decreasing $\overline{p}$ from \$5 to \$4. CP show that the average offer price increases almost monotonically with $\overline{p}$. In contrast, we find no response to an increase in the number of random posted prices that a subject must "beat".

4.3. The First Price auction (FP-1)

If subjects are rational decision-makers whose only possible systematic mistake is to misconceive the BDM for a FP mechanism and they submit optimal FP offers, we should expect them to recognize the FP card correctly. Under this scenario, the hypothesized mapping for the FP game is $(G = \{FP\}, \alpha(FP|FP) = 1\}$, and risk neutral participants will have offer prices centered around $s^*(FP) = 3.50$.

Hypothesis 3. Offer prices in FP are consistent with the theoretically optimal bids of utility maximizers participating in the first price version of the \$2 card mechanism.

Result 3. The distribution of FP offers is *not* consistent with risk neutral agents. However, it is consistent with risk averse bidding with noise.

The mean of the FP-1 distribution is \$3.30 and its median is \$3.25. All tests reject that the distribution is centered on the risk neutral optimum offer of \$3.50 (p-values < 0.03) (the same finding holds for FP-2). Furthermore, only 14% of subjects submit an offer price near the risk-neutral optimum in FP-1, and only 10% do so in FP-2. Thus, Hypothesis 3 is rejected. The FP data are not consistent with the postulated behavior of FP misconceived subjects with risk-neutral preferences. However, the data are consistent with the optimal strategy of risk averse FP optimizers (Harrison, 1990) (the appendix further develops the analysis under risk aversion).

4.3.1. Comparing offer prices in FP and BDM

A comparison of FP and BDM presents mixed evidence in support of the FP-GFM mapping.

Hypothesis 4. Under the FP-GFM mapping, offer prices in the BDM will be lower than in FP, with increased mass around \$2.00.

Hypothesis 4 holds under both risk neutrality and risk aversion as long as $s^*(FP) > s^*(BDM)$. We begin with supporting evidence for the hypothesis.

Result 4. Offer prices are lower in the BDM than in FP.

Inspection of the two cumulative distributions of Fig. 2 (panel 1) shows that the distributions are consistent with the hypothesis. Table 1 and statistical tests also confirm that offer prices in BDM1-1 are lower than in FP-1. The mean of FP-1 is \$3.30 while that of BDM1-1 is \$2.93. The medians are similarly ordered: \$3.25 vs. \$3.00. The statistical tests of Tables 2–4 show that the equality of the two distributions is rejected at the 5% significance level in two of the three tests. This being said, this result must be interpreted carefully in light of the next results where the FP treatment helps us characterize the behavior of players in BDM1. This ultimately provides evidence against Hypothesis 4.

Result 5. The proportions of subjects with offer prices near \$2 in BDM1-1 and FP-1 are not different.

The optimal offer price in the BDM1-1 treatment is \$2. As previously noted, a total of 7.9% of our subjects offer $\$2 \pm \$0.05$ in this treatment. As they behave according to the prediction of game theory, CP suggest that many of these correctly understand the BDM and bid optimally. However, in a model of optimal bidding with idiosyncratic errors, one might expect to observe offer prices around \$2 independently of the subject's optimal strategy if the variance is sufficiently large. This is all the more possible given that both our data and CP's show a significant amount of heaping at \$0.50 intervals (this is easily spotted in all CDF plots). Simply stated, one cannot readily infer that offers near \$2 in a BDM sample are from subjects of the BDM type.

Having the FP treatment can help identify BDM types by providing a key comparator: the proportion of offer prices near \$2 in the FP-1 treatment. The proportion of \$2 offers observed in FP provides a benchmark for the proportion of \$2 offers expected in the BDM distribution that are generated by FP-misconceived subjects. If FP and FP-misconceived subjects are less likely to make offer prices around \$2 than BDM types, the presence of BDM types in the BDM sample will result in a greater proportion of \$2 offers in the BDM sample than in the FP sample.

In our results, 8.2% of FP-1 subjects offer within 5 cents of \$2, while 7.9% offer in the same range in BDM1-1. These proportions are statistically indistinguishable ($p = 0.9323$). The equal proportions of \$2 offers in the BDM and FP groups leads to a rejection of the hypothesis that the there is a distinct group of BDM optimizers in our BDM sample. It is important to note that this test does not rely on specifying a mapping for FP subjects. The critical assumption is that subjects in FP and

FP-misconceived subjects make similar offers, and that the proportion of their offers near \$2 is less than that of BDM types. Hence, having the data for the misconceived game (FP) is valuable since it provides identifying power for the presence of BDM optimizers in the BDM treatments.

It should be noted that the proportion of \$2 offers is higher in CP's data than ours. Hence, there may very well be a detectable proportion of BDM types in their experiment. However, without a FP treatment to provide a benchmark, it is difficult to evaluate this claim. Looking back at Fig. 1, we note that the mass of \$2 offers in all CP treatments is roughly comparable to the mass at other integer values (\$3 and \$4) and not much greater than at half-integer values (\$2.5 and \$3.5). There is arguably little that is noteworthy about the proportion of subjects near \$2 but the main point is a methodological one. Obtaining data for the misconceived games in G_g (i.e. FP) can help assess the validity of the hypothesized game form mapping.

Taken together, Results 4 and 5 rule out the FP-GFM mapping as a candidate explanation for the observed offer prices. The inability to detect a distinct group of optimal BDM offers in BDM1 suggests that if the FP-GFM mapping is correct, it must be that $\alpha(BDM|BDM) \approx 0$ and $\alpha(FP|BDM) \approx 1$. In turn, this implies that the distributions for FP-1 and BDM-1 should be similar, but Result 4 rejects this equality.

4.4. Simultaneous completion of both tasks (WS)

In the within-subject treatment, subjects were presented with both the BDM1 and FP cards simultaneously. Instructions clearly warned subjects that the mechanisms for selling the two cards were different but they were not told what the difference was. This treatment helps understand whether misconception is influenced by the set of tasks performed at once. The guiding hypothesis comes from the prior assumption of task independence (Assumption A.2):

Hypothesis 5. Under the FP-GFM mapping, the distributions of offer prices for FP1 and WS-FP are similar; and the distributions for BDM1 and WS-BDM are also similar.[15]

The evidence rejects both of these predictions. Under standard theory, risk neutral subjects performing two payoff-independent tasks should choose the same action whether they perform the tasks separately or simultaneously. This is captured by Assumption A.2. Hypothesis 5 applies this logic as a formal null hypothesis for the WS sub-treatment. Without task interaction the distribution for WS-BDM should be the same as BDM1-1, and WS-FP the same as FP-1.

Result 6. For both FP and BDM, offer prices change when subjects complete both tasks simultaneously.

For the BDM, the Mann-Whitney *U*-test of the hypothesis that the BDM1-1 and WS-BDM samples come from the same distribution is rejected ($p < 0.031$). The Pearson's Chi-Squared test also rejects that the medians are equal ($p < 0.051$). For FP, we also weakly reject the equality of the variances ($p < 0.055$). These findings suggest that presenting both tasks simultaneously resulted in a loss of task independence, with the greatest effect observed on the BDM cards. Observing differences in the distributions when the optimal strategies remain constant suggests that, within the confines of the GFR paradigm, both the BDM and FP mappings change as we go from the single to the simultaneous tasks. This presents strong evidence that any mapping of the form characterized in Definition 1 cannot satisfy Assumption A.2. If Definition 1 is maintained, then Assumption A.2 has to be relaxed to be consistent with the data.

Perhaps a more insightful way of looking at the evidence is to directly compare the two distributions of offer prices in the WS treatment. We first compare the aggregate distributions without pairing the subject's two choices (the paired data are explored below).

Result 7. When subjects complete the BDM and FP tasks simultaneously, the resulting distributions of offer prices do not differ across tasks.

The distributions of offers in the WS-BDM and WS-FP cannot be distinguished statistically. None of the tests across these distributions lead to a rejection of the null hypothesis that they are equal.[16] Fig. 3 shows the CDFs of WS-BDM and WS-FP plotted together.

Looking at these results, one may conclude that subjects tend to behave more or less identically in both, and in a manner more consistent with FP than BDM. This result is favorable to the FP-GFM hypothesis only under the assumption that the probability of being a BDM optimizers is too low to allow detection of the difference (i.e. if $\alpha(FP|BDM)$ is close to one). If BDM optimizers were present in significant numbers, we should reject the hypothesis that the WS-FP and WS-BDM distributions are statistically the same.[17]

[15] This hypothesis does not hold with risk averse subjects. See the analysis with risk averse subjects in the Appendix.

[16] The two distributions are statistically indistinguishable from that of FP-1, but statistically different from BDM1-1 at the 5% level.

[17] The results of this treatment, should be contrasted with Bartling et al. (2015) who show that it is possible to reduce GFM (and induce subjects to reveal their true preference) using the strategy method. The strategy method decomposes the BDM task into simpler sub-tasks that appear to help comprehension. Their result suggests that framing may be at play. In our case, combining the FP and BDM tasks appears to decrease rather than enhance comprehension.

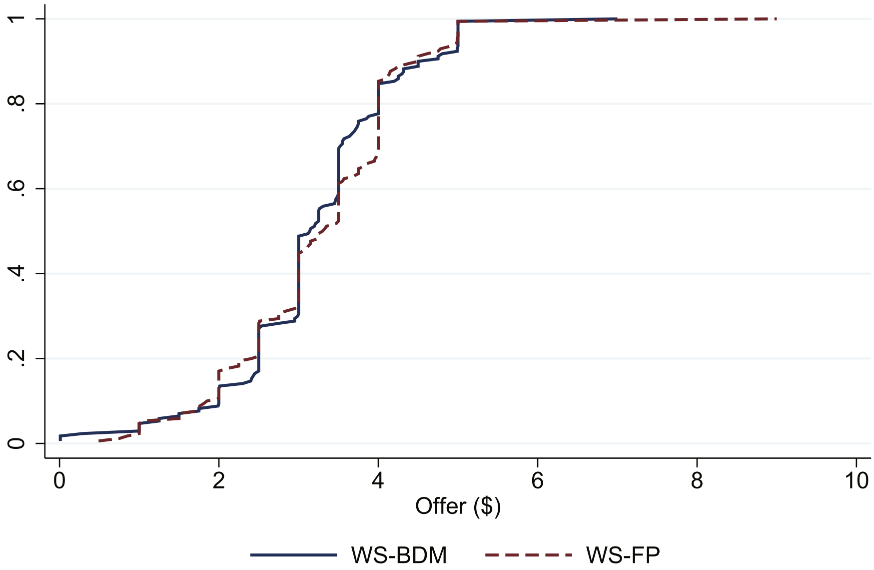

Fig. 3. CDF Comparison of WS Sub-treatments.

4.5. Further results

The evidence so far fails to support Hypotheses 2 and 5 and offers mixed evidence in favor of Hypothesis 4. It is difficult to say whether the problem lies with Definition 1, Assumptions A.1-A.2, the FP-GFM Mapping itself, or a combination of these. Results 4 and 5 show that the FP-GFM mapping cannot rationalize subjects' choices in the BDM and FP treatments. Result 2 can be rationalized without Assumption A.1 and Results 6 and 7 can be rationalized without Assumption A.2. Taken together, these results indicate that the FP-GFM mapping is not robust to treatment variations. Before embarking in a broader discussion of misconception in the BDM, it is useful to bring out two additional observations regarding idiosyncratic errors in the WS treatment.

4.5.1. Joint distribution of individual offers in WS: independent decision errors

Our previous analysis of the WS data took them as independent samples. Now, reconsider the data as matched pairs of bids submitted by individual subjects. Fig. 4 plots the paired offer prices in WS-BDM and WS-FP. All data points are weighted by frequency: larger circles indicate a larger number of subjects who submitted offer prices at the center of the circle on the graph.

Two observations are readily made. First, a detectable proportion of subjects offers similar amounts in both sub-treatments (i.e. near the 45° line). More precisely, 20.6% offered exactly the same amount in both, increasing to 24% when including offers within 5 cents of each other. Second, when offers are not equal, subjects are just as likely to offer more in WS-BDM than they are in WS-FP. This is observable from the roughly equal weights of points above and below the 45% line. Overall, 43.5% of subjects offered a higher amount in WS-FP than in WS-BDM and 35.9% offered more in the WS-BDM than in WS-FP. A test of proportions fails to reject that the proportion of subjects offering a higher amount in WS-BDM is equal to the proportion who offers a higher amount in WS-FP. Those who offered a higher amount in the WS-FP offer on average $1.113 more (median is $0.915) than in the WS-BDM. Those who offered more in the WS-BDM exceed their FP choice by $1.037 (median is $0.950). Equality of these values cannot be rejected either.

It is curious that the vast majority of subjects are just as likely to submit a higher price in one treatment as they are to offer less in it. A theory-conforming risk neutral FP subject facing two identical FP auctions should submit identical offer prices for both cards. Yet, the large majority of subjects submit significantly different offers.

This is technically admissible in a model with idiosyncratic errors but the lack of correlation between the two offers requires a weak version of the bidding with errors model. In a strong version of the model with errors, a rational subject who thinks she is playing the same FP game twice would compute the optimal offer with an error and report the same offer on both cards. The independent errors suggest instead that subjects treat the two games similarly, but make separate computations for each. This is perhaps because they were told that the cards were different, but they failed to identify the difference. This is difficult to reconcile with a definition of GFM where agents are rational and where their only mistake is to misconceive one game for another. At a minimum, the result conveys a high degree of player uncertainty. This is quite possibly instruction dependent, but we should be concerned that subjects in the WS treatment appear to either suffer from deeper confusion than predicted by the FP-GFM mapping, or they perhaps have very weak optimizing skills.

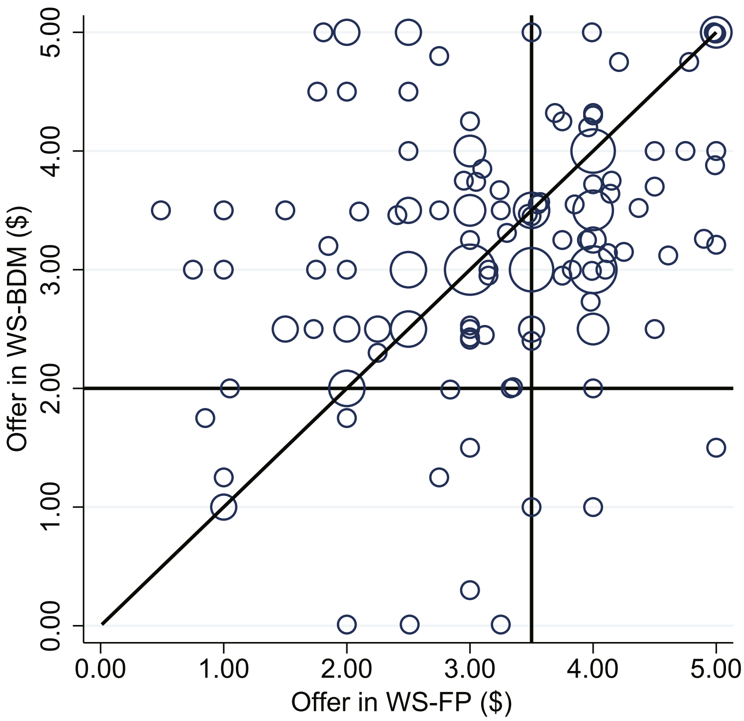

Fig. 4. Scatter plot of matched offers in WS treatment.

4.5.2. Revealed misconception: Increased confusion in WS

Following CP, we asked subjects to report their payoff after they observed the random posted price. A subject in a BDM treatment likely reveals its FP-GFM type (denote R-FPM) if she wins (i.e. makes an offer smaller than the random price) but incorrectly requests to be paid her own offer price. By symmetry, we introduce the concept of *revealed second price game form misconception* (R-SPM) if a subject in a FP treatment wins but incorrectly requests to be paid the random price. These revealed misconceptions are explicit errors that have the potential to reveal something new about subjects' understanding of the game.

Result 8. R-FPM is found in all BDM treatments. R-SPM is almost nonexistent in the FP treatment. R-SPM is found in the WS treatment.

The lower part of Table 1 presents evidence on R-FPM and R-SPM. As a benchmark, CP report that 11.8% of all subjects explicitly revealed a FP misconception on their card. Our corresponding rate in BDM1 is 13.4%. As these numbers accord quite well, it can be assumed that a similar proportion of subjects suffers from FP-GFM in both experiments.[18,19] Presenting both cards simultaneously could have reduced the rate of FP-GFM but it did not. At 9.9%, the rate of R-FPM in WS-BDM is not statistically different from the 6.1% observed in BDM1-1. Confronting subjects with both mechanisms simultaneously and prompting them to carefully consider the two cards did not affect the rate of R-FPM. On the other hand, R-SPM is statistically higher in WS-FP (7.6%) than in FP-1 (1.2%).

Doing both tasks together appears to have confused subjects rather than helped them properly recognize the two games. This suggests that GFM may be highly sensitive to context. It raises the specter that increased complexity lowers task comprehension and increases misconception. Confusion about the difference between the two mechanisms may also be responsible for the statistically indistinguishable distributions between the two cards.

[18] We also looked at "*possible misconception*" in reference to cases where a correct understanding of the rules cannot be asserted with certainty. These correspond to cases where a subject's offer prices were greater than the random posted prices on both cards. As such, it is possible that the subjects may have suffered from misconception, but this is not revealed because of the outcome. The proportion of subjects with possible R-FPM in BDM1 is 31.1%, similar to the 33.5% reported in CP.

[19] Interestingly, subjects make significantly fewer R-FPM mistakes in BDM2. For the data as a whole, we observe a 1.9% proportion of R-FPM. However, since BDM2 leads to fewer subjects winning at least one card, it is perhaps more appropriate to condition these statistics. In the BDM2, 53 subjects won at least one card (and thus could potentially reveal GFM). Of these, only 3 subjects (5.67%) made a single mistake each. In comparison, 22 subjects out of the total of 112 (19.6%) who won at least one card in BDM1 made at least one error. The proportions are significantly different ($p = 0.020$).

5. Discussion and conclusion

Testing for the robustness of a particular type of GFM across treatments requires some formalization of the theory proposed by Cason and Plott. Conceiving the theory as a mapping that describes a player's possible types (where a type is a game the player think she is playing) brings forth a formalization of the robustness of the mapping to treatment variations. In this work, we investigate whether the mapping is invariant to parameter changes and to the presence of a simultaneous payoff-independent game.

Equipped with this theoretical structure, we present new data to assess CP's finding that subjects participating in BDM experiments behave in a manner consistent with individuals being mainly from two types of rational optimizing agents: individuals who properly recognize the BDM game and others who misconceive it as a first price mechanism.

The results from our own benchmark BDM1 treatment broadly replicate Cason and Plotts findings. As such, they support the FP-GFM mapping hypothesis and its characterization of BDM players. There are also other elements of our data that, taken individually, are consistent with the mapping. For instance, Result 4 shows that offer prices in BDM1 are on average lower than in the FP treatment. Similarly, Result 7 indicates that the FP and BDM distributions in the WS treatment are statistically similar.

However, taken within the broader context of the entire set of data generated across all treatments, the evidence suggests that the FP-GFM mapping is not a robust characterization of BDM players. When Result 4 is revisited in light of the finding that BDM types are not statistically detectable (Result 5), lower offers in BDM1 than in FP are no longer consistent with a robust FP-GFM mapping where nearly all players are of the FP-misconceived type. And while the WS-FP and WS-BDM distributions are similar, the WS-BDM distribution is statistically different from the BDM1 data (Result 6). Since the two BDM cards were identical and since the addition of the WS-FP task does not change the optimal strategy for a risk-neutral subject in WS-BDM, the observed shift in the BDM offers can only be attributed to a violation of the assumption that the mapping is impervious to the addition of an independent task (this conclusion generalizes to risk averse subjects although the argument is slightly different). It follows that the mapping in the WS-BDM must be different than in the BDM1 (or that factors other than GFR errors are at play).

This being said, the most definitive evidence against the robustness of the FP-GFM mapping comes from comparing BDM2 to BDM1 (Result 2). If a significant proportion of BDM participants in both treatments were optimizing subjects who are FP-misconceived, we would observe lower offer prices in BDM2 than in BDM1. However, this change in parameter elicited no statistically detectable difference in price offers (and the direction of the change was opposite of that predicted). This is a clear rejection of the hypothesis that the FP-GFM mapping is robust.

The overall conclusion is that while there are many similarities between BDM offers and real or presumed misconceived first price offers, some of the systematic variations one would expect under the FP-GFM mapping are not present, while other observed variations are inconsistent with a robust FP-GFM mapping. Therefore, while we would reach the same conclusion as Cason and Plott if we inspected our BDM1 treatment alone, the accumulated evidence indicates that the FP-GFM mapping cannot explain key features of the data as a whole.

It is always possible that some other mapping may fit all of our data. However, it is beyond the scope of this paper to i) engage in a search for another mapping that could better explain the data from our treatments, to seek a more general theory of misconception than the one laid down in Definition 1 and Assumption A.1 and A.2; or, to explore framing or non-standard preferences. Within the confines of explanations based on GFM, it could be that: (1) the BDM or FP mappings (G, α) were mis-specified; (2) subjects misconception of games are context dependent (Assumption A.1 or A.2 are violated); or (3) subjects harbor misconceptions that are not captured in Definition 1. This opens the possibility for a wide range of speculation, but better yet would be to explore and test new hypotheses by developing clever new treatments to identify the origins of treatment effects.

It is important to emphasize that where GFM is suspected, experimentation across treatments forces researchers to test jointly for incentive effects and the possible change in mapping that different games or sets of instructions may engender. Since the general theory of GFR is non-falsifiable, one cannot reject misconceptions outright.

The formalized structure we developed to perform tests across treatments can also help reinterpret recent results like those of Fehr et al. (2015). They conduct within-subject experiments with the BDM in a replication of Plott and Zeiler (2005) to conclude that GFM cannot explain their data. As we made explicit by fleshing out the conceptual framework for treatment comparisons, such a conclusion cannot be reached since the GFM mapping could vary across different tasks performed by subjects.

At best, what can be tested for under the general theory of GFR is (i) the goodness of fit of the data to a hypothesized mapping and (ii) its invariance across treatments. Using identical games in different treatments (as we did by introducing the WS treatment to compare the data with single card treatments) can be used to block incentive effects, allowing clean tests of the mapping. More generally, testing for robustness requires assuming additional structure than just a GFM mapping, at least for the purpose of setting up null hypotheses. This is potentially quite challenging but experimental economics has at its disposal an array of design methods. Deploying clever experimental designs should help researchers obtain the data necessary to identify the source of statistical test failures when they arise.

This being said, some features of the BDM data are unsettling. In particular, our treatments suggest that the mapping is influenced by contextual information. It also appears that subjects have very poor optimizing skills, resulting in very noisy data. This seems strangely at odds with a theory of misconception constructed around optimal behavior in whatever game a

Table 6
Optimal offers for risk neutral, risk averse, misconceived and non-misconceived subjects.

	GFM Type	Game BDM1	Game FP	Game BDM2
Risk Neutral Subject[a]				
Single Task	$\alpha(BDM\|BDM)=1$	$ 2	$ 3.5	$2
or WS (two tasks)	$\alpha(FP\|BDM)=1$	$ 3.5	$ 3.5	$3
Risk Averse Subject with Risk Preference x[a]				
Single Task[b]	$\alpha(BDM\|BDM)=1$	$ 2	$s_1(x)\in(2,\ 3.5]$	$2
	$\alpha(FP\|BDM)=1$	$s_1(x)\in(2,\ 3.5]$	$s_1(x)\in(2,\ 3.5]$	$s_2(x)\in(2,\ s_1(x)]$
WS (two tasks)[c]	$\alpha(BDM\|BDM)=1$	$ 2	$s_3(x)\in(s_1(x),\ 3.5]$	n.a.
	$\alpha(FP\|BDM)=1$	$s_4(x)\in(s_1(x),\ 3.5]$	$s_4(x)\in(s_1(x),\ 3.5]$	n.a.

[a] $(s_1(x),\ s_2(x),\ s_3(x),\ s_4(x))$ are optimal bids for a subject with risk preferences x. The optimal offer in the BDM1 is $2 if the agent has weak preferences over distributions ordered by first order stochastic dominance.
[b] The bid $s_1(x)$ is the optimal offer price for a risk averse agent who performs a FP auction once. A risk averse agent offers less in the BDM2 than in the BDM1.
[c] Risk decreases in the two-task case (WS treatment) because the agent receives two independently distributed payoffs. A risk averse agent offers a higher amount (equal for both tasks under GFM) relative to what she would offer in the FP auction alone ($s_3(x)>s_1(x)$ and $s_4(x)>s_1(x)$).

subject may have recognized. Implementing procedures that have been shown to reduce errors in the BDM (Bartling et al., 2015; Noussair et al., 2004; Quercia, 2016) and understanding why they are partially effective may help discern whether sub-optimal behavior in the BDM is due to misconception, non-standard preferences, or framing; but the question remains open.

Appendix A. General analysis of GFM with risk averse subjects

This Appendix extends the analysis to risk averse subjects and shows that the main conclusions reached in the text carry over. The analysis is more complex because unobserved risk preferences add a new source of subject heterogeneity. A formal theoretical framework is used to demonstrate that the FP-GFM mapping cannot explain some key features of the bid distributions even when one allows for arbitrary risk averse preferences. Table 6 and Propositions 1 and 2 of this appendix derive properties of the optimal strategies under the FP-GFM mapping with risk averse subjects. Section A.3 shows that the results presented in Section 4 of the paper violate Proposition 1 and some of the implications of Proposition 2.

We first discuss how risk aversion changes the optimal strategy in our treatments. To isolate the effect of risk aversion, we initially ignore possible misconceptions. The analysis is then extended to consider risk aversion and GFM simultaneously.

A.1. Risk aversion, no misconception

We begin with an analysis of risk aversion alone. Subjects correctly recognize the game: $G_g=\{g\}$. Denote by $X\subset\Re$ the set of possible risk preferences amongst subjects and $H(x)$ the associated population CDF. For example, an element of X could be a coefficient of absolute or relative risk aversion and we assume that X includes a value for risk neutral preferences. It will be useful for our analysis of the WS treatment to denote by $s^*(g|x,\ g')$ as the optimal strategy (under standard theory) in game g for a player of type x who plays games g and g' simultaneously. In Table 6, the games played are BDM1, BDM2 and FP and the strategies are bids. The optimal strategy for a single game is $s^*(g|x)$. With risk averse subjects and random payoffs, we typically have $s^*(g|x,\ g')\neq s^*(g|x)$. For the sake of simplicity, assume that there is a unique optimal strategy. Applying the theoretical prediction to each type, and aggregating across types using $H()$, we obtain a predicted distribution of optimal strategy. If subjects play games g and g', we denote the distribution of predicted strategies for game g by $F^*(s^*|g,\ g',\ H)$. We could add a random error to the individual strategies without affecting the main results.

The experimenter observes subjects' strategies for each game played. With two games, we denote by $F(s|g,\ g')$ the observed distribution of strategies for game g when subjects play games g and g'. When subjects recognize both games correctly, the inference problem is stated as follows: does there exist a type distribution $H()$ such that the predicted strategy distribution matches the observed strategy distribution?[20]

$$F(s|g,g')=F^*(s^*|g,g',H).$$

With a single game, this identity becomes $F(s|g)=F^*(s^*|g,H)$. Take, for example, the case of BDM1 in Table 6. The case under consideration corresponds to the lower panel (entitled 'Risk Averse'), line $\alpha(BDM|BDM)=1$, and column 'BDM1'. In this case, there is no distribution $H()$ that can rationalize our data. The predicted bid distribution calls for a mass of subjects at $2 under any distribution of risk preference. But the experimental distribution shows that subjects typically submit an

[20] More generally, one could match the joint distribution of bids in g and g'. Considering joint distributions does not add insight in our treatments.

offer price greater than their induced value. Adding a random bidding error does not change the conclusion because offer prices are not centered around \$2 as predicted by Table 6. Something else is needed to rationalize the distribution.

A.2. Risk aversion and misconception

We follow CP and assume that some subjects misconceive the BDM for FP: they behave according to the FP-GFM mapping. A subject's type is now two dimensional: risk preference and GFM. Denote by $F^*(s^*|g, H, \alpha)$ and $F^*(s^*|g, g', H, \alpha)$, the optimal distributions of strategies in a population with type distributions ($H()$, $\alpha()$), who respectively play game g, or both g and g'.

The main difference between GFM and correct GFR is that the researcher has two degrees of freedom to rationalize the observed bid distribution: $H()$ and the mapping $(\alpha(), G_g, G_{g'})$, which we denote just $\alpha()$ for simplicity. Under GFM, the inference problem is stated as follows: does there exist a pair of functions ($H()$, $\alpha()$) such that the predicted strategy distribution matches the observed strategy distribution?

$$F(s|g, g') = F^*(s^*|g, g', H, \alpha) \quad (1)$$

The optimal strategy, $s^*(g|x, g')$, is unchanged under GFM. What is new is the expectation that subjects behave optimally for the game they believe they are playing. GFM introduces the game mapping as a new source of player heterogeneity. GFM puts much structure on the data. The only free parameters that can rationalize the observed strategies are the functions $H()$ and $\alpha()$. Establishing robustness across multiple treatments that vary (g, g') means having a single pair of functions ($H()$, $\alpha()$), where the function $\alpha()$ can only depend on the game type (FP or BDM), while the function H cannot depend on the game type.

A.2.1. Comparing BDM1 and BDM2

To illustrate, take the case of BDM1 and BDM2. Under the null hypothesis of a constant mapping (see Assumption A.1): $\alpha(BDM1|x) = \alpha(BDM2|x) = \alpha(BDM|x)$. We say that some subjects are misconceived (they behave according to the FP-GFM mapping) if $\alpha(FP|BDM, x) > 0$ for some risk preference type x that is present in the population (formally, for some x such that $dH(x) > 0$).

Proposition 1. *Take any distribution of risk preference and assume that some subjects behave according to the FP-GFM mapping. We have $F(s|BDM2) > F(s|BDM1)$ for some $s \in S_0 \neq \emptyset$.*

The proposition states that the distribution of bids for BDM2 will lie to the left of the one for BDM1. The proof follows from Table 6. A misconceived risk neutral or type-x risk averse subject bids less in BDM2 than in BDM1 since $3 < 3.5$ and $s_2(x) < s_1(x)$. If misconception is present, we obtain that S_0 must contain $[s_2(x), s_1(x)]$ (and [3,3.5] if agents are risk neutral). Proposition 1 implies that the mean bid should be lower under BDM2 than BDM1 when risk aversion is allowed.

A.2.2. Joint analysis of FP, BDM1 and WS

GFM is also instructive in the analysis of the WS treatment because we can compare the distributions from the WS treatment to one another, as well as to their single task analogues (BDM1 and FP). However, risk aversion changes the analysis because misconception may influence strategies through an indirect channel. Even if a type x has no misconception about game g ($\alpha(g|g, x) = 1$), we must admit the possibility that a misconception for game g' can change the optimal offer price in game g when the two games are played simultaneously.

To illustrate, take line WS and column FP in Table 6. A subject bids $s_3(x)$ in *FP* if she does not misconceive the BDM and $s_4(x)$ if she does. Thus, Hypothesis 5 in the text does not necessarily hold with risk aversion. Stated differently, assumption A.2 does not imply treatment independence of strategies in the face of risk aversion. Instead, we have to leverage different implications of the theory to analyze the WS treatment when some subjects are risk averse.

Proposition 2. *Assume misconceptions take the form of the FP-GFM mapping*[21] *and some subjects are risk averse. For some s, we have:*

$$F(s|FP, BDM) < Min(F(s|FP), F(s|BDM, FP)) < Max(F(s|FP), F(s|BDM, FP)) < F(s|BDM).$$

Proof. Take a given type x. This type is misconceived with probability $\alpha(FP|BDM, x)$. The observed bids on FP alone, BDM alone, FP played together with BDM, and BDM played together with FP are respectively:

$$s_{FP}(x) = s_1(x)$$
$$s_{BDM}(x) = (1 - \alpha(FP|BDM, x))2 + \alpha(FP|BDM, x)s_1(x)$$
$$s_{FP|BDM}(x) = (1 - \alpha(FP|BDM, x))s_3(x) + \alpha(FP|BDM, x)s_4(x)$$
$$s_{BDM|FP}(x) = (1 - \alpha(FP|BDM, x))2 + \alpha(FP|BDM, x)s_4(x)$$

[21] GFM happens only from BDM toward FP, that is, $\alpha(FP|BDM, x) \geq 0$ and $\alpha(BDM|FP, x) = 0$.

Inequalities $s_1(x) < s_4(x)$ implies $s_{BDM}(x) < s_{BDM|FP}(x)$. Inequality $s_1(x) > 2$ implies $s_{FP}(x) > s_{BDM}(x)$. Inequality $s_3(x) > 2$ implies $s_{FP|BDM}(x) > s_{BDM|FP}(x)$. Inequalities $s_3(x)$, $s_4(x) > s_1(x)$ imply $s_{FP|BDM}(x) > s_{FP}(x)$. Putting these results together, we obtain,

$$s_{BDM}(x) < Min(s_{FP}(x)), s_{BDM|FP}(x)) < Max(s_{FP}(x), s_{BDM|FP}(x)) < s_{FP|BDM}(x).$$

This inequality holds for any x. The next step is to aggregate strategy distributions. We have, for example,

$$F(s|FP) = \int_X I(s_{FP}(x) \leq s)dH(x)dx = \int_X I(s_1(x) \leq s)dH(x)dx.$$

$$\begin{aligned} F(s|BDM) &= \int_X I(s_{BDM}(x) \leq s)dH(x)dx \\ &= \int_X (1 - \alpha(FP|BDM, x))I(2 \leq s)dH(x)dx + \int_X \alpha(FP|BDM, x)I(s_1(x) \leq s)dH(x)dx. \end{aligned}$$

Since $s_1(x) > 2$ for any x, we obtain $F(s|FP) < F(s|BDM)$. Aggregating similarly across types x delivers the remaining inequalities. □

Proposition 2 implies that in the most general case with risk aversion and GFM:

$$3.5 > E(FP|BDM) > Max(E(FP), E(BDM|FP)) > Min(E(FP), E(BDM|FP)) > E(BDM) > 2.$$

Three benchmark cases are insightful. With no risk aversion and no misconception, we obtain the most restrictive prediction

$$E(FP|BDM) = E(FP) = 3.5 > 2 = E(BDM|FP)) = E(BDM).$$

With risk aversion and without misconception, we obtain

$$3.5 > E(FP|BDM) > E(FP) > E(BDM|FP)) = E(BDM) = 2.$$

With risk neutrality and misconception, we obtain[22]

$$E(FP|BDM) = E(FP) = 3.5 > E(BDM|FP)) = E(BDM) > 2.$$

With risk neutrality, distribution of strategy in a given game is independent of the other games the subject is playing. This explains why the top panel in Table 6 holds for both the single task case and the WS (two tasks) case.

A.3. *Application to treatments* BDM1, BDM2, FP *and* WS

We apply Proposition 1 and 2 and an additional prediction derived from Table 6 to our four treatments. Our main findings are as follows:

1. Take BDM1 and BDM2 jointly. Offers in each treatment taken separately imply that misconception must be present. But Proposition 1 then implies that $E(BDM2) < E(BDM1)$ which is rejected (see Tables 1 and 4). There do not exist a pair of functions $H()$ and $\alpha()$ such that condition (1) holds for BDM1 and BDM2.
2. Take BDM1 and FP jointly. According to Table 6 offer prices should be identical in FP and BDM1 if all subjects were misconceived (subject type x bids $s_1(x)$ in both treatments). This is inconsistent with the finding that $3.5 > E(FP) > E(BDM1) > 2$ (see Table 1 and Fig. 2 (left panel)). We conclude that some subjects are not misconceived. If that's the case, however, Table 6 implies that there should be more bids at $2 in BDM1 than in FP1. This is not the case.
3. Take WS-BDM, WS-FP, FP and BDM1 jointly and assume that some, but not all, subjects are misconceived. Proposition 2 implies that $E(FP|BDM) > E(BDM|FP)$ and $E(FP|BDM) > E(FP)$ which are both rejected (see Tables 1 and 4). Note, however, that the other inequalities from Proposition 2, $3.5 > E(FP|BDM)$ and $Min(E(FP), E(BDM|FP)) > E(BDM) > 2$, are not rejected in Tables 1 and 4.

To sum up, the FP-GFM mapping cannot explain bidding strategies when all treatments are taken jointly, even if we admit any distribution of risk aversion across subjects. It is true that GFM and risk aversion can explain bidding strategies in each treatment in isolation, and the CP benchmark holds with risk averse subjects. For each treatment, there exist a pair of functions $H()$ and $\alpha()$ such that the evidence from Table 1 and Figs. 2 and 3 is consistent with Eq. (1). Thus, GFM passes the test when each game is considered separately. But this is not sufficient to ascertain that the FP-GFM mapping is robust, and this is where the comparison of treatment variations is particularly useful and instructive.

If the FP-GFM mapping is robust and preferences are distributed equally across treatment groups, the resulting data should be consistent with the incentive effects or the other assumptions of a robust mapping for a single distribution of heterogeneity. But this is not the case. Clearly, the subject's choices reveal much misconception amongst subjects. However, as we have shown here, the data cannot be explained by a single mapping, even when risk aversion is incorporated into the analysis.

[22] Under GFM and RN, we obtain, $F(s|FP, BDM) = F(s|FP) < F(s|BDM, FP) = F(s|BDM)$ since $s_{FP}(x) = s_{FP|BDM}(x) = 3.5 > s_{BDM|FP}(x) = s_{BDM}(x) = 2$. This case is not interesting because it cannot explain FP bids.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2019.01.003.

References

Bartling, B., Engl, F., Weber, R.A., 2015. Game form misconceptions are not necessary for a willingness-to-pay vs. willingness-to-accept gap. J. Econ. Sci. Assoc. 1 (1), 72–85.
Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behav. Sci. 9 (3), 226–232.
Berry, J., Fischer, G., Guiteras, R.P., 2018. Eliciting and utilizing willingness to pay: evidence from field trials in Northern Ghana. Working Paper.
Binmore, K., Shaked, A., 2010. Experimental economics: where next? J. Econ. Behav. Org. 73 (1), 87–100.
Bohm, P., Lindén, J., Sonnegård, J., 1997. Eliciting reservation prices: Becker–DeGroot–Marschak mechanisms vs. markets. Econ. J. 107 (443), 1079–1089.
Camerer, C.F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., et al., 2016. Evaluating replicability of laboratory experiments in economics. Science 351 (6280), 1433–1436.
Camerer, C.F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B.A., Pfeiffer, T., et al., 2018. Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. Nat. Hum. Behav. 2 (9), 637.
Cason, T.N., Plott, C.R., 2014. Misconceptions and game form recognition: challenges to theories of revealed preference and framing. J. Polit. Econ. 122 (6), 1235–1270.
Chang, A.C., Li, P., 2017. A preanalysis plan to replicate sixty economics research papers that worked half of the time. Am. Econ. Rev. 107 (5), 60–64.
Chou, E., McConnell, M., Nagel, R., Plott, C.R., 2009. The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person guessing game. Exp. Econ. 12 (2), 159–179.
Duvendack, M., Palmer-Jones, R., Reed, W.R., 2017. What is meant by "replication" and why does it encounter resistance in economics? Am. Econ. Rev. 107 (5), 46–51.
Fehr, D., Hakimov, R., Kübler, D., 2015. The willingness to pay–willingness to accept gap: a failed replication of plott and zeiler. Eur. Econ. Rev. 78, 120–128.
Freeman, D., Manzini, P., Mariotti, M., Mittone, L., 2016. Procedures for eliciting time preferences. J. Econ. Behav. Org. 126, 235–242.
Georganas, S., Levin, D., McGee, P., 2017. Optimistic irrationality and overbidding in private value auctions. Exp. Econ. 20 (4), 772–792.
Harrison, G.W., 1990. Risk attitudes in first-price auction experiments: a Bayesian analysis. Rev. Econ. Stat. 541–546.
Horowitz, J.K., 2006. The Becker–DeGroot–Marschak mechanism is not necessarily incentive compatible, even for non-random goods. Econ. Lett. 93 (1), 6–11.
Isoni, A., Loomes, G., Sugden, R., 2011. The willingness to pay willingness to accept gap, the endowment effect, subject misconceptions, and experimental procedures for eliciting valuations: comment. Am. Econ. Rev. 101 (2), 991–1011.
Kagel, J.H., Levin, D., 1993. Independent private value auctions: bidder behaviour in first-, second-and third-price auctions with varying numbers of bidders. Econ. J. 103 (419), 868–879.
Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. J. Polit. Econ. 1325–1348.
Karni, E., Safra, Z., 1987. Preference reversal and the observability of preferences by experimental methods. Econometrica 675–685.
Maniadis, Z., Tufano, F., List, J.A., 2014. One swallow doesn't make a summer: new evidence on anchoring effects. Am. Econ. Rev. 104 (1), 277–290.
Noussair, C., Robin, S., Ruffieux, B., 2004. Revealing consumers' willingness-to-pay: a comparison of the BDM mechanism and the Vickrey auction. J. Econ. Psychol. 25 (6), 725–741.
Plott, C.R., Zeiler, K., 2005. The willingness to pay-willingness to accept gap, the endowment effect, subject misconceptions, and experimental procedures for eliciting valuations. Am. Econ. Rev. 95 (3), 530–545.
Plott, C.R., Zeiler, K., 2011. The willingness to pay willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations: reply. Am. Econ. Rev. 101 (2), 1012–1028.
Quercia, S., 2016. Eliciting and measuring betrayal aversion using the bdm mechanism. J. Econ. Sci. Assoc. 2 (1), 48–59.
Rutström, E.E., 1998. Home-grown values and incentive compatible auction design. Int. J. Game Theory 27 (3), 427–441.
Vecchio, R., Annunziata, A., 2015. The effect of monetary versus nonmonetary endowment on WTP in BDM auctions. Appl. Econ. Lett. 22 (2), 137–140.