Truth, Honesty, and Strategic Interactions

Bernabe, Angelique and Hossain, Tanjim and Yu, Haomiao

Analysis Group, University of Toronto, Ryerson University

29 September 2021

Truth, Honesty, and Strategic Interactions *

Angelique Bernabe,[†] Tanjim Hossain,[‡] Haomiao Yu[§]

September 29, 2021

Abstract: We experimentally investigate how introducing the concept of *truth* in the natural context of a game affects player behavior using two games. Two players simultaneously make reimbursement claims for a damaged product, where players' payoffs depend only on their claims but *not* on the true price. Both games are dominance-solvable, and one of them has a strictly dominant strategy equilibrium, which many participants easily identified. Yet, claims in our experiments are significantly affected by the price. Analyzing the role of truth on participants' choices, we show that one needs strategic considerations and preferences for honesty to explain the results.

JEL Classification: C72, C91, D91

Keywords: Preference for Honesty, Truth-telling in Games, Traveler's Dilemma, Dominance-solvable Games, Strictly Dominant Strategy, Level-k thinking

---

*Please direct all correspondence to tanjim.hossain@rotman.utoronto.ca. We thank Navin Kartik, David Freeman, Subhasish Chowdhury for stimulating comments and references, and seminar/conference participants at the CEA Annual Conference, University of Bath, and University of Toronto for their useful feedback. Financial support from SSHRC Insight Grants No. 502502 and No. 435-2019-1069 is gratefully acknowledged.

[†]Analysis Group, Montreal, QB H3B 0G7, Canada.

[‡]Department of Management-UTM and Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6, Canada.

[§]Department of Economics, Ryerson University, Toronto, ON M5B2K3, Canada.

1 Introduction

How does a concept of truth in the natural context of a game affect player behavior in that game? The true underlying state is assumed away or ignored in many commonly analyzed games, including the famous Prisoner's Dilemma. In real-life situations, for example, whether to confess about a crime may depend on whether the person had committed the crime, even if that consideration does not affect the outcome. Another example would be an ultimatum game-type scenario. In particular, even if a retailer has complete bargaining power over a buyer and knows the buyer's willingness to pay, she may consider her cost of acquiring the product in making a price offer. Such behavior may arise from some preference for being consistent with the truth. The truth may also affect a player's behavior by affecting her beliefs regarding the other player's behavior. In this paper, we run a set of laboratory experiments utilizing a variation of the well-known game *Traveler's Dilemma* (Basu, 1994) and a new game that uses the same basic setup to investigate how the truth affects the behavior of players.

In these two games, two travelers simultaneously choose a reimbursement claim for a souvenir that an airline company damaged. The airline company is not aware of the price of the souvenir and, hence, uses the claims to determine each traveler's reimbursement amount. Unlike in the traditional Traveler's Dilemma, both travelers in our games are aware of the common purchase price, which represents the truth. In the game *Regular*, the reimbursement function follows that of Traveler's Dilemma. If the claims are the same, both travelers receive their claims. Otherwise, the traveler making the lower claim receives a reward on top of her own claim, while the traveler making the higher claim gets the lower claim minus a penalty. In the game *Upward*, the travelers still get their claims if their claims are the same. If not, the traveler making the higher claim receives a reward on top of her own claim, and the traveler making the lower claim gets the lower claim minus a penalty instead. *Regular* is dominance-solvable, and, under common knowledge of rationality, both travelers make the lowest possible reimbursement claim in the unique equilibrium. In *Upward*, for both travelers, it is a strictly dominant strategy to claim the highest possible reimbursement claim, regardless of their purchase price. This constitutes the unique equilibrium of *Upward* as long as travelers maximize their payoffs, even if they have incorrect beliefs or limited power of cognitive reasoning.

Our experimental results from *Regular* suggest that the truth affects player behavior in a large way, where many players' claims seem to be based on the price, and other players seem to respond to that. This behavior may arise due to a preference for honesty among some players and other players responding to that. However, any action by a player can be rationalized based on her beliefs about the other player's actions in *Regular*. Moreover, iterated reasoning to reach a dominance-solvable outcome is complex and depends on players' cognitive skills and their beliefs about the other players. The observed results, thus, may also arise if such complexities coax some players to simply choose claims around the price and others to respond to that. On the other hand, such complexities are not present in solving the game *Upward*. Even though there is a large action set, there is a strictly dominant strategy. Nonetheless, we observe that many players who play the dominant strategy when the price is kept unknown switch to claims equaling the price or slightly more in the game *Upward*. Such behavior is hard to explain without a preference for the truth. The two games together suggest that while almost a quarter of the players exhibit behavior consistent with a strong preference for making an honest claim, a large proportion of the players also behave strategically. Moreover, an equally large proportion seems to take both preferences for honesty and strategic considerations into account.

In our experiments, participants play either the *Regular* game or the *Upward* game in multiple periods under different prices of the souvenir, including a period where the price is not provided. We find that the average reimbursement claim across all periods in *Regular* is much higher than the equilibrium prediction of the minimum possible claim. In *Upward*, the average claim is somewhat lower than the equilibrium prediction of the maximum possible claim. In both games, participants' choices of claims are affected by the price. We characterize the claims of 81% of the participants' in *Regular* as price sensitive, and this share is 60% in *Upward*. While participants forming beliefs about the other player's choices based on the price may explain price sensitivity in *Regular*, such beliefs cannot explain it in *Upward*, where choosing the maximum possible claim is a strictly dominant strategy to report. Reimbursement claims are closer to the price under *Regular* than under *Upward* at all price levels. Thus, the truth affects player behavior in both games, and this effect is stronger in *Regular*.

Digging deeper into the claims under different prices for each participant, we find significant variation in behavior across participants. We use these variations to explore how the truth affects player behavior in games. This allows us to characterize the participants into different categories. The existence of a strictly dominant strategy makes the categorization of participants in *Upward* simpler, and we discuss the behavior of the participants in that game first. In the game *Upward*, a sizable fraction of the participants choose the dominant strategy of the maximum claim in the period where the price is not explicitly mentioned and then choose claims equaling the price. Such behavior is hard to explain with beliefs about the other player's actions or cognitive biases and limitations. One would require some preferences for making a truthful claim to explain such behavior. Overall, almost a quarter of the participants choose claims at the price or slightly higher, and we categorize them as *Honest*. Around 30% of the participants in *Upward* choose the dominant strategy claim in all periods and are categorized as *Strategic*. Almost all of the remaining participants make claims between the price and the dominant strategy. Such behavior can be explained by an objective function that tries to maximize a convex combination of the material or monetary payoff from the game and a psychological payoff from truth-telling.

In the game *Regular*, categorizing participants is more complex as all but the highest possible claim can be a best response to some action by the other participant. However, we find only around 2% of the participants make the lowest possible claim, which is the equilibrium strategy under full rationality, in all periods. Rather, almost 8% of the participants claim right below the price, which can constitute equilibrium strategies under limited reasoning if players are believed to anchor their claims on the price. Note that we do not see any such behavior in *Upward*, which is consistent with the fact that choosing the dominant strategy, *i.e.*, the maximum claim, is still a player's best response to the other player choosing the price as her claim. Almost 31% of the participants in *Regular* make claims that equal the price or are slightly higher. Following *Upward*, these participants are categorized as *Honest*. Nonetheless, this can also be explained by a naive strategy of just choosing the price without any preference for honesty. Finally, claims of almost a third of the participants in *Regular* suggest that they believe that the other player will choose a reimbursement claim greater than the price. Distributions of claims from the two games together indicate that a large share of participants exhibit some preference for choosing the price as their claim. The presence of a price leads to a large share of claims around the price, specifically, in *Regular*. As a result, claims are better coordinated, tacitly, in *Regular*. This results in a higher payoff, on average, when a price is mentioned relative to when no price is mentioned. Hence, in a game like Traveler's Dilemma, where strategic considerations may lead to Pareto inferior outcomes for players, preferences for honesty may actually increase their payoffs by making it easier to coordinate around the true state.

The main goal of our experiments is to explore how a concept of truth affects player behavior in games. On the one hand, there is a large literature on theoretical explorations of games where players have some preferences for being honest or telling the truth (e.g., Graetz et al., 1986; Erard and Feinstein, 1994; Matsushima, 2008; Kartik, 2009; Chen, 2011). On the other hand, while many papers provide experimental evidence of preferences for truth-telling (see Rosenbaum et al. (2014) for a survey and Abeler et al. (2019) for a meta-analysis), typically they are conducted in an individual choice setting. To be sure, some experimental studies investigate preferences for honesty in game settings, too; See Gneezy (2005), Charness and Dufwenberg (2006) and subsequent papers including Gneezy et al. (2020) and Besancenot and Vranceanu (2020), for example. However, in the games considered in those papers, only one player can choose to behave honestly or not. On the contrary, both players know the truth or the price in our games *Regular* and *Upward*, and their (material) payoff functions – which are not affected by the truth – are publicly known. As a result, it is easier to identify a preference for honesty and a preference for material payoffs in our experimental setting. Additionally, unlike in many of the experiments that follow the setup from Fischbacher and Föllmi-Heusi (2013), the researcher also knows the truth in our experimental setting. Instead of relying on statistical inference as in those papers or on comparison of behavior under different treatments (Mazar et al., 2008; Jiang, 2013; Potters and Stoop, 2016), our framework allows us to *cleanly* identify preferences for honesty by directly comparing players' actions with the objective truth that is commonly known to the players. Thus, we offer a new experimental paradigm for studying preferences for honesty that is different from the two most common frameworks—the die rolling paradigm (Fischbacher and Föllmi-Heusi, 2013) and sender-receiver games (Gneezy, 2005; Hurkens and Kartik, 2009).

Our experiments should not be viewed as a test of Traveler's Dilemma. Rather we use the dilemma as the base for one of our games, *Regular*, to investigate how the truth affects people's behavior. It is, nonetheless, worth noting that even though the price of the souvenir is mentioned in Basu (1994) and Basu (2007), the price is rarely mentioned or discussed in the experimental explorations of Traveler's Dilemma.[1] This is perhaps because the equilibrium is not affected by prices with rational players and under most alternative solution concepts considered in the literature. However, it is common in practice that a traveler would know the price of the souvenir where the truth may affect their claim decisions. While strategic considerations may enter into the traveler's choice of reimbursement claim, the price provides a benchmark for an honest claim. It is, thus, essential to see how a traveler's choice for the reimbursement claim may be affected by the truth that may capture preferences for honesty, in addition to strategic considerations. Our experimental design and results answer such questions and also illustrate the importance of experimentally exploring the truth in games where the truth is naturally defined.

The rest of the paper is organized as follows. In Section 2, we lay out the experimental design of the paper after describing the games played in the experiment in detail. We present our empirical findings in Section 3. In Section 4, we suggest some explanations for our experimental finding and provide a categorization of the participants. Section 5

[1] Upon the completion of this paper, we found that Umbhauer (2019) independently tested how a commonly known price may affect claims in Traveler's Dilemma using a classroom experiment. They view prices as merely viewed as focal points and do not investigate strategic behavior based on the price. On the other hand, our goal is to utilize commonly-known prices to investigate how truths may affect players' decisions in both Traveler's Dilemma and a new game with a dominant strategy equilibrium.

concludes the paper. We present the experimental instructions in the Appendix.

2 Experimental Design

Our main experiment revolves around two 2-player games. We will first describe the experimental setup and the structure of the games, and discuss the payoff functions and the equilibria in detail. Then, we will describe our experimental procedure step-by-step and present participant information.

2.1 Experimental Setting and the Two Games: *Regular* and *Upward*

Suppose two travelers who do not know each other happened to have bought the same souvenir during their trips to the same country. Upon returning from a vacation, each of them found that her souvenir was ruined because of the airline's negligence. Each of them independently reports the damage to the airline. Realizing that the souvenirs are identical and not knowing the exact price, the airline asks the travelers to report the price of the souvenir for their reimbursement claims and explains to them how the compensation would be determined. This serves as the premise for the two games that the participants play in our experiments.

The reimbursement scheme that the airline adopts in the game *Regular* is as follows: if the claims are equal, both travelers receive their claim exactly, and if they are unequal, the lower claim is considered the base for reimbursement with the traveler making the lower claim receiving a small bonus and the one making a higher claim receiving a small penalty. This reimbursement scheme is exactly the same as that in the well-known Traveler's Dilemma game. The game *Upward* uses the same basic setup, but with a different reimbursement scheme. Specifically, if the claims are unequal, the traveler making the higher claim receives her own claim plus a small bonus, and the traveler making the lower claim receives her own claim minus a small penalty. In both games, we explicitly consider the price of the souvenir and assume that each player is aware of the price of the souvenir as would be natural in practice.

We now present the games formally. Suppose the two players indexed by $i \in \{1, 2\}$ and they both paid a price of p for the souvenir. To get reimbursed for the price, player i can make a claim of $c_i \in C \equiv \{\underline{c}, \underline{c}+1, ..., \bar{c}-1, \bar{c}\}$ where $\underline{c}$ and $\bar{c}$ are positive integers that are, respectively, the lower and upper limits of the price and also of acceptable claims. Let $m_i^A(c_i, c_j)$ be the reimbursement that player i receives in the game $A \in \{R, U\}$, if players i and j claim c_i and c_j respectively. Here R and U represent games *Regular* and

Upward, respectively. The reimbursement amounts are presented below:

$$m_i^R(c_i, c_j) = \begin{cases} c_i + k, & \text{if } c_i < c_j, \\ c_i, & \text{if } c_i = c_j, \\ c_j - k, & \text{if } c_i > c_j; \end{cases} \quad m_i^U(c_i, c_j) = \begin{cases} c_i - k, & \text{if } c_i < c_j, \\ c_i, & \text{if } c_i = c_j, \\ c_i + k, & \text{if } c_i > c_j, \end{cases} \tag{1}$$

where $k \geq 2$ is a commonly known integer.

Assuming both players are fully rational, the Nash equilibria of the two games can easily be characterized. In *Regular*, making a smaller claim than the other player is rewarded. Player i's best response to player j making a claim of c_j is to claim $max\{\underline{c}, c_j - 1\}$. As a result, the game is dominance-solvable. In the unique Nash equilibrium, both players will make a claim of $\underline{c}$, independent of the price of the souvenir. On the other hand, in the game *Upward*, making a higher claim than the other player is rewarded. As a result, it is a dominant strategy to make a claim of $\bar{c}$. Consequently, both players making a claim of $\bar{c}$, independent of the value of p, is the unique Nash equilibrium. This leads to the following observation:

Claim 1. *Each game has a unique Nash equilibrium. Independent of the price, both players claim $\underline{c}$ in Regular and both players claim $\bar{c}$ in Upward.*

The game *Regular* is dominance-solvable, leading to a unique equilibrium that is independent of the truth – the price in this context, when the players are rational. However, this may not hold when players have incorrect beliefs, have limited cognitive abilities, or have preferences beyond material payoffs. For example, any reimbursement claim within $\underline{c}$ and $\bar{c} - 1$ can be rationalized based on a player's belief about the other player's strategy. If players have limited cognitive ability according to level-k (Nagel, 1995; Stahl and Wilson, 1995) or cognitive hierarchy models (Camerer et al., 2004), a player's chosen strategy would depend on her cognitive level, the behavior of the level-0 players, and also the distribution of cognitive levels, for a cognitive hierarchy model. With such alternative solution concepts, the price may affect the outcomes by affecting some players' beliefs or actions. On the other hand, a preference for taking actions closer to the truth in addition to a preference for material payoff may also make equilibrium behavior price-dependent in *Regular*. Thus, if the true price affects player behavior in *Regular*, that may potentially be explained by some sort of preference for honesty, as well as behavioral game-theoretic explanations.

In the game *Upward*, claiming $\bar{c}$ strictly dominates any other action in terms of material payoffs. Even though the game has a large action space, the dominant strategy makes a payoff maximizing player's task very simple. Such a player will choose the highest possible claim of $\bar{c}$, independent of her beliefs about other players. Thus, under most alternative solution concepts, such as level-k or cognitive hierarchy models, the price will not have any impact on strategic players. In particular, any level-k player with $k > 0$ will claim $\bar{c}$. However, a preference for taking actions closer to the truth will make the behavior of strategic players in *Upward* price-dependent. This suggests that if the true price affects player behavior in *Upward*, some preference for behavior close to the truth may be necessary to explain that.

2.2 The Experimental Procedure

This pen and paper experiment was conducted at Ryerson University in Canada between January 20 and February 7 of 2020.[2] We ran 38 sessions that were run six days a week, with two to four sessions a day. We recruited participants from Ryerson University, the University of Toronto, and the surrounding community using campus email lists and flyers around Ryerson University.

We started a session by showing instructional videos to the participants, followed by a comprehension test with three questions about the rules of the main exercise of the experiment where participants made reimbursement claim for a damaged souvenir for a number of periods. This was followed by an explanation of the correct answers to ensure participants' perfect understanding of the rules. Then the participants completed the main reimbursement claim exercise. Next, we elicited participants' perceptions about social norms regarding a number of dishonest acts. Finally, they answered some questions about their demographic characteristics. They were also asked to describe their playing strategies in the main part and report the number of questions they got correct in the comprehension test. The experiment featured two treatments – *Regular* and *Upward*. The two treatments differed only in terms of the game in the main part of the experiment. All participants within a session received the same treatment. A session lasted slightly more than 30 minutes, on average. To ensure consistency throughout all sessions, we played instructional videos using clear graphics to explain the rules as well as written instructions. Participants had the opportunity to ask questions if the rules were unclear. In addition, all sessions were run by the same researcher following a script. In total, 425 participants participated in this experiment, 212 in *Regular* and 213 in *Upward*. No participant participated in more than one session. The experimental exercises are described in greater detail below. The instructions used in the experiment are presented in the appendix.

[2]This experiment took place before the COVID-19 pandemic, when travel was not restricted, and participants could easily relate to air travel and claiming compensation from an airline company.

Main Exercise: Reimbursement Claim

In the reimbursement claim elicitation exercise, there were six periods. In each period, a participant played a reimbursement claim game, as described in Section 2, according to the treatment. In treatment *Regular*, participants played the game *Regular* six times and they played the game *Upward* six times in treatment *Upward*. In all periods, the claims had to be an integer between 20 and 100 and the penalty size was 5. That is, $\underline{c}$, $\bar{c}$, and k equaled 20, 100, and 5, respectively. We mentioned in the instructions that, in some periods, a price will be mentioned, and in other periods no price will be mentioned. However, we did not mention in which periods a price would be mentioned and what the potential prices might be. During the exercise, we provided no price in the first period. In the following five periods, prices of 25, 40, 60, 75, and 95 were used. These prices were presented in a randomly chosen sequence for each participant. Participants received no feedback regarding the outcomes of the games during the session. After all the sessions were completed, a participant was randomly matched with another participant who received the same treatment but was potentially in a different session. For this pair of participants, we matched the periods according to their prices to determine the outcome of the games. Participants were paid for a randomly chosen game. All participants were aware of this payment scheme. The currency of the game was denominated in "Rye$".

Social Norm Elicitation

After the reimbursement claim exercise, we elicited the participants' perception of social norms surrounding specific dishonest acts using methods similar to those in Krupka and Weber (2013). They were presented with four scenarios describing a dishonest act. For each act, they had to report how socially inappropriate that act was considered on a 5-point scale, ranging from *socially inappropriate* to *socially appropriate*. For each question, if a participant chose the most common answer among all the participants' answers across all sessions, she received Rye$10 and received nothing otherwise. Thus, a participant had incentives to guess how all participants would view the social appropriateness of the scenarios. Participants' views of social acceptability of different dishonest acts provide us with a measure of their preferences for honesty in an individual setting.

For each dishonest act, we constructed a dummy variable to indicate whether the participant considered the act broadly socially appropriate. Specifically, this dummy variable equaled 1 if the participant chose the act to be *value neutral*, *somewhat socially appropriate*, *socially appropriate* and 0 if she chose *somewhat socially inappropriate*, *socially inappropriate*. The percentages of participants for whom the dummy variable equaled 1 are presented in Table 1, for each treatment. The difference between the two treatments is not statistically significant for any of the four acts. Among the four dishonest acts, lying to the airline about a reimbursement claim is considered broadly socially appropriate most often, by almost 28% of the participants. This is followed by not paying for a cheap item when the item is mistakenly not charge, followed by not paying the subway fare, and followed by not paying for an expensive item, which is considered appropriate by slightly more than 9% of the participants. This table shows that participants predominantly perceived lying about reimbursement claims to be socially inappropriate. This would support a view that the reimbursement claims they made prior to this exercise reflect their perception about the appropriateness of being truthful in such claims.

Table 1: Percentage of Participants Finding An Act Broadly Socially Appropriate

Act	*Regular*	*Upward*	*p*-value of Difference
Not paying for a tomato in a grocery store	21.2%	19.7%	0.701
	(0.41)	(0.399)	
Not paying for leather gloves in a clothing store	9.9%	8.5%	0.604
	(0.299)	(0.279)	
Not paying subway fare	13.7%	14.6%	0.796
	(0.344)	(0.353)	
Lying about the price of a souvenir to an airline	28.8%	27.2%	0.724
	(0.454)	(0.446)	
Observations	212	213	

Notes: Standard deviations are in parentheses.

Comprehension Test

Recall that the participants completed a comprehension test before completing the reimbursement claim exercise. After eliciting social norms, we asked the participants to report the number of questions in the comprehension test they got correct and were paid Rye$5 for each correct question. Note that they were paid based on their claim here, independent of the actual number of questions they got correct. Thus, a participant was incentivized to report that she got all questions in the comprehension test correct, independent of the truth.

In the comprehension test, a participant was asked to calculate payoff according to the reimbursement rule specified in the main game under three scenarios. In the *Regular* treatment 31.1% of the participants solved all three questions correctly, while 77.5% of the participants did so in the *Upward* treatment. This indicates that the rules for *Upwards* were relatively simpler to understand. Nonetheless, as we re-explained the rule in each session and confirmed that all the participants understood the payoff scheme, we are confident that participants understood the rules of the game in both treatments before completing the reimbursement claim exercise. This difference in initial comprehension, however, provides the participants in *Regular* more room to dishonestly exaggerate about their performance in the comprehension test. Unsurprisingly, 32.7% of the participants in

Regular exaggerated their performance in the comprehension test, as opposed to 13.3% in *Upward*. When we look at only those who correctly solved two or fewer questions, these percentages become closer and reverse in their relative magnitude—47.2% and 59.6% in the *Regular* and *Upward* treatments, respectively. In other words, 49.7% of the participants who had the opportunity to benefit by lying about their performance in the comprehension test chose not to do so.

Questionnaire and Payment

Participants were also asked to explain their reasoning regarding how they chose their compensation claim in the main experiment. They were paid Rye$10 to answer this question. Using their answers, we created two dummy variables, one indicating whether they mention the price and one whether they mention other player's strategy or behavior. The first one suggests whether the price was a determining factor for their claims and the second one indicates whether they incorporated strategic concerns. They may have mentioned both price and other players or neither of them. Percentages of participants who mentioned the two factors in explaining their choices of compensation claims are presented in Table 2. The table shows that participants in the *Regular* treatment were more likely to consider both the price and the other player's behavior than in the *Upward* treatment. This is consistent with the fact that while the optimal strategy for a player in the *Regular* game depended on her beliefs about the other player's strategy, there was a dominant strategy for each player in the *Upward* game. We control for these two factors in our empirical analysis.

Table 2: Decision-making Factors Mentioned in the Questionnaire

	Regular	*Upward*	*p*-value of Difference
Price	75.9%	61.0%	<.001
	(0.428)	(0.489)	
Behavior of the other player	38.7%	21.1%	<.0001
	(0.488)	(0.409)	
Observations	212	213	

Notes: Standard deviations are in parentheses.

At the end of this part, we collected demographic information from participants such as gender, age, student status, etc. Table 3 summarizes these characteristics under the two treatments. The table shows that the participants in the two treatments were balanced other than in terms of their gender composition. This happened just by chance as the two treatments were run in the same time period, and a session was randomly assigned to a treatment. In our empirical analysis, we control for the gender of a participant to mitigate this imbalance.

Table 3: Individual Characteristics

	Regular	*Upward*	p-value of Difference
Female	49.5%	38.5%	0.022
	(0.501)	(0.488)	
Age (in years)	21.61	21.03	0.282
	(6.686)	(4.091)	
Student	93.9%	95.8%	0.382
	(0.24)	(0.202)	
Class year in university	2.072	2.123	0.684
	(1.269)	(1.195)	
Engineering major	29.2%	29.6%	0.940
	(0.456)	(0.457)	
Business major	46.2%	41.8%	0.358
	(0.5)	(0.494)	
Had taken game theory	10.0%	10.4%	0.879
	(0.316)	(0.322)	
Observations	212	213	

Notes: Standard deviations are in parentheses.

After all sessions were completed, earnings in Rye$ from all parts of the experiment were calculated for all the participants. This was then converted to Canadian $ (CA$) at the rate of CA$1 for Rye$10. Participants were paid electronically using Amazon, Tim Horton's, or Starbucks gift cards based on their total income for the experiment within a month of the experiment. The average payment was CA$16.05. In the remainder of the paper, we will suppress the currency denominator Rye$ for prices and reimbursement claims.

3 Experimental Results

Our experimental design allowed us to elicit participants' natural leanings towards the "truth" in non-strategic settings, as described in the previous section. Having found significant evidence of honest behavior in terms of perception of social norms and reporting of personal performance, we next explore whether participant behavior is affected by the presence of truth in two-player games. This allows us to investigate how strategic considerations interact with potential preferences for truth or honesty. Our main variable of interest is the amount of reimbursement claim in each of the six periods of the main exercise. First, we will present descriptive statistics of average reimbursement claims and then investigate the determinants of these claims. Moreover, by comparing a reimbursement claim to the associated price of the souvenir, we will describe how "honest" claims are under different prices. While we refer to claims closer to the price as more honest, we are agnostic about the sources of honest behavior such as preferences for acting closer to the truth, strategic responses to different kinds of beliefs, cognitive limitations, etc. Next, we will investigate the distributions of the reimbursement claims across participants under the two treatments. Finally, we will analyze the payoffs of the two treatments and investigate whether knowing the price of the souvenir leads to implicit coordination between players.

3.1 Average Reimbursement Claims

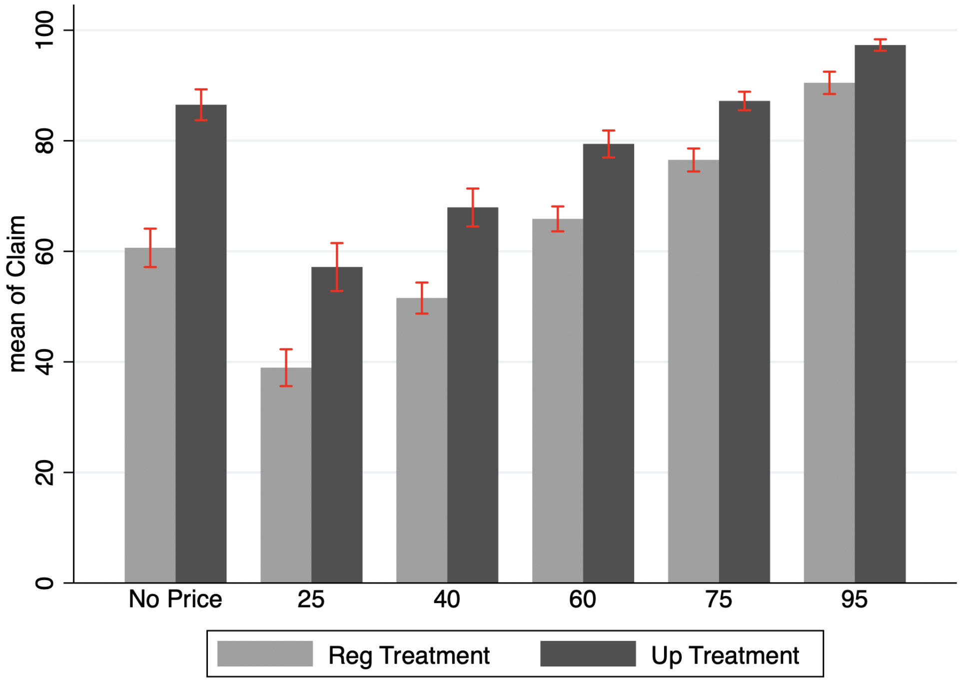
Equilibrium claims are very different across the two treatments. However, within a treatment, it is not affected by the presence or absence of a price and the price itself. Equilibrium claims are 20 and 100, respectively, for *Regular* and *Upward* treatments. Figure 1 presents average claims by participants by treatment and price. Recall that no information was provided about the price of the souvenir in the first period of the main exercise. In the following five periods, the prices were 25, 40, 60, 75, and 95, with the price sequence randomly chosen.

Average period 1 claims are 60.63 and 86.52 in *Regular* and *Upward*, respectively. Hence, claims are closer to the equilibrium when the game has a dominant strategy equilibrium. The figure also illustrates that average claims in periods where a price is mentioned are increasing in the price under both treatments. Moreover, when prices are 25 or 40, average claims are below period 1 average claim under both treatments. For a price of 60, the average claim is higher than the average claim in period 1 under *Regular*, but the opposite is true under *Upward*. For prices of 75 and 95, however, average claims are above the average claim in period 1 under both treatments.[3] Under *Upward*, the average claim in the period without a price is greater than the average of claims in the periods with a price. The opposite is true under *Regular*.

In the existing literature on Traveler's Dilemma experiments, where prices are not mentioned, reimbursement claims depend on the size of the penalty and rewards (Capra et al., 1999; Goeree and Holt, 2001; Basu et al., 2011). Specifically, when the reward/penalty amount is small relative to the range of possible claims, as is the case in our experiment, the distribution of claims is highly concentrated close to the highest possible claim. However, when the reward/penalty amount becomes rather large, players tend to make claims close to the theoretical prediction of the minimum claim. In our *Regular* treatment, participants chose around the average of possible claims in period 1, even though the reward/penalty amount is small. This may have resulted from the fact that while no price was mentioned in that period, the concept of a true price is still salient in *Regular*, unlike in classic Traveler's Dilemma experiments. Moreover, average claims in the fol-

[3] However, the difference is not statistically significant for the price of 75 in *Upward*.

Figure 1: Average Reimbursement Claims by Treatment and Price



lowing periods show that different prices can lead to different participant behavior even when the payoff structure of the game is kept constant.

Figure 1 shows that claims are affected by the price of the souvenir and are increasing in the price under both treatments. Nonetheless, the average claims are not independent of the underlying game. Average claims across all six periods are 64 and 79.26 under *Regular* and *Upward*, respectively. Moreover, for any price level and also when no price is mentioned, claims are higher under *Upward* than under *Regular*, on average. The difference in average claims between the two treatments is the largest in period 1, where no price was mentioned. This difference is decreasing in the price.

Average claims are much closer to the equilibrium under the *Upward* game, which has a strictly dominant strategy equilibrium. When no price is mentioned, very few participants chose the equilibrium action of 20 in *Regular*, while 61% of the participants chose the equilibrium action of 100 in *Upward*. Overall, participants chose the equilibrium claims 4.9% and 44.1% times under *Regular* and *Upward*, respectively.

To investigate the share of participants who seem to incorporate the price of the souvenir in their reimbursement claims, we compute the correlation between the reimbursement claim and the price of the souvenir across the five periods with a price for a participant. If the correlation coefficient for a participant is between -0.2 and 0.2 or she chooses the same claim under all prices, we label her as *not price-sensitive*. If the correlation coefficient is 0.8 or above, we label her as (positively) *price-sensitive*. In the *Regular* treatment, 9.9% of participants are *not price-sensitive*, 81.1% are *price-sensitive*, and 9% do not belong to either category. In the *Upward* treatment, 29.1% participants are *not price-sensitive*, 59.6% are *price-sensitive*, and 11.3% of the participants do not belong to either category.

3.1.1 Determinants of Reimbursement Claims

In Table 4, we further examine how prices and other factors affect claims made for the souvenir. We present fixed effects regressions of reimbursement claims for the two treatments separately, with results from *Regular* in the first two columns and results from *Upward* in the last two columns. In columns (1) and (3), we use the price of the souvenir as a regressor and do not include first-period observations as there was no price mentioned in that period. We control for period 1 claim by including it as a regressor. In the regressions we present, the impact of price is assumed to be linear. If we include quadratic and cubic terms, they are mostly statistically insignificant and do not improve the fit of the regression. Hence, we just present the linear specification in column (1). To allow for the non-linear impact of price, we include observations from all periods and use five dummy variables for the five price levels in columns (2) and (4).

We also control for genders of participants, two dummy variables that indicate whether they reported incorporating the price or other participants' behavior was a factor in their choice of claims, and whether they consider lying about the price of a souvenir socially appropriate in all the regressions in this table. We do not present results from regressions that control for any of the other social norm indicators in this subsection because social perception regarding lying about the price of a souvenir measures participants' beliefs about the level of honesty of others. Moreover, none of the other indicators typically have a significant impact on any of these regressions. This is consistent with the finding by Janezic (2020) that social preferences have a rather limited correlation with preferences for lying.

The table confirms the findings based on the average claims reported above. Columns (1) and (3) suggest that while the price had affected reimbursement claims in both treatments, the impact is greater in *Regular*. The columns also suggest that a participant's claim in period 1 partially explains her claims in subsequent periods. Allowing for the non-linear impact of prices and including period 1 in the sample does not change the conclusions, as evidenced in columns (2) and (4). We note that allowing for the non-linear impact of prices does not improve the fit, even when we restrict attention to the exact samples as in columns (1) and (3). We also find that participants who mentioned price as a decision-making factor, on average, submit a lower claim than those who did

Table 4: Price Sensitivity of Reimbursement Claims

	Claim *Regular* (1)	Claim *Regular* (2)	Claim *Upward* (3)	Claim *Upward* (4)
Price	0.731***		0.568***	
	(0.0250)		(0.0296)	
Price=25		-21.67***		-29.35***
		(1.706)		(2.023)
Price=40		-9.085***		-18.59***
		(1.531)		(1.723)
Price=60		5.241***		-7.099***
		(1.489)		(1.424)
Price=75		15.90***		0.681
		(1.655)		(1.313)
Price=95		29.85***		10.80***
		(1.792)		(1.324)
Claim in period 1	0.341***		0.364***	
	(0.0430)		(0.0476)	
Female	1.057	-2.671	-0.729	-2.464
	(1.647)	(2.066)	(1.814)	(1.994)
Mentioned price as a decision-making factor	-5.302** (2.470)	-6.378** (2.975)	-14.50*** (1.944)	-13.03*** (2.400)
Mentioned other participants as a decision-making factor	-0.955 (1.876)	2.095 (2.155)	3.112 (2.098)	4.245* (2.309)
Considers lying about price broadly socially appropriate	-0.0407 (2.033)	2.059 (2.615)	2.185 (1.821)	2.682 (2.121)
Constant	4.767	65.39***	20.71***	93.79***
	(3.168)	(3.282)	(4.868)	(2.803)
Observations	1060	1272	1065	1278
Adjusted R^2	0.599	0.420	0.502	0.368

Notes: Robust standard errors, clustered at the participant level, are presented in parenthesis and *, **, and *** denote p<0.10, p<0.05, and p<0.01, respectively. All regressors other than “Price”and “Claim in period 1”are dummy variables.

not mention price. In *Regular*, this observation is consistent with a price-based shift in a participant's belief about the other player's actions. However, such a belief-based explanation does not apply for *Upward* where optimal action for a participant, who only cares about material payoffs, does not depend on her beliefs. The constant in column (2) suggests that even when no price is mentioned, the participants chose a claim that is very far from the equilibrium prediction of 20 and is close to the middle of the range of allowable claims. On the other hand, the constant in column (4) suggests that the average claim in *Upward*, when no price is mentioned, is close to the equilibrium choice of 100 for participants who did not mention the price as a decision-making factor. Having a price increased the claim only for the price of 95. These observations, coupled with the large impact of price on participant choices, suggest that reimbursement claims are affected by a preference for honesty and truth-telling in *Upward*.

3.1.2 Deviation from the Price

To describe the relationship between the price of the souvenir and the corresponding reimbursement claim, we define the difference in absolute terms between the price and the reimbursement claim as the "size of dishonesty" for a given period. We do not posit that this variable is necessarily determined by a preference for honesty and use it only to quantify the relation between prices and reimbursement claims. Table 5 present results from regressing the deviation of the claim from the price on the price and other regressors. Results from the *Regular* treatment are in the first two columns, and results from the *Upward* treatment are in the last two columns. Table 4 showed that, under both treatments, participants submitted a higher claim when the price was higher. It is still possible that the distance between the price and the claim would increase as price increases, especially in *Regular*, where strategic considerations will pull the claims towards the lowest possible claim of 20. However, Table 5 shows that absolute deviation from price is decreasing in price under both treatments. Thus, participants chose more honest claims as the price moved closer to the upper limit of 100. These results suggest that a tendency towards making a claim closer to the price dominates the strategic pull towards the minimum possible claim in *Regular*.

The table provides some other insights. Columns (1) and (3) show that larger claims in period 1 are associated with greater dishonesty under both treatments. As is the case with claims in Table 4, we do not see any gender effect for the difference between price and claims either. Participants who mentioned the price as a factor in their decision-making behave in a more honest fashion under both treatments, and this effect is larger in *Upward*, where strategic considerations pull a participant claiming towards the highest possible amount. On the other hand, participants who considered lying about the price

Table 5: Determinants of the Size of Dishonesty

	Deviation from price *Regular* (1)	Deviation from price *Regular* (2)	Deviation from price *Upward* (3)	Deviation from price *Upward* (4)
Price	-0.113***		-0.406***	
	(0.0279)		(0.0309)	
Price=40		-1.467*		-3.887***
		(0.856)		(1.259)
Price=60		-4.094***		-11.64***
		(1.143)		(1.466)
Price=75		-5.406***		-18.85***
		(1.513)		(1.757)
Price=95		-7.915***		-27.98***
		(1.960)		(2.216)
Claim in period 1	0.160***		0.256***	
	(0.0349)		(0.0348)	
Female	-0.648	-1.909	-1.595	-2.617
	(1.402)	(1.521)	(1.694)	(1.833)
Mentioned price as a decision-making factor	-8.045***	-8.742***	-17.13***	-17.63***
	(1.983)	(2.122)	(1.713)	(1.961)
Mentioned other participants as a decision-making factor	2.138	3.165**	1.797	2.704
	(1.523)	(1.600)	(1.949)	(2.095)
Considers lying about price broadly socially appropriate	3.255*	3.999**	2.519	2.994*
	(1.692)	(1.795)	(1.573)	(1.748)
Constant	13.54***	20.88***	31.89***	42.88***
	(2.870)	(2.733)	(3.617)	(2.770)
Observations	1060	1060	1065	1065
Adjusted R^2	0.165	0.109	0.419	0.365

Notes: Robust standard errors, clustered at the participant level, are presented in parenthesis and *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. All regressors other than "Price"and "Claim in period 1"are dummy variables.

of a souvenir socially appropriate are less honest. These suggest that some participants have a predisposition towards honesty, which affects their choices in a strategic game.

Our findings regarding participants' choices of reimbursement claims, so far, are summarized below:

Result 1. *i) Under both treatments, reimbursement claims are price sensitive. However, the claims are clearly different between the two treatments.*
ii) Payoffs from the games and preferences for honesty both seem to affect participants' choice of reimbursement claims.
iii) Reimbursement claims appear to be more honest or closer to the price under Regular than under Upward.

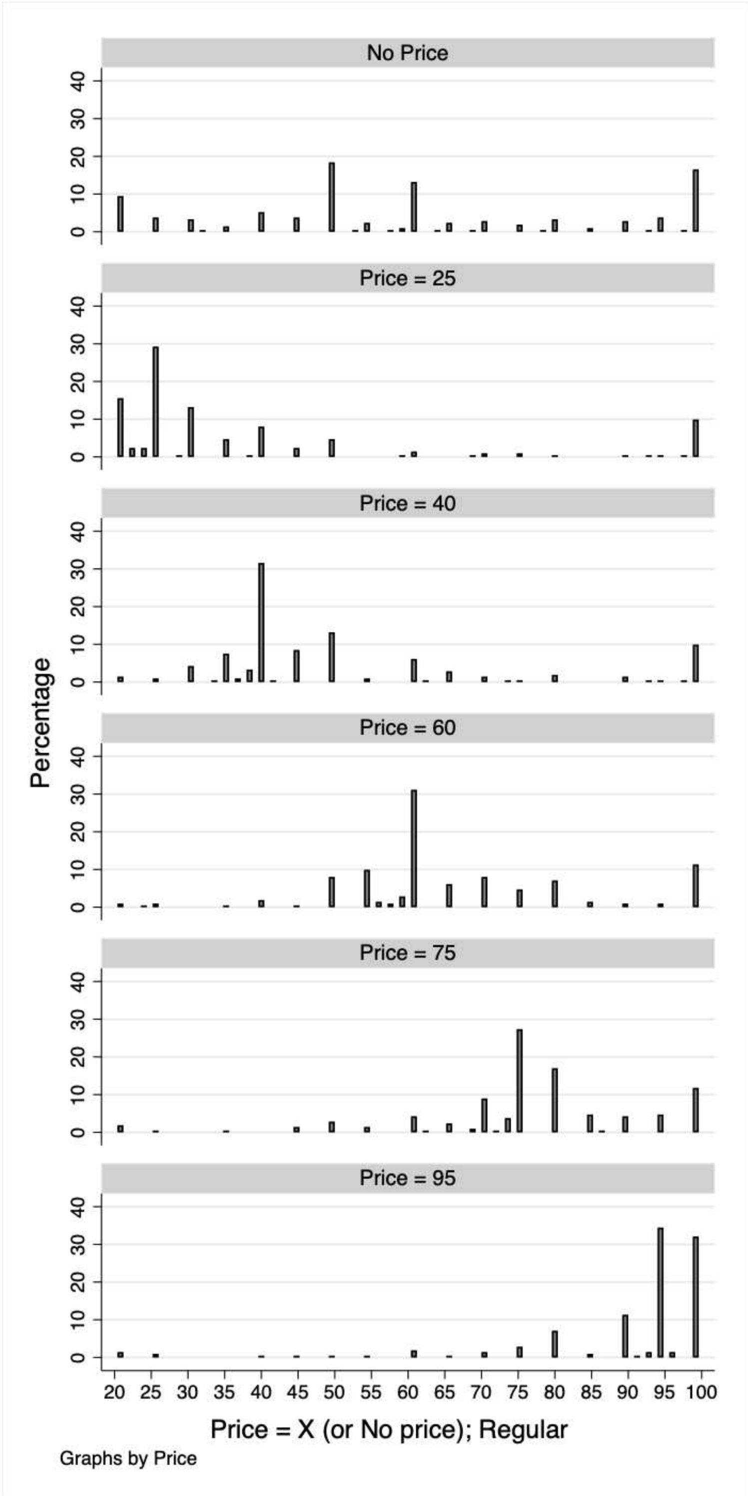
3.2 Distributions of Claims

While participants do not follow the equilibrium strategies and tend to be price-sensitive on average, that is not true for all participants. We see a noticeable level of dispersion in reimbursement claims, for a given price, across participants. Figures 2 and 3 present histograms of reimbursement claims submitted by participants for each price level (including period 1, where no price is mentioned) for treatments *Regular* and *Upward*, respectively. We use bins of size 2 for these histograms. The histograms clearly show that the distributions of claims are affected by both the underlying game and the price and absence thereof. This is consistent with the findings above.

Under *Regular*, the claim distributions are quite nuanced. When there is no price, around 54% of the claims are for 20, 50, 60, or 100. The remaining claims are scattered around the range of possible claims. While this period is similar to the one-shot standard Traveler's Dilemma setting with a small-sized reward/penalty amount in Goeree and Holt (2001), the claim distributions are rather different. Whereas claims are highly concentrated around the highest possible claim in their paper, only around a quarter of the claims are of 90 or greater amount in our period 1. Even when the value of the price is not provided, many participants may have made a conjecture about the price in our experiment, given that prices are explicitly mentioned in our instructions and participants were aware that they would learn the price in some periods. Indeed, we see more than 30% of the claims to be either 50, which is halfway in between 0 and the maximum possible claim, or 60, which is the average of possible claims.

When a price is mentioned, the majority of claims are concentrated around the price. Specifically, around 40% of claims are between the price and the price plus 5. On the other hand, around 10% of the claims are within price minus 1 and price minus 5. For prices lower than 95, around 10% claims are for 100. When the price equals 95, this

Figure 2: Histogram of Reimbursement Claims in *Regular* by Price



share is much greater. Over the six periods, fewer than 5% of the claims are for the Nash equilibrium prediction of 20. When no price is mentioned, this share is around 8.5%, and it jumps to more than 15% when the price is 25. However, for the other four price levels, this share hovers around 1%. Overall, the distributions of claims become more concentrated towards the upper portion of the range as the price increases.

For *Upward*, under all prices and in the period where no price was mentioned, we see a large number of claims of 100, which is the dominant strategy. This share is larger in the periods without a price and with a price of 95, at 61% and more than 67%, respectively. At the other prices, this share ranges between 30% and 40%. When there is no price, claims that are smaller than 100 are scattered across the entire range with small spikes at 50, 60, and 80. When there is a price, a large number of claims of 100 are followed by a slightly smaller number of claims near the price. Note that almost all claims near the price are at or slightly above the price, and very few are below the price. While a preference for taking actions closer to the price may lead to an optimal choice between the price (inclusive) and 100, it will not lead to a claim below the price.[4] The bottom five panels of Figure 3 suggest a pattern: About a third of the participants choose claims around 100, and the remaining claims are mostly distributed between the price and 100. Thus, the distribution shifts to the right as the price increases.

Our findings from this subsection can be summarized as:

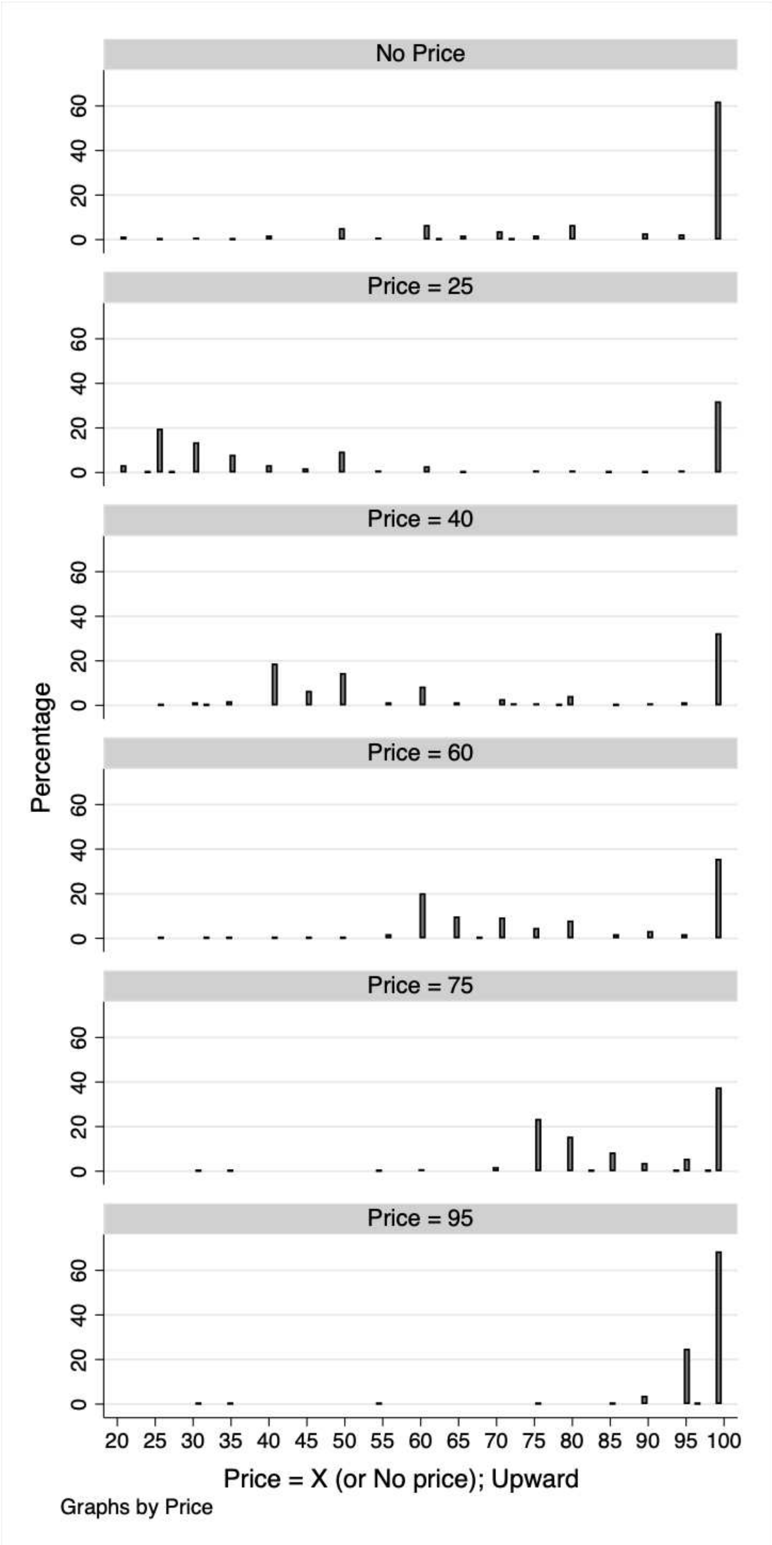
Result 2. *i) Under both treatments, a large percentage of participants choose claims equaling the price or slightly above.*
ii) While a noticeable percentage of participants choose claims right below the price in Regular, such claims are very uncommon in Upward.
iii) The highest possible claim of 100 is three times as common in Upward, compared to Regular.

3.3 Coordination Benefits of Mentioning a Price

Figure 2 shows that, in *Regular*, claims are more dispersed in period 1 than in subsequent periods. In the periods with a price, the majority of participants make claims near the price. Thus, when a price is mentioned, prices may act as a coordination device in this game. Fixing the sum of claims by the two players, the total payoffs of the two players in *Regular* is the highest when the claims are equal. As a result, controlling for the claim, players will enjoy greater payoffs, on average, when their claims are better coordinated. On the other hand, there is no such coordination benefit in *Upward*, where the sum of

[4] Preference for honesty among some participants may lead to claims right below the price as a strategic response in *Regular*. Consistent with that, we find a noticeable share of claims right below the price in that treatment.

Figure 3: Histogram of Reimbursement Claims in *Upward* by Price



payoffs always equals the sum of the claims. In this subsection, we explore how the presence and absence of prices affect participants' payoffs.

To calculate the final payments to participants for each period, we randomly matched each participant to another participant from the same treatment and paid them for one period that was randomly chosen. We can use this match to investigate how the price tacitly leads to coordination between participants. First, we look at the difference in claims between the two matched participants in different periods. In *Regular*, the average difference between claims of the two matched partners is 29.33 in period 1, where no price is mentioned and is 18.79, on average, in periods 2 to 6, where a price is mentioned. The difference is these differences statistically significant at the 0.01% level. On the other hand, in *Upward*, these differences are 22.74 and 20.58 in period 1 and periods 2 to 6, respectively. The difference in these differences is not statistically significant at any conventional level. Thus, the presence of a price brings the reimbursement claims closer in *Regular*, but there is no such effect in *Upward*.

Does this tacit coordination in *Regular* have an impact on payoffs? We calculate a participant's payoff from each period to test whether this coordination in *Regular* led to increased payoffs. We find that participants earn a much higher average payoff in periods 2 to 6 than in period 1 in *Regular*. This difference is not explained by the difference in the claims they make in those periods. Specifically, the average claims are 60.63 and 64.68 in period 1 and periods 2 to 6, respectively. However, the average payoffs are 44.95 and 55.15 in period 1 and periods 2 to 6, respectively. Such discrepancy is not observed in *Upward*. In fact, the average payoffs in periods 2 to 6 are lower than in period 1, and this difference can be explained by differences in average claims between periods with a price (77.81) and the period without a price (86.52).

We further explore the impact of having a price on payoffs in Table 6. The table presents fixed effects regressions of calculated payoffs from each period on whether the period included a price while controlling for the participant's claim in that period and other characteristics. Columns (1) to (3) present results for the treatment *Regular*, and columns (4) to (6) present results for *Upward*. Payoffs in period 1 of *Regular* are significantly lower than payoffs in the periods where a price is mentioned. This effect becomes slightly smaller when we control for the claim in that period, but still accounts for around 7.8. Given an average payoff of 53.45 across the periods in *Regular*, this difference represents an economically significant effect. As expected from the averages, payoffs are higher in period 1 relative to other periods in *Upward*. However, when we control for the claim, as seen on columns (5) and (6), payoffs in period 1 are actually lower by a small, but statistically significant, amount. Finally, columns (3) and (6) suggest that participants who considered price in their decision-making received higher payoffs in *Regular* by a large amount. Such participants in *Upward*, however, had a small negative impact on their payoffs. This is consistent with our finding that being price-sensitive leads to better coordination in *Regular*, but does not have much of an impact in *Upward*.

Table 6: Determinants of Payoffs

	Payoff *Regular* (1)	Payoff *Regular* (2)	Payoff *Regular* (3)	Payoff *Upward* (4)	Payoff *Upward* (5)	Payoff *Upward* (6)
Period 1	-10.20***	-7.836***	-7.780***	8.711***	-0.740***	-0.680***
	(1.296)	(1.212)	(1.222)	(1.490)	(0.227)	(0.232)
Claim		0.584***	0.597***		1.085***	1.078***
		(0.0298)	(0.0272)		(0.0048)	(0.0047)
Female			0.680			-0.223
			(1.566)			(0.418)
Mentioned price as a decision-making factor			4.441**			-0.818**
			(2.001)			(0.405)
Mentioned other participants as a decision-making factor			-2.395			0.632
			(1.607)			(0.432)
Considers lying about price broadly socially appropriate			-1.014			0.620
			(1.796)			(0.455)
Constant	55.15***	17.40***	14.025***	77.78***	-6.631***	-5.812***
	(0.719)	(1.367)	(2.440)	(1.336)	(0.442)	(0.521)
Observations	1272	1272	1272	1278	1278	1278
Adjusted R^2	0.024	0.422	0.431	0.014	0.980	0.981

Notes: Robust standard errors, clustered at the participant level, are presented in parenthesis and *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. All regressors other than "Claim"are dummy variables.

These findings can be summarized in the following result:

Result 3. *In the treatment Regular, price-sensitive behavior leads to participants' claims to be closer to each other. This results in an increase in payoffs when the price of the souvenir is mentioned to participants.*

4 Preference for Honesty and Categorization of Participants

Distributions of the reimbursement claims show that many participants' choices are affected by the prices. Moreover, there is noticeable heterogeneity in choices across participants.This section first discusses how results from both *Regular* and *Upward* treatments can be explained. Specifically, we first explore whether level-k models are enough to explain our findings and then offer an alternative model by considering the heterogeneity of players' preferences. Using this model, we categorize the participants into different types based on their choices across periods.

4.1 Explaining our Experimental Findings

As observed in the previous section, claims under both treatments are quite heterogeneous. Even though theoretical predictions do not depend on the price, distributions of reimbursement claims very much depend on the price in our data. In *Regular*, while a majority of claims are within 5 of the price, claims are scattered around the entire range of possible claims (20 to 100). Note that any claim between 20 and 99 can be rationalized by some beliefs about the other player's actions. Moreover, as finding out the Nash equilibrium requires iterated reasoning, participants with limited cognitive ability may not reach the equilibrium of the lowest claim. Thus, the price may act as a focal point or a naïve anchor. Similarly, one may argue that players' instinctive choices of the price, the truth, may reflect their level-0 reasoning in a level-k based model. However, such behavior can also arise from a preference for honesty. Thus, it is hard to pin down a specific explanation for our findings in *Regular*. About 31% of the participants in that treatment choose the price or right above it as their reimbursement claims in at least four out of the five periods with a price. Almost 15% of the claims are of 100, which can be considered to be another common choice of naïve players. On the other hand, around 10% of the claims can be considered as strategic responses to level-0 players choosing the price or a focal number such as 50, 60, or 100 for a reasonable value of k.[5] Hence, our results cannot be explained very well by level-k models without making level-0 types very flexible and very large.[6]

It is even harder to explain our experimental results from the treatment *Upward* using level-k models. Making the highest possible claim of 100 is the strictly dominant strategy in that game. This means that any level-k player with $k > 0$ should choose 100 her claim. The presence of a dominant strategy also implies that players should choose 100 independent of their beliefs about the other player. In our experiment, 61% of the participants chose 100 when no price was mentioned. However, once there is a price, around 18% of them switch to claiming exactly the price. Among those who chose below 100 in period 1, around 22% claim the price when there is one. A significant portion of those who chose 100 in period 1 start claiming the price or right above occurs, even when we look at only participants who got all three comprehension questions correct. It is

[5] Arad and Rubinstein (2012) finds that behavior of most participants can be explained by cognitive levels of 3 or lower.

[6] Tong and Freeman (2021) estimate that around a third of the participants in their data, which they consider to be on the high side, can be considered level-0 players with a flexible definition for non-strategic play.

hard to argue that these participants do not understand that claiming 100 is the optimal strategy. Thus, participants choosing the price as their claims indicate that they strongly prefer making honest claims. On the other hand, to explain our findings with a level-k model, we have to assume that about 70% of the participants are level-0 and a large group of them go from being max-choosing level-0 in period 1 to price-choosing level-0 as soon as the price is mentioned. Such an explanation seems rather hard to believe. Hence, we consider a preference-based model that can explain the behavior of a large group of participants, especially in *Upward*. An alternative justification on why our experimental results differ significantly from what the theory predicts as in Claim 1, and argue that a preference for honesty can explain the behavior of a large set of participants, especially in *Upward*.

Formally, we may model *Regular* and *Upward* as Bayesian games to incorporate a preference for honesty. Suppose player i is characterized by her type $t_i \in T$, a common type space. We define player i's utility if her type is t_i in the game $A \in \{Regular, Upward\}$ to be

$$u_i^A(c_i, c_j; t_i) = (1 - \lambda(t_i))m_i^A(c_i, c_j) - \lambda(t_i)(c_i - v_i(p))^2, \quad (2)$$

where the material payoff m_i^A is defined as in Equation 1, $\lambda : T \to [0, 1]$ measures the level of honesty by determining how much weight player i puts on making an honest claim and $v_i(p)$ represents how player i values a souvenir purchased at price p. Under this consideration, below is a simplified example. Note that as shown later, our empirical findings can be better explained when we relax the strong assumptions in the example.

Example 1. Suppose players can be one of two types with $T \equiv \{strategic, honest\}$. Let us assume that player i believes that the other player j is *strategic* with probability τ and is *honest* with probability $1 - \tau$. This belief is not affected by her own type. As is the case in our experiment, let $p > \underline{c} + 2$ and $k > 2$. We also assume that: (1) $v_i(p) = p$; (2) $\lambda(strategic) = 0$ and $\lambda(honest) = 1$, for simplicity. With these simplified assumptions, this setup is equivalent to assuming that some players are *honest* types who choose the price as their claim and the remaining players maximize their expected payoffs based on their beliefs about the type of the other players. In the game *Upward*, there will be a unique Bayesian Nash equilibrium where an honest player claims p, and a strategic player claims $\bar{c}$. The equilibrium in *Regular*, however, depends on a *strategic* player's beliefs regarding τ. If τ is high enough, *strategic* players will make the lowest possible claim. For lower values of τ, they will choose a claim slightly lower than the price. This suggests that even a *strategic* player may behave in a price-sensitive manner in equilibrium. Moreover, if we restrict attention to pure strategy equilibria, only the lowest possible claim and claims one or two units below the price can be expected from *strategic* players in *Regular*.

While the above example discusses player behavior with extreme values of $\lambda(t_i)$ with only two types of players, other behavior can be observed under different assumptions. For example, if $\lambda(t_i)$ is strictly in between 0 and 1 for a player, she may choose a claim in between the price and maximum possible claim in *Upward*. However, she will not choose a claim lower than the price for any value of $\lambda(t_i)$. Equilibrium predictions will be more nuanced in *Regular* as there is no dominant strategy in that game. In particular, almost any possible claim can be an optimal choice depending on player i's beliefs about the other player's types, her own level of honesty ($\lambda(t_i)$), and how she values the souvenir ($v_i(p)$). This example also illustrates that when players have a preference for honesty, we may observe that claims closer to the price are on either side of p in *Regular*. However, we may only observe claims equaling or slightly above p, but not right below p in *Upward*. This is consistent with our findings in Section 3.2.

4.2 Categorization

Based on the observations in Sections 3.1 and 3.2, and aided by the preference as in Equation 2 above, we categorize the participants into different types. For this, we use participants' reimbursement claims in the five periods where they are clearly informed of the price of the damaged souvenir (periods two to six).

We first discuss different types in the treatment *Regular*. Example 1 illustrates a case where an *honest* player will make a claim equaling the price. A *strategic* player's optimal claim will depend on her belief regarding τ, which is the probability that the other player is *strategic*. Specifically, she will claim 20 or an amount right below the price if the probability is low or high, respectively. We consider a claim that equals 20 to be strategic with a high τ. Moreover, we consider a claim that is between $p-5$ and $p-1$ (inclusive) to be strategic with a low τ. This allows for a strategic response as described in the above example as well as a response by a level-k player with reasonable levels of $k>0$ when the level-0 players choose the price. We consider claims within p and $p+5$ (inclusive) to be honest. This captures the fact that many participants mentioned that they claimed 5 or 10 above the price to compensate for the hassle the airline put them through when we asked participants how they chose their claims.[7] Finally, a claim lower than $p-5$ is considered a low claim, and a claim above $p+5$ is considered a high claim.

Based on the above terminology, we label a participant as *Honest* if her claim is considered honest in at least four out of the five periods with a price. While these participants can also be considered as level-0 players who chose the price or slightly above as their claim, answers by some of these participants in the questionnaire suggest that honesty, being truthful, or social appropriateness were important in their decision-

[7]Equation 2 reflects this when a player's valuation for the souvenir $v(p)$ equals $p+e$ for some positive e.

making.[8] Those who claim 20 in four or more periods are labeled as *Strategic High*. They correspond to those strategic players who believe the other is most likely to be strategic too. Those who make a claim within $p-5$ and $p-1$ in at least four periods are labeled as *Strategic Low*. They correspond to strategic players who believe that the other is likely to be honest. In *Regular*, almost 10% of the participants choose the maximum claim of 100 in four or more periods. We label these participants as *Max Claimer*. Among the remaining participants, we identify those who consistently claim higher or lower than the price. We label those who made a claim weakly below p in at least four periods with a low claim in at least two periods as *Low Claimer*. Similarly, those who claim weakly above p in at least four periods with a high claim in at least two periods are labeled as *High Claimer*. These two types of participants' behavior can be explained by beliefs regarding the other players' claims with or without a preference for honesty. The remaining participants did not consistently follow a pattern in making their claims across the five periods. We label them as *Inconsistent*.

The left panel of Table 7 presents the share of participants belonging to these seven types. Very few of the participants are strategic with a belief that the other player is very likely to be strategic. Fewer than 8% of the participants are strategic with the belief that the other participant is very likely to be honest. The largest group of participants are categorized as *Honest*. *Max Claimer* and *Inconsistent* participants do not behave strategically or in a consistent manner relative to the *honest* behavior. We may consider them, comprising of a quarter of the participants, as level-0 players.

Categorization of participants is simpler in *Upward*, as the game has a strictly dominant strategy. In a level-k model, all players with $k > 0$ will choose the maximum claim of 100, which is the Nash equilibrium strategy. However, other claims can be explained when players have a preference for honesty. Example 1 suggested that while a *strategic* player will claim 100, an *honest* player will claim the price. A player who exhibits an intermediate level of preference for honesty can be represented by Equation 2 with $\lambda(t_i) \in (0, 1)$. Such a player will choose a claim above p but below 100. We label participants who claim 100 in at least 4 periods as *Strategic*. Participants who claim within p and $p+5$ in at least 4 periods are labeled as *Honest*, as was done in *Regular*. Among the remaining participants, those who claim weakly above p in at least four periods with a high claim in at least two periods are labeled as *High Claimer*. The remaining participants are labeled as *Unexplained* as their claims in the five periods cannot be explained by a preference for honesty, even allowing for different beliefs about the other player's actions.

The right panel of Table 7 presents the share of participants belonging to the four types

[8] Some examples of their answers include: "In my opinion, it is socially inappropriate to claim for more money than the value of my souvenirs," "I chose to be honest with the price," and "I didn't want to lie to the airline company, that would be socially inappropriate."

in *Upward*. More than 30% of the participants are categorized as *Strategic*. Slightly below a quarter of the participants are *Honest*. The largest share of participants are categorized as *High Claimer*, which represents an intermediate level of preference for honesty. Thus, the majority of the participants show some preference for honesty. We would like to note that a significant share of participants categorized as *Honest* or *High Claimer* make a claim of 100 in the first period and also correctly answer all three comprehension questions. This bolsters the claim that the participants categorized as *Honest* or *High Claimer* are indeed motivated by some preference for honesty. Fewer than 7% participants' behavior in *Upward* is not well explained by a preference for honesty, and one may consider them as players who choose their actions randomly.

Table 7: Categorization of Participants

Regular		*Upward*	
Strategic High	0.9%	*Strategic*	29.6%
	(0.097)		(0.457)
Strategic Low	7.5%	*Honest*	23.9%
	(0.265)		(0.428)
Honest	30.7%	*High Claimer*	39.9%
	(0.462)		(0.239)
Max Claimer	8.5%	*Unexplained*	6.6%
	(0.279)		(0.248)
Low Claimer	11.8%		
	(0.323)		
High Claimer	24.1%		
	(0.428)		
Inconsistent	16.5%		
	(0.372)		
Observations	212		213

Notes: Standard deviations are in parentheses.

4.2.1 Differences Across Genders

Now we briefly discuss whether there is much difference between males and females regarding their claiming behavior. Table 8 presents average claims and percentages of participants categorized as *Honest* and as *Strategic* for the two treatments separately.[9] In *Regular*, males made higher claims and were more likely to be *Strategic* and less likely to be *Honest*. However, the difference is statically significant at the 5% level only for the likelihood of being *Strategic*. The other two differences are barely statistically significant at the 10% level. Slightly more females were either *Strategic* or *Honest* than males, but

[9] We include both *Strategic High* and *Strategic Low* as *Strategic* for *Regular*. Note that three participants did not identify as male or female, and we exclude them from this table.

this difference is not statistically significant. In *Upward*, on the other hand, males made higher claims and were more likely to be *Strategic* than females. Males and females were equally likely to be *Honest*. However, females were one and a half times more likely to be a *High Claimer* than males, suggesting that more females exhibit an intermediate level of preference for honesty than males do. Both of our treatments clearly suggest that male participants were more likely to be *Strategic* than females. This is consistent with the finding in Buchan et al. (2008) that men viewed investment games more strategically than women did.

Table 8: Differences in Claims and Types Across Genders

Regular	Males	Females	p-value of Difference
Claims	65.92	62.26	.0932
	(17.919)	(13.143)	
Honest	24.8%	35.2%	.0985
	(0.434)	(0.48)	
Strategic	12.4%	4.8%	.0489
	(0.331)	(0.214)	
Observations	105	105	
Upward	Males	Females	p-value of Difference
Claims	81.32	76.31	.0261
	(16.909)	(14.054)	
Honest	22.3%	25.6%	.5834
	(0.418)	(0.439)	
Strategic	38.5%	15.9%	.0004
	(0.488)	(0.367)	
High Claimer	33.1%	51.2%	.0085
	(0.472)	(0.503)	
Observations	130	82	

Notes: Standard deviations are in parentheses.

5 Conclusion

In this paper, we explore how the presence of truth in the natural context of a game affects player behavior using two 2-player games. We provide evidence that both strategic considerations and preferences for honesty affect participants' choices in our games. For some participants, strategic behavior may also be constrained by limited cognitive skills and beliefs about the other participants. In these games, the key is that the payoff functions are publicly known and are not affected by the truth. Our paper offers a game setting where the truth is known to all participants and the researcher so that experimentally, preferences of honesty and strategic considerations can be easily separated. We hope that our setup will provide a useful experimental paradigm in the study of preferences for honesty, especially in a game setting.

Our finding suggests that the truth may impact player choices in other game situations such as Prisoner's Dilemma and ultimatum game-type scenarios as discussed earlier. Based on our findings, one may speculate that we will see more instances of both players cooperating when they are innocent and both players defecting when they are guilty. While the truth affects player behavior, how it affects depend very much on the strategic considerations of the underlying game. The two games investigated in this paper have very similar contexts, but the strategic play resulting from the games are very different. Thus, they may provide insights on a broad set of games where strategic considerations and preferences for the truth may interact. In *Upward*, we provide a game with a strictly dominant strategy with a rich set of possible actions. This game may provide a useful tool for experimental investigation of dominant strategy games, both in the presence and absence of a true state.

In many real-life scenarios, multiple people are asked to report about a true state, and the final outcome depends on the report but not the true state as the true state remains unknown or unverifiable to the entities designing the payoffs. A game like *Regular* may represent, in addition to the damage claim scenario, an arms race where the true state is the military strength of nations. Similar situations may arise for online retailers concerning promotions of common products in holiday seasons or clearance of out-of-season products where procurement costs of the inventory represent the truth. We illustrate that the truth may affect players' choices even if payoff variations from different outcomes do not depend on the true state. On the other hand, the game *Upward* may represent applicants for grant funding where applicants know how much money they need, and assessors may decide on the grant awards based on the demand of multiple grant applicants. Assessors may consider the requested amount of funding as an indicator of how deserving a project is and reward projects that ask for larger amounts. Similar situations may arise for financial aid applications, allocation of emergency funds for a pandemic, and so on.

Finally, our paper also contributes to the large literature on Traveler's Dilemma. While the price of the souvenir is mentioned in the initial theoretical exploration, it is neglected in the experimental explorations even though it would be natural for travelers to know the price. Thus, it is of interest to investigate if and how the existing findings may change if the participants are presented with the price. In particular, one may ask how reward or penalty amounts (Capra et al., 1999; Goeree and Holt, 2001; Basu et al., 2011) or response times (Rubinstein, 2007, 2016) would affect participants as the price is also varied. This may allow one to test preferences for honesty further. We leave all these for future work.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond, "Preferences for truth-telling," *Econometrica*, 2019, *87* (4), 1115–1153.

Arad, Ayala and Ariel Rubinstein, "The 11-20 Money Request Game: A Level-k Reasoning Study," *American Economic Review*, 2012, *102* (7), 3561–3573.

Basu, Kaushik, "The traveler's dilemma: Paradoxes of rationality in game theory," *The American Economic Review*, 1994, *84* (2), 391–395.

_ , "The traveler's dilemma," *Scientific American*, 2007, *296* (6), 90–95.

_ , Leonardo Becchetti, and Luca Stanca, "Experiments with the traveler's dilemma: welfare, strategic choice and implicit collusion," *Social choice and welfare*, 2011, *37* (4), 575.

Besancenot, Damien and Radu Vranceanu, "Profession and deception: Experimental evidence on lying behavior among business and medical students," *Journal of Economic Behavior & Organization*, 2020, *179*, 175–187.

Buchan, Nancy R, Rachel TA Croson, and Sara Solnick, "Trust and gender: An examination of behavior and beliefs in the Investment Game," *Journal of Economic Behavior & Organization*, 2008, *68* (3-4), 466–476.

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

Capra, C Monica, Jacob K Goeree, Rosario Gomez, and Charles A Holt, "Anomalous behavior in a traveler's dilemma?," *American Economic Review*, 1999, *89* (3), 678–690.

Charness, Gary and Martin Dufwenberg, "Promises and partnership," *Econometrica*, 2006, *74* (6), 1579–1601.

Chen, Ying, "Perturbed communication games with honest senders and naive receivers," *Journal of Economic Theory*, 2011, *146* (2), 401–424.

Erard, Brian and Jonathan S Feinstein, "Honesty and evasion in the tax compliance game," *The RAND Journal of Economics*, 1994, pp. 1–19.

Fischbacher, Urs and Franziska Föllmi-Heusi, "Lies in disguise—an experimental study on cheating," *Journal of the European Economic Association*, 2013, *11* (3), 525–547.

Gneezy, Uri, "Deception: The role of consequences," *American Economic Review*, 2005, *95* (1), 384–394.

_ , Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen, "Bribing the self," *Games and Economic Behavior*, 2020, *120* (5), 311–324.

Goeree, Jacob K and Charles A Holt, "Ten little treasures of game theory and ten intuitive contradictions," *American Economic Review*, 2001, *91* (5), 1402–1422.

Graetz, Michael J, Jennifer F Reinganum, and Louis L Wilde, "The tax compliance game: Toward an interactive theory of law enforcement," *Journal of Law, Economics, & Organization*, 1986, *2* (1), 1–32.

Hurkens, Sjaak and Navin Kartik, "Would I lie to you? On social preferences and lying aversion," *Experimental Economics*, 2009, *12* (2), 180–192.

Janezic, Katharina A., "Heterogeneity in lies and lying preferences," *Working paper*, 2020.

Jiang, Ting, "Cheating in mind games: The subtlety of rules matters," *Journal of Economic Behavior & Organization*, 2013, *93*, 328–336.

Kartik, Navin, "Strategic communication with lying costs," *The Review of Economic Studies*, 2009, *76* (4), 1359–1395.

Krupka, Erin L and Roberto A Weber, "Identifying social norms using coordination games: Why does dictator game sharing vary?," *Journal of the European Economic Association*, 2013, *11* (3), 495–524.

Matsushima, Hitoshi, "Role of honesty in full implementation," *Journal of Economic Theory*, 2008, *139* (1), 353–359.

Mazar, Nina, On Amir, and Dan Ariely, "The dishonesty of honest people: A theory of self-concept maintenance," *Journal of marketing research*, 2008, *45* (6), 633–644.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 1995, *85* (5), 1313–1326.

Potters, Jan and Jan Stoop, "Do cheaters in the lab also cheat in the field?," *European Economic Review*, 2016, *87*, 26–33.

Rosenbaum, Stephen Mark, Stephan Billinger, and Nils Stieglitz, "Let's be honest: A review of experimental evidence of honesty and truth-telling," *Journal of Economic Psychology*, 2014, *45*, 181–196.

Rubinstein, Ariel, "Instinctive and cognitive reasoning: A study of response times," *The Economic Journal*, 2007, *117* (523), 1243–1259.

_ , "A typology of players: Between instinctive and contemplative," *The Quarterly Journal of Economics*, 2016, *131* (2), 859–890.

Stahl, Dale O. and Paul W. Wilson, "Enforcement of contribution norms in public good games with heterogeneous populations," *Games and Economic Behavior*, 1995, *10* (1), 218–254.

Tong, Hanh T and David J Freeman, "Anchors of strategic reasoning in the traveler's dilemma," *Journal of Economic Behavior & Organization*, 2021, *191*, 28–38.

Umbhauer, Gisèle, "Traveler's dilemma: how the value of the luggage influences behavior," Working Paper, Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg 2019.

Appendix

A: Experimental Instructions

Instructions for the Main Exercise

In this economic experiment, we explore how people decide the amount of money to claim from an airline company that damaged their goods. This session will last around 30 minutes. If you follow the instructions carefully and make good decisions, you can earn up to CA$20. All the participants in this session have been recruited in the same way and are reading the same instructions. You must not communicate with any other participants or discuss the details of the experiment with anyone during or after the session. The session will consist of two exercises, which are described below. In both exercises, we will use a currency called Rye$. Your earnings from both exercises will be in terms of Rye$. The Rye$ that you earn in this experiment will be exchanged into CA$ at the rate of CA$1 for Rye$10. You will be paid via a gift card denominated in CA$.

Suppose that you bought a souvenir on a foreign trip and the airline company broke the souvenir during the flight. You may be informed of the price you paid for the souvenir. It turns out that another traveler on the same flight also bought the same souvenir at the same price and the airline broke that piece too. Suppose that, according to the airline company's policy, you and the other traveler can claim between Rye$20 and Rye$100 to compensate you for the price of the souvenir. The airline company does not know the price that you or the other traveler paid for the souvenir. It will consider your and the other traveler's claims together to determine the compensations according to the following rule:

Only the payoff functions, as described below, were different between the two games: *Regular* and *Upward*

(**Payoff Function for *Regular***)

1) If you make the <u>lower</u> claim, then you will receive Rye$5 more than <u>your own claim</u> (lower of the two claims).
2) If you make the <u>higher</u> claim, then you will get Rye$5 less than <u>the other traveler's claim</u> (lower of the two claims).
3) If you both make the <u>same</u> claim, then you will get the <u>amount you claimed</u>.

(**Payoff Function for *Upward***)

1) If you make the higher claim, then you will get Rye$5 more than your own claim (higher of the two claims).
2) If you make the lower claim, then you will receive Rye$5 less than your own claim (lower of the two claims).
3) If you both make the same claim, then you will get the amount you claimed.

The other traveler's compensation will also be determined according to the above rule. Your task is to choose the amount of money that you want to claim from the airline company.

Decision-making in Each Round and Differences between Rounds There will be six rounds in this exercise. Each round lasts 30 seconds. In each round, you may learn the price you paid for the souvenir. This price may differ across periods. You will choose your compensation claim for the broken souvenir in every round. For simplicity, your claim can only be an integer (no decimals) between Rye$20 and Rye$100.

Ending the Round At the end of each round, the researchers will collect your answer sheet. Your income will be calculated at a later time, once the research assistant has digitized all answer sheets. You will earn an amount based on your Rye$ earnings from one randomly chosen round between rounds 1 and 6. You will be matched randomly with a participant who participated in any session of this experiment. This participant will play the role of the other traveler. Your earning will be based on the compensation claims made by you and that participant in the same round (same scenario). Your earning in Rye$ will be converted into money at the rate of CA$1 for Rye$10.

Reimbursement Claim Entry Form

The following sentence did not appear in Period 1 The price you paid for the souvenir is [insert price in Rye$].
You can claim between Rye$20 and Rye$100 as compensation and your claim can only be an integer (no decimals). What amount of money do you claim to the airline company?
YOUR CLAIM: Rye$_____.

Social Norm Elicitation

In this exercise, you will be presented with four multiple-choice questions and one open-ended question. Each of the multiple-choice questions describes a scenario. Please rate the social appropriateness of the act described in the scenario on a 5-point scale.

Payment: We will first find the most common answer among all the participants in this experiment (approximately 500 participants) for all questions. For each of questions 1 to 4, you will receive Rye$10 if your answer is the same as the most common answer and Rye$0 otherwise.

You will be asked another question after the four social appropriateness questions. You will be paid Rye$10 for answering this question, independent of what your answer is.

For questions 1 to 4, please mark the option that best represents how socially appropriate the act described in the question is.

Scenarios for the four questions:

1. At a grocery store, the cashier did not charge a customer for one item in their basket (a tomato). The customer decides not to tell the cashier and keep the "free" item.
2. There is no customer service representative at a TTC subway station. A TTC user walks past the empty booth and does not pay for his fare.
3. A traveller lies about the price of the souvenir to the airline company on the claim form.
4. At a clothing store, the cashier did not charge a customer for one item in their basket (leather a pair of leather gloves). The customer decides not to tell the cashier and keep the "free" item.

Rate the social appropriateness:

a. socially inappropriate
b. somewhat socially inappropriate
c. value neutral
d. somewhat socially appropriate
e. socially appropriate