Transparency in Social Responsibility: Investigating Consumers’ Behaviors and Motives

Tim Kraft  
Darden School of Business, University of Virginia, Charlottesville, VA, KraftT@darden.virginia.edu

León Valdés, Yanchong Zheng  
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, lvaldes@mit.edu, yanchong@mit.edu

January 18, 2016

We design an incentivized human-subject experiment to study the impact of supply chain transparency on consumers’ valuations of a firm’s social responsibility practices. Lower transparency captures higher uncertainty in the outcomes of the firm’s social responsibility efforts. To deepen our understanding of the behavioral motives behind consumers’ valuations, we investigate (i) whether and when consumers’ valuations can be attributed to indirect reciprocity (i.e., consumers rewarding a firm for its active efforts); and (ii) how heterogeneity in consumers’ prosocial orientations (i.e., willingness to sacrifice one’s own benefit to help others) impacts the roles transparency and indirect reciprocity play in consumers’ valuations. Our results provide insights into when a company can benefit from improved transparency and how a company can better communicate its social responsibility practices to consumers. We observe that consumers are willing to pay a higher price under a higher level of transparency. In addition, there exists an important interplay among transparency, indirect reciprocity, and consumers’ prosocial orientations. High prosocial consumers do not exhibit indirect reciprocity. Their valuations are primarily driven by the social outcome (e.g., the worker’s pay) and the level of transparency, rather than by the knowledge about the firm’s effort. In sharp contrast, indirect reciprocity has a strong positive effect on low prosocial consumers’ valuations when transparency is high. However, as transparency decreases, we first observe a negative effect of indirect reciprocity on low prosocial consumers’ valuations, and then indirect reciprocity disappears. We complement our incentivized experiment with a product choice study and an attitudinal survey to further demonstrate the generalizability of our findings to practice.

Key words: Transparency; indirect reciprocity; prosocial orientation; social responsibility; behavioral economics; experimental economics

1. Introduction

In 2007, Patagonia launched “The Footprint Chronicles,” a website dedicated to giving consumers visibility into Patagonia’s supply chain (Patagonia 2014). The Footprint Chronicles includes detailed information regarding the social responsibility initiatives Patagonia has undertaken with its suppliers. Similarly, in 2013 Nestlé introduced a QR code on the packaging for Kit-Kats. Consumers can scan the code to obtain in-depth information about the social impacts of the product (Nestlé 2013). Websites and smart labels such as these provide consumers with unique insights into the social impacts of a company’s supply chain. However, how much consumers value such
transparency remains an open question for companies. In this paper, we employ an incentivized human-subject experiment to examine whether and when transparency impacts consumers’ valuations of a company’s social responsibility practices in its upstream supply chain.\footnote{We follow the European Commission’s definition of social responsibility as “[companies integrating] social and environmental concerns in their business operations and in their interaction with their stakeholders on a voluntary basis” (Dahlsrud 2008). We specifically focus on “social” concerns in social responsibility.} To better understand the drivers behind consumers’ valuations, we investigate the roles consumers’ reciprocal motives and prosocial orientations play in shaping their behavior under different levels of transparency.

A main objective of supply chain transparency is to provide companies with extensive visibility into their own supply chains. In a recent KPMG survey of senior executives, half of the executives interviewed admitted that “[their companies] know very little about partners beyond [their immediate tier 1 partners], let alone their entire network” (KPMG International 2013). Gaining visibility into a supply chain, however, can be a costly and time-consuming endeavor. Thus, many companies are at a crossroad in determining how transparent they want to make their supply chains (Marshall et al. 2016). For example, Marks & Spencer reports that it is working with “customers and stakeholders to identify what information they consider to be important about where and how M&S products are produced” (Marks & Spencer 2015, p. 12). Our research investigates whether there exists a potential revenue benefit for companies with a more transparent supply chain. We focus on transparency with respect to a company’s visibility into the social responsibility practices in its supply chain, while taking the disclosure of information to consumers as given. Examining the extent to which consumers value such transparency represents a critical first step in understanding the level of detail consumers desire when a company communicates its social responsibility practices.

In practice, companies often communicate detailed information about their efforts to improve social responsibility in their supply chains (e.g., the number of audits conducted). The disparity in transparency between companies is instead often reflected in the precision of information regarding the outcomes of the companies’ efforts. For example, both H&M and PUMA exert efforts to improve worker wages in their upstream supply chains and both communicate these efforts in their corporate social responsibility reports. However, the level of detail presented regarding the outcomes of these efforts is very different. While H&M measures the actual wages paid to the workers (H&M 2014), PUMA only assigns facility wage grades to its suppliers, with a higher grade implying better compliance with PUMA’s standards (PUMA 2014). Similarly, Unilever monitors and publicizes detailed working conditions in its suppliers’ factories (Oxfam 2013), whereas Carrefour demonstrates minimal information about the realized outcomes of its efforts to improve worker
well-being in its supply chain (Carrefour 2014). Motivated by examples such as these, we define transparency as the extent to which the outcomes of a company’s social responsibility efforts are precisely known to the company and consumers. In a supply chain with high transparency, both the company and consumers observe the outcomes of the company’s efforts to improve social responsibility in its supply chain. Conversely, in a supply chain with low transparency, even the company is uncertain of the outcomes of its efforts. In our design, we take the level of transparency as a characteristic of the supply chain and examine how consumers’ valuations of social responsibility vary across different levels of supply chain transparency.

To enhance our understanding of consumers’ valuations, we design our experiment to investigate the behavioral motives underlying consumers’ decisions. Various social preferences can motivate a consumer to value social responsibility, such as altruism (e.g., Levine 1998, Andreoni and Miller 2002), inequality aversion (e.g., Fehr and Schmidt 1999, Bolton and Ockenfels 2000), and reciprocity (e.g., Fehr et al. 1998, Nowak and Sigmund 1998). We focus on the preference of indirect reciprocity. Following Alexander (1987, p. 5), indirect reciprocity is defined as “the return from a social investment in another...from someone other than the recipient of the beneficence.” Thus, indirect reciprocity arises when individuals help those who have helped others. In a social responsibility context, consumers motivated by indirect reciprocity would be willing to reward a company for its active engagement in improving practices in its supply chain (e.g., ensuring reasonable compensation to workers). Such reciprocal motives have been suggested as important drivers of consumers’ valuations for social responsibility (e.g., Trudel and Cotte 2009, Nielsen 2013). We investigate whether and how the presence of indirect reciprocity interacts with the level of supply chain transparency. Our findings offer insights on when a company can benefit from emphasizing the effort it actively exerts to maintain a socially responsible supply chain.

It has been shown that consumers are heterogeneous in their tendency to care about social responsibility (e.g., García-Gallego and Georgantzís 2011). Therefore, it is important to understand how the effects that transparency and indirect reciprocity have on consumers’ valuations are impacted by individual heterogeneity. We focus on one particular aspect of heterogeneity—consumers’ prosocial orientations. We define prosocial orientation as the extent to which an individual is willing to sacrifice his/her own benefit to improve the payoff of a person that he/she directly interacts with. Prior studies have attributed the consumption of socially responsible products to values such as benevolence, equality, and responsibility (Vermeir and Verbeke 2006). Individuals who adhere to these values tend to have higher prosocial orientations. Hence, there is a natural connection between a person’s prosocial orientation and his/her attention to social responsibility.

Based on the above three key aspects, we address the following research questions: (i) How does transparency affect consumers’ valuations of a company’s social responsibility practices? (ii)
Are consumers motivated by *indirect reciprocity* in their valuations? (iii) How do transparency and indirect reciprocity interact with each other to influence consumers’ valuations? (iv) How do the answers to these three questions vary for consumers with different *prosocial orientations*? To address these questions, we design a three-player game with the following social context: A Worker has helped a Firm to make a product, and the Firm would like to sell the product to a Consumer. To capture our focus on social responsibility in the upstream supply chain, we design the wealth statuses of the players (via initial endowments) such that the Worker is a disadvantaged party compared to the other two players. In our laboratory experiment, the responsible treatment of the Worker is operationalized through monetary transfer from the Firm to the Worker (see Bartling et al. 2015, Pigors and Rockenbach 2015 for a similar approach). The Consumer’s valuation of social responsibility is measured by eliciting his willingness-to-pay (WTP) for the product given the Firm’s decision on the monetary transfer to the Worker. We incentivize Consumers’ decisions to elicit their *truthful* WTP. Employing an incentivized experiment is important because evidence exists that what individuals claim they would do with respect to social responsibility differs from their actual purchase behavior. For example, Devinney et al. (2010, p. 112) observe that “individuals either purposely overstate their social credentials or just want to look good in surveys, making it nearly impossible to believe what they say about their social proclivities.”

Our main results and contributions are as follows. We develop a micro-level understanding of consumers’ behaviors and motives in their valuations of social responsibility under different levels of information uncertainty. While researchers have used incentivized experiments in the lab (e.g., Rode et al. 2008, Krishna and Rajan 2009) and the field (e.g., Kimeldorf et al. 2006, Hainmueller et al. 2015) to study consumers’ valuations of social responsibility, none of these works examine the impact of information uncertainty on consumers’ valuations. Within the sustainable operations management (OM) literature, there is a growing interest in studying how the type and nature of the information presented to individuals affect their behavior. For example, Muthulingam et al. (2013) show that the order in which energy-saving recommendations are presented to small and medium firms significantly affects the resulting adoption rate of such recommendations. Roels and Su (2013) demonstrate that what type of comparison information a social planner should present to households regarding their energy consumptions (e.g., versus neighbors’ consumptions) depends on whether individuals are ahead-seeking or behind-averse. Kalkanci et al. (2013) compare mandatory versus voluntary disclosure of a company’s social and environmental impacts and show that voluntary disclosure is perceived more positively than mandatory disclosure. We differ from the above body of research by studying a new aspect of information – transparency regarding the outcome of a company’s social responsibility practices in its upstream supply chain.
Our research advances our understanding of consumer behaviors and motives under information transparency (and the lack thereof) by demonstrating three key results. First, we determine that improved transparency has a positive impact on consumers’ valuations. To further strengthen the applicability of this insight from our incentivized experiment to practice, we design and administer a product choice study (in the postexperiment survey) that considers broader social responsibility contexts and multiple product categories. Our results show that the positive value of improved transparency generalizes to the product choice setting. Second, we demonstrate an interaction effect between transparency and consumers’ reciprocal motives. In particular, we observe that indirect reciprocity exists and has a positive effect on consumers’ WTP only when the level of transparency is high. In contrast, indirect reciprocity disappears when the level of transparency is low. These results offer the first evidence on how uncertainty in the outcome of one’s action affects indirect reciprocity in a three-player setting. By doing so, we also add to the emerging literature that examines the effect of outcome uncertainty on social preferences primarily based on two-player dictator games (e.g., Brennan et al. 2008, Krawczyk and Le Lec 2010).

Third, by analyzing individual heterogeneity, we highlight an important interplay among transparency, indirect reciprocity, and individuals’ prosocial orientations in determining consumers’ valuations. Specifically, we show that the presence of indirect reciprocity and how it interacts with transparency significantly depend on consumers’ prosocial orientations. High prosocial consumers do not exhibit reciprocal motives. Instead, their valuations of social responsibility are primarily driven by the level of transparency and the outcome of the company’s action (rather than the fact that the company has actively exerted effort). In contrast, reciprocal motives have a strong positive effect on low prosocial consumers’ WTP when the level of transparency is high. However, for lower levels of transparency, low prosocial consumers may be motivated to either punish low effort levels by the company or reduce their WTP due to a perceived decrease in responsibility. These interactions among transparency, indirect reciprocity, and prosocial orientation suggest ways in which companies can benefit by adapting their communication strategies to the prosocial orientations of the targeted consumers. For example, if the targeted population has a low prosocial orientation, then emphasizing efforts can be beneficial for a company, but only if the outcomes of these efforts are transparent.

2. Experimental Design and Procedure
Our experimental design consists of three parts: the Consumer Purchase Game, two control tasks to elicit individuals’ prosocial orientations and risk preferences, and a postexperiment survey. The design of each part is grounded in well-established approaches in the experimental and behavioral sciences literatures. In what follows, we discuss the dynamics of each part as well as the experimental procedure. Detailed justifications of key design choices are deferred to Appendix A. Results found in the online appendix are referenced as O.X.
2.1. The Consumer Purchase Game

We design the Consumer Purchase Game as a three-player game with the following roles: a Firm (she), a Consumer (he), and a Worker. At the beginning of the game, the three players are given a background story that the Worker has helped the Firm to make a (hypothetical) product, and the Firm wants to sell the product to the Consumer. The Firm, the Consumer, and the Worker are initially endowed with 160, 120, and 20 tokens, respectively. The initial endowments are explicitly designed to capture differences in the players’ wealth statuses, particularly the wealth differences between the Worker and the other two players (see Appendix A). The Firm receives an additional provisional payment of 120 tokens. The Firm earns this provisional payment only if she manages to sell the product to the Consumer. To prevent the introduction of pre-experiment biases related to specific practices or products, we call the above three roles Players A, B, and C in the experiment and refer to a generic product with no specific features.

The sequence of events is as follows. First, the Firm decides how much from the provisional 120 tokens she is willing to use to generate a payment to the Worker. Let \( e \) denote the Firm’s decision. The potential values of \( e \) range from 0 to 120 in increments of 20; i.e., \( e \in \{0, 20, 40, 60, 80, 100, 120\} \). Second, the Consumer states the maximum price that he is willing to pay to the Firm for the product, given the Firm’s decision. We denote the Consumer’s decision as his willingness-to-pay (WTP). The Worker does not make any decisions. If the product is sold, then the Worker receives a payment \( w \) that depends on \( e \). After the Firm and the Consumer make their decisions, Nature chooses the product price, \( p \), from a uniform distribution between 1 and 120 tokens. If \( p \) is lower than or equal to the Consumer’s WTP, then the product is sold. In this case, the final payoffs to all players (including initial endowments) are as follows: (i) the Consumer pays the price of the product to the Firm and earns a payoff of \( \pi_C = 120 - p \); (ii) the Firm receives the provisional 120 tokens plus the price of the product minus her decision, earning a payoff of \( \pi_F = 160 + 120 + p - e \); and (iii) the Worker receives the payment and earns a payoff of \( \pi_W = 20 + w \). If instead, the price \( p \) is strictly higher than the Consumer’s WTP, then the product is not sold and the players earn their initial endowments: \( \pi_C = 120, \pi_F = 160, \) and \( \pi_W = 20 \). Figure 1a summarizes the above sequence of events.

In the Consumer Purchase Game, we use monetary transfer from the Firm to the Worker to operationalize social responsibility (see Appendix A). In addition, we use a random price to elicit the Consumer’s truthful WTP. This approach is known as the Becker-DeGroot-Marschak mechanism (Becker et al. 1964) and is a common technique in the experimental economics literature (e.g., Klos

\(^2\)Making players’ payoffs dependent on their own decisions is a common way of incentivizing decisions in human-subject experiments. In the sustainable OM literature, we are aware of only one behavioral study that uses an incentivized scheme (Agrawal et al. 2015).
et al. 2005, Halevy 2007). We also apply the strategy method (e.g., Fehr and Fischbacher 2004, Falk et al. 2008) to obtain the Consumer’s decision: the Consumer states his WTP for each of the seven values of $e$ that the Firm may choose while the Firm is choosing $e$ (i.e., before knowing the actual value of $e$). This way, we extract the WTP decision for every possible value of $e$ for each Consumer in our sample.\(^3\) After the Firm and the Consumer make their decisions, the players’ final payoffs are jointly determined by the value of $e$ chosen by the Firm and the corresponding WTP stated by the Consumer.

We manipulate the Consumer Purchase Game in two dimensions in our experiment. First, to study the effect of transparency on the Consumer’s WTP, we manipulate the relationship between the Firm’s decision $e$ and the payment to the Worker $w$. Specifically, we define $w$ as follows:

$$w = \begin{cases} 
0 & \text{if } e = 0, \\
\text{Uniformly distributed on } [e - s, e + s] & \text{if } e > 0,
\end{cases}$$

where the parameter $s$ models the level of transparency. We examine three different values of $s$, leading to three Transparency conditions: $s = 0$ (High Transparency), $s = 10$ (Medium Transparency), and $s = 20$ (Low Transparency). The condition of $s = 0$ represents high transparency because the payment to the Worker if the product is sold is exactly equal to the Firm’s decision

---

3 One may question whether using the strategy method primes the Consumer to state his WTP in a monotonic fashion as he goes through all possible $e$ values in a table. We conduct additional sessions in which each $e$ value is presented one at a time and on separate screens in a random order. We confirm that Consumers’ WTP decisions are not statistically different between implementing the strategy method with a table and with a random-order presentation. Thus, any priming effects the strategy method may induce are likely to be minimal in our data. See Appendix O.5.2 for further details.
Conversely, $s = 20$ represents low transparency because given $e$, there is still large uncertainty regarding the actual payment to the Worker.\(^4\) Note that in all three Transparency conditions, the expected payment to the Worker is equal to the Firm’s decision.

The Firm’s decision in our design models the Firm’s *effort* to improve social responsibility in its supply chain. The Consumer always observes the Firm’s decision (i.e., effort) because we focus on transparency regarding the *outcome* of the Firm’s effort, as opposed to transparency regarding the effort itself. In a supply chain with high transparency, both the Firm and the Consumer know exactly the Firm’s effort and the resulting payment to the Worker (i.e., the outcome). In a supply chain with low transparency, however, both the Firm and the Consumer only know the possible range of payment to the Worker even though the Firm’s effort is visible.\(^5\)

To study the effect of *indirect reciprocity* on the Consumer’s WTP, the second dimension we manipulate in our experiment is the process by which the decision $e$ is selected. We compare two Selection conditions: the *Decision* condition versus the *Random* condition. Under the Decision condition, the Firm chooses $e$ as discussed above (see Figure 1a). In contrast, under the Random condition, Nature chooses $e$ uniformly and randomly from the feasible set \{0, 20, 40, 60, 80, 100, 120\}, and the Firm automatically accepts the chosen value (see Figure 1b). Since the Firm does not actively choose $e$ in the Random condition, a higher or lower value of $e$ in this condition cannot be interpreted by the Consumer as the Firm being more or less responsible for the Worker. In other words, the Consumer’s WTP in the Random condition is driven by preferences other than indirect reciprocity. Conversely, the Consumer’s WTP in the Decision condition is driven by both these preferences *and* indirect reciprocity. Therefore, differences observed in WTP between these two conditions measure indirect reciprocity. This approach is well established in the literature for studying direct reciprocity (e.g., Blount 1995, Charness and Haruvy 2002, Falk et al. 2008). We extend it to measure indirect reciprocity. The three Transparency conditions combined with the two Selection conditions yield a $3 \times 2$ factorial design and a total of six treatments. This design allows us to study the individual effects of transparency and indirect reciprocity on WTP, as well as their interaction effect.

\(^4\) Our definitions of high, medium, and low transparency are in relative terms and chosen solely to generate directional insights. The values of $s$ that we have selected do not map to a specific product or scenario.

\(^5\) Transparency may also improve the visibility of a company’s social responsibility effort and hence, consumers’ trust toward the company. In practice, companies can often rely on a number of third-party certifications (e.g., the Fair Trade Label) or information providers (e.g., the Good Guide App) to establish the credibility of their practices and therefore, improve consumer trust. Our design is also relevant to contexts outside of social responsibility. For example, transparency has become a critical issue for charitable organizations and non-profits seeking donations from donors (Waters 2008). Our insights can help these organizations to improve their fundraising through better communication of their initiatives.
2.2. Control Tasks: Prosocial Orientation and Risk Preference

We include two control tasks in our experiment to measure Consumers’ prosocial orientations and risk preferences. We are interested in studying the Consumer’s prosocial orientation because it naturally connects with motives that induce a person to care about social responsibility. We also examine how risk preference may affect the Consumer’s decision since we study transparency with regard to uncertainty in the outcome of the Firm’s decision.

In our experiment, we define prosocial orientation as the extent to which an individual is willing to sacrifice his/her own benefit to improve the payoff of a person that he/she directly interacts with. To measure the Consumer’s prosocial orientation, we employ a variation of the dictator game (Forsythe et al. 1994) as follows. The dictator and the recipient (called Players 1 and 2 in the task) are initially endowed with 120 and 20 tokens, respectively. The Consumer from the Consumer Purchase Game is assigned the role of the dictator and is asked to choose the number of tokens, \( a \), from the set \( \{0, 20, 40, 60, 80, 100, 120\} \) that he is willing to use to generate a payment to the recipient. After the Consumer makes his decision, the recipient receives a random payment \( t \) that is uniformly distributed on \([a - s, a + s]\) if \( a > 0\), and exactly equal to zero otherwise. The value of \( s \) used in the dictator game is set equal to the value of \( s \) that the Consumer faced in the Consumer Purchase Game. Matching these two values ensures that we measure the Consumer’s prosocial orientation under the same level of uncertainty regarding outcomes. Given this design, the Consumer’s decision \( a \) as the dictator measures his prosocial orientation. A higher value of \( a \) implies that the Consumer is more willing to improve the recipient’s payoff at his own cost. Hence, he is more prosocial. We analyze whether transparency and indirect reciprocity have different effects on the WTP decisions made by Consumers with high versus low prosocial orientations.

The second control task measures the Consumer’s risk preference. To do so, we employ the multiple price list method (Holt and Laury 2002). Participants are presented with ten pairs of lotteries. For each pair, they are asked to choose which lottery, A or B, they would prefer to play (see Table B.1, Appendix B). For lottery A (B), the two possible outcomes are 60 and 74 (6 and 140) tokens. The expected values of both lotteries increase as one goes down the list, with lottery A (B) having a higher expected value in pairs 1–4 (5–10). By design, each participant is expected to either always prefer lottery B, or switch his/her preference from lottery A to lottery B only once; that is, preferring lottery A up to a certain pair and preferring lottery B thereafter. Risk-neutral participants (who maximize their expected payoffs) will choose lottery A for pairs 1–4 and lottery B for pairs 5–10. Risk-averse participants, however, could prefer to play lottery A even in cases where its expected value is lower than that of lottery B. Therefore, the pair at which a participant switches preference from A to B reflects his/her risk preference. To incentivize their choices, participants are informed that one of the ten pairs of lotteries will be randomly selected.
to be played by the computer at the end of the session. They will earn additional income based on the outcome of the lottery they choose in the randomly selected pair.

2.3. Postexperiment Survey

After participants complete both the Consumer Purchase Game and the control tasks, they are asked to complete a postexperiment survey. The survey consists of the following four parts: (i) questions about participants’ decisions in the Consumer Purchase Game, (ii) a product choice study, (iii) an attitudinal survey, and (iv) a demographic survey. Part (i) is used to obtain insights on how participants make their decisions in the game. In part (ii), we employ a simplified conjoint design (Feinberg et al. 2012) to examine how our results in the game may apply to broader social responsibility topics and product categories. Participants are presented with multiple pairs of products with different characteristics and are asked to state which product in each pair they prefer to buy. We vary two characteristics between the two products in a pair: the product price and a message attached to each product regarding the outcomes of the company’s social responsibility practices. The message is vague (representing low transparency) for one product and precise (representing high transparency) for the other. We measure whether a participant is willing to pay a premium for the product with a precise message and if so, how much of a premium he/she is willing to pay. We discuss the design and analysis of the product choice study in more detail in §5.1.

Part (iii) of the postexperiment survey contains a series of binary-choice and 5-point Likert scale questions to help us better understand participants’ general attitudes toward social responsibility and how these attitudes relate to their everyday purchase decisions. The participants’ responses provide additional data to further measure their prosocial orientations beyond the dictator game described in §2.2. Finally, part (iv) is used to obtain demographic and background information about our participants.

2.4. Experimental Procedure

We conducted all experimental sessions in the computer laboratories of two large universities. The sessions encompassed a 3 (Transparency: High vs. Medium vs. Low) × 2 (Selection: Decision vs. Random) factorial design; i.e., six treatments in total. All six treatments were run at both universities.

All sessions followed the same procedure. In each session, participants completed the following tasks in the same order: (i) two rounds of the Consumer Purchase Game, (ii) the dictator game and the risk preference task, and (iii) the postexperiment survey. Participants were given the details of the next task only after they completed the previous one. This included informing them of the occurrence of round 2 of the Consumer Purchase Game only after they completed round 1. We followed this procedure to prevent participants from altering their current decisions in
anticipation of future decisions. In addition, the outcomes of the tasks were not presented to the participants until they had completed all tasks. This procedure eliminated the possibility that outcomes from previous tasks affected current and future decisions, including income effects and changes in generosity due to past outcomes (see Appendix A).

At the beginning of round 1 of the Consumer Purchase Game, participants were randomly and anonymously matched into groups of three. Within each group, participants were randomly assigned to one of three roles: Firm (Player A), Consumer (Player B), or Worker (Player C). Participants were informed of the treatment conditions they were in and played the Consumer Purchase Game as described in §2.1. After participants completed round 1 of the game, they were told that they would be playing the Consumer Purchase Game for a second and final time. The computer performed the following role reassignments for round 2. The Consumer participants in round 1 were assigned the role of the Firm in round 2. The Firm participants in round 1 were assigned the role of the Consumer in round 2. The Worker participants in round 1 remained in that role in round 2. Therefore, two thirds of all participants played the role of the Consumer once and the role of the Firm once (in two different rounds); one third of all participants played the role of the Worker twice. All participants were then randomly and anonymously rematched into groups of three. They were informed that they were randomly and anonymously assigned to a new group, but they were not informed of how the roles were reassigned. The Transparency condition used in round 2 was different than the one used in round 1, whereas the Selection condition (i.e., Decision or Random) was the same in both rounds. The order of Transparency conditions used in the two rounds was randomized across sessions. Given a Selection condition, each of the six possible orders of Transparency conditions was implemented in 2 to 4 sessions.

After two rounds of the Consumer Purchase Game, participants who had played the Consumer role in either round performed the two control tasks: the dictator game and the risk preference task (both described in §2.2). Hereafter, we refer to these participants as the Consumer participants. In the dictator game, the Consumer participants were informed that they had been randomly and anonymously paired with another participant and they were asked to make decisions as the dictator (Player 1). In each pair, the level of uncertainty about the recipient’s actual payment was designed to match the Transparency condition that the dictator faced when he/she was the Consumer in the Consumer Purchase Game. For the risk preference task, the participants made individual decisions. Half of all Consumer participants played the dictator game before performing the risk preference task, whereas the other half followed the reverse order. After both the Consumer Purchase Game and the control tasks were completed, all participants were shown the outcomes of all tasks, including their decisions, the decisions of the participants with whom they were matched, the realized price of the product and the realized payment to the Worker in each round of the Consumer
Purchase Game, and the resulting payoffs. Finally, all participants completed the postexperiment survey described in §2.3. See Appendix O.6 for the sample instructions and postexperiment survey.

The experimental tasks were implemented using the z-Tree software (Fischbacher 2007). During a session, participants were not allowed to talk to each other. They made decisions and interacted with each other only through computer terminals. Before each task, participants read detailed instructions on the screen and answered a set of practice questions. They had to answer all practice questions correctly before proceeding to the next task. They were also provided with a reference sheet for the instructions.\textsuperscript{6} A total of 198 participants played the role of the Consumer. All of the participants in the study were students; 80.8% of them were undergraduates and the remaining 19.2% were graduate students. In addition, 61.8% of them were female, and the average age was 21.7 years old (with a standard deviation of 4.8 years).\textsuperscript{7} Participants were provided with monetary compensation based on their total payoffs in all tasks. Every 16 tokens were worth 1 U.S. dollar. Participants earned an average of $29.56, with a minimum of $20 and a maximum of $40. Each session lasted on average 90 minutes. Table 1 summarizes the number of Consumer participants in each treatment and the corresponding treatment conditions.

<table>
<thead>
<tr>
<th>Selection condition</th>
<th>Transparency condition</th>
<th>No. of Consumer participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision</td>
<td>High</td>
<td>26</td>
</tr>
<tr>
<td>Decision</td>
<td>Medium</td>
<td>29</td>
</tr>
<tr>
<td>Decision</td>
<td>Low</td>
<td>31</td>
</tr>
<tr>
<td>Random</td>
<td>High</td>
<td>38</td>
</tr>
<tr>
<td>Random</td>
<td>Medium</td>
<td>34</td>
</tr>
<tr>
<td>Random</td>
<td>Low</td>
<td>40</td>
</tr>
</tbody>
</table>

Note. Consumer participants refer to participants who played the Consumer role in either round 1 or round 2 of the Consumer Purchase Game.

3. Hypotheses

Our experimental design allows us to analyze how Consumers’ WTP for a socially responsible product is affected by three key factors: (i) the level of transparency about the payment to the Worker; (ii) Consumers’ indirect reciprocity toward the Firm; and (iii) Consumers’ prosocial orientations. Note that if Consumers only care about their own payoffs, then they will always state WTP = 0 regardless of the value of $e$. However, an extensive literature has shown that individuals often

\textsuperscript{6} Participants’ responses in the postexperiment survey confirm that they understood the dynamics of the game. For example, they understood that choosing WTP = 0 would maximize the Consumer’s payoff, that the Worker was more likely to receive the payment $w$ if the WTP was higher, and that a larger $e$ chosen by the Firm would give the Worker a larger share of the gain from selling the product. When asked whether they would change their decisions if they were to play the game again, a vast majority of the participants stated that they would not.

\textsuperscript{7} We verify that our experimental results remain unchanged after controlling for participants’ demographic factors. See Appendix §O.5.1 for more details.
care about others’ well-being in addition to their own payoffs (for a review, see Fehr and Schmidt 2006). For example, inequality aversion may motivate individuals to reduce the payoff differences between themselves and others (Fehr and Schmidt 1999, Bolton and Ockenfels 2000). Within the social responsibility literature, there also exist a number of studies showing that consumers’ WTP for a product and/or perceptions of a company often increase if the company demonstrates better practices (e.g., Mohr et al. 2001, Sen and Bhattacharya 2001, De Pelsmacker et al. 2005). Therefore, in our experiment, we expect Consumers’ WTP to be increasing in the value of $e$. Building on this expected behavior, we next discuss our main hypotheses regarding the effects of transparency and indirect reciprocity. Recall from §2.1 that the Random condition is designed solely to isolate the effect of indirect reciprocity on Consumers’ WTP. Therefore, our hypothesis regarding transparency only examines Consumers’ WTP in the Decision condition; our hypothesis regarding indirect reciprocity examines Consumers’ WTP in both the Decision and Random conditions.

We first consider how transparency regarding the actual payment to the Worker impacts Consumers’ WTP. Under low and medium transparency, the exact payment that the Worker receives if the product is sold is unknown. A limited number of studies show that the existence of uncertainty in the outcome of one’s actions reduces one’s willingness to contribute to others’ welfare. For example, Brock et al. (2013) compare the standard dictator game with one where the payment to the recipient is uncertain and the expected payment is equal to the amount the dictator transfers. They observe that the dictators transfer significantly lower amounts to the recipients in the presence of uncertainty. Exley (2015) measures participants’ valuations of lotteries when the outcomes of the lotteries determine the payment they would receive versus when the outcomes determine the donation a charity would receive. The author shows that the participants’ valuations are more adversely affected by the lotteries’ risk level in the latter case. Researchers have also shown that in public goods games, uncertainty about the effects of one’s actions can induce more self-interested behavior (Budescu et al. 1995, Gangadharan and Nemes 2009, Levati et al. 2009). Given these results, we develop the following hypothesis.

**Hypothesis 1.** Given the Firm’s decision $e$, Consumers’ willingness-to-pay increases as the level of transparency increases.

Our second hypothesis examines how indirect reciprocity affects Consumers’ WTP. We first distinguish indirect reciprocity from direct reciprocity. Direct reciprocity occurs when an individual treats someone else in the same manner as that person has treated him/her. Conversely, indirect reciprocity occurs when an individual treats someone else in the same manner as that person has treated a third party. Direct reciprocity has been studied extensively in the literature (e.g., Fehr et al. 1998, Charness and Rabin 2002, Cox 2004). The most common setting used to study direct
reciprocity is the gift exchange game (Akerlof 1982). For example, Charness (2004) shows that the marginal increase in an employee’s effort given a unit increase in his wage is larger when the wage increase is actively chosen by the firm than when it is randomly determined. This phenomenon is attributed to direct reciprocity.

Although less studied than direct reciprocity, the existence of indirect reciprocity has also been documented in the literature (e.g., Seinen and Schram 2006, Engelmann and Fischbacher 2009, Stanca 2009). Stanca (2009) is most closely related to our paper. The author studies the following two-stage game. In the first stage, a participant (the sender) chooses an amount $x$ that he/she wants to send to another participant (the recipient). In the second stage, a third participant (i.e., different from the sender and the recipient) observes the amount $x$ chosen by the sender and chooses an amount $y$ that he/she wants to transfer to the sender. The author finds that the amount $y$ chosen in the second stage positively correlates with the amount $x$ sent in the first stage. He attributes this result to indirect reciprocity. Our work differs from Stanca (2009) in at least two important aspects: (i) we investigate indirect reciprocity under different levels of transparency; and (ii) we disentangle indirect reciprocity from other social preferences by comparing WTP between the Decision and Random conditions.

In our Consumer Purchase Game, indirect reciprocity occurs if the Consumer rewards or punishes the Firm when he perceives the Firm acting responsibly or irresponsibly toward the Worker. We extend Charness’s (2004) approach to measure indirect reciprocity. That is, given a unit increase in the Firm’s decision, we consider indirect reciprocity to exist if the marginal increase in Consumers’ WTP is larger in the Decision condition than in the Random condition (see Figure 1 for the design of these two conditions). To the best of our knowledge, the current literature has not studied the impact of transparency (or uncertainty about how one’s action affects another’s payoff) on either direct or indirect reciprocity. Therefore, we make the following null hypothesis that indirect reciprocity exists regardless of the level of transparency.

**Hypothesis 2.** At each level of transparency, the same increase in $e$ yields a larger increase in Consumers’ willingness-to-pay in the Decision condition than in the Random condition.

In addition to the above hypotheses, our experimental design also allows us to investigate how the effects of transparency and indirect reciprocity on WTP vary depending on Consumers’ prosocial orientations. Among studies that examine individuals’ prosocial orientations, one paper that is most relevant to ours is de Kwaadsteniet et al. (2006). The authors experimentally study a common resource dilemma where a person’s usage of the common resource reduces the availability of the resource to others (i.e., a negative externality). They compare participants’ behavior when the size of the common resource is certain versus uncertain. They show that although all
participants increase their usage of the common resource when the size of the resource becomes uncertain, such an increase is much larger for participants with a low prosocial orientation than for those with a high prosocial orientation. While this result suggests that uncertainty induces more self-interested decisions among low prosocial individuals than among high prosocial ones, no research has examined the interactions among transparency, indirect reciprocity, and prosocial orientation. Investigating these interaction effects along with the above hypotheses offers insights into when consumers value transparency and how companies can better communicate their social responsibility practices to consumers.

4. Experimental Results

In this section, we discuss our experimental results. Our data was collected by running sessions in two universities, with two rounds of the Consumer Purchase Game being played in each session. We verify that (i) there are no significant differences in the decisions made by participants from the two universities in any treatment (two-sided Wilcoxon rank-sum test, \( p > 0.1 \)); and (ii) all of our results continue to hold if we only use data from round 1 of the Consumer Purchase Game. Therefore, we present our results based on all data collected. Hereafter, we refer to the decision \( e \) as the “effort” chosen by the Firm (in the Decision condition) or by Nature (in the Random condition). We also abbreviate indirect reciprocity as reciprocity to simplify our discussion.

Table 2 presents the summary statistics of Consumers’ WTP for all treatments. We observe that the average WTP is significantly higher than zero in all treatments and for all effort levels (one-sided Wilcoxon signed-rank test, \( p < 0.001 \)). As expected, we also observe that the average WTP is increasing in the effort levels. This result is confirmed by the significantly positive Pearson’s correlation between the effort levels and the Consumers’ WTP. This correlation ranges between 0.25 and 0.35 across treatments (one-sided \( t \) test, \( p < 0.001 \)).

To formally test our hypotheses, we estimate the following random-effects two-sided Tobit model with Consumers’ WTP as the dependent variable:

\[
WTP_{jk}^* = \text{Intercept} + \alpha_{TM} \cdot T_M + \alpha_{TH} \cdot T_H + \alpha_D \cdot D + \alpha_{DTM} \cdot D \cdot T_M + \alpha_{DTH} \cdot D \cdot T_H + \beta \cdot e_k + \\
+ \beta_{TM} \cdot T_M \cdot e_k + \beta_{TH} \cdot T_H \cdot e_k + \beta_D \cdot D \cdot e_k + \beta_{DTM} \cdot D \cdot T_M \cdot e_k + \beta_{DTH} \cdot D \cdot T_H \cdot e_k + \delta_j + \epsilon_{jk}.
\]  

The average WTP being significantly higher than zero even when \( e = 0 \) may be due to Consumers’ efficiency considerations; i.e., they prefer the product being sold to increase the total (expected) payoff for the group. The presence of efficiency considerations is well documented in the social preferences literature (e.g., Charness and Rabin 2002, Engelmann and Strobel 2004). To examine whether efficiency considerations influence our results, we perform two robustness checks. First, we conduct sessions for a variation of the Consumer Purchase Game where the 120 tokens are not provisional. Instead, the Firm starts with an endowment of 160 + 120 = 280 tokens. Second, we repeat our analysis by excluding the data from those Consumers who state a positive WTP when \( e = 0 \). We verify for both analyses that our results regarding the effects of transparency and reciprocity reported in §4 continue to hold. Hence, efficiency considerations likely have a minimal effect on our conclusions. See Appendix O.5.3 for more details.

Related to this observation, Mohan et al. (2015) show in a series of non-incentivized experiments that consumers’ purchase intentions and perceptions of a firm increase with lower CEO-to-worker pay ratios.
Table 2: Mean, Median, and (Standard Deviation) of Consumers’ Willingness-to-Pay

<table>
<thead>
<tr>
<th>Effort</th>
<th>Transparency condition</th>
<th></th>
<th></th>
<th></th>
<th>Transparency condition</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>52 [55] (41)</td>
<td>47 [45] (43)</td>
<td>37 [28] (39)</td>
<td>36 [22] (38)</td>
<td>45 [38] (41)</td>
<td>41 [32] (37)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. All values are significantly positive (Wilcoxon signed-rank test, \( p < 0.001 \)). In all treatment conditions, the correlation between effort and WTP is significantly positive (t test, \( p < 0.001 \)).

The variable \( WTP^* \) is the latent variable for WTP in the Tobit model. The subscript \( j \) is the index for a Consumer participant and \( k \) is the level of effort chosen. The variable \( \epsilon_k \) with \( k = 1, 2, \ldots, 7 \) represents the seven possible effort levels in the set \( \{0, 20, 40, 60, 80, 100, 120\} \). The dummy variables \( T_M \) and \( T_H \) indicate Medium and High transparency. The dummy variable \( D \) indicates the Decision condition. The interaction terms among the dummy variables are included as independent variables in Equation (1) to investigate the interaction effect between transparency and reciprocity. The term \( \delta_j \) represents the individual-specific error, and \( \epsilon_{jk} \) is the independent error across WTP decisions.

We apply a random-effects two-sided Tobit model for the following reasons. First, the dependent variable WTP is bounded between 0 and 120, with approximately 28% of our observations taking the value of 0. Therefore, corner solution outcomes occur in our data and using a Tobit model to correctly account for them is desirable (Wooldridge 2010, pp. 667–689). Second, due to our use of the strategy method, we have seven WTP observations from each Consumer, corresponding to the seven possible effort levels. We accommodate such repeated measures by employing the random-effects approach; i.e., we incorporate an individual-specific error term \( \delta_j \). Our choice of regression model is well established in the related literature (e.g., List 2006, Bolton et al. 2013). Because the variable WTP in our setting is not actually censored, we are not interested in studying the uncensored latent variable \( WTP^* \). Instead, we focus on the marginal effects of the key independent variables in Equation (1). Appendix C presents the estimation results of Equation (1).

4.1. Hypothesis 1: Does Consumers’ Willingness-to-Pay Increase with the Level of Transparency?

We begin by analyzing the effect of transparency on Consumers’ WTP in the Decision condition. Figure 2 demonstrates how the average WTP observed changes with the effort level in each Transparency condition. We observe that the average WTP under high transparency is higher than
that under medium or low transparency. Similarly, except when $e_k = 0$, the average WTP under medium transparency is higher than that under low transparency.

To formally test Hypothesis 1, we analyze the marginal effects of $T_M$ and $T_H$ in Equation (1), as well as their differences $\Delta_T \equiv T_H - T_M$ at all effort levels. The marginal effects of $T_M$ ($T_H$) show the changes in average WTP from low to medium (high) transparency. A positive value indicates that the average WTP is higher under medium (high) transparency than under low transparency. Similarly, the values of $\Delta_T$ show the changes in average WTP from medium to high transparency. A positive value indicates that the average WTP is higher under high transparency than under medium transparency.

Table 3 summarizes the marginal effects of $T_M$, $T_H$, and their differences, $\Delta_T$. We first observe that all marginal effects are positive, with the marginal effects of $T_H$ being statistically significant at all effort levels. Thus, we support Hypothesis 1 (WTP increases as transparency increases) when comparing high versus low transparency. Second, the values of $T_M$ are significant only at $e_k = 120$, and the values of $\Delta_T$ are never significant. Therefore, Hypothesis 1 is partially supported for medium versus low transparency, but it is not supported for high versus medium transparency. These results imply decreasing marginal increments in the Consumers’ WTP as transparency increases. That is, improving from low to medium transparency yields a larger increase in WTP than improving from medium to high transparency (at a high effort level). In summary, we obtain the following result:

**Result 1** As the level of transparency increases, Consumers’ willingness-to-pay increases with decreasing marginal increments.
### Table 3  Marginal Effects of $T_H$, $T_M$, and Their Differences, $\Delta_T$

<table>
<thead>
<tr>
<th>Transparency condition</th>
<th>$T_H$</th>
<th>$T_M$</th>
<th>$\Delta_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8.98 (6.49)*</td>
<td>2.52 (5.57)</td>
<td>6.46 (6.75)</td>
</tr>
<tr>
<td>20</td>
<td>10.44 (7.12)*</td>
<td>3.81 (6.29)</td>
<td>6.63 (7.48)</td>
</tr>
<tr>
<td>40</td>
<td>11.93 (7.74)*</td>
<td>5.34 (7.01)</td>
<td>6.60 (8.17)</td>
</tr>
<tr>
<td>60</td>
<td>13.39 (8.33)*</td>
<td>7.05 (7.71)</td>
<td>6.34 (8.81)</td>
</tr>
<tr>
<td>80</td>
<td>14.78 (8.89)**</td>
<td>8.91 (8.38)</td>
<td>5.87 (9.40)</td>
</tr>
<tr>
<td>100</td>
<td>16.04 (9.42)**</td>
<td>10.83 (9.01)</td>
<td>5.21 (9.90)</td>
</tr>
<tr>
<td>120</td>
<td>17.13 (9.90)**</td>
<td>12.73 (9.57)*</td>
<td>4.40 (10.31)</td>
</tr>
</tbody>
</table>

Note. Values shown are the marginal effects evaluated at the Decision condition; i.e., at $D = 1$ in Equation (1). Values in parentheses are the standard errors.

**: $p < 0.05$; *: $p < 0.1$; $p$ values are derived from one-sided $t$ tests.

### 4.2. Hypothesis 2: Does Reciprocity Exist at Each Level of Transparency?

Next, we address whether Consumers exhibit reciprocity in their WTP at each level of transparency. If reciprocity exists, then we also examine whether it has a positive or negative impact on Consumers’ WTP; i.e., whether it generates a higher or lower average WTP in the Decision condition than in the Random condition. Figure 3 shows the average WTP observed under the Decision and Random conditions for all three levels of transparency. Table 4 summarizes the marginal effects of $D \cdot e_k$ in Equation (1). A positive value indicates that the slope of WTP as a function of $e_k$ is larger in the Decision condition than in the Random condition and hence, reciprocity exists. We divide our discussion by Transparency condition.

**Figure 3  Effect of Reciprocity on Consumers’ Willingness-to-Pay (Experimental Data)**

We first examine the High Transparency condition. In Figure 3a, we observe that the average WTP increases in the effort level at a faster rate in the Decision condition than in the Random condition; i.e., the solid line has a steeper slope than the dotted line. This observation is confirmed by the marginal effects in the regression. As shown in Table 4 (column 2), under high transparency,
all marginal effects are significantly positive. Therefore, reciprocity exists and Hypothesis 2 is supported when transparency is high. In addition, the average WTP is significantly higher in the Decision condition than in the Random condition for $e_k \geq 100$ (one-sided Wilcoxon rank-sum test, $p < 0.1$). Hence, the presence of reciprocity generates a positive effect on Consumers’ WTP at high effort levels when transparency is high.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Marginal Effects of the Interaction Term $D \cdot e_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparency condition</td>
<td>High</td>
</tr>
<tr>
<td>Effort</td>
<td>0</td>
</tr>
<tr>
<td>&amp; 20</td>
<td>0.10 (0.04)***</td>
</tr>
<tr>
<td>&amp; 40</td>
<td>0.11 (0.04)***</td>
</tr>
<tr>
<td>&amp; 60</td>
<td>0.12 (0.04)***</td>
</tr>
<tr>
<td>&amp; 80</td>
<td>0.13 (0.04)***</td>
</tr>
<tr>
<td>&amp; 100</td>
<td>0.13 (0.04)***</td>
</tr>
<tr>
<td>&amp; 120</td>
<td>0.13 (0.04)***</td>
</tr>
</tbody>
</table>

Note. Values in parentheses are the standard errors.

\(**\*\*\*: p < 0.01; \*: p < 0.1; p values are derived from one-sided t tests."

We next consider the Medium and Low Transparency conditions. Our results show that the effect of reciprocity diminishes as the level of transparency decreases. First, observe for the Medium Transparency condition in Table 4 that although the marginal effects of $D \cdot e_k$ are positive at all effort levels, they are significant only when $e_k \geq 100$. In addition, the presence of reciprocity at $e_k \geq 100$ does not result in a significantly higher average WTP in the Decision condition than in the Random condition (two-sided Wilcoxon rank-sum test, $p > 0.3$ for all effort levels; see also Figure 3b). Second, observe for the Low Transparency condition in Table 4 and Figure 3c that the slopes of WTP as a function of effort are very similar between the Decision and Random conditions. Hence, reciprocity does not exist under low transparency.

To summarize, we find partial support for Hypothesis 2 and obtain the following result:

**Result 2** There exists an interaction between transparency and indirect reciprocity.

(i) Under high transparency, indirect reciprocity exists for all effort levels and has a positive effect on Consumers’ willingness-to-pay at high effort levels.

(ii) Under medium transparency, indirect reciprocity exists only at high effort levels and does not have an effect on Consumers’ willingness-to-pay.

(iii) Under low transparency, indirect reciprocity does not exist.
4.3. The Impact of Consumers’ Prosocial Orientations

Next, we investigate how Consumers’ prosocial orientations impact their WTP decisions. Recall from §2.2 that we employ the dictator game to elicit Consumers’ prosocial orientations. We first note that the distribution of the dictators’ decisions are very similar across all treatments (Kruskal-Wallis test, \( p = 0.73 \)). Regardless of the treatment, the first quartile, median, and third quartile are equal to 0, 20, and 40 tokens, with the average ranging between 22 and 24 tokens. Therefore, we use the median dictator decision among all Consumers (which is equal to 20) to classify two types of Consumers. Those participants who transferred 0 or 20 tokens are classified as low prosocial and account for 64% of the entire sample (127 participants). Those participants who transferred 40 or more tokens are classified as high prosocial and account for the remaining 36% of the sample (71 participants).

We next investigate the behavior of high and low prosocial Consumers with respect to our hypotheses. To do so, we first reestimate Equation (1) with the data from these two groups of Consumers separately. We then compute for each prosocial type the same marginal effects as before. The detailed regression results and marginal effects are reported in Appendix D.

We first confirm that the effect of transparency on Consumers’ WTP does not vary significantly between prosocial types. As in the pooled results, both types of Consumers state significantly higher WTP under high transparency than under low transparency, whereas their WTP does not differ significantly between high and medium transparency. When comparing medium and low transparency, high prosocial Consumers state significantly higher WTP under medium transparency for all positive effort levels, whereas low prosocial Consumers do so only when \( e_k \geq 100 \) (see Table D.4 in Appendix D). Nonetheless, we continue to observe that as transparency increases, both high and low prosocial Consumers’ WTP increases with decreasing marginal increments (i.e., Result 1).

The key difference in behavior that we observe between high and low prosocial Consumers is with regard to the role reciprocity plays in influencing their WTP decisions. Figures 4 and 5 present the high and low prosocial Consumers’ WTP decisions for each combination of Transparency and Selection conditions. As shown in Figures 4a–4c, high prosocial Consumers do not exhibit reciprocity in any of the Transparency conditions; i.e., the slope of the solid line is never significantly steeper than that of the dotted line (see Table D.5 in Appendix D). Hence, we do not find support for Hypothesis 2 among high prosocial Consumers. Our findings suggest that high prosocial Consumers’ WTP decisions are mainly driven by the expected payment to the Worker (i.e., the value of \( e \)) and the level of transparency. In other words, high prosocial Consumers are more concerned with observing a precise social responsibility outcome than knowing whether the outcome is due to the Firm’s active effort or determined by Nature.

\(^{10}\) We repeat our analysis by defining low (high) prosocial Consumers as those participants who transferred exactly 0 (40 or more) tokens in the dictator game. Under this classification, low (high) prosocial Consumers constitute approximately the lower-third (upper-third) of the distribution of dictator decisions from all Consumers. We observe similar results as reported in §4.3. See Appendix O.4 for more details.
In sharp contrast, low prosocial Consumers’ WTP is significantly affected by reciprocity. First, Figure 5a is very similar to Figure 3a. The solid line being steeper than the dotted line shows that reciprocity exists for low prosocial Consumers under high transparency (see Table D.5 in Appendix D). In addition, similar to Figure 3a, reciprocity leads to a significantly higher WTP at \( e_k = 120 \) in the Decision condition than in the Random condition \( (p < 0.1).^{11} \) Second, we continue

\[ ^{11} \text{We compare low prosocial Consumers’ WTP decisions between the Decision and Random conditions for a given effort in two aspects: (i) the proportion of zero values (\( \chi^2 \) test) and (ii) the distribution of non-zero values (one-sided Wilcoxon rank-sum test). This approach is appropriate because a high percentage of WTP decisions from low prosocial Consumers are equal to zero (38.8% versus only 9% from high prosocial Consumers; see Lachenbruch 2002, Delucchi and Bostrom 2004). Under medium and high transparency, the proportion of zero values is not significantly different, but the distribution of positive WTP decisions significantly differs between the Decision and Random conditions. Under low transparency, the proportion of zero values is significantly larger for the Decision condition, but the distribution of positive WTP decisions is not significantly different. The \( p \) values reported correspond to the significant results.} \]
to observe the presence of reciprocity for low prosocial Consumers under medium transparency (see Figure 5b). However, unlike the case of high transparency, reciprocity now has a negative effect on WTP that drives the average WTP at low effort levels to be lower in the Decision condition than in the Random condition \( (p < 0.1 \text{ for } e \in \{20, 40\}) \). Third, as shown in Figure 5c, reciprocity does not exist under low transparency as the slope of the solid line is not significantly higher than that of the dotted line. Hence, we partially support Hypothesis 2 among low prosocial Consumers. The following result summarizes our findings regarding Consumers’ prosocial orientations:

Result 3 There exists an interaction among transparency, reciprocity, and prosocial orientation.

(i) High prosocial Consumers do not exhibit indirect reciprocity under any level of transparency.

(ii) Low prosocial Consumers

(a) Exhibit significant indirect reciprocity under both high and medium transparency, with indirect reciprocity having a positive (negative) effect on their willingness-to-pay under high (medium) transparency at high (low) effort levels.

(b) Do not exhibit indirect reciprocity under low transparency.

Finally, we note that low prosocial Consumers demonstrate a unique behavior under low transparency. Their average WTP in the Decision condition is significantly lower than their average WTP in the Random condition at all positive effort levels except for \( e = 120 \) \( (p < 0.05 \text{ for } e \in \{20, 40, 60, 80\}, p < 0.1 \text{ for } e = 100; \text{ see Figure 5c}) \). This result is driven by a significantly larger fraction of low prosocial Consumers (9 out of 17 or 53%) stating WTP = 0 for all effort levels in the Decision condition under low transparency (versus 7 out of 28 or 25% in the Random condition under low transparency; one-sided \( \chi^2 \) test, \( p < 0.1 \)). We postulate that low prosocial Consumers’ behavior under low transparency is affected by the different levels of responsibility they perceive for themselves between the Decision and Random conditions. We discuss this phenomenon further in §4.4.2.

4.4. Discussion

In this subsection, we discuss potential behavioral explanations for (i) why Consumers state significantly lower WTP under low transparency than under high transparency, and (ii) why low prosocial Consumers state significantly lower WTP in the Decision condition than in the Random condition under low transparency.

4.4.1. Consumers’ Lower Willingness-to-Pay under Lower Transparency We first discuss why Consumers’ WTP is significantly lower under low transparency than under high transparency. Our finding suggests that Consumers’ valuations of social responsibility decrease when the resulting social outcomes are highly uncertain (i.e., when transparency is low). One possible
explanation is that our Consumer participants are risk averse and dislike paying for an uncertain outcome, even if the uncertainty only impacts another participant, such as the Worker. To investigate this hypothesis, we classify Consumer participants as risk averse or non-risk averse based on their decisions in the risk preference task (see §2.2). A participant is considered risk averse if he/she switches from lottery A to lottery B in this task at pair 6 or later of the multiple price list; otherwise he/she is classified as non-risk averse. Based on this classification, we analyze whether the effect of transparency on WTP differs between risk-averse and non-risk-averse Consumers (see Appendix O.1). We observe that only non-risk-averse Consumers state significantly lower WTP under low transparency than under high transparency. Conversely, the WTP decisions of risk-averse Consumers are not statistically different across transparency conditions. Hence, there is no evidence that the lower WTP observed under low transparency can be attributed to the Consumers’ risk aversion. Instead, we postulate that it is a result of Consumers’ self-serving biases.

A self-serving bias is an individual’s tendency to “place greater weight on information that is consistent with [his/her] preferences” (Babcock et al. 1996, p. 3). Under low transparency, a Consumer prone to a self-serving bias may use the presence of uncertainty to justify his/her selection of a lower WTP and hence, keep a higher payoff for him/herself. A study by Haisley and Weber (2010) documents the presence of self-serving biases in a dictator game context. They find that when the distribution of the recipient’s payoff is unknown, the dictator overestimates the recipient’s expected payoff and contributes less than when the distribution is known. By overestimating, the dictator justifies keeping a larger amount for him/herself.

A common approach to study whether a self-serving bias exists is to utilize Likert-scale questions to examine whether people tend to focus more on information that is consistent with their preferences or actions (e.g., Babcock and Loewenstein 1997, Cross et al. 2011, Paharia et al. 2013). Following this approach, we ask the Consumers to state on a 5-point Likert scale how much they agree or disagree with each of the following three statements: “Not knowing what the exact payment to Player C was, I mainly focused on Player C’s minimum possible [maximum possible] [average] payment when making my decision.” The italicized phrase is the only difference across the three statements. A higher score assigned to a statement means that the participant agrees with that statement more strongly. Based on their responses, we categorize our Consumer participants as follows. If a participant gives the “minimum possible” statement the highest score among all three statements, then the participant is classified as demonstrating a self-serving bias. Otherwise, the participant is classified as not demonstrating a self-serving bias. Under this categorization, 16 out of 30 (53%) Consumer participants in the Decision condition under low transparency demonstrate a self-serving bias.\footnote{We were not able to perform a similar analysis for Consumers in the Medium Transparency condition because the survey data was corrupted during one of the sessions. Consumers in the High Transparency condition did not evaluate these statements and hence, were not categorized.}

We were not able to perform a similar analysis for Consumers in the Medium Transparency condition because the survey data was corrupted during one of the sessions. Consumers in the High Transparency condition did not evaluate these statements and hence, were not categorized.
We then analyze if the effect of transparency on WTP depends on whether Consumers demonstrate a self-serving bias or not (see Appendix O.2). We observe that Consumers who demonstrate a self-serving bias state a significantly lower WTP under low transparency than under high transparency. In sharp contrast, the WTP decisions of Consumers who do not demonstrate a self-serving bias are not statistically different between high and low transparency. Therefore, our analysis suggests that the presence of a self-serving bias is a key contributor to the significant differences in WTP that we observe between high and low transparency.

4.4.2. Responsibility Alleviation by Low Prosocial Consumers under Low Transparency

We next discuss the observation that low prosocial Consumers state lower WTP in the Decision condition than in the Random condition under low transparency. This unexpected result may be related to Consumers’ perceptions of responsibility alleviation, a phenomenon that has been documented in other games examining social preferences (e.g., Charness 2000, Dana et al. 2007). In our context, when the Firm makes an active decision in the Decision condition, both the Firm’s and the Consumer’s decisions determine the final outcome of the game and the payoffs to all three players. Conversely, in the Random condition, the Consumer is the only player who makes an active decision to influence the final payoffs of all three players. Therefore, the Consumer is likely to feel more responsible for the Worker’s well-being in the Random condition than in the Decision condition. With a decreased sense of responsibility for the final outcomes, the Consumer’s low prosocial orientation motivates him to state a lower WTP in the Decision condition than in the Random condition when transparency is low. We indeed find evidence in the postexperiment survey that our participants perceived different levels of responsibility between the Decision and Random conditions. For example, a participant who played the Consumer role in the Random condition under low transparency stated that “being [the Consumer] made me feel pressured as the decision maker...that affects other people.” In contrast, another participant who played the Consumer role in the Decision condition under low transparency stated that “It is [the Firm] which decides whether to be fair or not with [the Worker].” Our observations are in line with Gneezy et al.’s (2010) results from a field experiment. The authors study consumers’ purchase decisions for souvenir photos in a setting where half of the sales revenue goes to a charity. They contrast the case when consumers name their own prices for the photos (i.e., a “pay-what-you-want” strategy) with the case when consumers observe listed prices. The authors find that consumers make larger donations to the charity under pay-what-you-want pricing and attribute this outcome to consumers’ perceptions of “shared social responsibility” under this strategy.

We also observe an interesting connection between the potential phenomenon of responsibility alleviation and the presence of self-serving biases. In particular, the observation that low prosocial
Consumers state lower WTP in the Decision condition than in the Random condition under low transparency occurs only for Consumers who exhibit a self-serving bias. Furthermore, among low prosocial Consumers who exhibit a self-serving bias, a significantly larger fraction of them in the Decision condition than in the Random condition state WTP = 0 for all effort levels (8 out of 10 or 80% versus 6 out of 18 or 33%; one-sided $\chi^2$ test, $p < 0.05$). We conjecture that the following behaviors are happening. On the one hand, low prosocial Consumers who exhibit a self-serving bias use the presence of uncertainty to justify very little contribution to increase the Workers’ payoffs. On the other hand, when they become the sole decision maker, the increased sense of responsibility counteracts such justification and motivates them to state a higher WTP. These exploratory observations suggest that the behaviors of low prosocial individuals are potentially more malleable and easily influenced by environmental factors (e.g., uncertainty and whether they are the only decision makers who can affect others’ payoffs). This malleability has been documented in studies of social dilemmas where one’s actions have negative externalities on others’ welfare (e.g., Roch and Samuelson 1997, Au and Kwong 2004). Further examining how low prosocial individuals’ behaviors are influenced by environmental factors within a social responsibility context could be a valuable future research direction.

5. Validation of Experimental Results

In this section, we validate the robustness of our key results. Specifically, we confirm the positive effect of improved transparency on Consumers’ WTP with the product choice study that considers broader social responsibility contexts and multiple product categories. We also demonstrate that our results regarding high and low prosocial Consumers are robust to an alternative classification of prosocial types based on participants’ responses in the attitudinal survey. Finally, we examine how the initial endowment of the disadvantaged party in our experiment (i.e., the Worker) impacts our results regarding the effects of transparency and reciprocity on Consumers’ WTP.

5.1. Product Choice Study: The Value of Transparency for Broader Social Responsibility Contexts and Multiple Product Categories

Our first validation examines how well our finding that improved transparency has a positive effect on Consumers’ WTP generalizes to broader social responsibility contexts and multiple product categories. Recall from §2.3 that we design and include in our postexperiment survey a product choice study in which participants are presented multiple pairs of products with different attributes and are asked to state which product in each pair they prefer to buy. We examine three product categories – coffee, t-shirts, and laptop computers – and five social responsibility topics – treatment of the company’s (suppliers’) employees, community development in regions where the company (suppliers) operates, and charitable donations. Each Consumer participant evaluates 12 pairs of
products divided into two 6-pair blocks, with each block consisting of a distinct product category and social responsibility topic. We manipulate two attributes within a 6-pair block: the product prices and the message attached to each product about the company’s social responsibility practices for one of the topics. Within a pair, one product is labeled with a vague message regarding the company’s practices (representing low transparency), while the other product is labeled with a precise message regarding practices (representing high transparency). The product with a precise message is priced equal to or greater than the product with a vague message. The 6-pairs of products within a block examine six different levels of price premiums for the precise message. See Appendix O.3 for more details regarding the design and implementation of the product choice study.

Our findings in the product choice study confirm and complement our experimental result that Consumers value improved transparency (i.e., supporting Hypothesis 1). First, we observe that for all three product categories and all five social responsibility topics, a vast majority of Consumers prefer products with more transparent information when they are not offered at a premium (see Figure O.2, Appendix O.3). Second, we analyze the average premium that Consumers are willing to pay for products with more transparent information. We calculate the premium a Consumer is willing to pay as follows. Let $p_v$ denote the price of a product with a vague message (which does not change across questions for a given product category). If the Consumer prefers the product with a precise message over the product with a vague message at all prices up to and including $\bar{p}$, then the premium that this Consumer is willing to pay is $(\bar{p} - p_v)/p_v \times 100\%$. Figure 6 summarizes the average premiums in percentages that Consumers are willing to pay for different product categories and social responsibility topics. We observe that our Consumer participants are on average willing to pay a premium ranging from 1.6% to almost 11% when more transparent information is provided.

Finally, to compare our results in the product choice study with those in the Consumer Purchase Game, we perform the following additional analysis. We define two groups of Consumers based on the premium they are willing to pay for products with a precise message in the product choice study, as calculated above. The “premium” group contains Consumers who are willing to pay a positive premium for both product categories and social responsibility topics they evaluated. Conversely, the “no-premium” group contains Consumers who are not willing to pay a premium for at least one product category and social responsibility topic. We then analyze whether the effect of transparency on WTP in the Consumer Purchase Game is stronger for the premium group versus the no-premium group (see Appendix O.3). We observe that Consumers in the premium group state a significantly higher WTP under high transparency than under medium or low transparency.

\[13\text{ If a Consumer always prefers the product with a vague message or he/she prefers the product with a precise message only when the prices are the same, then we treat the premium he/she is willing to pay as zero.}\]
for almost all effort levels. For Consumers in the no-premium group, the marginal effects of the transparency variables in the regression are positive but not significant (except for $T_M$ at $e_k \geq 100$). This analysis, together with our previous results on Consumers’ preferences in the product choice study, confirm that the positive impact of improved transparency on Consumers’ WTP observed in the Consumer Purchase Game generalizes to broader social responsibility topics and multiple product categories.\footnote{One may ask whether the participants’ preferences in the product choice study are influenced by the treatment conditions they faced in the experiment. We determine that such a spillover effect is unlikely based on the following analysis. For each product studied, we run a linear regression in which the dependent variable is the participants’ premiums and the independent variables include the Transparency and Selection conditions when they played the Consumer role, the round in which they played the Consumer role, and the Transparency condition when they played the Firm role. None of these regression models are statistically significant for any of the products ($F$ test, $p > 0.15$).}

5.2. Attitudinal Survey: An Alternative Classification of Prosocial Types

Next, we investigate whether our insights regarding high and low prosocial Consumers’ behaviors discussed in §4.3 are robust to an alternative classification of prosocial types based on Consumers’ responses in the attitudinal survey. This approach is closely related to “psychographic segmentation” in the marketing domain, where customers are segmented based on psychological dimensions “including activities, interests, opinions, needs, values, attitudes, and personality traits” (Wells 1975, p. 197). Psychographic segmentation is often done via attitudinal surveys (e.g., Hofstede et al. 1999, Straughan and Roberts 1999, Dickson 2001). We follow a similar approach to segment Consumers. First, we perform a factor analysis to identify the set of relevant attitudinal questions whose responses are internally consistent and reflect a participant’s prosocial orientation.\footnote{See Appendix O.4 for further details on the factor analysis and the set of relevant questions identified. The Cronbach $\alpha$ measure among the responses for these questions is 0.9.} Second, because the relevant questions contain both binary-choice (Yes or No) and 5-point Likert
scale questions, we code the binary choices into 0/1 scores (1 meaning Yes) and the 5-point scale responses into scores ranging from 0 to 1 with an increment of 0.25. We then take the average score among all relevant questions for a participant as the participant’s final score. Third, we use the median value of the distribution of final scores among all Consumers as the threshold to classify prosocial types: Consumers with a final score higher than or equal to the median are classified as high prosocial (101 out of 196 or 52%), whereas Consumers with a final score lower than the median are classified as low prosocial (95 out of 196 or 48%). The above procedure follows from methods commonly used in attitudinal studies in social science disciplines (e.g., Henry and Sears 2002, Zemack-Rugar et al. 2012).

Based on this alternative classification of high and low prosocial Consumers, we reestimate Equation (1) with the data from these two groups separately (see Table O.8, Appendix O.4). We observe that our insights regarding high and low prosocial Consumers discussed in §4.3 are in general robust to this alternative classification of prosocial types. First, both high and low prosocial Consumers state higher WTP under high transparency than under low transparency. Second, the presence and the effect of reciprocity on both types of Consumers’ WTP remain very similar to the observations summarized in Result 3. The only exception is that high prosocial Consumers exhibit some reciprocity but with a small magnitude that has no significant effect on their WTP.\[^[16]\]

5.3. Testing the Effect of Having a Disadvantaged Party

Finally, we examine whether and how the Worker’s initial endowment affects our results. This variation sheds light on how much the effects of transparency and reciprocity depend on the presence of a disadvantaged third party. To do so, we increase the Worker’s initial endowment from 20 to 80 tokens. This change improves the Worker’s financial position while preserving the order of wealth statuses among the Firm, Consumer, and Worker from our original design. We conducted sessions of this variation for both the Decision and Random conditions under high and low transparency (see Appendix O.5.4 for more details). In this new design, we observe a similar interaction between transparency and reciprocity as before; i.e., reciprocity exists under high transparency but does not exist under low transparency. However, we no longer observe a significant effect of transparency on Consumers’ WTP. Instead, Consumers state similar WTP between high and low transparency in the Decision condition. This loss of significance is due to the fact that under low transparency, Consumers in the Decision condition state a higher WTP in the new design than in our original experiment. To understand why, we examine from the postexperiment survey the extent to which these Consumers exhibit a self-serving bias. We observe

\[^[16]\text{Specifically, the marginal effects of } D \cdot e_k \text{ are smaller than 0.13 at all effort levels for high prosocial Consumers, compared to ranging from 0.10 to 0.29 for low prosocial Consumers.}\]
significantly fewer participants in the new sessions demonstrating a self-serving bias (only 2 out of 15 versus 16 out of 30 in our original experiment; one-sided $\chi^2$ test, $p < 0.05$). This observation confirms our earlier discussion that the presence of a self-serving bias contributes to a lower WTP under low transparency. Nonetheless, this observation hints that the benefit of transparency may diminish as the Worker becomes less disadvantaged. This preliminary result suggests an interesting future research direction to examine the value of transparency while varying the degree to which a third party is disadvantaged.

6. Implications and Conclusions

As the role of social responsibility in business continues to evolve, consumer demands are forcing companies to adapt their operations to meet the needs of a changing marketplace. Companies must address not only how to establish socially responsible practices throughout their supply chains but also how to demonstrate these practices to the public. In this paper, we employ an incentivized controlled laboratory experiment with human subjects to investigate how transparency, indirect reciprocity, and individual prosocial orientation jointly affect consumers’ valuations of a company’s social responsibility practices. To further demonstrate the potential implications of our findings to practice, we complement our experiment with a product choice study and an attitudinal survey. In what follows, we conclude the paper by highlighting two specific insights on how our findings relate to (i) when a company can benefit from improved transparency, and (ii) how a company can better communicate its social responsibility practices to consumers.\(^\text{17}\)

**Insight 1** Consumers value improved transparency regarding a company’s social responsibility practices.

Supply chain transparency is an emerging business challenge for companies. Increasing the transparency of a supply chain can be a costly and time-consuming endeavor (Doorey 2011). However, recent events, such as the Foxconn suicides in Shenzhen in 2010 and the Rana Plaza collapse in Bangladesh in 2013, have highlighted the need for companies to gain more visibility into their supply chains (Barboza 2010, Yardley 2013). As Patagonia’s Director of Environmental Strategy states, “transparency...is really becoming an expectation now. People want to know more about the supply chains making the products they’re buying” (Patagonia 2014). Our findings demonstrate that consumers are willing to pay more for a product when a company improves visibility into the social responsibility practices of its supply chain. This is particularly true if the improvement is from a non-transparent state, as the incremental increase in consumers’ willingness-to-pay diminishes as the level of transparency further improves.

\(^\text{17}\) We remark that our discussions in §6 are meant to demonstrate how our results can be used in practice rather than to offer definitive empirical proof of specific strategies.
Our findings in the Consumer Purchase Game represent an important first step in assessing the revenue benefit of improved transparency when a company communicates its social responsibility practices to consumers. The results from the postexperiment product choice study further demonstrate the generalizability of these findings. For all three product categories (coffee, t-shirt, and laptop computer) and all five social responsibility contexts (treatment of the company’s suppliers’ employees, community development in regions where the company [suppliers] operates, and charitable donations) we study, we observe that a vast majority of participants prefer products with more transparent messaging regarding a company’s social responsibility practices, and a considerable number of them are willing to pay a premium for these products. As companies face difficult decisions on whether to improve the transparency of their supply chains, our results underscore a potential revenue benefit from what is typically considered a costly undertaking.

**Insight 2** *Emphasizing the efforts exerted to improve social responsibility practices in its supply chain is beneficial for a company only when the outcomes of these efforts are transparent and when the company targets low prosocial consumers.*

Our results also highlight the type of information and the level of detail that consumers desire when a company communicates its social responsibility practices. We advocate that a company can benefit by tailoring its communication strategy to the type of consumers the company is targeting. We capture consumers’ heterogeneity by focusing on their prosocial orientations. If a company targets a consumer base with a diverse makeup of prosocial orientations, then it can benefit from emphasizing its social responsibility efforts only if it also communicates the outcomes of these efforts in a transparent fashion. This is because under high transparency, indirect reciprocity exists and has a positive effect on consumers’ willingness-to-pay. If a company targets consumers who naturally care about others’ well-being (i.e., high prosocial consumers), then attention should be placed more on communicating social responsibility outcomes through transparent information and less on communicating the amount of effort exerted. This is because high prosocial consumers are primarily concerned with observing the precise outcomes of a company’s social responsibility practices and do not exhibit reciprocal motives toward the company’s efforts. Finally, if the targeted population is more driven by self-interests (i.e., low prosocial consumers), then it is critical for the company to first ensure that the information being communicated is transparent. Without transparent information, low prosocial consumers are more inclined to either punish low effort levels by a company with a lower willingness-to-pay or justify a low willingness-to-pay by shifting responsibility for the third parties’ well-being onto the company.

In practice, companies can develop knowledge about the heterogeneity of their consumers (e.g., their different prosocial orientations) through psychographic segmentation, a common method used
within marketing to segment consumers based on their personalities and value systems. For example, leading market research firms, such as Cone Communications and Mintel Group, conduct consumer surveys to segment consumers based on their attitudes toward such topics as ethics and the environment, so that companies can then develop better strategies for targeting specific segments (e.g., Cone Communications/Echo 2013, Mintel Group 2015). In our study, we follow a similar approach and utilize an attitudinal survey to better understand our participants’ general opinions toward social responsibility. We show that classifying participants’ prosocial types based on their responses in the attitudinal survey yields similar results as classifying participants based on their decisions in the dictator game. This result further strengthens the relevance of our experimental findings regarding the behaviors of consumers with different prosocial orientations.

To conclude, our research yields valuable insights into the impact of transparency on consumers’ valuations of a company’s social responsibility practices, as well as the behavioral motives behind consumers’ valuations. We hope that these insights will motivate others to further study additional behavioral factors that can drive consumers’ and companies’ decisions in various social responsibility contexts.

Acknowledgments
The authors thank Atalay Atasu, Beatrice Boulu-Reshef, Vishal Gaur, Charles Holt, Elena Katok, Özalp Özer, Duncan Simester, the anonymous Associate Editor, and three anonymous reviewers for their constructive comments. The authors are thankful to the staff at the MIT Behavioral Research Lab and the VeconLab at the University of Virginia for their generous support in conducting the experimental sessions. The discussions during the authors’ presentations at the 2014 Behavioral Operations Conference (Cologne, Germany), 2014 Mostly OM Workshop (Beijing, China), 2015 Behavioral Operations Conference (Ithaca, NY), and research seminars at MIT, Stanford University, University of California, Los Angeles, the University of Texas at Dallas, and the University of Virginia were also beneficial.

Appendix A: Design Choice Justifications
Here we justify a few major choices in the design and implementation of our experiment.

Employing an economic game to operationalize social responsibility: Employing economic experiments with monetary transfers to study social preferences – the underlying motives for a consumer to care about social responsibility – is well established in the economics literature. Researchers have used simple economic games to derive important implications for applications that do not necessarily involve monetary transactions. For example, Ostrom et al. (1992) conclude that self-governance is possible via the study of a common pool resource game; Kosfeld et al. (2009) study the voluntary formation of institutions such as the Kyoto protocol through a public goods game; the ultimatum game (Thaler 1988, Cameron 1999), the dictator game (Hoffman et al. 1994, Dana et al. 2006), and the investment game (Berg et al. 1995, Buchan et al.
2008) have been widely used to study fairness considerations, altruism, and trust. Recently, Bartling et al. (2015) and Pigors and Rockenbach (2015) also use monetary transfers to operationalize socially responsible production in laboratory experiments.

**Providing a background story in the Consumer Purchase Game:** Whether to abstract away all specific contexts or to preserve contextual cues in an experiment is an ongoing debate among experimental economists and psychologists (e.g., Goldstein and Weber 1995, Ortmann and Gigerenzer 2000). Economists typically prefer a context-neutral frame to control for potential “home-grown” biases due to participants' different beliefs and experiences in a particular context (e.g., Hertwig and Ortmann 2001, p. 402). In contrast, psychologists typically believe it is necessary to preserve the natural environment that motivates the research (e.g., Cosmides and Tooby 1996). We choose a middle ground to leverage the merits of both approaches: providing a background story that reflects a supply chain context while using context-neutral terminologies in describing the game dynamics. This approach differentiates our design from other context-neutral economic games (e.g., dictator and ultimatum games) and ensures that we study the manifestation of fundamental preferences (e.g., indirect reciprocity) in the specific context of interest (i.e., social responsibility in the upstream supply chain).

**Different endowments for the three players in the Consumer Purchase Game:** We design the players’ endowments to mimic differences in wealth status among companies, consumers, and workers in a company’s upstream supply chain. Because our focus is on social responsibility in the upstream supply chain, we establish the presence of a disadvantaged party; i.e., the Worker player has a much lower endowment than either the Firm or the Consumer. Such a design is motivated by examples of workers in upstream supply chains (often in developing countries) being paid low wages; e.g., workers in China, Vietnam, and Bangladesh in the electronics and apparel industries. Within the literature, utilizing a zero or lower initial endowment to capture a disadvantaged party is common and often seen in ultimatum and dictator games (see Oosterbeek et al. 2004, Engel 2011 for reviews).

**Providing details of the next task only after the previous task is completed and providing feedback only after all tasks are completed:** We follow this procedure for two reasons. First, it has been documented in studies related to social preferences that a participant’s level of generosity may change in anticipation of future outcomes (Cojoc and Stolan 2014), or it may change if he/she can observe past outcomes (Gneezy et al. 2014). Second, providing feedback during the experiment may induce income effects. That is, knowing how much he/she has earned in earlier tasks (through feedback) may affect a participant’s decisions in later rounds (Merlo and Schotter 2003, Masclet et al. 2015). Both of these aspects can create undesirable order effects in our data. To minimize such order effects, we do not inform the participants of the details of the next task until they have completed the previous one, nor do we provide any feedback to the participants until the very end of the experiment. This approach is commonly used in the experimental literature (e.g., Bolton and Katok 1998, Bolton et al. 2005, Luhan et al. 2009).

**Appendix B: The Multiple Price List Used in the Risk Preference Task**

Table B.1 presents the multiple price list used in the risk preference task, with the last two columns showing the expected payoffs of both lotteries within a pair.
Table B.1 Lotteries Used in the Risk Preference Task

<table>
<thead>
<tr>
<th>Pair number</th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Lottery A’s expected value</th>
<th>Lottery B’s expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(10%, 74; 90%, 60)</td>
<td>(10%, 140; 90%, 6)</td>
<td>61.4</td>
<td>19.4</td>
</tr>
<tr>
<td>2</td>
<td>(20%, 74; 80%, 60)</td>
<td>(20%, 140; 80%, 6)</td>
<td>62.8</td>
<td>32.8</td>
</tr>
<tr>
<td>3</td>
<td>(30%, 74; 70%, 60)</td>
<td>(30%, 140; 70%, 6)</td>
<td>64.2</td>
<td>46.2</td>
</tr>
<tr>
<td>4</td>
<td>(40%, 74; 60%, 60)</td>
<td>(40%, 140; 60%, 6)</td>
<td>65.6</td>
<td>59.6</td>
</tr>
<tr>
<td>5</td>
<td>(50%, 74; 50%, 60)</td>
<td>(50%, 140; 50%, 6)</td>
<td>67.0</td>
<td>73.0</td>
</tr>
<tr>
<td>6</td>
<td>(60%, 74; 40%, 60)</td>
<td>(60%, 140; 40%, 6)</td>
<td>68.4</td>
<td>86.4</td>
</tr>
<tr>
<td>7</td>
<td>(70%, 74; 30%, 60)</td>
<td>(70%, 140; 30%, 6)</td>
<td>69.8</td>
<td>99.8</td>
</tr>
<tr>
<td>8</td>
<td>(80%, 74; 20%, 60)</td>
<td>(80%, 140; 20%, 6)</td>
<td>71.2</td>
<td>113.2</td>
</tr>
<tr>
<td>9</td>
<td>(90%, 74; 10%, 60)</td>
<td>(90%, 140; 10%, 6)</td>
<td>72.6</td>
<td>126.6</td>
</tr>
<tr>
<td>10</td>
<td>(100%, 74; 0%, 60)</td>
<td>(100%, 140; 0%, 6)</td>
<td>74</td>
<td>140</td>
</tr>
</tbody>
</table>

Note. The notation (x%, a; y%, b) represents a lottery consisting of an x% chance of earning a tokens and a y% chance of earning b tokens. Participants were not presented with the expected values of each lottery.

Appendix C: Regression Results for the Pooled Data

Table C.2 presents the regression results for Equation (1) using data from all Consumer participants.

Table C.2 Regression Results for Equation (1)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.13 (8.15)</td>
</tr>
<tr>
<td>$T_M$</td>
<td>-0.72 (12.04)</td>
</tr>
<tr>
<td>$T_H$</td>
<td>-1.68 (11.72)</td>
</tr>
<tr>
<td>$D$</td>
<td>-14.15 (12.58)</td>
</tr>
<tr>
<td>$D \cdot T_M$</td>
<td>6.88 (18.17)</td>
</tr>
<tr>
<td>$D \cdot T_H$</td>
<td>21.31 (18.21)</td>
</tr>
<tr>
<td>$e_k$</td>
<td>0.34 (0.03)**</td>
</tr>
<tr>
<td>$T_M \cdot e_k$</td>
<td>0.06 (0.05)</td>
</tr>
<tr>
<td>$T_H \cdot e_k$</td>
<td>-0.06 (0.05)*</td>
</tr>
<tr>
<td>$D \cdot e_k$</td>
<td>0.05 (0.05)</td>
</tr>
<tr>
<td>$D \cdot T_M \cdot e_k$</td>
<td>0.04 (0.08)</td>
</tr>
<tr>
<td>$D \cdot T_H \cdot e_k$</td>
<td>0.10 (0.08)</td>
</tr>
<tr>
<td>$\sigma_{\delta}$</td>
<td>48.10 (3.12)**</td>
</tr>
<tr>
<td>$\sigma_{\epsilon}$</td>
<td>19.43 (0.48)**</td>
</tr>
</tbody>
</table>

Note. Values in parentheses are the standard errors. ***: $p < 0.01$; *: $p < 0.1$; $p$ values are derived from one-sided t tests.

Appendix D: Regression Results and Marginal Effects by Prosocial Type

In this section, we present the regression results from Equation (1) and the corresponding marginal effects by Consumers’ prosocial orientations. To perform our analysis, we first reestimate Equation (1) for high and low prosocial Consumers separately. These regression results are presented in Table D.3. We then compute the marginal effects of $T_H$, $T_M$, and their differences, $\Delta_T$, and $D \cdot e_k$ for high and low prosocial Consumers. We present these results in Tables D.4 and D.5.
Table D.3  Regression Results for Equation (1), by Prosocial Type

<table>
<thead>
<tr>
<th>Variable</th>
<th>High prosocial</th>
<th>Low prosocial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (s.e.)</td>
<td>Estimate (s.e.)</td>
</tr>
<tr>
<td>Intercept</td>
<td>16.66 (8.02)**</td>
<td>-7.84 (11.19)</td>
</tr>
<tr>
<td>$T_M$</td>
<td>-4.81 (10.82)</td>
<td>-2.08 (17.76)</td>
</tr>
<tr>
<td>$T_H$</td>
<td>10.94 (11.90)</td>
<td>-6.30 (15.87)</td>
</tr>
<tr>
<td>$D$</td>
<td>-5.80 (11.02)</td>
<td>-27.90 (19.30)*</td>
</tr>
<tr>
<td>$D \cdot T_M$</td>
<td>19.01 (15.67)</td>
<td>5.15 (27.69)</td>
</tr>
<tr>
<td>$D \cdot T_H$</td>
<td>10.39 (11.65)</td>
<td>30.97 (26.61)</td>
</tr>
<tr>
<td>$e_k$</td>
<td>0.47 (0.05)**</td>
<td>0.26 (0.04)**</td>
</tr>
<tr>
<td>$T_M \cdot e_k$</td>
<td>0.09 (0.07)</td>
<td>-0.06 (0.07)</td>
</tr>
<tr>
<td>$T_H \cdot e_k$</td>
<td>-0.09 (0.08)</td>
<td>-0.04 (0.06)</td>
</tr>
<tr>
<td>$D \cdot e_k$</td>
<td>-0.02 (0.07)</td>
<td>0.01 (0.08)</td>
</tr>
<tr>
<td>$D \cdot T_M \cdot e_k$</td>
<td>21.56 (2.32)**</td>
<td>54.73 (4.91)**</td>
</tr>
<tr>
<td>$D \cdot T_H \cdot e_k$</td>
<td>-0.09 (0.10)</td>
<td>0.18 (0.11)**</td>
</tr>
<tr>
<td>$\sigma_e$</td>
<td>18.64 (0.69)**</td>
<td>19.27 (0.64)**</td>
</tr>
</tbody>
</table>

Note. Values in parentheses are the standard errors.

$^*$: $p < 0.01$; $^*$: $p < 0.05$; $^*$: $p < 0.1$; $p$ values are derived from one-sided $t$ tests.

Table D.4  Marginal Effects of $T_H$, $T_M$, and Their Differences, $\Delta_T$, by Prosocial Type

<table>
<thead>
<tr>
<th>Effort</th>
<th>$T_H$</th>
<th>$T_M$</th>
<th>$\Delta_T$</th>
<th>$T_H$</th>
<th>$T_M$</th>
<th>$\Delta_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>0</td>
<td>16.04 (9.20)**</td>
<td>10.19 (8.23)</td>
<td>5.85 (10.30)</td>
<td>8.32 (7.26)</td>
<td>0.84 (5.80)</td>
<td>7.48 (7.16)</td>
</tr>
<tr>
<td>20</td>
<td>17.52 (9.59)**</td>
<td>11.55 (8.76)*</td>
<td>5.78 (10.58)</td>
<td>10.46 (7.99)*</td>
<td>2.44 (6.60)</td>
<td>8.01 (8.11)</td>
</tr>
<tr>
<td>40</td>
<td>18.12 (9.88)**</td>
<td>12.62 (9.17)*</td>
<td>5.50 (10.75)</td>
<td>12.85 (8.73)*</td>
<td>4.49 (7.46)</td>
<td>8.36 (9.09)</td>
</tr>
<tr>
<td>60</td>
<td>18.45 (10.12)**</td>
<td>13.39 (9.50)*</td>
<td>5.06 (10.88)</td>
<td>15.47 (9.49)*</td>
<td>6.99 (8.37)</td>
<td>8.48 (10.08)</td>
</tr>
<tr>
<td>80</td>
<td>18.36 (10.36)**</td>
<td>13.85 (9.81)*</td>
<td>4.51 (11.00)</td>
<td>18.28 (10.27)**</td>
<td>9.92 (9.32)</td>
<td>8.36 (11.07)</td>
</tr>
<tr>
<td>100</td>
<td>17.89 (10.58)**</td>
<td>14.01 (10.11)*</td>
<td>3.88 (11.10)</td>
<td>21.23 (11.05)**</td>
<td>13.26 (10.29)*</td>
<td>7.97 (12.01)</td>
</tr>
<tr>
<td>120</td>
<td>17.05 (10.75)*</td>
<td>13.86 (10.36)*</td>
<td>3.19 (11.13)</td>
<td>24.27 (11.83)**</td>
<td>16.92 (11.27)*</td>
<td>7.34 (12.90)</td>
</tr>
</tbody>
</table>

Note. Values shown are the marginal effects evaluated at the Decision condition; i.e., at $D = 1$ in Equation (1). Values in parentheses are the standard errors.

$^*$: $p < 0.05$; $^*$: $p < 0.1$; $p$ values are derived from one-sided $t$ tests.

Table D.5  Marginal Effects of the Interaction Term $D \cdot e_k$, by Prosocial Type

<table>
<thead>
<tr>
<th>Effort</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td>Low</td>
<td></td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>0</td>
<td>0.05 (0.06)</td>
<td>-0.01 (0.06)</td>
<td>-0.05 (0.06)</td>
<td>0.09 (0.04)**</td>
<td>0.06 (0.05)</td>
<td>-0.04 (0.03)</td>
</tr>
<tr>
<td>20</td>
<td>0.05 (0.07)</td>
<td>-0.04 (0.07)</td>
<td>-0.05 (0.06)</td>
<td>0.10 (0.05)**</td>
<td>0.08 (0.05)*</td>
<td>-0.04 (0.04)</td>
</tr>
<tr>
<td>40</td>
<td>0.05 (0.08)</td>
<td>-0.07 (0.07)</td>
<td>-0.04 (0.07)</td>
<td>0.12 (0.05)**</td>
<td>0.11 (0.05)**</td>
<td>-0.04 (0.04)</td>
</tr>
<tr>
<td>60</td>
<td>0.04 (0.08)</td>
<td>-0.09 (0.07)*</td>
<td>-0.04 (0.07)</td>
<td>0.13 (0.05)**</td>
<td>0.13 (0.06)**</td>
<td>-0.04 (0.04)</td>
</tr>
<tr>
<td>80</td>
<td>0.04 (0.08)</td>
<td>-0.10 (0.07)*</td>
<td>-0.03 (0.07)</td>
<td>0.14 (0.05)**</td>
<td>0.16 (0.06)**</td>
<td>-0.04 (0.05)</td>
</tr>
<tr>
<td>100</td>
<td>0.03 (0.07)</td>
<td>-0.10 (0.07)*</td>
<td>-0.02 (0.07)</td>
<td>0.15 (0.05)**</td>
<td>0.18 (0.06)**</td>
<td>-0.04 (0.05)</td>
</tr>
<tr>
<td>120</td>
<td>0.02 (0.07)</td>
<td>-0.10 (0.06)*</td>
<td>-0.01 (0.06)</td>
<td>0.16 (0.05)**</td>
<td>0.20 (0.06)**</td>
<td>-0.03 (0.05)</td>
</tr>
</tbody>
</table>

Note. Values in parentheses are the standard errors.

$^*$: $p < 0.01$; $^*$: $p < 0.05$; $^*$: $p < 0.1$; $p$ values are derived from one-sided $t$ tests.
References


