The Differential Effects of CSR and CSI on Consumer Willingness to Pay: Implications for Service Providers and Retailers

By Sabine Benoit*, Julia Hartmann, Christina Sichtmann, Martin Wetzels

Service providers and retailers reselling branded have the discretion to set and adapt prices according to customers' willingness to pay (WTP). Research often notes markup effects, such that WTP increases in response to corporate social responsibility (CSR) and markdown effects, lowering their WTP for corporate social *ir*responsibility (CSI). Theory suggests attitude changes to (negative) CSI are stronger than to (positive) CSR, but the extent and whether this difference holds for WTP and across various product types are unknown. Using experimental data, an incentive-compatible measure, and an actual purchase, this article reports on three studies that show that consumers mark up WTP for CSR and mark down WTP for CSI. The differential effects arise across brands; compared with WTP for a competitor brand, the acceptable price of a focal CSR/CSI brand is marked down more than it is marked up. Comparing the WTP for a focal brand relative to the average CSR performance of that brand does not produce any within-brand differential effects The evidence also indicates a product type effect: Consumer WTP adaptation for CSR or CSI is stronger for utilitarian than for hedonic products. These findings have implications for service providers, retailers and manufacturing firms, as well as for further research.

INTRODUCTION

Corporate social responsibility (CSR) is of utmost strategic importance to appease and attract stakeholders (Kraus et al. 2021) and has become key priority in various sectors (e.g., Wei et al. 2018, Diallo et al. 2021). For instance, firms seek certification to ensure their products are being manufactured or provided in fair conditions (Smithers 2013), and service providers and retailers prefer products from responsible companies in their efforts to enhance customer attitudes such as satisfaction, loyalty or purchase intention (Bolton and Mattila 2015, Wei et al. 2018, Yeo et al. 2018). This is because consumers exhibit higher willingness to pay (WTP) for products produced in responsible manufacturing conditions (see Tully and Winer 2014 and Appendix A). As such it should improve service providers' or retailers' economic returns reselling those products (e.g., a Twinings tea in a restaurant or grocery store), though no prior research has established a reliable range of expected price premiums.

These insights give rise to two important questions. First, how do consumers react to both responsible *and* irresponsible behaviors in the supply chain? Providers must make assortment and pricing choices before, during, and after media reports of

Sabine Benoit is Professor of Marketing at Surrey Business School, University of Surrey, GU2 7XH Guildford, United Kingdom and Research School of Management at the ANU - Australian National University, Canberra ACT 0200, Australia, Email: S.Benoit@surrey.ac.uk
*Corresponding author

Julia Hartmann is Professor of Sustainability Management at EBS University for Business and Law, Burgstrasse 5, 65375 Oestrich-Winkel, Germany, E-Mail: Julia.hartmann@ebs.edu

Christina Sichtmann is Associate Professor of Marketing at the Chair of International Marketing, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria and Research Associate at Berner Fachhochschule, Brückenstraße 73, 2012 Bern, Switzerland, E-Mail: christina.sichtmann@univie.ac.at

Martin Wetzels is Professor of Marketing at EDHEC Business School, Lille Campus, 24, Avenue Gustave Delory CS 50411, 59057 Roubaix Cedex 1, France, Email: martin.wetzels@edhec.edu

scandals or brand-related CSR communication. If consumers are willing to pay higher prices for CSR, they similarly might reduce their WTP when confronted with evidence of corporate social irresponsibility (CSI). Studying both sides of the equation is of particular importance, because we anticipate substantial differences in consumers' interpretations of and reactions to negative CSI as opposed to positive CSR cues (Lange and Washburn 2012; Sen and Battacharya 2001), just as there are for information in general (Ludwig et al. 2013). We predict and test for differential effects, signaling a greater markdown for CSI than markup for CSR, both for a single brand (CSI-/CSR-related WTP for a focal brand versus WTP for the average CSR performance of that brand) and across brands (CSI-/CSR-related WTP for a focal brand versus WTP for a competitor brand with no CSR rating). Understanding these differential effects can inform service providers' and retailers' assortment and pricing decisions. It is also in line with scholars emphasizing that research around sustainability and responsibility in service settings should be a primary focus (Benkenstein et al. 2017).

Second, do different product types benefit (or suffer) equally from CSR or CSI? Companies adopt distinct marketing strategies for sustainable products across categories, such as including CSR claims solely in marketing communication (soft drinks) or adding them to product packaging too (coffee) (Nielsen 2015). Thus a product type effect might explain the varying price premiums for CSR that appear in prior literature (e.g., Tully and Winer 2014). For example, consumers might emphasize more CSR when they buy a utilitarian item like mineral water rather than a more experiential, hedonic product like pralines, because they focus on the different product attributes, including CSR (Dhar and Wertenbroch 2000). Therefore, considering the potentially differential effects of product types on consumers' WTP for products associated with CSR/CSI is of substantial interest for managers, because such information could help them adjust their price levels and margins appropriately (Gauri 2013). The effects of CSR efforts on company performance thus far have remained somewhat uncertain (Katsikeas, Leonidou and Zeriti 2016; Leonidou, Katsikeas, and Morgan 2013), and such information also could enable managers to justify their CSR investments (Lacey, Kennett-Hensel and Manolis 2015).

In response, we collect data about the outcomes of CSR and CSI in three experimental studies, in which we seek to determine the differential effects of WTP for CSR or CSI (Studies 1 and 2) and investigate the influences of product type (Study 3). In these vignette-based experiments, we use Becker, DeGroot and Marschak's (1964; abbreviated as BDM) lottery approach to measure WTP that requires actual buying behavior. This incentive-compatible method partly overcomes the limited capacity of hypothetical measures of WTP to reveal consumers' true values (Auger et al. 2003; Follows and Jobber 2000). To simulate point-of-sale situations, when consumers choose among alternative products, we introduce a BDM variant that explicitly accounts for competition across brands. With these studies, this article accordingly makes three main contributions to service, retail, and marketing literature.

First, prior research is *skewed towards CSR* and tended to investigate both *concepts in isolation*. More precicely, prior research focuses on understanding the effects of responsible firm behavior (CSR) on WTP (markup effect), without really addressing the consequences of irresponsible behavior (CSI) (markdown effect; Figure 1). In Studies 1 and 2, we investigate CSI markdown effects explicitly. And then even when previous studies include corporate irresponsibility, CSR and CSI have been studied in isolation. To the best of our knowledge, this study is the first to include both constructs in an incentive-compatible design, such that we can examine the *differential effects* on WTP within and across brands. With this approach, we gain new insights into consumers' reactions to CSR/CSI at the point of purchase.

Second, we detail a *product type effect* too. Service providers and retailers need to know which product categories in their assortment are likely to lead to the most pronounced WTP reactions, particularly because the assortment composition determines the firm's market position (Mantrala et al. 2009). Modern firms also are shifting toward more fine-tuned pricing strategies, such as dynamic pricing, which can account for external events (Grewal et al. 2011). Abundant research on WTP for CSR has produced inconsistent findings regarding various levels of price premiums (see Figure 1 and Appendix A; see also Tully and Winer 2014 or Wei et al. 2018). There may be several explanations for these inconsistencies, but one possibility is insufficient systematic attention to the different product categories (Tully and Winer 2014).

Third, our use of the BDM lottery approach (Becker et al. 1964) to measure WTP in a more realistic way represents another contribution to extant literature. For example, in Study 2, to account for the across-brand differential effect, we investigate consumers' WTP reactions when they have a choice between a brand that provides CSR/CSI information, as opposed to a brand for which no CSR/CSI information is available.

LITERATURE REVIEW

Criteria and Scope

With a meta-analysis, Tully and Winer (2014) reveal that people are willing to pay a positive, significant premium for socially responsible products (on average, 16.8%). They also find that most studies do not include incentive-compatible measurements for WTP and only investigate a single product category. In our own literature review, we also note an imbalance between research devoted to CSR versus CSI and a lack of comparisons of CSR and CSI, especially across different product categories (Figure 1, Appendix A). Several studies examine the link between WTP and either CSR or CSI separately, as we

detail in the next section. In Figure 1, we thus categorize prior studies according to two sets of criteria: whether the product or brand attributes investigated are responsible only or also irresponsible, and whether the method used to measure WTP is hypothetical (i.e., survey based), incentive compatible or based on purchasing data. Appendix A also indicates that of 36 studies identified, only 4 investigate WTP for two different product categories (utilitarian/hedonic), but they consistently use hypothetical WTP measures and do not focus on differences across product categories.

In our review, we only include studies that measure WTP and exclude those with different CSR outcome variables, such as profits (Leonidou, Katsikeas and Morgan 2013), customer relationship strength, word of mouth (Lacey, Kennett-Hensel and Manolis 2015), purchasing intentions (Auger et al. 2003; Auger et al. 2008; Auger et al. 2010; Sen and Battacharya 2001; Battacharya and Sen 2004), attitudes (Folkes and Kamins 1999; Madrigal and Boush 2008), or willingness to punish through activities like negative word of mouth (Lacey, Kennett-Hensel and Manolis 2015; Sweetin et al. 2013). Furthermore, we exclude studies that investigate constructs related to but distinct from CSR/CSI, such as organic or healthy products (e.g., van Doorn and Verhoef 2011). It would be beyond the scope of a single research project to include all such variables.

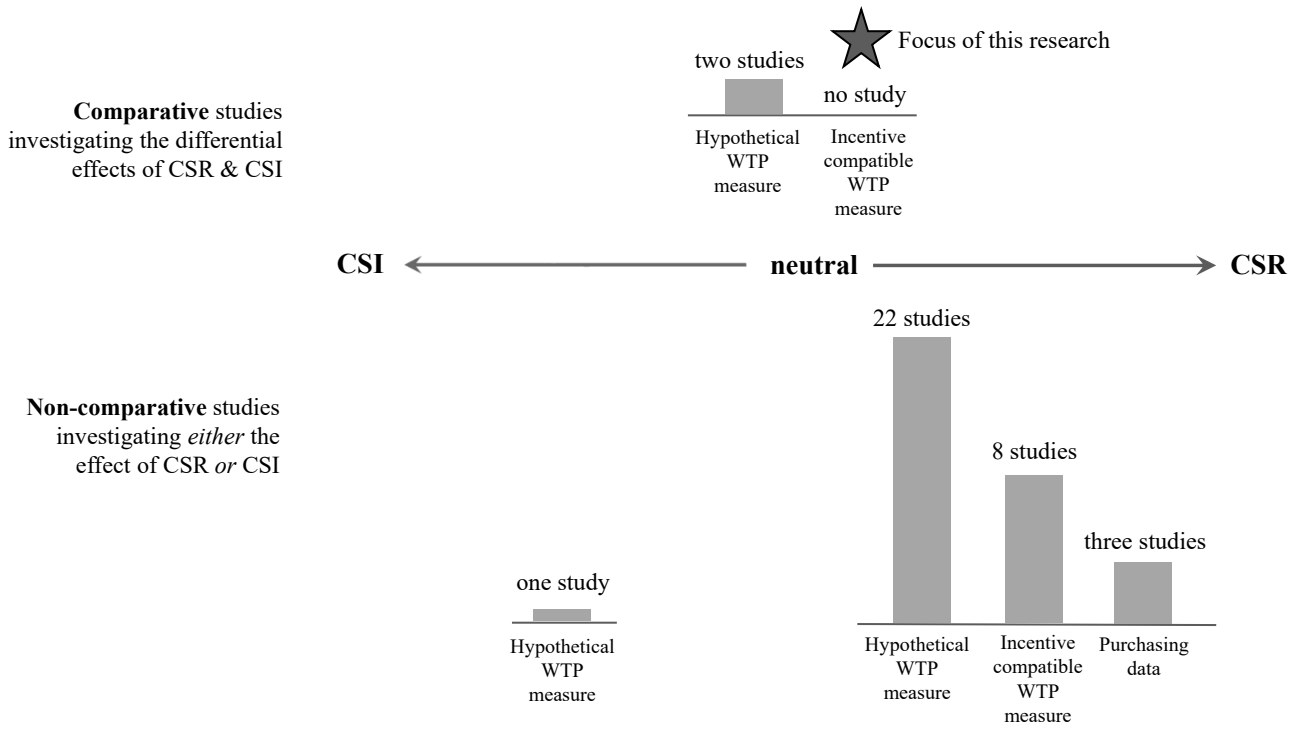

Figure 1: Results from previous literature

CSR and CSI

As Figure 1 shows, existing literature focuses on CSR rather than CSI, that is, on the positive effects of responsibility rather than the negative ones of a lack of responsibility. We believe that the two constructs need to be differentiated (Lange and Washburn 2012; Moosmayer 2012). For the purpose of our study, we define CSR as the strategies and practices of a company intended to advance environmental, social, or economic conditions without compromising its competitiveness (Aguinis and Glavas 2012; Porter and Kramer 2006). This definition implies proactive, responsible engagement by a company that takes CSR explicitly into account when making strategic decisions (Lin-Hi and Müller 2013).

The mere avoidance of harmful strategies and practices is not indicative of a company's CSR. Instead, avoidance behaviors are taken for granted and constitute what any company would be expected to do (Lin-Hi and Müller 2013). By analogy, our definition of corporate social irresponsibility (CSI) applies to firms that implement strategies and practices that potentially harm environmental, social, and economic conditions (Lin-Hi and Müller 2013; Murphy and Schlegelmilch 2013), such that they actively fail to act responsibly (Lange and Washburn 2012).

By including social (e.g., working conditions, fair remuneration) and environmental (e.g., low emissions, reduced waste, minimal resource consumption) practices for product design and manufacturing, our definitions also require that the products are manufactured by firms that perform well on the responsibility dimension. This understanding of CSR as involving all three elements of the triple bottom line (social, environmental, and economic) corresponds with extant views on sustainability (e.g., Connelly, Ketchen and Slater 2011; Hartmann and

Moeller 2014). We choose to employ the term CSR though, to link our work explicitly to the existing stream of literature on CSR and WTP.

For service providers and retailers this is relevant since we anticipate that (ir)responsible strategies and practices by manufacturers and their supply chains translate into product attributes that affect consumer perceptions at the point of sale. In keeping with equity theory (Walster, Walster and Berscheid 1978), these perceptions should influence the value consumers attribute to the product and thereby determine the price they are willing to pay for it (Auger et al. 2008). As Figure 1 reveals, most studies investigate responsible product attributes; only 2 of the 36 studies investigate the effects of both responsible and irresponsible attributes on WTP in a comparative way (Moosmayer 2012; Trudel and Cotte 2009; see Appendix A for details). Most studies conclude that responsible product attributes lead to higher WTP, and the few studies that also consider irresponsible product attributes find links to lower WTP (Jin, Smith and Cook 2012; Trudel and Cotte 2009). However, precisely how irresponsible product attributes determine consumer WTP continues to be poorly understood, which motivates our simultaneous examination of both responsible and irresponsible product attributes and WTP, in terms of markup, markdown, and differential effects. This is done in the hope that we can help explain the widely varying WTP increases (price premiums) noted in prior studies and give managers insights for price setting and assortment designs.

Product Type: Hedonic versus Utilitarian

Tully and Winer (2014) identify some evidence of product type effects in the relationship between CSR and WTP. In our effort to categorize existing literature according to product category, as well as our attempt to use priming tactics, we rely on the hedonic versus utilitarian product classification, which is a well-established categorization. Utilitarian and hedonic products elicit different consumption patterns (e.g., Dhar and Wertenbroch 2000), such that through the consumption of *hedonic* products, consumers primarily pursue an affective or sensory experience of aesthetic or sensual pleasure, fantasy, and fun (Hirschman and Holbrook 1982). The consumption of *utilitarian* products instead is cognitively driven, instrumental, and goal oriented, designed to accomplish a functional or practical task (Dhar and Wertenbroch 2000). Accordingly, classical consumer-related outcome variables, such as satisfaction, word of mouth, loyalty, patronage (Jones, Reynolds and Arnold 2006), provider preferences (Overby and Lee 2006), and value perceptions (Leroi-Werelds et al. 2014), differ by these product types. Because this evidence suggests that consumers react differently when purchasing hedonic versus utilitarian products, this variable must be accounted for in discussions of their WTP.

In Appendix A, we classify existing research on the impact of CSR/CSI on WTP by product category and find an approximately equal split between studies that feature utilitarian or hedonic products. Despite the importance of a comparison (Sen and Battacharya 2001), only three WTP studies examine the effect of products that we would consider representative of different product type categories (Auger and Devinney 2007; Moosmayer 2012; Thompson et al. 2010). Of these studies, none use incentive-compatible measures of WTP; for example, Auger and Devinney (2007) convert the probability of consideration, measured by a choice model, to estimate WTP. Moosmayer) and Auger and Devinney (2007) measure both product types with a single sample. Moosmayer (2012) manipulates social (ir)responsibility for one product type and ecological for the other, which limits the comparability of the product type effects. Finally, Thompson et al. (2010) take a value-added/non–value-added perspective and indicate a reward effect for the non–value-added (utilitarian) product category (plywood) but not for the value-added (hedonic) one (furniture).

Thus, though recent literature suggests a potential product type effect in the link between CSR and WTP (Tully and Winer 2014), no study has offered clear evidence. Our analysis offers some evidence that consumer reactions might be more pronounced for utilitarian products than hedonic ones. The single study that adopts BDM to assess what we consider a hedonic product (chocolate) revealed a price premium of around 30% (Didier and Lucie 2008; Appendix A). Other studies that use BDM to investigate the CSR–WTP link for a utilitarian product (orange juice) indicate price premiums of more than 50% (Bougherara and Combris 2009; Rode, Hogarth and Le Menestrel 2008; Appendix A). A few studies that include incentive-compatible measures for utilitarian products (tomatoes, cotton shirts) find positive price premiums between 11% and 28% (Alphonce and Alfnes 2012; Ellis, McCracken, and Skuza 2012; Appendix A).

WTP Measurement

In Figure 1, we also categorize existing literature according to the WTP measurement approach used in these studies (Appendix A). That is, WTP can be measured with real-life purchasing data collected at the point of sale, with incentive-compatible methods, or with hypothetical self-reports and surveys (Miller et al. 2011). Most studies use hypothetical methods using surveys to measure consumers' self-reported WTP. For example, contingent valuation approaches require respondents to state their WTP directly or indicate repeatedly whether they would buy a certain product at a given price. Conjoint analysis instead is designed to identify trade-offs across product attributes, such as price, quality, and related services (Backhaus et al. 2005), from which researchers can infer WTP.

The external validity of hypothetical and non-incentive compatible approaches is limited though, because consumers have little incentive to reveal their true WTP (Wertenbroch and Skiera 2002) and tend to over- or underestimate their real WTP, in accordance with a social desirability bias (Auger et al. 2003; Follows and Jobber 2000). Carrigan and Attalla (2001) postulate the "myth of the ethical consumer" who expresses higher (lower) attitudes toward products associated with CSR (CSI)

but does not behave accordingly. Because hypothetical measurements thus can result in inaccurate information (Miller et al. 2011), they may lead to improper conclusions. For example, Schreier and Werfer (2007) show that hypothetical measures produce WTP values three times higher than the actual price paid at the point of purchase. Therefore, it is understandable why Lacey, Kennett-Hensel, and Manolis (2015) conclude that few "real-world" CSR studies exist.

In response, and in an effort to support managers' profit-relevant decisions about consumers' "true" WTP for CSR, our study mimics a point of purchase and examines WTP derived from actual transactions with a BDM lottery (Becker, DeGroot, and Marschak 1964). Such techniques ensure incentive compatibility, because respondents must buy the product after bidding for it. Unlike methods based on self-reports, respondents have an incentive to state their true WTP (Wertenbroch and Skiera 2002), which reduces social desirability biases. Few studies to date have used incentive-compatible methods to measure a consumer's WTP for CSR attributes (Appendix A); to the best of our knowledge, none of them compares CSR with CSI.

In summary, this literature review maps prior research about the relationship of CSR/CSI with WTP. Figure 1 shows how it skews toward responsibility, as opposed to irresponsibility, and toward hypothetical WTP measures, rather than incentive-compatible ones. With three empirical studies, we aim to fill some of these gaps by comparing and contrasting WTP for products from firms that produce those products in responsible versus irresponsible conditions and by measuring WTP with an incentive-compatible method. Furthermore, we compare utilitarian versus hedonic products, in terms of their potential impact on the strength of consumers' reactions.

STUDY 1: WITHIN-BRAND DIFFERENTIAL EFFECT OF CSR/CSI ON WTP

Theory and Hypotheses

According to equity theory (Adams 1965; Walster, Walster and Berscheid 1978), consumers strive to reach equity in exchanges by balancing inputs and outputs. In a buying situation, the main input from the company is the product or service, and the consumer's input is the price paid. In keeping with prior research (e.g., Homburg, Koschate and Hoyer 2005), we assume that positive product features in general, and CSR in particular, lead consumers to perceive higher inputs from the company. Analogously, CSI should prompt perceptions of lower inputs from the company. To balance the input levels in this exchange, consumers should adjust their WTP at the point of sale, which will directly affect firms.

It is relevant to know not only how consumers react to responsible versus irresponsible firm behavior but also how they react relative to average CSR performance. That is, consumers may infer that the manufacturers of certain products perform neither particularly poorly nor particularly well, compared to others in the same industry. This average CSR rating offers a baseline expectation among consumers who buy the firm's products (i.e., average CSR as a hygiene factor; Lacey, Kennett-Hensel and Manolis 2015) and thus should not influence the product's value in the eyes of the consumer.

H1a: Consumers' WTP is higher for products by firms with positive CSR ratings as opposed to firms with average CSR ratings (markup effect).

H1b: Consumers' WTP is lower for products by firms with negative CSR ratings (CSI) as opposed to firms with average CSR ratings (markdown effect).

Prior literature has identified the "negativity bias" implying that humans react more strongly towards negative as opposed to positive information when making decisions (e.g., Lange and Washburn 2012; Moosmayer 2012). This negativity bias goes back to prospect theory (Kahneman and Tversky 1979) and its suggestions that consumers simplify their choices by comparing options against a status quo and thereby identifying them as either a positive shift (gain) or a negative shift (loss). The principle of loss aversion, or what we call the differential effect of negative events, implies that people perceive negative shifts as more severe than positive shifts and react more intensely to them (Kahneman and Tversky 1979), which then determines their perception of the equity in the exchange with a company.

Past research investigating the above in a CSR/CSI context shows that consumers react more strongly to irresponsible than to responsible firm behavior (e.g., Lange and Washburn 2012; Sweetin et al. 2013; Sen and Battacharya 2001). In their vignette-based study, Sen and Battacharya (2001) describe a company with a positive or negative CSR record, then ask respondents to evaluate that company. Considering only one brand in isolation, these consumers reacted with particular sensitivity to a negative CSR record (what we call CSI) and provided significantly lower company ratings, but not all consumers in that study offered enhanced company ratings if they read the positive CSR scenario. Considering these arguments and evidence, we formulate the following hypothesis:

H2: The increase in WTP for products by firms with a positive CSR rating is smaller than the decrease in WTP for products by firms with negative CSR ratings (CSI); that is, the markup effect of CSR is not as great as the markdown effect for CSI (within-brand differential effect).

Stimuli and Procedure

To assess H1 and H2, we employed a vignette-based experiment, followed by an incentive-compatible BDM lottery and a survey. The data collection procedure included three steps. First, participants read one of three vignettes, mimicking a newspaper article that contained information about the quality and CSR/CSI performance of a fictitious brand, "Granula." Similar to Luchs et al. (2010), we credited a hypothetical, independent consumer agency, the "Institute for Consumer Infor-

mation," for having published this evaluation. The quality of the product, described as good, stayed constant across manipulations and experimental groups. Although the same general wording provided the CSR/CSI information in each case, the manipulations featured e.g. different adjectives and verbs to identify CSR, average, and CSI firm behaviors (see Appendix B). We submitted the vignettes to a pretest with a student sample, which indicated that the CSI and average manipulations were not perceived as sufficiently different from each other. Therefore, we revised the newspaper article and expanded the section describing responsible, irresponsible, and average firm behavior to make sure the differences across groups were apparent.

In the revised versions, respondents in the experimental condition learned that the Institute for Consumer Information had rated companies according to their social and ecological responsibility and that Granula received a below-average value of 14 (out of 100) on the index, due to its very poor social and ecological performance (CSI manipulation); a fairly average rating of 62 (average manipulation); or an above-average score of 92 (CSR manipulation). They also learned that Granula had been criticized (CSI), mentioned (average), or praised (CSR) for its high (CSI), medium (average), or low (CSR) employee turnover, indicative of bad (CSI), normal (average), or good (CSR) working conditions. In addition, the description indicated that Granula did not exceed the minimum (CSI), did the average, or exceeded (CSR) the minimum requirements for energy consumption and waste disposal. Finally, Granula was implementing very short-term (CSI), mid-term (average), or long-term (CSR) contracts with bad (CSI), normal (average), or good (CSR) payment conditions for suppliers. In the bad payment conditions, suppliers were not able to plan ahead and had to wait a long time for reimbursement (CSI), whereas in the good payment conditions, they were able to plan and did not have to wait (CSR). Detailed information about the different manipulations appears in Appendix B. To avoid confounding factors, we chose CSR/CSI behaviors that did not directly affect product quality.

Second, after having read the newspaper article, respondents took part in a BDM lottery (Wertenbroch and Skiera 2002) so we could measure their WTP for the product. For the BDM lottery procedure, mimicking a buying situation, we comprehensively explained the process in person and informed participants that they might have to buy the product. These participants received a real product to evaluate (fruit drops), in packaging that clearly featured the brand name Granula (Appendix C). Then, participants were asked to state the maximum price they would pay for the product offered by Granula, the company they had just learned about in the newspaper article. A second price, called the "market price," was randomly drawn from an urn by the participant. If the stated WTP was higher than the market price, the participant would have been willing to pay a higher price than the actual market price. In this situation, the participant must buy the product and pay the market price. If the stated WTP was lower than the market price drawn from the urn, the product was considered too expensive, and the participant was not allowed to buy it. Thus, if participants overstated their true WTP, they increased their chances of winning the lottery but might have been forced to pay a higher price than they would have in a real buying situation. If participants understated their true WTP, they risked not getting the product for a price that they were willing to spend, and thus losing the lottery. This methodology ensures that participants have an incentive to reveal their true WTP (Wertenbroch and Skiera 2002). To check that the WTP measurement was valid and reliable, the respondents also indicated, after the procedure, whether the rules had been clear to them (see Appendix F).

Third, in the final step participants completed a questionnaire, which we used to ensure the quality of the data collection, including a credibility check for the manipulation, manipulation checks, self-reported demographic information, checks on the clarity of the BDM lottery or strategic bidding behavior, and face validity checks (see Appendix F). We deliberately collected these responses after the bid, so that participants could not reread the vignette or revise their bids. Appendix B details the experimental manipulations.

Sample and Quality Assessment

The research design is complex, requiring substantial time of respondents and necessary skills and expertise of interviewers, who must precisely follow the different steps of the experiment and correctly execute the BDM lottery. Because it also involves the actual sale of products, the data must be collected in time-consuming, rather costly, face-to-face interviews (Bougherara and Combris 2009). Therefore, we invested considerable resources to collect a unique field data set among consumers recruited in a pedestrian zone in a major German city. A professional market research firm incentivized the participants by offering a voucher for products (e.g., food, hygiene items), equivalent in value to about 3–5 Euros, that could be selected right after their successful participation in the study. Before starting the three-step study procedure, participants also received 2 Euros in cash, which they were asked to put into their wallets, with the explanation that the money could be used in the experiment, but if participants decided not to make use of those funds, they were allowed to keep them. This method mimics previous studies that rely on experimental auctions (e.g., De-Magistris, Del Giudice and Verneau 2015; Hustvedt and Bernard 2010; Pomarici and Vecchio 2014). We purposefully asked participants to put the funds into their wallets, to enhance the feeling of ownership before they entered the point of purchase for the BDM experiment. Finally, during the explanation of the BDM lottery, we again made it clear that the participants were under no obligation to buy the product.

The training of the interviewers occupied an entire day, to familiarize them with the specifics of the BDM lottery and the vignettes. One of the authors attended this training and observed the interviewers during the initial data collection

phase. Each interview took approximately 25 minutes. The final sample of 88 participants (Appendix E) is similar in size to other samples in our literature review that use incentive-compatible methods and field data (as opposed to data collected from students). To increase the quality, validity, and reliability of our study, we also performed several tests (Appendix F). We asked participants to rate the credibility of the newspaper article on a Likert-like scale (1 = "very unconvincing" to 7 = "very convincing"). The mean was 4.89 (SD = 1.56) and not significantly different across manipulations, suggesting that the vignettes were convincing ($F(2,85) = 1.72$, $p > .10$).

Next, to ensure that participants understood the CSR manipulation as intended, we asked them to rate Granula's CSR record: "The company Granula was evaluated higher than average by the Institute for Consumer Information with regard to its socially responsible behavior" (1 = "strongly disagree" and 7 = "strongly agree"). The mean values differed significantly across experimental groups ($F(2,85) = 92.46$, $p < .001$), such that participants in the CSR condition reported significantly higher means ($M_{CSR} = 6.22$) than those in the CSI ($M_{CSI} = 1.86$; $p < .001$) or average ($M_{average} = 4.00$; $p < .001$) groups. The means between the CSI and average groups also differed significantly ($p < .001$).

In addition, we tested whether quality perceptions differed significantly across the CSR manipulations; they did ($F(2,85) = 20.12$, $p < .001$). Respondents in the positive CSR group evaluated the product's quality as significantly higher ($M_{CSR} = 6.03$) than those in the CSI ($M_{CSI} = 3.93$; $p < .001$) or average ($M_{average} = 5.29$; $p < .001$) groups. Because the brand is fictitious, respondents could only rely on the provided CSR information to infer product quality, and in all three groups, quality was evaluated above the scale midpoint.

Finally, with several transparency and acceptability measures, we confirmed the validity and reliability of the elicited WTP values through BDM (Wertenbroch and Skiera 2002). As we show in Appendix F, respondents understood the BDM method. Those who agreed to participate generally did not understate their WTP to avoid buying the product. They also indicated a general interest in the product—in this case, fruit drops.

Results

We used regression with bootstrapped standard errors to assess the hypothesized effects of CSR and CSI on WTP, as reported in Table 1.[1] All the hypotheses are directional, so we use directional tests of significance (one-tailed) here and in the following studies, unless indicated otherwise. Participants in the CSR manipulation condition had a significantly higher WTP, and participants in the CSI manipulation had a significantly lower WTP, compared with those in the average condition. Overall, consumer WTP was –23.4% for CSI and +14.6% for CSR. With regard to the within-brand differential effect, such that the markdown effect might be greater than the markup effect for a single brand, we tested for the difference in WTP between these two manipulations by bootstrapping the difference between the coefficients based on 10,000 resamples, but the effect was not significant. That is, Study 1 does not support H2 regarding the within-brand differential effect between CSR and CSI.

Table 1. Regression results with WTP in %a as the dependent variable[1]

	Mean	β	SE	z statistic
H1a: CSR	14.6	.148*	.088	1.68
H1b: CSI	-23.4	-.240**	.080	-2.99
H2: Δβ CSR – CSI		-.092	.150	-0.62
R^2		.211		
Wald χ^2		25.23**		

Notes: Regression analyses with bootstrapped standard errors using 10,000 resamples; n=88.
** $p < .01$; * $p < .05$.
[a] Percentage difference compared with the CSR average manipulation.

Discussion

With Study 1, we confirm H1a and H1b, which predict markup and markdown effects. In our incentive-compatible method, we reveal that WTP for products from irresponsible firms is significantly lower than that for products from firms with average or responsible ratings. Consistent with equity theory and prior literature, consumers exhibited a higher WTP for products from companies behaving responsibly. However, we cannot confirm H2, which we based the negativity bias going back to prospect theory (Kahneman and Tversky 1979). Rather than finding that the markdown effect was stronger than the markup effect, our results indicate that consumers react similarly to negative and positive responsibility information about a firm. However, this result might reflect the small sample size, so we investigate the differential effect further in Study 2.

STUDY 2: ACROSS-BRAND DIFFERENTIAL EFFECT OF CSR/CSI ON WTP

Theory and Hypothesis

To confirm that the unexpected finding regarding H2 was not due to either the small sample size or the study design, which might not have reflected a real-life point-of-purchase scenario sufficiently well, we conducted Study 2 to include choice among different brands. Conceptually, the main difference in

1 We used bootstrapping for the standard errors of the coefficients because of the small number of participants (n = 88; CSI = 29, average = 27, CSR = 32). The comparatively low explained variance in WTP, indicated by R^2, arises because we only investigated one antecedent of WTP, whereas WTP also depends on other factors, such as gender, product quality, involvement, point of sale, and so on. Although CSR may provide additional value to a product, as we seek to determine, it certainly is not the only explanatory variable. For all three studies reported herein, we examine whether including gender as a covariate affects the findings; it does not. Therefore, to ensure the parsimoniousness of our models, we decided to report them without the gender covariate.

Study 2 is that participants receive two similar products from two different manufacturers: Granula and Livida, an additional fictitious brand. Respondents received CSR (CSI, average) information only about one manufacturer (Granula), with no such information provided for the second brand (Livida). This set-up is similar to real-life market situations, in which consumers who do not want to buy a product from one firm (e.g., due to its irresponsible behavior) have the options to lower their WTP or switch to a competitor, even if they don't know anything about the latter firm's CSR performance. In line with the negativity bias (Lange and Washburn 2012; Moosmayer 2012), when two equivalent products are available, compared with different or no information on CSR, the markup for products produced by responsible firms should be less intense than the markdown for products from firms that behave irresponsibly. Therefore, we hypothesize:

H3: When a competitor product is available, the markup effect for the focal product, by firms with a positive CSR rating, is smaller than the markdown effect for products by firms with negative CSR ratings (CSI) (across-brand differential effect).

Stimuli and Procedure

The design of Study 2 largely replicates Study 1, so here, we highlight only the differences. First, this experiment started with an equivalence test for the two products, herbal drops from the brands Granula and Livida. This test ensures that any differences in WTP for the two products are not due to study participants' different perceptions of the brands or the quality of the two products. Participants indicated, on 7-point Likert scales (1 = "do not agree at all" to 7 = "agree totally"), whether they liked the two brands Granula and Livida equally well, and if, in principle, they would have an equivalent tendency to buy products from Granula and Livida the next time they planned to buy drops. The results indicate that participants like the two brands equally well (M = 5.86, SD = .81) and would buy either of them (M = 5.99, SD = .77). That is, participants did not perceive any major differences between the two products before the manipulation that could have affected their WTP.

Second, the newspaper article was slightly adjusted to inform participants that the quality of the herbal drops offered by Granula and Livida was identical and high. Specifically, the updated version explains that the Institute for Consumer Information not only assesses the quality of the producers but also their CSR performance. However, for cost reasons, it would conduct this assessment only for a random subsample of firms, and to date, the information was only available for Granula but not Livida. The newspaper article continues to explain and manipulate the CSR (CSI, average) performance of Granula, as in Study 1 (Appendix B).

Third, the BDM lottery includes both brands to ensure incentive-compatible measures of both WTPs. Participants had to state the price they would be willing to pay for herbal drops from Granula and Livida. If they indicated a higher WTP for Granula than for Livida, it would suggest that they perceive drops from Granula as more valuable. If they instead indicated a lower WTP for Granula, they regard those drops as less valuable. With this WTP measurement, we can compare the WTP of a product with known CSR performance to the WTP of a product with unknown CSR performance (across-brand differential effect). Finally, the BDM concluded with a two-step lottery process. Participants first drew a brand name from an urn (Granula or Livida), then drew the market price for this brand. As in Study 1, participants had to buy the product and pay the market price if their stated WTP for the drawn brand was higher than the market price.

We conducted two pretests to affirm parts of this design. First, we tested perceived quality and liking of seven different brand names, to identify one that would be perceived as equivalent to Granula, such that the brand names of the two different companies would not influence WTP or brand preferences. According to 93 students from a central European university, Livida did not differ significantly from Granula on quality ($M_{Granula}$ = 3.52; M_{Livida} = 3.55, p = .873) or liking of the brand name ($M_{Granula}$ = 2.80; M_{Livida} = 2.87, p = .055). Second, we pretested the packaging for the two products among 41 employees at another university in Europe. Because these initial results indicated a significant difference ($M_{Granula}$ = 3.17, M_{Livida} = 3.76, p < .001), we adjusted the packaging by decreasing the variation in colors, such that the revised packages used black and dark grey (instead of black and blue), and making the font size and lettering identical.

Sample and Quality Assessment

The data collection followed the design of Study 1, implemented by the same professional market research firm in the same city. Recruitment of and incentives for the 138 participants also remained the same (see Appendix E). As Appendix F indicates, the manipulation checks and validity and reliability items provide a similar pattern of results as in Study 1.

Results

We first checked for a within-brand differential effect of CSI versus CSR on WTP, using the same analysis as in Study 1 (WTP for Granula as the dependent variable, between-subject design). The results of the regression analyses with bootstrapped standard errors are in Table 2. Consistent with Study 1, we find significant effects of the CSI and CSR manipulations on WTP (H1) but do not find support for the within-brand differential effect (H2); the difference in the WTP coefficients for WTP in the CSR versus CSI manipulation was not significant.

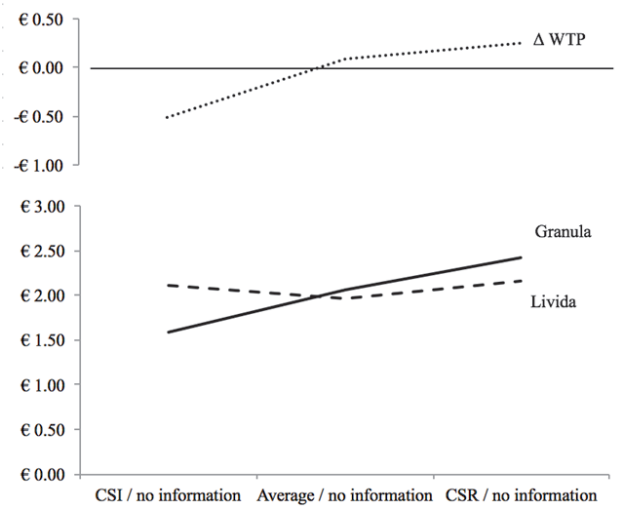

Figure 2. WTP for Granula and Livida across manipulation conditions

Table 2. Regression results with WTP in %a for Granula as dependent variable

	Mean	β	SE	zs statistic
H1a: CSR	19.17	.179**	.076	2.36
H1b: CSI	-21.46	-.228**	.066	-3.45
H2: Δβ CSR – CSI		-.049	.126	-0.39
R^2		.197		
Wald χ^2		41.26**		

Notes: Regression analyses with bootstrapped standard errors using 10,000 resamples; n=138.
** $p < .01$. * $p < .05$.
[a]Percentage difference compared with the CSR average manipulation.

However, Study 2 also differs from Study 1, in that the design includes a competitive product, so we can examine how WTP reactions to CSR-related information for Granula compare with WTP for a competitor product (Livida), for which no CSR information is provided. Figure 2 summarizes the mean WTP for Granula and Livida and the difference in WTP (ΔWTP = $WTP_{Granula} - WTP_{Livida}$) across the three manipulation conditions.

The lower part of Figure 2 reveals, as we expected, that respondents stated a lower (higher) WTP for Granula compared with Livida in the CSI (CSR) manipulation condition. In a second step, we calculated the difference in WTP for Granula and Livida (ΔWTP = $WTP_{Granula} - WTP_{Livida}$, upper part of Figure 2) and regressed this difference on the CSR manipulation (Table 3). The findings confirm H1a and H1b (markup and markdown effects). When confronted with negative information about the manufacturing conditions of one product, consumers seem to attribute significantly less value to it, compared with a product for which they have no information about the manufacturing conditions. The results therefore support H3; the difference in coefficients is significant. These across-brand differential effects mean that when a competitor product is available, consumers mark down their WTP in response to CSI information to a greater extent than they mark up their WTP in response to CSR information.

Table 3. Regression results with ΔWTP in %a as the dependent variable

	Mean	β	SE	z statistic
H1a: CSR	12.83	.085*	.035	2.40
H1b: CSI	-25.52	-.299**	.042	-7.06
H3: Δβ CSR – CSI		-.214**	.063	-3.40
R^2		.419		
Wald χ^2		81.21**		

Notes: Regression analyses with bootstrapped standard errors using 10,000 resamples; n=138.
** $p < .01$; * $p < .05$.
[a]Percentage difference compared with the CSR average manipulation.

Discussion

When consumers consider only one brand—perhaps because they are very loyal to it, are bound by a contract, or perceive switching barriers—they react similarly to negative and positive responsibility information about its products. We consider three potential explanations for this finding from our two studies. First, the negativity bias based on prospect theory refers to a loss or gain from the status quo, such that the loss is a negative shift from an existing status. The differential effect hypothesis rests on the principle of loss aversion, and in our experimental design, we created a purchase situation in which consumers state a WTP to buy a product, which means they "gain" a product and "lose" some financial amount. We framed the purchase of a product from an irresponsible firm, relative to one from an average firm, as a loss, whereas purchasing a product from a responsible firm, relative to the average one, was framed as a gain. Yet the loss associated with purchasing a slightly inferior product might have been unclear. Even though prior CSR literature suggests and finds some differential effects, they may be more prominent for outcome variables such as company attitudes or purchase intentions (Lange and Washburn 2012; Sen and Battacharya 2001) rather than for behavioral variables such as WTP. Second, our formulation of the CSR/CSI manipulation might be relevant. To avoid confounding factors, we solely referred to aspects unrelated to the product quality (e.g., supplier payment conditions). Thus, the product participants stated their WTP for something that was physically identical but had different imagined levels of CSR, which could increase or decrease the value of the product. In such a situation, the loss aversion principle might not be operative. Third, the basic choice of WTP as our dependent variable might be pertinent. Our studies did not include more extreme or alternative behavior, such as boycotting or negative word of mouth, which may be more likely for CSI than for CSR, according to extant theory.

When a competitor brand is available though, we find a differential effect. These findings likely account for consumer behaviors toward many products, especially those with low switching barriers like fast moving consumer goods. Thus, the WTP delta for no information versus CSI is greater than that for no information versus CSR (across-brand differential effect). It may seem possible that the average performance of Granula was perceived as rather negative, so the difference in the CSI condition, which we refer to as the markdown effect, is not particularly high. However, this explanation does not hold, because the two means are not significantly different ($M_{Granula\_average}$ = 2.05, SD = .109; $M_{Livida\_average}$ = 1.97, SD = .098, $p > .05$[2]). Therefore, a competitive product offered by a firm that consumers cannot evaluate in terms of its responsibility leads those consumers to react less strongly to positive information but more strongly to negative information.

STUDY 3: EFFECT OF PRODUCT TYPE ON THE RELATIONSHIP BETWEEN CSR/CSI AND WTP

Theory and Hypotheses

In line with the findings from our Studies 1 and 2 and equity theory (Walster, Walster and Berscheid 1978), we expect that the WTP for products from companies with responsible practices is generally higher than that for products by companies with irresponsible practices. But do these CSR markup and CSI markdown effects arise equally across different product categories? Prior studies indicate that hedonic and utilitarian products produce different decision-making patterns and outcome variables, as we detailed in our literature review. Generally, the consumption of hedonic products relates more to emotions and affective reasoning, whereas the consumption of utilitarian products is more cognitively driven (Dhar and Wertenbroch 2000), prompting variation in outcomes such as satisfaction, value, and repurchase intentions (e.g., Jones, Reynolds, and Arnold 2006; Overby and Lee 2006; Leroi-Werelds et al. 2014). We argue that the impact of CSR and CSI on WTP also may be stronger for utilitarian products than for hedonic ones.

Specifically, due to the more cognitive nature of the decision making for utilitarian products, it follows that consumers likely perceive the different product or brand attributes separately, such that each attribute contributes incrementally to the overall value and product choice (Sheth, Newman and Gross 1991). In contrast, the affect heuristic prompted by hedonic products leads consumers to rely primarily on affect to make choices (Slovic et al. 2002), because it is the more efficient way to make a judgment, in that either positive or negative markers (e.g., product attributes, brand) become a foundation for the outcome or overall evaluation. When people use an affect heuristic, they refer to their affect pool (to various degrees) to make a judgment that is independent of their cognitive judgment (Slovic et al. 2002). Thus, when relying on the affect heuristic, consumers do not evaluate different product attributes separately or trade them off against one another to achieve an overall value; they rather form an overall judgment through the heuristic. Accordingly, we posit that hedonic products lead consumers to form WTP judgments through the affect heuristic, but when making the same judgment for utilitarian products, consumers instead adopt an additive logic. Because CSR represents an additional positive feature of the product, it gets added to the value of utilitarian products, but it is overlaid by the affect heuristic for hedonic products. Following this logic, the effect of CSR/CSI on WTP should be stronger for utilitarian products than for hedonic ones.

H4: The impact of CSI/CSR on WTP is stronger for utilitarian products than for hedonic products (product type effect).

Stimuli and Procedure

To compare utilitarian and hedonic products in a single study, we needed a product category that could be perceived either way. Using the same product category also limits the influence of confounding factors, such as different types (e.g., food versus non-food), price levels, or usage situations (e.g., professional versus private usage). Therefore, the focal product in Study 3 is a common beverage, a fizzy mixture of juice with sparkling water. In the study context, it is neither hedonic nor utilitarian per se, so it might evoke either connotation. To manipulate the product type, we used different flavors and primed participants with a vignette-based experiment.

The fictional brand name of the beverage was Katoga. For the utilitarian product, we chose apple flavor, one of the most popular beverages in Germany, where we conducted the study (see Toytown Germany 2006). For example, its consumption was almost twice as high in 2012 as that for all soft drinks from the Coca-Cola company combined (Statista 2015). The bottle was printed with product information in black letters and a single font. For the hedonic product, we chose rhubarb flavor, which implies a more exotic product that has become popular more recently. For this bottle, the rhubarb product information was more colorful and used multiple fonts (see Appendix D).

We submitted the priming task to a pretest with a student sample. The students undertook one of two (hedonic/utilitarian) sentence completion tasks (which we describe subsequently, in relation to the main study), then indicated on a 7-point Likert scale (1 = "rather utilitarian" and 7 = "rather hedonic") whether they perceived the Katoga beverage as hedonic or utilitarian. The results of the pretest confirmed that participants in the hedonic priming condition perceived the apple beverage as relatively utilitarian and the rhubarb beverage as relatively hedonic ($Mean_{utilitarian}$ = 2.35; $Mean_{hedonic}$ = 5.37). These students also rated beverage quality and revealed no significant difference across the two priming groups, so differences in quality

2 Two-tailed test, because no direction is expected.

perceptions were unlikely to influence WTP ($Mean_{utilitarian}$ = 4.72; $Mean_{hedonic}$ = 4.60; F(1,180)=0.46, p>.10).

We manipulated the CSR and CSI conditions as in Study 1 but excluded the average condition, such that we devised a 2 x 2 between-subjects experimental design, in which respondents either bid for the utilitarian or the hedonic product and answered questions about the product from a firm with CSR or CSI practices. Participants were randomly assigned to one of the four conditions. To measure WTP, we again relied on a BDM lottery as an incentive-compatible measurement method.

The experiment involved four parts (Appendix C). Before the first step, the interviewer explained the procedure—and in particular, the BDM lottery—carefully to participants in a face-to-face setting. Then each respondent was assigned a product and took part in a priming task. Priming entails the temporary, internal activation of respondents, leading to mental representations that influence their response tendencies (Bargh and Chartrand 2000) and judgments (e.g., Jain, Mathur and Maheswaran 2009; van Doorn and Verhoef 2011). Various methods can prompt this internal activation, involving text (e.g., text unscrambling, memory, writing), pictures, or videos (Jain, Mathur and Maheswaran 2009). We used a sentence completion task, such that participants had to complete six sentences (see Appendix D), each of which indicated either hedonic or utilitarian consumption of fizzy drinks. For example, for the hedonic priming, one of the sentences started with "A fizzy drink smells particularly good, when…." For the utilitarian priming, a sentence started with "A fizzy drink effectively satisfies one's thirst, when…." The remaining three parts of Study 3 were identical to those of Study 1: Participants read the newspaper article referring to the fictitious brand Katoga, provided a bid for the presented beverage, participated in the lottery, and responded to a questionnaire that included manipulation checks and demographic items (see Appendices D and E).

Sample and Quality Assessment

The data collection relied on individual, face-to-face interviews, with the assistance of the same professional market research firm in the same major German city as in Studies 1 and 2. The recruitment and incentives for these 108 participants also were the same (see Appendix E). We included similar items for the manipulation checks and validity and reliability assessments; the pattern of results also remains similar (Appendix F).

Results

In H4, we predicted a stronger effect of the utilitarian manipulation on WTP; the results from a regression analysis using bootstrapping support this prediction. The main effect of CSI versus CSR priming is significant (M_{CSR} = 1.40 €, M_{CSI} = .97 €; observed coefficient β = .329, bootstrapped SE = .101; z = 3.24, p < .01), the main effect of the utilitarian (vs. hedonic) priming is significant ($M_{hedonic}$ = 1.23 €, $M_{utilitarian}$ = 1.10 €; observed coefficient β = -.252, bootstrapped SE = .084; z = -2.99, p < .01), and the interaction between the two experimental factors is significant (observed coefficient β = .306, bootstrapped SE = .100; z = 3.06, p < .05; overall model R^2 = .315). Whereas the difference in WTP for hedonic versus utilitarian products in the CSR condition is Δ0.02 € ($M_{CSR\text{-}hedonic}$ = 1.41 €, $M_{CSR\text{-}utilitarian}$ = 1.39 €), it reaches Δ0.25 € in the CSI condition ($M_{CSI\text{-}hedonic}$ = 1.08 €, $M_{CSR\text{-}utilitarian}$ = .83 €). By calculating the conditional marginal effects for product type priming for CSI (z = -3.02, p < .01) and CSR (z = -.22, p > .10), we also found that for CSI, the difference between hedonic and utilitarian product types is significant, whereas the difference for CSR is not significant.[3]

Discussion

The findings from Study 3 suggest a product-type effect when we seek to determine WTP for CSR/CSI, in line with the findings of a meta-analysis by Tully and Winer (2014), who note "some evidence" that WTP is affected by CSR for different types of products, such as food, electronics, wood, and clothing. They call for systematic investigation of this product type effect; in response, our study helps close the gap, by showing that hedonic and utilitarian products are subject to different levels of markup and markdown effects for CSR and CSI. Specifically, the markup effect of CSR is higher for utilitarian products than for hedonic products. Figure 2 shows that the CSR and CSI price response functions are much farther apart for the utilitarian than for the hedonic product and thus the reactions are more dispersed.

CONCLUSION

Results and Discussion

We tested whether consumers would exhibit a markup effect in their WTP for products from firms that manufacture under responsible conditions and a markdown effect in their WTP for products from firms that produce under irresponsible conditions (H1, markup and markdown effect), as well as whether the markdown effect is stronger for brands considered in isolation (H2, within-brand differential effect) or in comparison with competitor brands (H3, across-brand differential effect). With our incentive-compatible WTP measurement, which reduces social desirability bias, we confirm H1a and H1b but not H2 in Studies 1 and 2, then confirm H3 in Study 3. That is, we find a significant markup effect for CSR, which produces a 14.6%

3 We employed the WRS package in R (R Core Team 2016) and robust statistics (Wilcox 2012). Using the function t2waybt from the R package WRS with a bootstrap-t method and 10,000 resamples, we obtained significant main (product priming and CSR vs. CSI) and interaction (all p < .05) effects. Moreover, specifying the appropriate contrasts in the function linconb from the R package WRS using a bootstrap-t method with 10,000 resamples, we found a significant effect for the difference between hedonic and utilitarian products across CSI and CSR (t = 2.27, p < .05).

higher WTP than the WTP of the average CSR group. Furthermore, we uncover a markdown effect for CSI, with a 23.4% lower WTP (Appendix G). Yet with regard to the hypothesized differential effect, in both studies and when a brand appears in isolation, consumers react similarly to irresponsible and responsible firm practices, despite previous literature predicting a differential effect (Lange and Washburn 2012). When consumers have a choice and comparatively state two WTPs, for a focal company with (ir)responsible practices and for a firm that provides no information about CSR, we find a stronger markdown effect for CSI than markup effect for CSR. This study is the first to provide multiple angles and empirical evidence about such differential effects and affirm that when the CSR performance of a brand is considered in isolation, such as by a loyal customer, no differential WTP effect exists for CSR/CSI. Finally, we find empirical support for the predicted product type effect on the relationship between CSR/CSI and WTP. Notably, the markdown effect for CSI is stronger for utilitarian products, materialized in a significantly lower WTP. For firms, it therefore is meaningful to categorize products in their assortments and adapt their pricing strategies accordingly, because the effect of CSR/CSI on WTP differs with assortment categories.

Theoretical Implications

These findings offer following main contributions to theory and literature on WTP for CSR and CSI. First, as illustrated in Figure 1, most previous studies investigate the effect of *more* responsibility (CSR) on WTP; considerably less attention has centered on the effect of irresponsible firm behaviors (i.e., CSI) on WTP. Our findings thus underline the importance of conducting more research into the consequences of irresponsibility, as opposed to focusing just on CSR. Related to this we contribute by studying both phenomena together. In accordance with the negativity bias, we predicted two differential effects, both within and across brands. We found no empirical evidence for a within-brand differential effect, but the differential effect does arise when consumers state their WTP for two products in a comparison setting. This is in line with prior literature (Lange and Washburn 2012) and supports evidence of differential effects in other contexts (Ludwig et al. 2013). Our findings encourage more research into these differential effects to determine when a negativity bias holds and when it does not.

Next, by taking the product category into consideration, we aimed to shed light on the partly conflicting results associated with CSR/CSI–WTP links (no, negative, and positive effects), as well as the vastly different price premiums (e.g., single-digit figures to greater than 60% price premium). In comparing the WTP effects for utilitarian versus hedonic products, our first theoretical contribution is establishing that product type helps explain the different levels of WTP variations. That is, WTP, measured in an incentive-compatible way, varies for different levels of CSR and CSI, as well as for different types of products. Theories about the CSR–WTP relationship thus should address the influences of product characteristics and its consumption situation. These points underline the importance of further research and theorizing in this area.

Lastly, studying sensitive topics such as the behavioral reactions to CSR can be difficult, because consumers are reluctant to reveal their true attitudes and behaviors, which may not be perceived as socially desirable (Podsakoff et al. 2003). Yet calls for more thorough, realistic assessments of consumers' WTP for CSR persist (e.g., Auger and Devinney 2007; van Doorn and Verhoef 2011). Unlike most studies, we did not employ hypothetical measures for WTP but instead used incentive-compatible measurement approaches in all three studies that require real buying behavior within a BDM lottery. Thus, our measurement approach responds to calls for more real-world CSR research studies (Lacey, Kennett-Hensel and Manolis 2015). In line with prior recommendations (Wertenbroch and Skiera, 2002), we perceive that such approaches achieve more meaningful results about WTP.

Beyond this we introduced a variant of the BDM lottery approach that considers consumers' WTP reactions in the presence of a competitor's product. This setting is even more realistic; it represents the situation at the point of sale, when consumers must choose among products and compare brands. This adapted BDM variant seems particularly useful when consumers do not have price anchors in mind, as was the case in our experiment with fictitious brands and as would be realistic for product innovations. Our finding of different results related to the differential effects also shows that amending the method makes a difference. Thus, we encourage other researchers to use this updated version of BDM to ensure a realistic design and accurate results.

Managerial Implications

Firms face markup or markdown effects when brands in their assortment are associated with CSR or CSI. Existing findings are mixed about whether and how much consumers are willing to pay more (or less) for products from companies that behave (ir)responsibly. Furthermore, practitioners lack insights into product type effects. Our findings shed light on these questions, suggesting some important implications for service, retail and industry managers.

Consumers mark down or mark up their WTP when they realize that firms are engaged in CSR or CSI. The price premium for products associated with responsible firms (CSR) is 14.6% (Study 1) or 17.5% (Study 2)—a substantial amount that can support firms operating on thin margins (Gauri 2013). Prior findings based on incentive-compatible measures reveal similarly high price premiums (Alphonce and Alfnes 2012; Didier and Lucie 2008; Ellis, McCracken and Skuza 2012; van Doorn and Verhoef 2011) or even higher ones (Bougherara and Combris 2009; Rode, Hogarth and Le Menestrel 2008) (Appendix A). Consumers also mark down their WTP in response to products from a firm with irresponsible practices (see Figure 1). A

service provider or retailer offering a product that is associated with an irresponsible firm (e.g. a tea or soft drink in a restaurant or retail store) should expect a markdown of around 23.4% (Study 1) or 22.3% (Study 2). These results can also help industry managers justify their CSR investments and support higher resell prices for products linked to CSR.

In both Studies 1 and 2, the markdown effect was higher than the markup effect, yet in neither was the difference (within-brand differential effect, H2) significant. In contrast, we find a differential effect when a competitor brand is available (across-brand differential effect, H3) and consumers must state their WTP for both products. These findings can inform managers' pricing and assortment strategies for products that induce switching barriers. Two assortment strategies then can be differentiated. A defensive assortment strategy aims to avoid products that might be associated with irresponsibility and potential markdown effects. Such a strategy would help firms hedge against the risk of being "punished" for manufacturers' irresponsibility, which represents a pertinent threat (Hartmann and Moeller 2014). An offensive assortment strategy instead aims to enhance markup effects for products from firms associated with responsible behavior to capture value from manufacturers' CSR investments. On the basis of prior literature and the differential effects predicted in H2 and H3, it might seem as if managers should prioritize defensive over offensive assortment strategies (e.g., Leonidou, Katsikeas and Morgan 2013). However, our results instead suggest prioritizing the defensive strategy only for products that also invoke substantial choice perceptions. If consumers feel as if they do not have a choice (e.g., for complementary products like toothbrush heads or vacuum cleaner bags, for products that induce high brand loyalty), we instead recommend adopting both strategies simultaneously. That is, when the consumer has a lot of choices, the risks of choosing a potentially "dangerous" assortment with products from manufacturers that might fail to behave responsibly increase. In adjusting their WTP (markup and markdown effect), consumers in low choice situations likely react with similar intensity to both CSR and CSI, whereas with more choice, consumers react more strongly when confronted with negative information about a brand. Still, according to extant findings (Lange and Washburn 2012; Sen and Battacharya 2001), managers may expect a differential effect for attitudinal variables, independent of the choice.

Limitations and Further Research

These results should be viewed in light of the research limitations. The products analyzed were fast moving consumer goods and food products; additional research might investigate other products and categories, or else test a product typology that encompasses more heterogeneity (e.g., functional, emotional, social, or health values). We manipulated the two fizzy drinks to make one appear more exotic (rhubarb) and thus emphasize a hedonic nature. But this manipulation may have affected our results if the two drink flavors are associated with different buying frequencies, because distinct buying frequencies influence consumer behavior (Cai and Aguilar 2013). Furthermore, in real-life situations, consumers might adopt various strategies when they confront CSI, such as ending the relationship or participating in boycotts (e.g., Klein, Smith and John 2004; Neilson 2010). These effects are not captured by our research, so further studies could examine other outcome variables as well. Our research design involved an actual product sale, so the data collection had to be conducted through face-to-face interviews (Bougherara and Combris 2009), which took 25 minutes each. We put enormous effort into our data collection (three separate BDM lotteries, field data, Study 1 88 participants; Study 2 138 participants; Study 3 108 participants). Yet we still must note the limitations of these relatively small samples, as are common to BDM studies (Bougherara and Combris 2009; Hustvedt and Bernard 2010). Finally, this study was conducted in a single European country, so replications in other countries are needed to establish the generalizability of our findings.

Beyond these avenues for further research, we note again that most studies focus on CSR, even though our research shows that CSI can produce even more pronounced reactions. Therefore, we recommend that researchers focus on the price premiums that potentially can be reached with CSR, as well as the markdowns that brand manufacturers as well as service providers and retailers that sell their products might suffer if they manufacture their products irresponsibly. This study gives rise to several questions about the "dark side of responsibility." It seems particularly important to investigate the differential effects, by which consumers seemingly might react more intensely to CSI than to CSR (Lange and Washburn 2012; Sen and Battacharya 2001), though we find only partial support for this effect. We encourage more research to investigate the conditions and outcome variables for which managers can expect differential effects.

REFERENCES

Adams, J. Stacy (1965). "Inequity in Social Exchange." in *Advances in Experimental Social Psychology,* Leonard Berkowoitz (ed). New York: Academic Press.

Aguinis, Herman and Ante Glavas (2012). "What We Know and Don't Know About Corporate Social Responsibility: A Review and Research Agenda." *Journal of Management* 38 (4), 932-68.

Alphonce, Roselyne and Frode Alfnes (2012). "Consumer Willingness to Pay for Food Safety in Tanzania: An Incentive-Aligned Conjoint Analysis." *International Journal of Consumer Studies* 36 (4), 394-400.

Auger, Pat, Paul Burke, Timothy M. Devinney, and Jordan J. Louviere (2003). "What Will Consumers Pay for Social Product Features?" *Journal of Business Ethics* 42 (3), 281-304.

Auger, Pat and Timothy Devinney (2007). "Do What Consumers Say Matter? The Misalignment of Preferences with Uncon-

strained Ethical Intentions." *Journal of Business Ethics* 76 (4), 361-83.

Auger, Pat, Timothy M. Devinney, Jordan J. Louviere, and Paul F. Burke (2008). "Do Social Product Features Have Value to Consumers?" *International Journal of Research in Marketing* 25 (3), 183-91.

--- (2010). "The Importance of Social Product Attributes in Consumer Purchasing Decisions: A Multi-Country Comparative Study." *International Business Review* 19 (2), 140-59.

Backhaus, Klaus, Robert Wilken, Markus Voeth, and Christina Sichtmann (2005). "An Empirical Comparison of Methods to Measure Willingness to Pay by Examining the Hypothetical Bias." *International Journal of Market Research* 47 (5), 543-62.

Bargh, John A. and Tanya L. Chartrand (2000). "The Mind in the Middle: A Practical Guide to Priming and Automaticity Research." Pp 253-85 in *Handbook of Research Methods in Social Psychology,* H. Reis and C. Judd (eds). New York: Cambridge University Press.

Battacharya, C. B. and Sankar Sen (2004). "Doing Better at Doing Good: When, Why, and How Consumers Respond to Corporate Social Initiatives." *California Management Review* 47 (1), 9-24.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak (1964). "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9 (3), 226-32.

Benkenstein, Martin, Manfred Bruhn, Marion Büttgen, Christiane Hipp, Martin Matzner and Friedemann W. Nerdinger (2017). "Topics for Service Management Research – A European Perspective", *SMR - Journal of Service Management Research*, 1 (1), 4–27.

Bolton, Lisa E. and Anna S. Mattila (2015). "How Does Corporate Social Responsibility Affect Consumer Response to Service Failure in Buyer–Seller Relationships?" *Journal of Retailing* 91 (1), 140-53.

Bougherara, Douadia and Pierre Combris (2009). "Eco-Labelled Food Products: What Are Consumers Paying For?" *European Review of Agricultural Economics* 36 (3), 321-41.

Cai, Zhen and Francisco X. Aguilar (2013). "Meta-Analysis of Consumer's Willingness-to-Pay Premiums for Certified Wood Products." *Journal of Forest Economics* 2013 (15-31).

Carrigan, Marylyn and Ahmad Attalla (2001). "The Myth of the Ethical Consumer - Do Ethics Matter in Purchase Behaviour?" *Journal of Consumer Marketing* 18 (7), 560-78.

Connelly, Brian, David Ketchen, and Stanley Slater (2011). "Toward a 'Theoretical Toolbox' for Sustainability Research in Marketing." *Journal of the Academy of Marketing Science* 39 (1), 86-100.

De-Magistris, Tiziana, Teresa Del Giudice, and Fabio Verneau (2015). "The Effect of Information on Willingness to Pay for Canned Tuna Fish with Different Corporate Social Responsibility (Csr) Certification: A Pilot Study." *Journal of Consumer Affairs* 49 (2), 457-71.

Dhar, Ravi and Klaus Wertenbroch (2000). "Consumer Choice between Hedonic and Utilitarian Goods." *Journal of Marketing Research* 37 (1), 60-71.

Diallo, Mbaye Fall, Norchène B. D. Mouelhi, Mahesh Gadekar, and Marie Schill (2020). "CSR Actions, Brand Value, and Willingness to Pay a Premium Price for Luxury Brands: Does Long-Term Orientation Matter?" *Journal of Business Ethics,* 169 (2), 241–260.

Didier, Tagbata and Sirieix Lucie (2008). "Measuring Consumer's Willingness to Pay for Organic and Fair Trade Products." *International Journal of Consumer Studies* 32 (5), 479-90.

Ellis, Joan L., Vicki A. McCracken, and Nathan Skuza (2012). "Insights into Willingness to Pay for Organic Cotton Apparel." *Journal of Fashion Marketing & Management* 16 (3), 290-305.

Folkes, Valerie S. and Michael A. Kamins (1999). "Effects of Information About Firms' Ethical and Unethical Actions on Consumers' Attitudes." *Journal of Consumer Psychology* 8 (3), 243-59.

Follows, Scott B. and David Jobber (2000). "Environmentally Responsible Purchase Behaviour: A Test of a Consumer Model." *European Journal of Marketing* 34 (5/6), 723.

Gauri, Dinesh K. (2013). "Benchmarking Retail Productivity Considering Retail Pricing and Format Strategy." *Journal of Retailing* 89 (1), 1-14.

Grewal, Dhruv, Kusum L. Ailawadi, Dinesh Gauri, Kevin Hall, Praveen Kopalle, and Jane R. Robertson (2011). "Innovations in Retail Pricing and Promotions." *Journal of Retailing* 87, S43-S52.

Hartmann, Julia and Sabine Moeller (2014). "Chain Liability in Multitier Supply Chains? Responsibility Attributions for Unsustainable Supplier Behavior." *Journal of Operations Management* 32 (5), 281-94.

Hirschman, Elizabeth C. and Morris B. Holbrook (1982). "Hedonic Consumption: Emerging Concepts, Methods and Propositions." *Journal of Marketing* 46 (3), 92-101.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2005). "Do Satisfied Customers Really Pay More? A Study of the Relationship between Customer Satisfaction and Willingness to Pay." *Journal of Marketing* 69 (2), 84-96.

Hustvedt, Gwendolyn and John C. Bernard (2010). "Effects of Social Responsibility Labelling and Brand on Willingness to Pay for Apparel." *International Journal of Consumer Studies* 34, 619-26.

Jain, Shailendra Pratap, Pragya Mathur, and Durairaj Maheswaran (2009). "The Influence of Consumers' Lay Theories on Approach/Avoidance Motivation." *Journal of Marketing Research* 46 (1), 56-65.

Jin, Yao, Ronn J. Smith, and Laurel Aynne Cook (2012). "A Race to the Bottom? Consumer Responses to Human Rights Performance." *AMA Marketing & Public Policy Academic Conference Proceedings* 22, 61-62.

Jones, M. A., K. E. Reynolds, and M. J. Arnold (2006). "Hedonic and Utilitarian Shopping Value: Investigating Differential Effects on Retail Outcomes." *Journal of Business Research* 59, 974-81.

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2), 263-91.

Katsikeas, Constantine, Constantinos Leonidou, and Athina Zeriti (2016). "Eco-Friendly Product Development Strategy: Antecedents, Outcomes, and Contingent Effects." *Journal of the Academy of Marketing Science* DOI 10.1007/s11747-015-0470-5.

Klein, Jill Gabrielle, N. Craig Smith, and Andrew John (2004). "Why We Boycott: Consumer Motivations for Boycott Participation." *Journal of Marketing* 68 (3), 92-109.

Kraus, S., M. Cane, M. and D. Ribeiro-Soriano (2022). "Does doing good do well? An investigation into the relationship between consumer buying behavior and CSR", *Economic Research-Ekonomska Istraživanja*, forthcoming.

Lacey, Russell, Pamela A. Kennett-Hensel, and Chris Manolis (2015). "Is Corporate Social Responsibility a Motivator or Hygiene Factor? Insights into Its Bivalent Nature." *Journal of the Academy of Marketing Science* 43, 315-32.

Lange, Donald and Nathan T. Washburn (2012). "Understanding Attributions of Corproate Social Irresponsibility." *Academy of Management Review* 37 (2), 300-26.

Leonidou, Constantinos, Constantine Katsikeas, and Neil Morgan (2013). "'Greening' the Marketing Mix: Do Firms Do It and Does It Pay Off?" *Journal of the Academy of Marketing Science* 41 (2), 151-70.

Leroi-Werelds, Sara, Sandra Streukens, Michael Brady, and Gilbert Swinnen (2014). "Assessing the Value of Commonly Used Methods for Measuring Customer Value: A Multi-Setting Empirical Study." *Journal of the Academy of Marketing Science* 42 (4), 430-51.

Lin-Hi, Nick and Karsten Müller (2013). "The Csr Bottom Line: Preventing Corporate Social Irresponsibility." *Journal of Business Research* 66 (10), 1928-36.

Luchs, Michael G., Rebecca Walker Naylor, Julie R. Irwin, and Rajagopal Raghunathan (2010). "The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference." *Journal of Marketing* 74 (5), 18-31.

Ludwig, Stephan, Ko de Ruyter, Mike Friedman, Elisabeth C. Brüggen, Martin Wetzels, and Gerard Pfann (2013). "More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates." *Journal of Marketing* 77 (1), 87-103.

Madrigal, Robert and David M. Boush (2008). "Social Responsibility as a Unique Dimension of Brand Personality and Consumers' Willingness to Reward." *Psychology & Marketing* 25 (6), 538-64.

Mantrala, Murali K., Michael Levy, Barbara E. Kahn, Edward J. Fox, Peter Gaidarev, Bill Dankworth, and Denish Shah (2009). "Why Is Assortment Planning So Difficult for Retailers? A Framework and Research Agenda." *Journal of Retailing* 85 (1), 71-83.

Miller, Klaus M., Reto Hofstetter, Harley Krohmer, and Z. John Zhang (2011). "How Should Consumers' Willingness to Pay Be Measured? An Empirical Comparison of State-of-the-Art Approaches." *Journal of Marketing Research* 48 (1), 172-84.

Moosmayer, Dirk C. (2012). "Negativity Bias in Consumer Price Response to Ethical Information." *Business Ethics: A European Review* 21 (2), 198-208.

Murphy, Patrick E. and Bodo B. Schlegelmilch (2013). "Corporate Social Responsibility and Corporate Social Irresponsibility: Introduction to a Special Topic Section." *Journal of Business Research* 66 (10), 1807-13.

Neilson, Lisa A. (2010). "Boycott or Buycott? Understanding Political Consumerism." *Journal of Consumer Behaviour* 9 (3), 214-27.

Overby, Jeffrey W. and Eun-Ju Lee (2006). "The Effects of Utilitarian and Hedonic Online Shopping Value on Consumer Preference and Intentions." *Journal of Business Research* 59 (10/11), 1160-66.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003). "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." *Journal of Applied Psychology* 88 (5), 879-904.

Pomarici, Eugenio and Riccardo Vecchio (2014). "Millennial Generation Attitudes to Sustainable Wine: An Exploratory Study on Italian Consumers." *Journal of Cleaner Production* 66 (3), 537-45.

Porter, Michael E. and Mark R. Kramer (2006). "Strategy & Society: The Link between Competitive Advantage and Corporate Social Responsibility." *Harvard Business Review* 84 (12), 78-92.

R Core Team (2016). "A Language and Environment for Statistical Computing." https://www.R-project.org/. Accessed 19th Dec. 2016.

Rode, J., R. M. Hogarth, and M. Le Menestrel (2008). "Ethical Differentiation and Market Behavior: An Experimental Approach." *Journal of Economic Behavior & Organization* 66, 265–80.

Schreier, M. and J. Werfer (2007). "Auktionen Versus Lotterien: Ein Empirischer Vergleich Zur Messung Von Zahlungsbereitschaften." *Die Betriebswirtschaft: DBW* 67 (1), 22-40.

Sen, Sankar and C. Battacharya (2001). "Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility." *Journal of Marketing Research* 38 (2), 225-43.

Sheth, Jagdish N., Bruce I. Newman, and Barbara L. Gross (1991). "Why We Buy What We Buy: A Theory of Consumption Values." *Journal of Business Research* 22 (2), 159-70.

Slovic, Paul, Melissa Finucane, Ellen Peters, and Donald G. MacGregor (2002). "The Affect Heuristic." Pp 397-420 in *Heuristics and Biases: The Psychology of Intuitive Judgment,* T. Gilovich, D. Griffin and D. Kahneman (eds). New York: Cambridge University Press.

Smith, Jacquelyn (2013). "The Companies with the Best Csr Reputations." https://www.forbes.com/sites/jacquelynsmith/2013/10/02/the-companies-with-the-best-csr-reputations-2/#789d253134ff. Accessed 25th September 2017.

Smithers, Rebecca (2013). "Growing Taste for Fairtrade Shows It's Britain's Cup of Tea." in *The Guardian,* London, UK: The Guardian.

Statista (2015). "Pro-Kopf-Konsum Von Softdrinks Von Coca-Cola in Ausgewählten Ländern Weltweit in Den Jahren 1991 Bis 2012 (in 8 Flüssigunzen)." http://de.statista.com/statistik/daten/studie/170099/umfrage/pro-kopf-verbrauch-von-getraenken-des-coca-cola-konzerns/. Accessed 18 January 2016.

Sweetin, Vernon H., Lynette L. Knowles, John H. Summey, and Kand S. McQueen (2013). "Willingness-to-Punish the Corporate Brand for Corporate Social Irresponsibility." *Journal of Business Research* 66 (10), 1822-30.

Thompson, Derek W., Roy C. Anderson, Eric N. Hansen, and Lynn R. Kahle (2010). "Green Segmentation and Environmental Certification: Insights from Forest Products." *Business Strategy & the Environment* 19 (5), 319-34.

Toytown Germany (2006). "A Brief History of Apfelschorle." http://www.toytowngermany.com/forum/topic/45623-a-brief-history-of-apfelschorle/. Accessed 20 January 2016.

Trudel, Remi and June Cotte (2009). "Does It Pay to Be Good?" *MIT Sloan Management Review* 50 (2), 61-68.

Tully, Stephanie M. and Russell S. Winer (2014). "The Role of the Beneficiary in Willingness to Pay for Socially Responsible Products: A Meta-Analysis." *Journal of Retailing* 90 (2), 255-74.

van Doorn, Jenny and Peter C. Verhoef (2011). "Willingness to Pay for Organic Products: Differences between Virtue and Vice Foods." *International Journal of Research in Marketing* 28 (3), 167-80.

Walster, E., G. Walster, and E. Berscheid (1978). *Equity: Theory and Research*. Boston: Allyn and Bacon, Inc.

Wei, Wei, Gaeul Kim, Li Miao, Carl Behnke and Barbara Almanza (2018). "Consumer inferences of corporate social responsibility

(CSR) claims on packaged foods." *Journal of Business Research*, 83, 186–201.

Wertenbroch, Klaus and Bernd Skiera (2002). "Measuring Consumers' Willingness to Pay at the Point of Purchase." *Journal of Marketing Research* 39 (2), 228-41.

Wilcox, Rand (2012). *Modern Statistics for the Social and Behavioral Sciences: A Practical Introduction*. Boca Raton, FL: Taylor & Francis.

Yeo, Amy Chu-May, Sky Xiu-Mei Lee, and Steve Carter, S. (2018). "The influence of an organization's adopted corporate social responsibility constructs on consumers' intended buying behavior: A Malaysian perspective." *Social Responsibility Journal*, 14(3), 448–468.

Appendix A: Overview of empirical studies of WTP for CSR

No.	CSR/ CSI	Product type *Product*	Data/Method	Relationship CSR/CSI → WTP	WTP	Study
1	CSR	Utilitarian *tomatoes*	Incentive-compatible	CSR → higher WTP	+11%	Alphonce, R., and Alfnes, F. (2012). Consumer willingness to pay for food safety in Tanzania: an incentive-aligned conjoint analysis. International Journal of Consumer Studies, 36, 394-400.
2	CSR	Utilitarian *kitchen garbage bags*	Real purchase data	CSR → lower WTP	-2.7% to -3.7%	Anstine, J. (2000). Consumers' willingness to pay for recycled content in plastic kitchen garbage bags: a hedonic. Applied Economics Letters, 7, 35-40.
3	CSR	Hedonic *coffee*	Real purchase data	CSR → higher WTP for CSR customers	n/a	Arnot, C., Boxall, P. C., and Cash, S. B. (2006). Do ethical consumers care about price? A revealed preference analysis of fair trade coffee purchases. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 54, 555-565.
4	CSR	Utilitarian & hedonic *soap, athletic shoes*	Hypothetical	CSR → mixed effect on WTP	n/a	Auger, P., and Devinney, T. (2007). Do what consumers say matter? The misalignment of preferences with unconstrained ethical intentions. Journal of Business Ethics, 76, 361-383.
5	CSR	Hedonic *coffee*	Hypothetical	U-shaped relationship CSR → WTP	+47% to +57%	Basu, A. K., and Hicks, R. L. (2008). Label performance and the willingness to pay for Fair Trade coffee: A cross-national perspective. International Journal of Consumer Studies, 32, 470-478.
6	CSR	Not specified *not specified*	Hypothetical	CSR → higher WTP	n/a	Becchetti, L., and Rosati, F. C. (2007). Global social preferences and the demand for socially responsible products: Empirical evidence from a pilot study on fair trade consumers. World Economy, 30, 807-836.
7	CSR	Utilitarian *toilet paper*	Real purchase data	CSR → higher WTP	+13% to +18%	Bjørner, T. B., Hansen, L. G., and Russell, C. S. (2004). Environmental labeling and consumers' choice—an empirical analysis of the effect of the Nordic Swan. Journal of Environmental Economics & Management, 47, 411-434.
8	CSR	Utilitarian *orange juice*	Incentive-compatible	CSR → higher WTP	+50% to +55%	Bougherara, D., and Combris, P. (2009). Eco-labelled food products: What are consumers paying for? European Review of Agricultural Economics, 36, 321-341.
9	CSR	Hedonic *music CDs*	Hypothetical	CSR → higher WTP	n/a	Carter, R. E. (2009). Will consumers pay a premium for ethical information? Social Responsibility Journal, 5, 464-477.
10	CSR	Utilitarian *canned tuna*	Incentive-compatible	CSR → higher WTP	n/a	De-Magistris, T., Del Giudice, T., and Verneau, F. (2015). The effect of information on willingness to pay for canned tuna fish with different corporate social responsibility (CSR) certification: A pilot study. Journal of Consumer Affairs, 49, 457-471.
11	CSR	Hedonic *coffee*	Hypothetical	CSR → higher WTP	+10%	De Pelsmacker, P., Driesen, L., and Rayp, G. (2005). Do consumers care about ethics? Willingness to pay for fair-trade coffee. Journal of Consumer Affairs, 39, 363-385.
12	CSR	Hedonic *men's dress shirt*	Hypothetical	CSR → higher WTP	+36%	Dickson, M. A. (2001). Utility of no sweat labels for apparel consumers: Profiling label users and predicting their purchases. Journal of Consumer Affairs, 35, 96-119.
13	CSR	Hedonic *chocolate*	Incentive-compatible	CSR → higher WTP	+27 to +33%	Didier, T., and Lucie, S. (2008). Measuring consumer's willingness to pay for organic and Fair Trade products. International Journal of Consumer Studies, 32, 479-490.
14	CSR	Utilitarian *cotton apparel (T-Shirt)*	Incentive-compatible	CSR → higher WTP	+28%	Ellis, J. L., McCracken, V. A., and Skuza, N. (2012). Insights into willingness to pay for organic cotton apparel. Journal of Fashion Marketing & Management, 16, 290-305.
15	CSR	Hedonic *athletic shoes*	Hypothetical	CSR → higher WTP	+1% to +10%	Feldman, P. M., and Vasquez-Parraga, A. Z. (2013). Consumer social responses to CSR initiatives versus corporate abilities. Journal of Consumer Marketing, 30, 100-111.
16	CSR	Utilitarian *cotton shirts*	Hypothetical	CSR → higher WTP	+17% to +19%	Ha-Brookshire, J. E., and Norum, P. S. (2011). Willingness to pay for socially responsible products: Case of cotton apparel. Journal of Consumer Marketing, 28, 344-353.
17	CSR	Hedonic *woolen gloves*	Hypothetical	CSR → higher WTP	+2% to +3%	Hustvedt, G., Peterson, H. H., and Chen, Y.-J. (2008). Labelling wool products for animal welfare and environmental impact. International Journal of Consumer Studies, 32, 427-437.
18	CSR	Utilitarian *t-shirts*	Incentive-compatible	CSR → higher WTP	n/a	Hustvedt, G., and Bernard, J. C. (2010). Effects of social responsibility labelling and brand on willingness to pay for apparel. International Journal of Consumer Studies, 34, 619-626.

No.	CSR/ CSI	Product type *Product*	Data/Method	Relationship CSR/CSI → WTP	WTP	Study
19	CSI	Hedonic *apparel industry*	Hypothetical	CSI → lower WTP	n/a	Jin, Y., Smith, R. J., and Cook, L. A. (2012). A race to the bottom? Consumer responses to human rights performance. AMA Marketing & Public Policy Academic Conference Proceedings, 22, 61-62.
20	CSR	Utilitarian *organic cotton*	Hypothetical	CSR → higher WTP	n/a	Lin-Hi, N., and Müller, K. (2013). The CSR bottom line: Preventing corporate social irresponsibility. Journal of Business Research, 66, 1928-1936.
21	CSR	Hedonic *coffee*	Hypothetical	CSR → higher WTP	+3% to +4%	Loureiro, M. L., and Lotade, J. (2005). Do fair trade and eco-labels in coffee wake up the consumer conscience? Ecological Economics, 53, 129-138.
22	CSR	Utilitarian *apples*	Hypothetical	CSR → higher WTP	+5%	Loureiro, M. L., McCluskey, J. J., and Mittelhammer, R. C. (2002). Will consumers pay a premium for eco-labled apples? Journal of Consumer Affairs, 36, 203-220.
23	CSR	Utilitarian *bananas*	Hypothetical	CSR → higher WTP	n/a	Mahé, T. (2010). Are stated preferences confirmed by purchasing behaviours? The case of fair trade-certified bananas in Switzerland. Journal of Business Ethics, 92, 301-315.
24	CSR	Hedonic *athletic shoes*	Hypothetical	CSR → higher WTP (mixed findings)	none to +11%	Marquina, P., and Morales, C. E. (2012). The influence of CSR on purchasing behaviour in Peru and Spain. International Marketing Review, 29, 299-312.
25	CSR	Utilitarian *fruit/vegetables detergent/cleaner Packed food/drink meat/fish clothing electrical goods*	Hypothetical	CSR → higher WTP	+11% to 13%	McGoldrick, P. J., and Freestone, O. M. (2008). Ethical product premiums: Antecedents and extent of consumers' willingness to pay. International Review of Retail, Distribution & Consumer Research, 18, 185-201.
26	CSR	Utilitarian *fresh produce*	Hypothetical	CSR → higher WTP (mixed findings)	none to more than 20%	Misra, S. K., Huang, C. L., and Ott, S. L. (1991). Consumer willingness to pay for pesticide-free fresh produce. Western Journal of Agricultural Economics, 16, 218-227.
27	CSR & CSI	Utilitarian & hedonic *mobile phones* *athletic shoes*	Hypothetical	CSR → higher WTP CSI → lower WTP	+11% to +19% -24% to -41%	Moosmayer, D. C. (2012). Negativity bias in consumer price response to ethical information. Business Ethics: A European Review, 21, 198-208.
28	CSR	Utilitarian *yogurt*	Hypothetical	CSR → higher WTP	n/a	Perrini, F., Castaldo, S., Misani, N., and Tencati, A. (2010). The impact of corporate social responsibility associations on trust in organic products marketed by mainstream retailers: A study of Italian consumers. Business Strategy & the Environment, 19, 512-526.
29	CSR	Hedonic *wine*	Hypothetical/ incentive-compatible	CSR → higher WTP	n/a	Pomarici, E., and Vecchio, R. (2014). Millennial generation attitudes to sustainable wine: an exploratory study on Italian consumers. Journal of Cleaner Production, 66, 537-545.
30	CSR	Not specified *not specified*	Incentive-compatible	CSR → higher WTP	+24% to +65%	Rode, J., Hogarth, R. M., and Le Menestrel, M. (2008). Ethical differentiation and market behavior: An experimental approach. Journal of Economic Behavior & Organization, 66, 265–280.
31	CSR	Hedonic *restaurant*	Hypothetical	CSR → higher WTP (mixed findings)	none to more than 10%	Namkung, Y., and Jang, S. (2017). Are consumers willing to pay more for green practices at restaurants? Journal of Hospitality & Tourism Research.
32	CSR	Hedonic *restaurant*	Hypothetical	CSR → higher WTP (mixed findings)	none to more than 30%	Schubert, F., Kandampully, J., Solnet, D., and Kralj, A. (2010). Exploring consumer perceptions of green restaurants in the US. Tourism & Hospitality Research, 10, 286-300.
33	CSR	Hedonic *fashion apparel*	Hypothetical	CSR → higher WTP	n/a	Shen, B., Wang, Y., Lo, C. K. Y., and Shum, M. (2012). The impact of ethical fashion on consumer purchase behavior. Journal of Fashion Marketing & Management, 16, 234-245.
34	CSR	Utilitarian & hedonic *plywood, furniture*	Hypothetical	CSR → higher WTP	none to +17%	Thompson, D. W., Anderson, R. C., Hansen, E. N., and Kahle, L. R. (2010). Green segmentation and environmental certification: Insights from forest products. Business Strategy & the Environment, 19, 319-334.
35	CSR & CSI	Utilitarian & hedonic *coffee, cotton t-shirts*	Hypothetical	CSR → higher WTP CSI → lower WTP	+17% -29%	Trudel, R., and Cotte, J. (2009). Does it pay to be good? MIT Sloan Management Review, 50, 61-68.
36	CSR	Utilitarian *wood products*	Hypothetical	CSR → higher WTP	n/a	Vlosky, R. P., and Ozanne, L. K. (1999). A conceptual model of US consumer willingness-to-pay for environmentally certified wood products. Journal of Consumer Marketing, 16, 116-136.

Note: The table is based on our own evaluation of the description in the paper of whether the above products are hedonic or utilitarian. In some cases, the classification was rather straightforward: Garbage bags, for example, are clearly utilitarian (Anstine 2000). In other cases, we made the classification dependent on the manipulation in the respective study. For example, in the Ha-Brookshire and Norum's (2011) study, apparel is utilitarian because it is a basic cotton t-shirt, but in another study by Shen et al. (2012) we classify apparel as hedonic because Shen et al. (2012) describe the apparel as "fashionable clothing".

Appendix B: Data Collection for Study 1, Granula Drops

Part 1: CSR, average and CSI Manipulations

In a recent study, the Institute for Consumer Information assessed 15 producers of *[product category]*. The *[product category]* of most producers received the highest ratings "very good" or "good" on quality and taste. Among these is *[company name]*, a company whose *[product category]* earned above average evaluations even within the good evaluations on quality and ingredients. All the ingredients were also tested in terms of their safety and classified as unobjectionable.

In addition to quality assessments, the Institute for Consumer Information also evaluates producer performance in the areas of social and environmental responsibility. Companies are evaluated based on what they do to develop employees, how fairly they treat suppliers and customers, and whether they help to protect the environment. The index to assess a company's Corporate Social Responsibility ranges from 0 (=very irresponsible) to 100 (=very responsible).

Here, *[company name]* performed *[very poorly/ average/ very good]* and received a *[below-average value of 14 / fairly average value of 62 / above-average value of 92]* on the index. With this result *[company name]* was ranked *[amongst the top positions / amongst the middle positions / amongst the bottom positions]* compared to other companies.

It was *[criticized/mentioned/praised]* that *[company name]* had a *[high/average/low]* employee turnover, which indicates *[bad/ normal/good]* working conditions. In addition, *[even though/ __ / __] [company name]* is compliant with regulations for environmental protection [__ / and / and] the company [does *not exceed the minimum / does meet the average / does exceed]* requirements for (e.g., energy consumption or waste disposal). Lastly, *[company name]* implements *[very short-term/ mid-term/ very long-term]* contracts with *[bad/normal/ good]* payment conditions for suppliers: Suppliers *[cannot plan ahead/can plan ahead within a normal range / can plan ahead]* and *[not have to wait a long time / have to wait an average amount of time / have to wait a long time]* for reimbursement.

Part 2: BDM Lottery

[Oral explanation of the BDM method by the interviewer]

Please indicate the price you are willing to pay for the Granula drops.

[After the participant states the WTP, the second price is drawn randomly from an urn and compared with this WTP. The comparison determines whether the participant must buy the product or not.]

Product image:

Part 3: Survey

Questions related to measures to ensure quality, validity, and reliability and demographics (see Appendix F).

Appendix C: Data Collection for Study 2, Granula and Livida Drops

Part 1: Product Equivalence Test

Two questions checked for perceived differences in the two products:

1) I like the brands of Granula and Livida equally well.
2) Generally, in case I buy herbal drops, I would equally prefer to buy Granula or Livida.

Part 2: CSR, average and CSI Manipulations

Differences with Study 1 are underlined.

In a recent study, the Institute for Consumer Information assessed 15 producers of *[product category]*. The *[product category]* of most producers received the highest ratings "very good" or "good" on quality and taste. Among these are the manufacturers *Granula* and *Livida*. Both companies earned above average evaluations on quality and ingredients. All ingredients were also tested in terms of their safety and classified as unobjectionable.

In addition to quality assessments, the Institute for Consumer Information also evaluates the environmental and social responsibility of some manufacturers. As this evaluation is very cost intensive, the institute evaluates only a random sub-sample

of manufacturers each year. Companies are evaluated based on what they do to develop employees, how fairly they treat suppliers and customers, and whether they help to protect the environment. This year, Granula was evaluated in terms of its environmental and social responsibility, but Livida was not part of the sample.

Granula received a *[very poorly/ average/ very good]* evaluation. The index to assess a company's Corporate Social Responsibility ranges from 0 (=very irresponsible) to 100 (=very responsible). Granula received a *[below-average value of 14 / fairly average value of 62 / above-average value of 92]* on the index. With this result, Granula was ranked *[amongst the top positions / amongst the middle positions / amongst the bottom positions]* compared to other companies.

It was *[criticized/mentioned/praised]* that Granula had a *[high/ average/low]* employee turnover, which indicates *[bad/normal/ good]* working conditions. In addition, *[even though/ __ / __]* Granula is compliant with regulations for environmental protection [__ / and / and] the company [does *not exceed the minimum / does meet the average / does exceed]* requirements for (e.g., energy consumption or waste disposal). Lastly, Granula implements *[very short-term/ mid-term/ very long-term]* contracts with *[bad/normal/good]* payment conditions for suppliers: Suppliers *[cannot plan ahead/can plan ahead within a normal range / can plan ahead]* and *[not have to wait a long time / have to wait an average amount of time / have to wait a long time]* for reimbursement.

Part 3: BDM Lottery

[Oral explanation of the BDM method by the interviewer]

Please indicate the maximum prices you are willing to pay for the Granula and Livida drops.

[After participants stated their WTP and completed the questionnaire, the brand and its market price were drawn randomly from an urn and compared with the WTP for this brand of each participant. The comparison determined whether the participant had to buy the product or not.]

Product image:

Part 4: Survey

Questions related to measures to ensure quality, validity, and reliability and demographics (see Appendix F).

Appendix D: Data Collection for Study 3, Katoga Beverages

Part 1: Hedonic and Utilitarian Priming

Please complete the following sentences. [Pictures supported the hedonic and utilitarian priming. In the middle of the page three hedonic or utilitarian pictures relating to the context were shown.]

Hedonic Priming:

- Drinking a fizzy drink with rhubarb is stimulating because…
- If I want to treat myself with something alcohol free and special in a restaurant, I would order the following fizzy drink…
- When I buy fizzy drinks for guests on a special occasion, then…
- Fizzy drinks smell particularly good when…
- I very much enjoy fizzy drinks when…
- Very special fizzy drinks are fun because…

Utilitarian Priming:

- Drinking an apple fizzy drink is extremely advisable when…
- The primary function of fizzy drinks is…
- A fizzy drink effectively satisfies one's thirst when
- In everyday life, a fizzy drink is especially convenient for…
- Drinking fizzy drinks is extremely worthwhile because…
- An apple fizzy drink is a good beverage in order to…

Part 2: CSR and CSI Manipulations

Refer to Part 1 of Study 1.

Part 3: BDM Lottery

Refer to Part 2 of Study 1.

Product Images:

Part 4: Survey

Refer to Part 3 of Study 1. Questions related to measures to ensure quality, validity, and reliability and demographics (see Appendix F)

Appendix E: Sample description

Criterion	Category	Study 1		Study 2		Study 3	
		Granula Drops		Granula & Livida Drops		Katoga Beverages	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Gender	Male	35	39.77	57	41.6	61	56.4
	Female	53	60.23	80	58.4	47	43.5
Age group	18 – 29 years	23	26.14	62	44.9	27	25.0
	30 – 39 years	28	31.82	33	23.9	22	20.4
	40 – 49 years	18	20.45	24	17.4	32	29.6
	50 – 59 years	11	12.50	13	9.4	23	21.3
	Beyond 60 years	8	9.09	6	4.3	4	3.7
Education	Certificate of secondary education	9	10.23	16	11.6	7	6.5
	General certificate of secondary education	23	26.14	42	30.4	21	19.6
	A-levels	32	36.36	49	35.5	31	29.0
	University degree	23	26.14	31	22.5	47	43.9
	Other	1	1.14	0	0.0	1	0.9

Appendix F: Measures to ensure quality, validity, and reliability

Measure	Item	Study 1, Granula Drops	Study 2, Granula & Livida Drops	Study 3, Katoga Beverages
		Mean (SD)	Mean (SD)	Mean (SD)
Credibility of the newspaper article	The newspaper article is very convincing.[1]	4.89 (1.56)	5.35 (1.40)	4.44 (1.71)
Manipulation check for product type priming	From my point of view, the [company & product] rather satisfies…[2]			2.55 (0.84) utilitarian priming 5.75 (0.63) hedonic priming ANOVA F = 505.34; $p \leq .01$
Manipulation check for CSR manipulation	The company received an above average evaluation from the Institute for Consumer Information in the area of CSR.[1]	1.86 (0.95) for CSI 4.00 (1.82) for average 6.22 (0.83) for CSR ANOVA F = 92.46; $p \leq .001$	Manipulation check item 1 1.48 (0.70) for CSI 4.44 (0.90) for average 6.78 (0.42) for CSR ANOVA F = 3.83; $p \leq .001$	1.41 (0.86) for CSI 6.22 (0.75) for CSR ANOVA F = 950.86; $p \leq .001$
Perception of product quality	The [company & product] is of high quality.[1]	5.11 (1.56) for total sample 3.93 (1.56) for CSI 5.29 (1.38) for average 6.03 (0.93) for CSR	5.54 (1.27) for total sample 4.67 (1.38) for CSI 5.56 (0.80) for average 6.37 (1.00) for CSR	5.35 (1.79) for total sample 4.69 (1.81) for CSI 6.06 (1.45) for CSR
Clarity of the BDM method	From my point of view, the instructions for the purchasing game were clear and understandable.	6.55 (0.64)	6.34 (0.84)	6.70 (0.79)
	I was aware that my offer is binding and that may have to the buy the [product].[1]	6.13 (1.38)	5.90 (1.12)	6.38 (1.30)

Measure	Item	Study 1, Granula Drops Mean (SD)	Study 2, Granula & Livida Drops Mean (SD)	Study 3, Katoga Beverages Mean (SD)
Strategic bidding behavior	I am not interested in a [product type] now. Therefore, my offer was so low that I would not have to the buy the [product].[1]	2.68 (1.77)	1.58 (1.04)	1.99 (1.80)
Face validity	I am highly interested in [product type].[1]	4.72 (1.52)	3.18 (1.57)	5.47 (1.39)
	Correlation with the price respondents usually pay for a similar product	r = .57**	r = .44**	r = .38**

**$p < .01$. *$p < .05$. [1] 1 = do not agree at all; 7 = fully agree, [2] 1 = a basic need; 7 = a need for pleasure.

Appendix G: WTP across studies

		Study 1	Study 2		Study 3		
		Granula Drops	Granula & Livida Drops		Katoga Beverages		
			Granula	Livida	Hedonic	Utilitarian	Total
CSI (EUR)	**Mean**	**1.57**	**1.59**	**2.11**	**1.08**	**.83**	**.97**
	SD	0.52	0.51	0.62	0.38	0.25	0.35
	n	29	44	44	31	25	56
Average (EUR)	**Mean**	**2.05**	**2.06**	1.97			
	SD	0.67	0.76	0.68			
	n	27	48	48			
CSR (EUR)	**Mean**	**2.35**	**2.42**	**2.16**	**1.41**	**1.39**	**1.40**
	SD	0.71	0.73	0.62	0.39	0.36	0.37
	n	32	46	46	28	24	52
Total (EUR)	**Mean**	**2.00**	**2.03**	**2.08**	**1.23**	**1.10**	**1.17**
	SD	0.71	0.75	0.64	0.42	0.41	0.42
	n	88	138	138	59	49	108
Changes in WTP in %	Average-CSR	14.6	17.5				
	Average-CSI	-23.4	-22.3				

Keywords:

Corporate social responsibility (CSR), corporate social irresponsibility (CSI), sustainability, willingness to pay (WTP), price setting