Emotion, Norms, and Anchors:
Three Investigations On Consumer Decisions Under Elective Pricing

By

Minah Hong Jung

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Business Administration
in the
Graduate Division
Of the
University of California

Committee in charge:
Professor Leif D. Nelson, Chair
Professor Clayton R. Critcher
Professor Don A. Moore
Professor Gabriel Lenz

Spring 2015

Abstract

Emotion, Norms, and Anchors:
Three Investigations On Decisions Under Consumer Elective Pricing
By

Minah Hong Jung

Doctor of Philosophy in Business Administration
University of California, Berkeley
Professor Leif D. Nelson, Chair

This dissertation documents consumers' decision-making when they have an opportunity be both maximally selfish and kind toward others. Using Consumer Elective Pricing, we explore how social forces operate in influencing consumers' decisions. A set of three investigations examines how emotion, perceptions of social norms, and anchoring as judgment heuristics operate in influencing consumers' decisions. We find that each of the three social forces uniquely shapes consumers' behavior under elective pricing. Consumers are sensitive to the presence of charitable giving but largely insensitive to the scope of their giving when they pay what they want and a portion of their payment goes to charity (Chapter 2). When paying forward, consumers infer a higher level of kindness in others when they are informed about others' kind behavior, raising their own payments to match their perceptions of social norms (Chapter 3). Inconsistent with the prior relevant research on anchoring effects, we find that consumers are not significantly influenced by payment anchors unless anchors are hypothetically presented in lab settings, have no actual financial consequences, or have large distributive (we also term subjective) gaps (Chapter 4). In a series of field and lab studies, this set of three investigations sheds light on the psychology that enforces the deviations from the "rational" economic decisions.

ACKNOWLEDGMENTS

Graduate school at the University of California, Berkeley represents one of the happiest and most rewarding periods of my life. I attribute this to not just the school and program, but also to many people who have contributed to this peculiarly fortunate time. The next few words in turn represent my sincere attempt at articulating my gratitude towards them.

One of the best decisions in my life was to come to Cal to work with Leif D. Nelson. Leif is not just my academic advisor and colleague, he is also an ideal mentor, and most importantly, a lifelong friend. His intellectual support has been the guiding light in my development as a researcher. Our conversations on research have been highlights of my graduate school career, and I look forward with great excitement to our future collaborations. I am especially grateful that Leif has taught me the importance of academic integrity and honesty. He has challenged and advanced the methodological standards, and his calling to rectify issues in identification of truth has fundamentally elevated not just the methodological standards of my research, but also of the field. My accomplishments during my time at Cal are attributable to Leif's selfless devotion to my development and success. I am incredibly proud to be his student, and can only wonder about possible ways in which I could reciprocate. Leif is my best friend. He has been most kind, generous, and superbly fun to spend time with. He has also taught me to appreciate college radio stations and baseball. I will continue to enjoy his outstanding music selections and miss his presence every time I watch baseball.

Clayton Critcher is devastatingly talented. His logic and intuition in human psychology are simply riveting. Clayton is also the hardest working person I know. As an advisor, Clayton has been extremely generous with his time in guiding my research and career. As a collaborator, he has been patient and resourceful. Clayton is also one of the warmest and closest friends that I have. I treasure my time hanging out with Clayton, eating his wonderful home cooked meals, and drinking cocktails in the United Clubs in various cities' airports.

I am also grateful to Don Moore. Don has been an incredibly reliable source of intellectual support for the past five years. His critique and comments on my work have helped me to become a better researcher. I also thank Gabe Lenz for his warm support and availability.

Ayelet and Uri Gneezy are wonderful to work with. They have been exceptionally helpful, warm, and generous. And their dry humor is exquisite.

My work would not have been possible without Silva Kurtisa, the manager of the Behavioral Lab at the Haas School of Business. I admire her remarkable efficiency. I also sincerely appreciate our friendship during the past five years. I will miss her crude honesty and our trips to the Cal stadium.

I attribute my psychological well-being to my loving fellow graduate students at Cal. Weiyi Ng has been the most nurturing presence. He has been with me in all of my anxious and happy moments, teaching me statistics, cooking amazing food, watching silly TV shows, singing, and playing piano. I could not ask for a better friend than Weiyi. Kristina Schoemmel entered my hectic graduate school life like a fairy of light. I cherish our times in Berkeley,

Barcelona, and Copenhagen. I thank Donatella Taurasi for her incredible warmth, Italian meals, and aesthetic tastes. I thank Jenn Logg and Liz Tenney for our sisterhood bond. I will miss our hikes in immaculate Big Sur. Alice Moon, Michael O'Donnell, Muping Gan, and Fausto Gonzalez have been exceptionally helpful academically and most delightful socially. Sheng Li is one of the most resourceful and kindest people I know. Angus John Hildreth has been a loyal and caring friend. I thank Scott Roeder and Hannah Perfecto for sharing the marketing PhD student moments. I am thrilled about my continued friendships with these astute researchers in the coming years.

Tony Gannon has been the most grounding source of my adult life. I am extremely fortunate to have known him and to remain as his friend. I admire his moral standards, appreciation for, and passion to support other lives in struggle. No words can describe how I feel about our remarkable growth since we met in college.

My parents are incredible in trusting me and granting my independence early on. Although I left home as a teenager, they have been with me in everything I do. Their creativity, strength, and love have been the ceaseless force in my life.

I have been fortunate to work with the talented and hard-working undergraduate students in the Science of Giving and Appreciating Lab. No doubt that my work is also their work. These young students have encouraged me to aim for the ideal and to stay optimistic. I sincerely appreciate their genuine curiosity in research. I feel incredibly blessed to be their friend and witness their maturity into young adulthood.

I thank the National Science Foundation Graduate Fellowship and the Greater Good Science Center for their generous financial support.

Table of Contents

CHAPTER 1: INTRODUCTION

Consumers feel pain when paying more than they want to. Hence, they often make efforts to find good deals to pay as little as possible. As this behavior sounds common and familiar, there are also cases in which people willingly pay much more than they have to. For example: consider Humble Bundle. Humble Bundle is an online media retailer that allows its customers to pay any amount they want for a bundle of media products (mostly videogames) downloadable at the company's website. Customers can also decide how much of their payment goes to charity. Considering that many of the company's online videogames are available for free, most customers at Humble Bundle evidently pay more than they have to. Yet since its launch in 2010, Humble Bundle has quickly turned into a multi-million dollar business and retains the "pay what you want" business model. What could have contributed to this surprising success?

This phenomenon proves to be more curious if we consider certain assumptions that underlie extant theories of human decision-making. One such is that people make economic decisions that ultimately benefit themselves (Smith, 1776). The pursuit of material self-interest is considered dominant norms in economic relationships (Fiske, 1992; Heyman & Ariely, 2004). But as shown in the case of Humble Bundle case, this clearly does not constitute the full story: indeed, people often sacrifice their economic self-interest for others. And evidently, this "kind" behavior occurs even in anonymous online commercial exchanges. Further investigation is needed: there are evidently social forces beyond material and rational self-interest that influence consumer behavior.

This dissertation comprises a set of three investigations on consumer decision-making in the midst of the conflict between their own material self-interests and their concerns toward others' welfare. While the relevant literature suggests that multiple social forces guide decisions in social and economic environments, I identify and focus on three main social forces guiding consumers' decisions: emotion, social norms, and anchoring. The three chapters featured here detail our investigations of how each of these social forces operates in influencing consumers' behavior in various commercial exchanges.

These examinations heavily employ the setting of Consumer Elective Pricing (CEP). CEP is a pricing system in which customers can decide how much to pay for things (often including zero). The most well-known form of consumer elective pricing is pay-what-you-want (PWYW), as instanced in the above example of Humble Bundle. CEP has been increasingly gaining traction by both companies and academic communities since its major public exposure through a British rock band, Radiohead, in 2007. Radiohead allowed its fans to pay any price (including zero) they wanted for a then newly released album and the band later claimed that the promotion was highly successful. Since then, an increasing number of non- and for-profit companies have adopted PWYW or other variants of CEP. CEP as such, not only provides an interesting setting and arena to test our theories, but is also a pricing model that is gaining prominence and importance in the market.

CEP also provides several advantages for the study of consumer behavior. One of the challenges in studying the operation of social forces is that different types of social forces often influence behavior simultaneously. Good research aims to separate, highlight and identify these multiple social forces; CEP affords this in the following ways.

Firstly, CEP offers a real-world commercial setting in which consumers experience a tension between pursing self-interest and their concerns toward others' welfare. That is, CEP allows us to study how social forces influence real-world economic decision-making.

Secondly, the malleability of CEP allows for the implementation of variants of its basic pay-what-you-want form; these variants effectively separate out relevant social forces of interest. This is employed in Chapters 2 and 3 with two variants of CEP: shared social responsibility and pay-it-forward. While these variants remain financially identical to "Pay What You Want", the most basic form of CEP, the we will see that simple slight modifications will sufficiently transform the nature of commercial exchanges to specifically test the operation of the construct of interest.

Finally, CEP can be operationalized in both lab and field settings, enabling us to draw from the strengths and compensate for the weaknesses of both techniques. In the lab, the implementation of CEP allows us to test the operation of social forces in controlled setting, which nicely accommodates the findings from field experiments that naturally comes with much noise from the real-world settings. Furthermore, with an increasing number of companies using CEP, there are more opportunities for researchers to collaborate in large-scale field experiments. This research fully takes advantage of all of these benefits of CEP.

The description of each chapter follows; these essays feature slightly modified versions of the three papers from their most recently submitted or published manuscripts.

Chapter 2 features the first of the three papers, *Signaling Virtue: Charitable Behavior under Consumer Elective Pricing*, a joint work with Leif D. Nelson, Ayelet Gneezy, and Uri Gneezy currently under review at *Marketing Science*. In this chapter, we investigate what influences consumers' decisions under Shared Social Responsibility (SSR), a variant of consumer elective pricing under which consumers pay what they want and a portion of their payment goes to charity. Through four field experiments, we aim to better understand consumers' quantitative and qualitative concerns that guided their decisions under SSR.

Studies 1 and 2 of Chapter 2 tests consumers' quantitative sensitivity toward their charitable contribution under SSR by varying the size of the portion going to charity. Consistent with Andreoni's theory of impure altruism, the main findings from these two studies highlight the operation of emotion in influencing consumers' charitable decisions; consumers pay more when any portion of their payments goes to charity but they are largely insensitive to the size of their charitable contribution, paying similar amount whether 1% or 99% of their payments went to charity.

Replicating the findings in Gneezy et al. (2012), customers in Studies 1 and 2 are less likely to purchase when a portion goes to charity than none went to charity. Furthermore, as with their payment amount decision, consumers' purchase likelihood was also insensitive to the size of charitable contribution under SSR; they are equally likely to buy when 1% goes to charity than 99% goes to charity. These results seem to suggest the *negative* emotion that consumers experience under SSR; they feel pressured pay more under SSR and tend to opt-out entirely.

Studies 3 and 4 in Chapter 2 examined whether or not the findings in Studies 1 and 2 were purely driven by selection bias. To do so, we removed selection based on SSR pricing information in Studies 3 and 4 by informing customers about SSR after they decided to buy a product. Furthermore, in Study 4 we tested the operation of social pressure by having some customers pay anonymously and the rest paying directly. The results of these two studies suggest that neither selection-bias or direct social pressure is entirely driving the findings in Studies 1 and 2. The four field studies show that the pattern of consumers' charitable behavior is largely consistent with Andreoni's impure altruist explanation; consumers pay more when a portion of their payment goes to charity because they experience a positive emotion (a "warm-glow") but

are largely insensitive to how much of their payment goes to charity, which reflects a typical on-off operation of an emotional stimuli, the presence of charity in our context.

Chapter 3 features the second paper, *Paying More When Paying for Others*, another joint effort with Leif D. Nelson, Ayelet Gneezy, and Uri Gneezy, published at *Journal of Personality and Social Psychology* in 2014. In this paper, we consider a stimulus that clearly differs from charitable giving in Chapter 2 but produces a similar effect. Instead of charitable giving that evokes the operation of emotion, we consider information about others' behavior to evoke the operation of social norms in consumers' decision-making. We operationalize information of others' behavior by using pay-it-forward. Pay-it-forward (PIF) is a variant of elective pricing that contains our descriptive social norm manipulation; under PIF, customers are told that they have an opportunity to pay forward to someone who will come later, while their product has been paid for by another customer who came earlier. While the control pricing, PWYW, lacks any information about others' behavior, PIF contains information about other customers are being kind but this information is not explicit as to how kind other are (i.e., how much they paid). In a series of field and lab studies, we investigate the influence of descriptive social norms on consumers' decisions under elective pricing by comparing their behaviors under two elective pricing systems, PIF and PWYW, and predicted that people would pay more under PIF than they would under PWYW.

The first four studies in Chapter 3 are field experiments that compared customers' behavior under PIF to those under PWYW in both non-profit and for-profit settings. The results of these four studies are consistent with our basic prediction that people pay more under paying forward than paying what they want in real world settings. The next four lab studies aim to test what explains the PIF behavior in the controlled settings. Studies 7 and 8 are particularly informative about the underlying process. When participants are not explicitly informed about others' behavior, they pay more under PIF than they did under PWYW, but when they are told about how much others paid, the effect is eliminated. Our mediation analysis suggests that participants' perception of others' payments fully mediates the relationship between the pricing condition and payment amounts. As suggested by our reviewers, the last three online studies aim to rule out potential compounds and to test the generalizability of the PIF effect.

Although PIF does not have a charitable giving component, it increases customers' payments compared to those who pay what they want. This effect seems to be driven by people's (inaccurate) belief about others' behavior; under PIF people tend to think that others are paying more than they do and they raise their payment to match their perceptions of other's behavior. And the effect is eliminated when people are explicitly informed how much others pay, suggesting that people are following explicit descriptive norms of others' behavior when available.

Chapter 4 presents our third paper, *Anchoring In Payment: Evaluating A Judgmental Heuristic In Field Experimental Settings*, a collaborative work with Leif D. Nelson and Hannah Perfecto. This paper is currently under review at *Journal of Marketing Research*. This paper is somewhat similar to the paper featured in Chapter 2 such that our manipulations involve quantitative variations, but instead of varying the magnitude of charitable contribution, we vary the magnitude of payment anchors of different types (e.g., default, suggested, average price) to test how these anchors influence payments. In addition, the studies in this chapter use PWYW or pure donation and do not involve other variants of elective pricing such as SSR or PIF.

In 13 large-scale field and four lab experiments we tried to better understand how different numeric information influences (or does not influence) how much people pay under

elective pricing. We test a variety of numeric values as anchors. Across the studies, these anchors vary widely in how far apart they are from each other in both absolute (e.g., $3 and $8 are $5 apart) and distributive terms (e.g., 23% of the people pay up to $3,whereas 48% of people pay up to $5, so they are about 25% apart in payment distribution). Inconsistent with the findings in the existing literature on anchoring, we find null or weak anchoring effects unless the gap between two anchors in the distribution of payments is quite large, when anchors do not have financial consequences to decision-makers, or when responses were collected in laboratory settings. We find that the anchoring effect is often much smaller and even null when people are actually paying for things under elective pricing.

The three chapters crystalizes my long rigorous attempts to understand both consumer decisions and its underlying psychology. As with most investigations in social sciences, this dissertation leaves ample room for alternative explanations for the observed phenomena. I seek to make better predictions.

As the following chapters as well as the materials presented in appendix clearly show, we made our predictions, data collection method, statistical analyses to be as transparent as possible by following the guidelines laid out by Simmons, Simonsohn, and Nelson (2011) as well as by utilizing public resources like the Center for Open Science (see http://osf.io) to pre-register our studies. One of the consequences of following these methodological guidelines is that it requires additional time and financial resources to test predictions in tolerably powered studies. During this process, however, I learned that these efforts are absolutely *minimal* in substantiating our findings. As with many scientists and the human subjects in our studies, I am at risk of my own biases toward incentives (i.e., to publish) that misguide my actions in claiming what is true or not in my research even without intending to. Although there seems no perfect solution to preempt such self-deception, properly powered replications have been absolutely worthwhile and necessary in guiding my research.

CHAPTER 2: SIGNALING VIRTUE UNDER CONSUMER ELECTIVE PRICING

Abstract

Four field experiments examined the quantitative and qualitative forces influencing behaviors under consumer elective pricing called "shared social responsibility" (SSR, Gneezy, Gneezy, Nelson, & Brown, 2010). Under SSR consumers can pay what they want and a percentage of their payment goes to support a charitable cause. Customers in our experiments were sensitive to the presence of charitable giving, paying more when a portion of their payment went to charity (Studies 1-4), but were largely insensitive to what portion of their payment went to charity (Studies 1 and 2). To test possible explanations we examined how consumers' qualitative concerns to signal a positive image influenced their decisions and found that neither self-selection into paying (Studies 3 and 4) nor social pressure (Study 4) explained higher payments under SSR.

Corporations worldwide spend millions of dollars every year on corporate social responsibility (CSR) programs (Mohin 2012; Social Investment Forum, 2009). Although consumers strongly support socially responsible companies, they are often skeptical about the actual delivery of CSR promises (Reputation Institute's CSR RepTrak 2013 Study) and might be suspicious of ulterior motives (David, Kline, and Dai, 2005; Friestad and Wright, 1994). In part because skepticism and distrust may explain low returns of firms' CSR investments, some companies have adopted an extremely risky CSR strategy relying entirely on consumers' fairness and generosity.

Humble Bundle, for example, is an online media distributor that allows its customers to pay any price for their product, and to allocate their payment to content developers or charitable organizations, as well as Humble Bundle itself. In this manner, Humble Bundle employs CSR and managed to build a successful business around it: Since launching in 2010, the company has grossed over $50 million, over $20 million of which went to support charitable causes. Customers are not exploiting Humble Bundle; rather they appear to be rewarding it. Why are customers paying more than they have to?

In this paper we investigate potential social forces that contribute to the success (or failure) of CSR programs in consumer elective pricing settings. In consumer elective pricing the buyer can pay any price for a good or service; self-interest behavior is captured by paying the minimum price allowed. Paying more than the minimum, therefore, reflect the operation of more complex social forces than simple selfish behavior. The most well-known form of consumer elective pricing is pay-what-you-want (PWYW). In its simplest form, PWYW offers a customer is offered a good or service for which she decides how much to pay (often including $0). Recent years have seen a rise in academic and media attention on PWYW. In 2007, the well-known band Radiohead sold an album using PWYW, with over 1.2 million downloads, and with many of those customers electing to pay enough to make the album highly profitable for the band (Mohamud 2007). Following this success, other independent artists and companies have tried to mimic this approach.

Shared Social Responsibility (SSR; see Gneezy, Gneezy, Nelson, & Brown 2010) is a slightly modified version of PWYW that offers a mix of PWYW and CSR, by allocating a percentage of customers' payments to charity. This model and its variants have been recently adopted by an increasing number of companies (e.g., StoryBundle, Indie Game Stand, MacWorld). Considering the obvious financial risks of SSR, which are often amplified by the anonymous nature of online transaction settings, the success of some of these companies is intriguing.

Research on consumers' behavior under PWYW has shown that customers feel obligated to fairly reciprocate sellers by paying something even when they do not have to (Kim et al. 2009, Regner and Barria 2009; Schmidt, Spann, and Zeithammer 2012; Regner and Riener 2012; Mak, Zwick, and Rao 2010).

Fewer studies have directly examined behaviors under SSR. It is likely that fairness and reciprocity concerns operate under both PWYW and SSR, but that SSR involves additional forces. In a large field experiment at a popular amusement park, Gneezy et al. (2010 and 2012) compared behaviors under a fixed price, CSR, PWYW, and SSR for souvenir photographs taken during customers' roller-coaster rides. Their results show that customers paid five times more under SSR than under PWYW. Yet, customers in the SSR condition were fifty percent less likely to purchase a photo. Even with a decreased purchase rate, SSR was more profitable than either the PWYW or fixed price treatments.

The present research investigates the forces that guide consumers' selfish versus generous behaviors under consumer elective pricing. We investigated how consumers' signaling motivations—quantitative and qualitative—influenced their decision to buy and conditional on buying how much to pay in SSR setting.

Quantitative concerns of signaling virtue: Consumers' sensitivity to a portion of payments that goes to charity

PWYW involves financial risk: Customers can harm a firm by paying zero or very little. Hence, adopting any form of SSR pricing signals the company's prosocial commitment. An additional signal, in the form of the magnitude of the charitable contribution, could further influence consumers' decision of whether to purchase and how much to pay. By testing customers' response to different charitable allocations in a field setting we gained insights regarding the psychological process underlying their decisions. In addition, our experiments also highlight the financial implications for firms' socially responsible pricing strategy because even if allocating a large percentage to a charity increases payments, it also cuts into firms' profits. Below we outline predictions and explanations for consumers' behavior under SSR with varying charitable percentages.

If people were purely selfish, their predicted behavior is simple: Consumers would pay nothing in the PWYW pricing, regardless of the amount going to charity. A selfish consumer will pay a positive amount under PWYW only if she gets some external benefit from paying, such as signaling their prosocial identity to others (Greenwald and Breckler 1985, Schlenker 1986, Fehr, Brown, Zehnder 2009, Milinski, Semmann, and Karmbeck 2002). If people's seemingly kind behavior is primarily driven by their extrinsic signaling motivation, making payment observable or a product with an explicit prosocial signal (e.g., a charity logo on a product) will increase giving.

For non-selfish consumers, there are two main reasons why people pay more when giving to charity. One is pure altruism: Pure altruists in our experiments would be motivated by how much the charity will receive. Under this assumption, their behavior would depend on the net benefit provided to the charity. As the charitable percentage in SSR increases, the return on giving increases. Alternatively, consumers would want to reward the company based on their perceptions of company's generosity, risk in adopting SSR, or their enhanced shopping experience in purchasing a product and helping charity simultaneously (Gneezy et al. 2010). Whether customers pay to benefit the charity or to compensate a prosocial company, the pure altruist account predicts that customers will pay more as the firm gives higher percentages to charity.

The second reason why people would pay more when giving to charity is impure altruism (Andreoni 1990, 2002). This notion is based on a combination of pure altruism with the "warm glow" that people *feel* when they do something good (e.g., giving money to charity). According to this account, a consumer's positive emotion from giving is the impure part of the utility function. The larger the amount that goes to charity under SSR, the better the consumer will feel. But the marginal utility of the positive emotion diminishes as the size of that giving increases. That is, an impure altruist will feel a greater positive emotion when she gives any amount to charity relative to none, then will experience a smaller increase in positive emotion as the size of her giving increases. Depending on the parameters of the positive feeling from simply giving relative to the size of that giving, impure altruism results in more Note that this prediction assumes a *scope-insensitive consumer*. There is evidence suggesting that for some stimuli people are sensitive to presence or absence, but largely insensitive to magnitude, a tendency referred to

as scope insensitivity (Frederick and Fischhoff 1998, Hsee and Rottenstreich 2004). For example, Karlan and List (2011) tested the effectiveness of matching grants on charitable giving and found that a grant match substantially increases giving but increasing the matching ration from 1:1 to 1:3 did not increase giving.

Similarly, customers might be sensitive to the presence of charitable giving under SSR, paying more when any portion went to charity, but insensitive to the size of the charitable percentage, paying similar amounts regardless of the size of the charitable percentage. If customers were purely scope insensitive, then we would expect similar behavior when 1% or 99% of a payment was going to charity. Hence, the prediction under the assumption of impure altruism is that giving even a very small portion (e.g., 1%) of payment to charity would significantly increase the payment amount under SSR. But a substantial increase (e.g. 1% vs. 50%) in the portion of charitable giving would have a relatively small effect on giving when comparing to the effect of giving any from zero percent to charity.

Qualitative concerns of signaling virtue: What explains higher payments under SSR?

Recent evidence on PWYW (Gneezy et al, 2012) shows that self-image plays an important role in individuals' non-selfish behavior in markets; consumers have basic desires to feel good about who they are (Baumeister and Leary 1995, Bodner and Prelec 2002, Dunning 2007) and to similarly look kind and fair to others (Greenwald and Breckler 1985, Schlenker 1986, Fehr, Brown, Fehnder 2009, Milinski, Semmann, and Karmbeck 2002). Although these social goals push some consumers to show the kindly behavior that may signal their generosity, they also have the potential to push others to exit the situation altogether (e.g., Gneezy et al. 2010; 2012).

The success of a firm's SSR strategy may depend considerably on how SSR appeals to different segments of customers; a more "charitable type" of customers, or those who want to burnish their generous image, voluntarily engage in an SSR transaction and pay generous amounts (Gneezy et al. 2012). The rest, who are likely to feel that the "appropriate" price is more than what they wish to pay, would likely avoid the transaction altogether. As a result, the customers whose behavior under SSR is observed might be merely a reflection of a selection effect; those who choose to buy under SSR are those who are willing to pay more. That is, SSR might exclude those who feel that the "right" price under SSR is more than they want to pay.

These predictions are also in line with the models of Bodner and Prelec (2003) and Benabou and Tirole (2006, 2011); a consumer's self-image is assumed to be a game a person plays with herself. The consumer would like to see herself to be prosocial, but does not know her true preferences. Under this model, a consumer who is willing to pay only a low amount, may be more likely to buy the product under PWYW than under SSR, because in the latter case paying little would register as a negative signal to the self-image. The comparison of consumers' behavior in Studies 1 and 2 in which people could choose whether to purchase, with those in Studies 3 and 4 in which consumers are faced with the pricing option only after deciding to purchase, can help us in learning the importance of selection by participants who opt-out once they see the charity component.

OVERVIEW OF THE STUDIES

Four field studies examine the quantitative and qualitative forces influencing payments under PWYW and SSR. First two studies were aimed to test our pure vs. impure altruist accounts for consumers' quantitative sensitivity in charitable giving. We first considered how customers respond to elective pricing offers that have differential signaling value to self and others. By selling reusable fabric shopping bags at an organic grocery store, we tested how consumers responded to a very small charitable percentage under SSR (Study 1). In this study we also manipulated whether the product publicly displayed the charity versus the store logo to test whether consumers' signaling motive influenced their level of prosociality. In Study 2, we sought to further separate our two competing accounts by using a wider range of charitable allocations. In addition, we sold reusable shopping bags again at a traditional supermarket to increase the generalizability of the results from Study 1. In the last two studies we tested whether high payments under SSR was largely due to the attraction of a different—more generous—consumer segment by removing selection bias in consumers' purchasing decisions. In Study 3 we set up a coffee stand and randomized participants to see different charitable allocations (0%, 10%, or 50% going to charity) after they have already decided to purchase coffee. Finally, we looked whether high payments in the SSR setting is driven by consumers' self vs. social-image concerns by manipulating anonymity at a campus doughnut stand in a setting that has been similarly stripped of a selection effect (Study 4).

Study 1: A Trivial Charitable Percentage Under SSR and a Public Signal of Generosity

Study 1 varied the percentage allocated to charity (0%, 1%, or 50%) to identify which account best describes customers' quantitative concerns. As we argued above, a selfish customer will pay zero in all cases. The *pure altruist* will pay more with any increase in charitable allocation. The *impure altruist* and the *scope-insensitive* customer will pay substantially more at 1% than at 0%, but making a comparatively smaller increase in payment as that contribution increases to 50%.

In addition, in Study 1 we looked at how an external signal of generosity influenced behavior. Prior research suggests that people are frequently more generous when they can signal their prosocial identity to others (Glazer and Konrad, 1996; Gneezy et al., 2012). If people care about expressing their generous social identity, they might be more likely to purchase and pay more for a product that allows them to signal prosociality quite explicitly..But there is also some contrasting evidence suggesting that an extrinsic reward or an external prosocial signal could crowd out an intrinsic or self-signaling motivation (Bem 1972; Gneezy et al. 2012). To test how an explicit signal of charitable giving influenced consumers' behavior, in this study we varied the (observable) logo printed on our products to be either a commercial or charitable logo.

Method

In this field study we collaborated with a grocery store and an animal protection charity. The former is a well-known San Francisco based organic vegan cooperative owned and operated by its workers. We sold reusable fabric shopping bags in front of the store for nine days (from 9am to 6pm) in June and July 2012. Shoppers (N = 12,394) who entered the store saw store signs indicating a randomly assigned pricing condition. In the PWYW pricing condition, the sign read "Take a Bag, Pay What You Want". In the two SSR conditions, signs read, "Take a bag, Pay What You Want, [1% or 50%] of what you pay goes to [name of the charity]." We manipulated the public signal of generosity by varying the logo printed on the shopping bags to be either the grocery store logo or the charity logo (see Supplemental Materials for the images of the signs

and bags/logos). Note that we did not sell shopping bags with the charitable logo in the PWYW condition in which zero percentage went to charity, as this might have promoted complaints that the store was taking advantage of the charity by taking the entire revenue from the sale[1]. Therefore, we had five experimental conditions in this study.

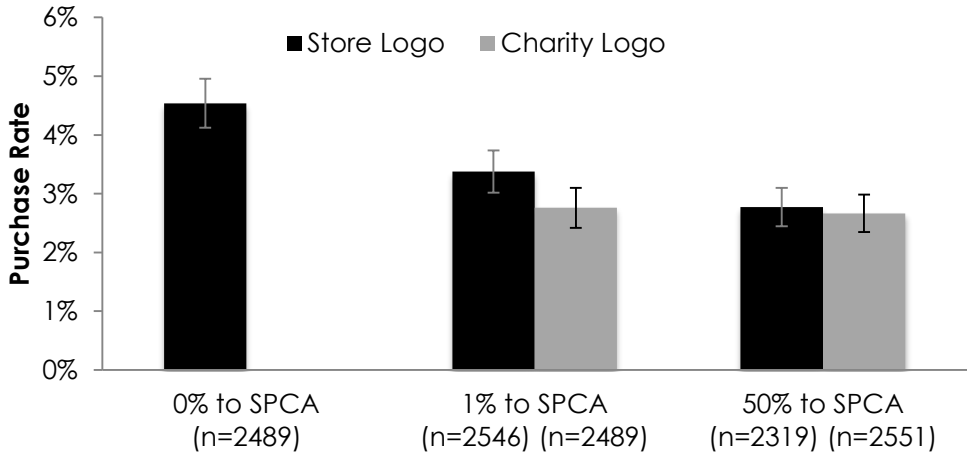
We set up tables in front of the two main entrances to the store[2]. Shopping bags were displayed on the table next to the pricing sign. In Studies 1 and 2, we randomized the five experimental conditions with every fifty shoppers entering the store. The research apprentices in the field followed the pre-randomized order of conditions throughout the entire experiment. We used this randomization strategy to control for time of day and day of week effects. We recorded the number of people who passed by our table, each customer's payment amount, the group size, the number of purchased bags per group, and easily observable demographic information (i.e., gender, ethnic background, and estimated age). In this and the following three studies, we predetermined the sample size and intermittently examined the data to ensure that the experiment was operating smoothly. We did not make any decision to continue or stop the experiment based on interim analyses.

Results and Discussion

For our analysis, we treated each individual bag sold per customer as our unit of analysis, with average payment-per-person (per bag) as the dependent variable. To derive purchase rates, we divided the number of transactions by the number of total passerby per condition. Controlling for the day of transaction did not significantly influence the direction or significance of our results. Therefore, we do not further discuss this variable.

Replicating Gneezy et al. (2012), people in all SSR conditions were less likely to buy a shopping bag than in the PWYW condition ($\chi^2(1, N=12{,}394) = 18.24$, $p < .001$). Notably, this was true even when only one percent of the payment went to charity; customers were significantly less likely to buy when 1% of their payment went to charity than 0% did, (3.08% vs. 4.54%), $\chi^2(1, N = 7354) = 10.13$, $p = .001$. Purchase rates between the 1% and 50% conditions did not differ (3.08% vs. 2.72%), $\chi^2(1, N=9{,}905) = 1.17$, $p = .279$. These results are consistent with the *scope-insensitive consumer* account; when deciding whether or not to purchase, consumers seem to care more about the presence of charitable giving but less about how much of their payment was forwarded to charity.

Figure 2.1. Purchase Rates in Study 1. Error bars reflect ± SE of the means.



Our logo manipulation did not affect purchase rates (3.08% for the store logo bags vs. 2.71% for the charity logo bags), $\chi^2(1, N = 9{,}905) = 1.19$, $p = .275$, suggesting that the social-signaling potential of one's generosity does not influence purchase decisions.

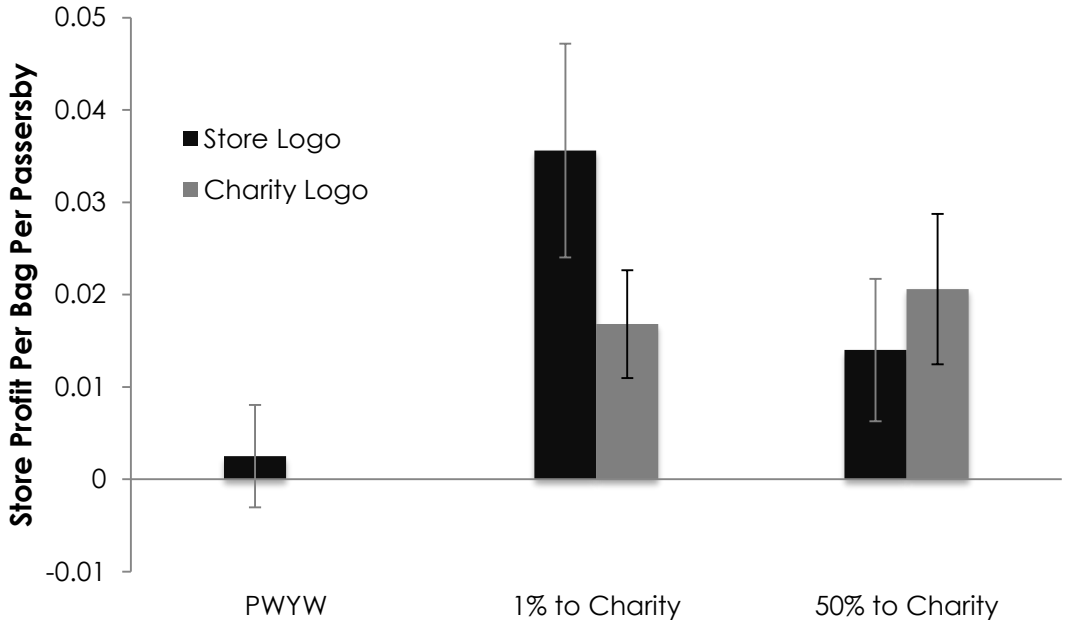
Thirty-five people asked that all of their payment to be forwarded to the charity and did not take a bag in return (i.e., they wanted to make a pure charitable donation), and were therefore excluded from our analyses of payment amount conditional on buying[3]. Pricing (PWYW, SSR with 1%, or SSR with 50%) significantly influenced average payments, F (2, 397) = 21.34, $p <$ .001; customers paid more as the charitable contribution increased. Customers paid more when either 1% or 50% of their payment went to charity than zero percent went to charity ($M_{0\%}$ = \$1.40 vs. $M_{Charity\%}$ = \$3.07), $t(398) = 4.55$, $p < .001$. As shown in Figure 2.2 customers paid more even when only one percent of their payment went to charity than when zero percent went to charity (*Ms* =\$2.24 vs. \$1.40), $t(261) = 3.19$, $p = .002$. Customers also paid more when 50% of their payment went to charity than when 1% went to charity (*Ms* =\$3.98 vs. \$2.24), $t(285) = 3.97$, $p < .001$). Consistent with the pure altruist account, customers in Study 1 paid more when the charitable percentage increased.

Figure 2.2. Mean Payment per Reusable Shopping Bag in Study 1. Error bars reflect ± SE of the means.

Because we did not sell bags with a charity logo in the PWYW condition, we ran a 2 (pricing: SSR with 1% or SSR with 50%) × 2 (logo type: a charitable logo or a commercial logo) ANOVA. Pricing significantly affected average payments; people paid more when a higher versus low percentage went to charity ($M_{SSR\text{-}1\%}$= \$2.24 vs. $M_{SSR\text{-}50\%}$= \$3.98), F (1, 283) = 16.22, $p < .001$). The main effect of the logo type ($M_{Store\ logo}$= \$3.07 vs. $M_{Charity\ Logo}$= \$3.11), F (1, 283) = .01, $p = .924$, and the interaction term of pricing and logo type was not significant, F (1, 283) = 1.24, $p = .266$. These results suggest that customers pay more when a high percentage goes to charity, but that their payments are not affected by the (social) signaling potential of their purchase.

To account for the effect of selection, we analyzed the data with profit per passerby as a dependent measure. We first deducted the cost per reusable bag ($1.35) from the average payment per bag and entered zero payments for pure donations. As shown in Figure 2.3 and in Tables 1.9 and 1.10 (in Supplemental Materials), the 1% SSR condition was more profitable than the PWYW condition, ($M_{0\%}$ = $0.003 vs. $M_{0\%}$ =$0.027), $t(7353) = 2.39$, p =.017. The 50% SSR condition was marginally more profitable than the PWYW condition, (*Ms* = $0.003 vs. $0.017), $t(7525) = 1.68$, p =.093. Profitability in the 1% SSR and 50% SSR conditions did not differ, (*Ms* = $0.027 vs. $0.017), $t(9904) = 1.07$, p =.284.

Figure 2.3. Store Profit per Passerby in Study 2. Error bars reflect ± SE of the means.



In our analysis of the charitable surplus per passerby, we included pure donations. Predictably, the charitable surplus per passerby increased substantially as the charitable portion increased, ($M_{1\%}$ = $0.01 vs. $M_{50\%}$ =$0.07), $t(9904) = 6.41$, p <.001.

In Study 1 we investigated how customers respond to the magnitude of charitable giving in the SSR setting. Consistent with previous investigations of consumer elective pricing, people reliably pay more when part of the payment will benefit charity, suggesting that there are perceived costs associated with paying little when payment is linked to a social cause (i.e., supporting a charitable cause). Furthermore, consistent with the *pure altruist* account, customers in Study 1 paid more as the charitable portion under SSR increased.

Prior research shows that people behave more prosocially when their behavior is publicly observed. People rarely give to charities anonymously (Glazer and Konrad 1996) but they are frequently more generous when their behavior is observed and tend to give more when doing so can boost their social prestige (Harbaugh 1998). Yet, customers in our studies were equally likely to purchase a bag with a charitable signal, as they were to purchase a bag with a commercial signal (i.e., without a charitable signal). In addition, once they have decided to buy a bag, customers did not pay more for a product that would publicly signal their generosity to others. These results suggest that even if people can extract more value from the reputational benefits, they are apparently not willing to pay more for the benefits.[4]

Study 2: The Scope Sensitivity in Charitable Giving and Competition

We had two goals for this study. First, we thought that it was critical to establish the reliability and generalizability of Study 1, so we aimed to replicate the critical conditions from that experiment. Second, in order to better discriminate the two accounts for consumers' prosocial motives, we added two additional allocations (99% and 100%).

Selling reusable shopping bags at a different location, we tried to replicate the same three SSR offers (PWYW with 0%, 1%, or 50% of payment going to charity) and tested two additional charitable levels: 99% and 100%. Following on the discontinuity observed in purchase likelihood in Study 1 when moving from 0% to 1% charitable contribution, Study 2 investigated whether there is a similar discontinuity when SSR shifts from a nearly entirely charitable to a purely charitable purchase (i.e., 99% and 100% to charity).

To test the robustness of our finding with respect to different populations, we conducted Study 2 in a different grocery store, collaborating with a traditional supermarket located in Oakland, California that sells typical products available in most large grocery chain stores in the U.S., and attracts local residents with modest income.

Importantly, the timing of Study 2 (November 2012 until February 2013) enabled us to observe the influence of another naturally occurring variable. On January 1, 2013, half way through our study, the Alameda Country's Reusable Bag Ordinance went into effect, mandating stores to charge $0.10 for every paper shopping bag. Following the language of residents at the time, we will simply refer to this as the "bag law." This policy change functionally changed the environment for our product by increasing the price of the alternative shopping bag competitor from $0 to $.10. We discuss the impact that this change could have on both purchase rates and purchase prices below.

Method

We sold reusable grocery bags in front of a large supermarket from 12pm to 5pm for 10 days in 2012 and 13 days in 2013. We determined the dates and duration of the study based on a few practical grounds. First, the grocery store has already partnered with other non-profit organizations, so we coordinated our experiment dates to avoid overlap and potential confounds. Second, because this study required the minimum of four research assistants (two at each entrance door) at all times, both the timing and duration of our experimental sessions depended on their availability. Third, to observe the impact of the bag law on consumers' behavior, we wanted to conduct the study for sufficient numbers of days before and after January 1st, 2013. We ordered 800 bags and decided to continue data collection until the bags were sold out.[5]

We set up two tables and covered the traffic at both entrances. The procedure of Study 2 was similar to the one used in Study 1 with two modifications. First, we sold only one type of bag with the store logo. Second, we tested five different charity percentages: PWYW (0% going to charity) and four SSR offers with 1%, 50%, 99%, and 100% of payment going to charity. We randomized experimental conditions after every 100 customers entered the store by switching the signs (the order of conditions was randomized ahead of time) and recorded the same information as in Study 1.

Results and Discussion

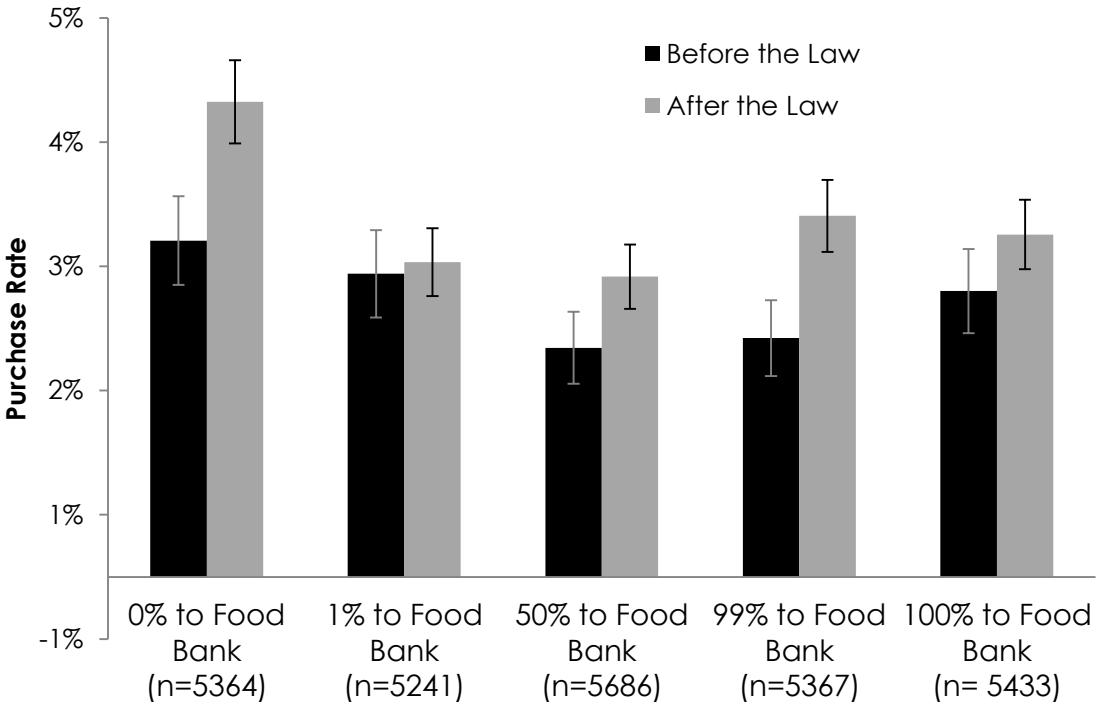
Customers (N = 27,091) saw a sign indicating the randomized condition before entering the supermarket; the bags were displayed on the table. We excluded 10 purchases made by the store manager, researchers' friends, and acquaintances that were informed about the purpose of our study. As in Study 1, we conducted our analysis using each bag sold per customer (total

number of bags sold = 714) as our unit of analysis and the average payment per bag as a dependent variable[6]. We also compared the purchase rates across all conditions. Finally, we compared the pattern of results obtained before and after the introduction of the bag law.

Purchase Likelihood

Consistent with the results from Study 1, people were significantly less likely to purchase a bag when any portion (aggregating the four SSR conditions) went to charity than when no portion did (3.4% vs. 2.45%; χ^2 (1, N = 27,092) = 15.51, $p < .001$) in both 2012, (2.71% vs. 2.12%), χ^2 (1, N = 10,150) = 3.12, $p = .081$, and 2013, (3.82% vs. 2.65%), χ^2 (1, N = 16,942) = 13.10, $p < .001$.

Figure 2.4. Purchase Rates in Study 2. Error bars reflect ± SE of the means.



People were much more likely to purchase a bag in 2013 after the enactment of the bag law (χ^2 (1, N = 27,092) = 9.33, $p = .002$) than they did in 2012. Before the bag law was enacted, customers were slightly more likely to buy a bag under PWYW (0% to charity) than at each of the different levels of SSR before the bag law was enacted. But none of those comparisons were statistically significant. However, after the bag law, customers were significantly more likely to purchase a bag at 0% charitable contribution than at 1% (χ^2 (1, N = 6,610) = 8.94, $p = .002$), 50% (χ^2 (1, N = 6,812) = 11.22, $p < .001$), 99% (χ^2 (1, N = 6,633) = 4.39, $p = .038$), or 100% (χ^2 (1, N = 6,742) = 6.06, $p = .014$). Purchase rates did not differ across the four SSR conditions before and after the bag law.

Consistent with the results in Study 1, people were sensitive to the presence of charity; they were less likely to purchase a bag, and paid more for it, when any portion of their payment went to charity. But they were largely insensitive to what proportion of their payment going to

charity, particularly after the bag law. Furthermore, people were much more likely to purchase a bag after the law was enacted but their overall payment decreased substantially after the law.

Purchase Price

In this analysis, we excluded 85 cases of pure donations—customers who donated money but declined to take a bag.[7] The seven cases of pure donation in the PWYW condition seem somewhat puzzling. It is possible that some customers in Studies 1 and 2 saw one sign as they entered and saw a different one as they exited or they are repeat customers who were exposed to a charity condition first and bought a bag and then saw the PWYW condition but still believed that some of their payment went to charity. Our study cannot completely rule out these possibilities. Replicating the results of Study 1, customers paid more when a portion of their payment went to charity ($M_{Charity\%}$ = \$2.90 vs. $M_{No\ Charity}$ = \$0.98, $t(712) = 5.76$, $p < .001$; See Figure 2.5). This was true when comparing PWYW to every level of SSR, $5.44 < t\text{'s} < 6.09$. Controlling for the date and time variables did not change the direction or significance of the results. But inconsistent with the results in Study 1, there were no significant differences in payment amounts between any pair of the four SSR conditions, all $t\text{'s} < 1.98$. Note that this pattern of results supports *the impure altruist* and *scope-insensitive consumer* account, and seems less consistent with the *pure altruist* account.

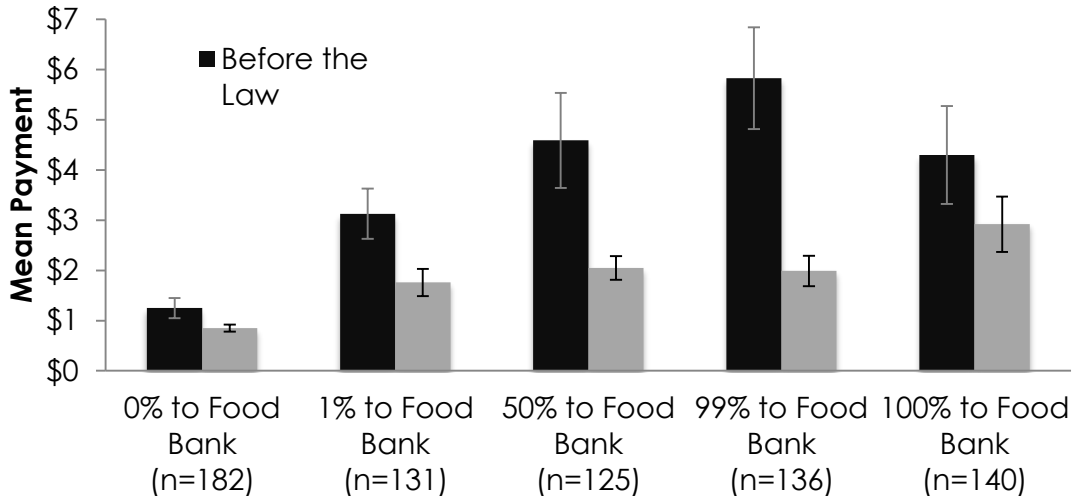
In 2012, before the enactment of the law, payments reflected some sensitivity to the level of charitable contribution. Customers paid more when 1% went to charity ($M_{1\%}$= \$3.13 vs. $M_{No\ Charity}$ = \$1.26, $t(101)=3.69$, $p < .001$), slightly, but not significantly, more again with a 50% contribution ($M_{1\%}$ = \$3.13 vs. $M_{50\%}$ = \$4.59, $t(85) = 1.42$, $p = .160$), and slightly, but not significantly, more again with a 99% contribution ($M_{50\%}$ = \$4.59 vs. $M_{99\%}$ = \$5.83), $t(77) = .90$, $p = .373$. Payments in the pure charity condition (100% to charity) were lower and quite similar to those in the 50% condition ($M_{100\%}$ = \$4.30). Payments did not differ significantly across 50%, 99%, and 100% conditions (see Supplemental Materials for test statistics for these comparisons).

These trends disappeared after the bag law was introduced in 2013. First, customers paid substantially more before the bag law was introduced than after (M_{2012} = \$3.82 vs. M_{2013} = \$1.92), $F(1, 704) = 38.80$, $p < .001$. The pricing manipulation also significantly influenced payments, $F(4, 704) = 12.89$, $p < .001$. Importantly, the effect of pricing was qualified by a significant interaction between the enactment of the bag law and the price, $F(4, 704) = 3.82$, $p = .004$. After the bag law, at every level of positive charitable contribution, people paid more than when 0% went to charity, $3.88 < t\text{'s} < 5.69$, but as can be seen in Figure 2.5, there were no additional differences between the SSR conditions.

The bag law requires customers to purchase a paper bag if they need a bag and do not have one. Consequently, the bag law increased purchase likelihood for the reusable fabric bags we were selling. The effect on payments were trickier to forecast, and as described above, somewhat difficult to interpret. We further discuss the bag law effect in the General Discussion.

Consistent with the results in Study 1, customers in Study 2 paid more even when only a trivial portion (1%) of their payment went to charity. The results of the two experiments were less consistent with respect to how different levels of charitable contribution influenced behavior. In Study 1 customers paid more as the contribution increased, whereas in Study 2 customers paid about the same regardless of the size of the charitable portion under SSR. We observe similar inconsistencies in consumers' scope-sensitivity in Studies 3 and 4. We lay out some of the potential explanations for these conflicting results in the General Discussion.

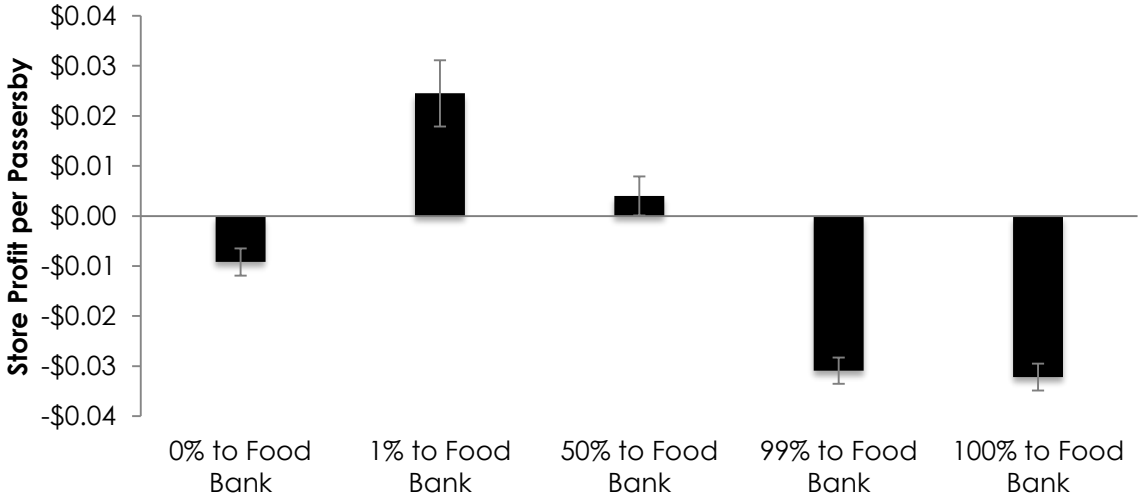
Figure 2.5. Mean Payment per Reusable Shopping Bag in Study 2. Error bars reflect ± SE of the means.



Profit per Passerby

As in Study 1, we analyzed store profit per passerby to account for selection. As shown in Figure 2.6, after deducting the charitable portion of payments and the full cost of bags, the sale of reusable bags were unprofitable for the store when charitable contributions were 0% (because payments failed to cover the cost of bags), 99% SSR (because 1% of revenue could not cover the cost of bags), or 100% SSR (because all revenue went to charity). The 50% SSR condition was significantly more profitable than the PWYW condition ($M_{0\%}$ = -$0.27 vs. $M_{50\%}$ = $0.18, $t(305)$ = 2.58, p = .010), though the profit per bag in this condition did not differ from $0, $t(124)$ = 1.01, p = .314. Finally, the 1% SSR with condition was the most profitable among the five pricing conditions ($M_{1\%}$ = $0.98 vs. $M_{50\%}$ = $0.18), t (254) = 2.57, p = .011, and the profit per bag per passerby in this condition was significantly higher than $0, $t(130)$ = 3.90, p < .001.

Figure 2.6. Store Profit per Passerby in Study 2. Error bars reflect ± SE of the means.



As in Study 1, we analyzed the charitable surplus including 86 cases of pure giving. The charitable surplus again increased as the size of charitable giving under SSR increased, $F(3, 21716) = 12.16$, $p < .001$.

While the results of Study 1 show some support for *the pure altruist account* (i.e. customers paid significantly more as the charitable contribution increased from 1% to 50%), *the impure altruist* and *scope-insensitive consumer* accounts better explain the pattern of consumers' behavior in Study 2.

Studies 1 and 2 offer general characterization of consumers' quantitative concerns under SSR. The next two field studies further investigate consumers' qualitative concerns in signaling their prosocial identity that might explain higher payments under SSR. First, we tested consumers' behavior in the absence of selection bias to test whether higher payments under SSR were purely driven by a particular kind of customers who choose to signal their generosity. In fact, the language of our accounts highlights the selection concerns: perhaps our manipulations are selecting for special kind of customers (e.g., impure altruists). If we want to generalize, we need to eliminate that potential selection effect. Second, we tested whether or not generosity under SSR is primarily explained by social pressure when customers' payments can be observed by others in the absence of selections.

Study 3: Randomized Charitable Giving and Sensitivity to the Scope of Charitable Giving

The results of Studies 1 and 2 might be driven by selection bias. That is, SSR merely attracts a higher spending segment and/or customers who wish to signal their generous identity by purchasing a product that benefits a charitable cause. Accordingly, we conducted Study 3 to remove this potential selection bias by randomly assigning customers to price (SSR v. PWYW) *only after* they had indicated that they wanted to purchase the product.

If higher payments under SSR are driven primarily by customers' qualitative concerns to signal their generous identity, then behaviors might change when customers are unaware of an opportunity to be generous when buying products. Study 3 aimed to assess whether selection effects fully explained observed differences in customers' behavior between PWYW and SSR.

In addition, as in Studies 1 and 2, we varied the percentage of customers' payments going to charity, which allowed us to test whether the *impure altruist* and *scope-insensitive consumer* account explained consumers' behavior in the absence of selection bias.

Lastly, we used coffee as our product in this study. It is plausible that a grocery bag that we used in Studies 1 and 2 attracted a particular kind of customers who care particularly about environmental issues. A souvenir photo such as the one used in Gneezy et al. (2010, 2012) is a personalized durable good, the value of which is restricted almost exclusively to the subject of the photo (and perhaps peculiarly devoted friends or family). Accordingly, it is plausible that such a personalized item activated a particular kind of altruism. Coffee, however, is far less likely to trigger any peculiar patterns of selfishness or selflessness and has a more generic, broader appeal.

Method

"Ola's Corner," a gourmet coffee vendor in the bay area, specializes in rare African coffee and tea and, among a few other locations, operates coffee stands at local farmer markets. We sold Ola's coffee at a farmers' market in Jack London Square in Oakland, California, using the existing infrastructure (i.e., the tent, signs, and carafes) that Dr. Ola used every weekend, but we completely replaced his staff with our research assistants. Customers (N = 193 individuals or

157 groups of individuals) in this study saw a sign advertising "Ola's Exotic African Coffees and Teas" and decided to buy coffee. To make sure that we had at least 50 transactions per condition, we used prior sales records and determined that three weekends would be sufficient for obtaining the minimum sample size goal of 50 transactions per condition. We ran the study on three consecutive Sundays in April and May 2012 approximately from 8:30 am to 2:30 pm. Note that we did not record the number of passerby in Studies 3 and 4 because we assumed that customers' purchase likelihood will not differ across randomized conditions (i.e., presumably, people who wanted to coffee approached our shop in all conditions).

We told customers who approached the stand asking for a cup of coffee that the price of their coffee would be determined by chance. Each customer was instructed to draw a folded piece of paper out of a tall opaque box. For groups, one person from the group drew a piece of paper, which contained our pricing manipulation. The price in the PWYW (0%) condition said, "today you can pay any price you want for a cup of Ola's coffee." In the 10% SSR condition the paper read, "today you can pay any price you want for a cup of Ola's coffee. 10% of your payment will benefit the Berkeley East Bay Humane Society." The 50% SSR condition was identical except for the stated percentage.

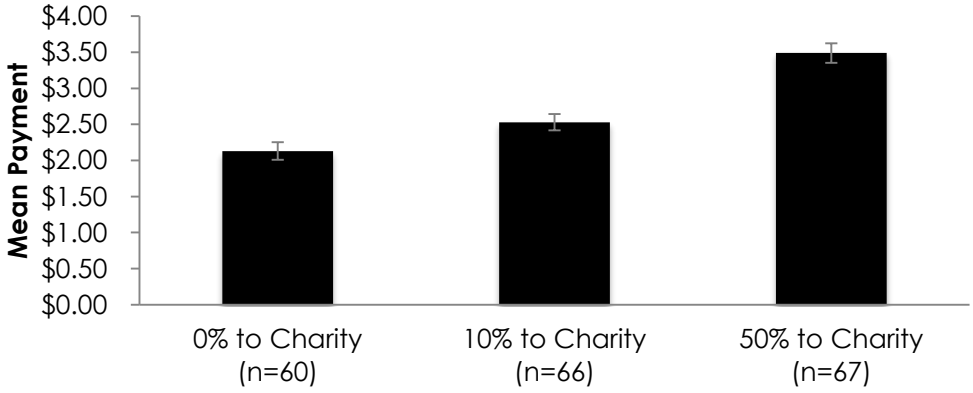
We asked customers to read out loud what was written on the paper and show it to the cashier (a research assistant)[8]. We recorded each customer's payment amount, the time of the transaction, the group size, the number of cups purchased per group, and immediately obvious demographic information as in previous studies.

Results and Discussion

Dr. Ola requested that we only sell coffee at pay-what-you-want prices but sell tea at fixed prices ($2 dollars for a small and $3 for a large). Therefore, customers who wanted tea were not exposed to our manipulation and tea sales were not recorded[9]. We conducted our analysis with each coffee cup sold as unit of analysis and the average payment per cup per person as a dependent variable[10]. Analyzing the data with group transaction as unit of analysis does not change the direction or significance of results. Controlling for the date variable did not change the direction or significance of the results. Alternative specifications are reported in Supplemental Materials, and all data and materials are available for download.

As shown in Figure 2.7, customers paid more for coffee when 10% of their payment went to charity than when 0% went to charity (*M*s= $2.53 vs. $2.13, t (124) = 2.25, p = .026), and paid more when 50% went to charity than when 10% went to charity ($M_{50\%}$ = $3.49 vs. $M_{10\%}$ = $2.53, t (131) = 3.18, p = .002).

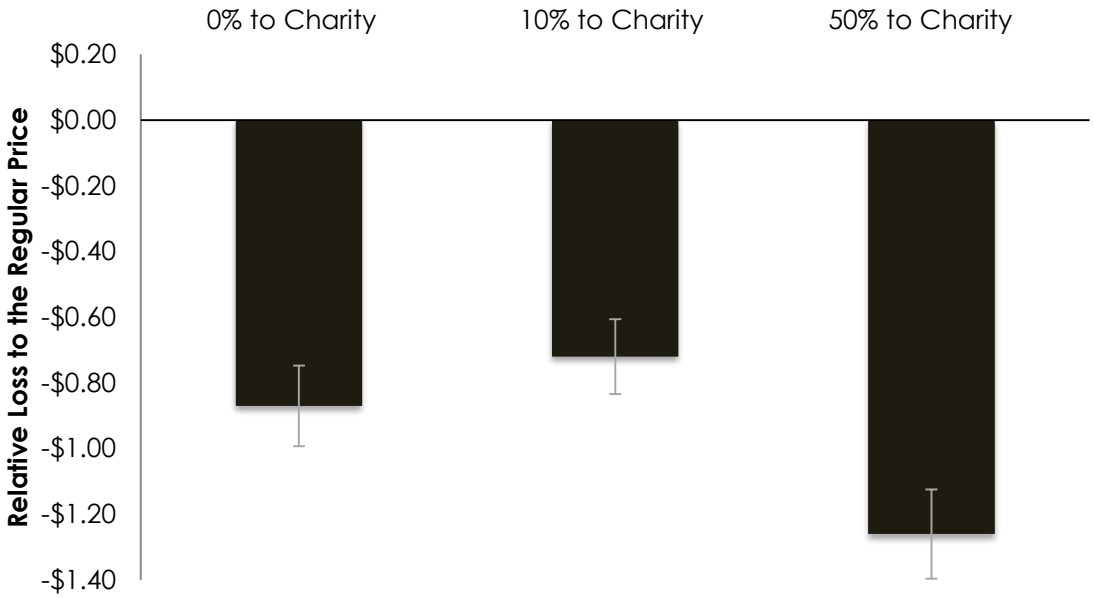
Figure 2.7. Mean Payment per Cup of Coffee in Study 3. Error bars reflect ± SE of the means.



Even in the absence of selection, customers paid more when any portion of their payments went to charity than when no portion did, and paid more when a higher portion went to charity, supporting the *pure altruist* accounts.

Dr. Ola typically charges three dollars for a large cup of coffee. As shown in Figure 2.8, after deducting the fixed price from the average payment of each condition, none of the three PWYW or SSR pricing was as profitable as the regular fixed price.

Figure 2.8. Relative Loss to the Regular Fixed Price per Cup of Coffee ($3) in Study 3. Error bars reflect ± SE of the means.



Note. The y-axis is the average payment per cup of coffee after deducting three dollars, the price the owner of the coffee shop was typically charging before the experiment.

These results suggest that selection bias cannot fully account for the higher payments observed under SSR versus PWYW. Customers decided to buy coffee because they simply wanted coffee[11]. Presumably, customers expected a typical fixed price and differentially compensated the company as it gave a higher amount to charity. It is also possible, however, that such a response was due to the social pressure from having to interact with the cashier or caring about other shoppers who might be observing their payment. If such concerns are driving higher payments under SSR, then allowing customers to pay anonymously should reduce payments. In other words, customers, especially those who would have opted out if they had known about the SSR pricing, might pay less if they were paying anonymously. Study 4 tested people's behavior under PWYW and SSR pricing when their payments were observed or anonymous in the absence of the selection effect.

Study 4: Sensitivity to the Scope of Charitable Giving and the Anonymity of Payment

Study 4 employed the same experimental paradigm as one in Study 3 with two modifications. First, we manipulated anonymity of payment; some customers paid directly to a cashier (as in Study 3), whereas others paid anonymously by placing their payment in an

envelope and placing it in a box located out of the cashiers' sight. Second, rather than gourmet coffee at a farmer's market, we sold glazed doughnuts at a UC Berkeley campus location that had high traffic of students and residents of Berkeley.

Method

In collaboration with Dream Fluff Donuts, a popular doughnut vendor in Berkeley, we sold glazed doughnuts during eleven days in September and October 2013. Our charity partner was the same as in Study 3. Our doughnut stand was located near Sather's Gate, a UC Berkeley campus location from 11am to 4pm during weekdays. People saw our sign that read "Dream Fluff Donuts" with a picture of three glazed doughnuts (see Figure 2.4.1 in Appendix). Once customers expressed interest in buying a doughnut, they were instructed to draw a folded piece of paper out of an opaque box, which contained the pricing manipulation. Customers were exposed to one of six conditions in a 3 (price: PWYW, 10% SSR, or 50% SSR) X 2 (payment method: anonymous or public) between-subjects design. For example, customers in the anonymous 50% SSR read, "Today, you can pay what you want for a Dream Fluff Donut! 50% of what you pay goes to [the name of charity]. (Your payment will be entirely *anonymous*.)" In the anonymous payment conditions, customers were provided with an envelope in which they were asked to put their payment, and drop it in a box located a few feet away from the table, but out of sight. We recorded the same information as in Study 3.

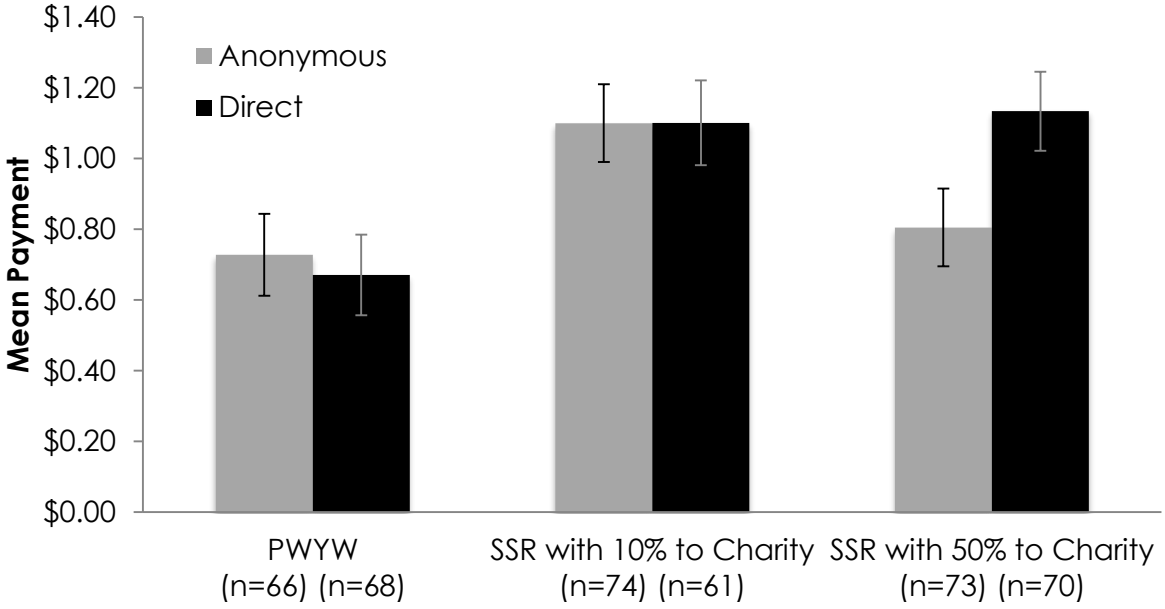
Results and Discussion

537 individuals (418 groups) approached our doughnut stand and expressed interest in purchasing a doughnut.[12] We determined our sample size so that we have at least 50 transactions per condition. Accordingly, based on the first week's sales records, we decided to conduct the study for eleven days. We considered the data with doughnut sold per person as unit of analysis (the same analysis with doughnut sold per group did not change the direction or statistical significance of the results). We report alternative specifications in Supplemental Materials, and all data and materials are available for download. We excluded 14 transactions (doughnut sold per person) because customers knew one of the experimenters. These customers were identified as they approached the stand and were logged as non-transactions before drawing from the box. We also excluded 10 cases in which group members selected separate pieces of paper from the box and were therefore exposed to different experimental condition or did not select a piece of paper for price[13]. Finally, we excluded 37 cases in which people decided not to purchase after selecting a piece of paper[14]. These exclusions left 412 doughnuts sold for our analysis. Our primary dependent variable was the average payment per doughnut per person.

We submitted customers' payments per doughnut to a 3 (price: PWYW, 10% SSR, or 50% SSR) X 2 (payment method: anonymous vs. direct) between-subjects ANOVA. As predicted, the main effect of Price was significant, $F(2, 405) = 6.01$, $p = .002$. Neither the main effect of Anonymity, $F(1, 405) = .89$, $p = .346$, or the interaction between Price and Anonymity was significant, $F(2, 405) = 1.76$, $p = .173$.

As shown in Figure 2.9, we partially replicated the results from Study 3. Considering only the direct payment condition (the payment method used in Study 1), we observe that people paid more when 10% went to charity than when 0% went to charity ($M_{SSR\text{-}10\%/Direct}$ = \$1.10 vs. $M_{PWYW/Direct}$ = \$0.67), $t(129) = 2.87$ $p = .005$). However, unlike in Study 3 we did not observe any difference in payments between 10% and 50%, ($M_{SSR\text{-}10\%/Direct}$ = \$1.10 vs. $M_{SSR\text{-}50\%/Direct}$ = \$1.13), $t(129) = -.17$ $p = .867$.

Figure 2.9. Mean Payment per Donut Sold in Study 4. Error bars reflect ± SE of the means.



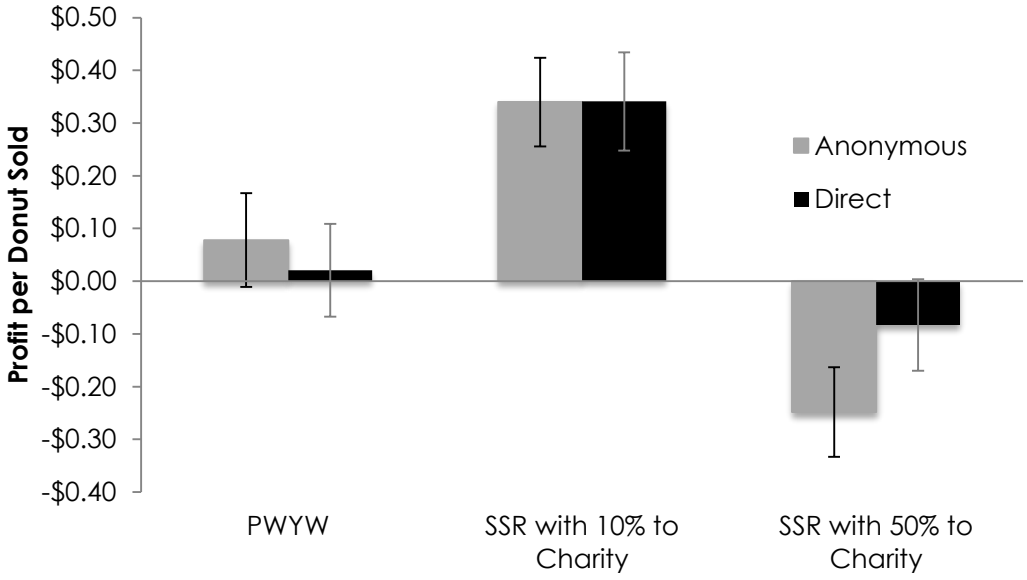
When paying anonymously, people paid more in the 10% SSR than the PWYW condition, ($M_{PWYW/Anonymous}$ = \$0.74 vs. $M_{SSR\text{-}10\%/Anonymous}$ = \$1.10), $t(136) = -2.38$ $p = .019$, but they did not pay more when 50% went to charity than when 10% did. In fact, people paid slightly less when 50% of their payment went to charity ($M_{SSR\text{-}50\%/Anonymous}$ = \$0.80 vs. $M_{SSR\text{-}10\%/Anonymous}$ = \$1.10), $t(144) = -1.81$ $p = .073$, or no portion went to charity, ($M_{PWYW/Anonymous}$ = \$0.73 vs. $M_{SSR\text{-}50\%/Anonymous}$ = \$0.80), $t(136) = -.52$ $p = .602$. We are reluctant to read too much into the latter effect, since it was neither predicted nor statistically significant.

Although we largely removed selection bias by randomly assigning customers to different pricing conditions, some customers (n=37) chose not to purchase after they draw a piece of paper from the box, which allowed another level of selection. We tried to address the selection concern with two additional analyses. First, assuming that those 37 people would want to pay zero, we reanalyzed the data by entering zero for their payments. Second, we assumed that these people could be sufficiently pressured to pay a "reasonable" amount. Accordingly, we entered the median payment, one dollar, for their payments. The results from the both analyses were consistent with our main analyses that excluded these transactions: People paid more when a larger portion went to charity, $5.51 < Fs < 6.19$, anonymity did not influence their payment decisions significantly, $.79 < Fs < 1.16$, and the interaction of these two variables were not significant, $.87 < Fs < 1.79$. Although these analyses are based on simple assumptions about how those who selected out of transaction could have behaved, it seems unlikely that selection bias fully explained our main results.

We interpret these results as providing two important insights. First, when customers were paying directly, we saw a replication of the main finding of Study 1; people pay more under SSR even when self-selection is eliminated. Second, this pattern is not significantly moderated by anonymity, suggesting that payments under SSR cannot be solely attributed to social pressures, as anonymity did not significantly lower payments.

We paid Dream Fluff Donuts $0.65 per donut. As shown in Figure 2.10, per doughnut profits (after deducting the charitable contribution and donuts' cost) was substantially larger in the 10% SSR than profit in the PWYW and 50% SSR conditions.

Figure 2.10. Profit per Donut Sold Condition in Study 4. Error bars reflect ± SE of the means.



Note that this study did not replicate the result of Gneezy et al. (2012) in which restaurant diners paid significantly more when paying anonymously than when paying directly to the manager. We can only speculate about the difference, but it is worth noting that the restaurant in Gneezy et al. (2012) contained only patrons who had chosen the restaurant, presumably because it uses pay-what-you-want pricing. Perhaps that self-selected population is particularly sensitive to self- and social-signaling pressures of that situation.

The combined the results of Studies 3 and 4 suggest that selection bias cannot fully explain higher payments under SSR. Instead, our data shows that an individual choosing to make a purchase adjusts his or her payment when a charity is a partial beneficiary. The results were more ambiguous about the extent to which customers are sensitive to the size of the charitable contribution. Finally, it is important to note that while the design of Studies 3 and 4 allowed us to eliminate a self-selection effect, it prevented us from observing differences in purchase likelihood.

GENERAL DISCUSSION

In theory, all consumers should want the bargains implicit in a pay-what-you-want price. If the same consumer also wanted to support a charity, then the bundling of SSR would be something of an improved offering. Nevertheless, in combination, these two motives are unappealing to consumers. Our research considers how these competing motives influence decisions in an environment that offers a diagnostic opportunity to be generous while paying any price for a good. Four field studies varied the charitable percentage in order to quantitatively assess how the charity was operating on behavior. All studies revealed that consumers are significantly influenced by whether or not some percentage was going to charity, but all studies also revealed that they were comparatively less sensitive to differences in how much was going

to charity. Customers pay more for a reusable shopping bag even when a trivial portion (i.e., one percent) of their payment goes to charity. As that percentage rises to 50% there is a mild increase in payment (Study 1), and then no increase in payment as that percentage increases to 100% (Study 2). In summary, Studies 1-4 suggest that consumers are largely *scope-insensitive*; in both payment amount and purchase likelihood, consumers were sensitive to the presence of charity but largely insensitive to how much of their payment went to charity.

Our studies investigated behavior with a qualitative opportunity to look good along with a quantitative means for showing it. Study 3 considered whether higher SSR payments were merely the result of a selection effect; that is, only the most generous customers decide to buy, and the resulting purchase prices are therefore higher. Selection was not the whole story, however, as SSR effects persisted even if the prices were presented only after the decision to buy (Study 3). A subsequent study additionally manipulated whether customer payments were privately determined (i.e., an anonymous payment) or public (i.e., a direct payment). Again, people paid more when some percentage went to charity, but that tendency was largely unaffected by whether or not the payment was public. These results suggest that higher payments under SSR are not driven either by a particular segment of market choosing to buy a SSR product or social pressure to look generous to others.

Alternative Explanations For Consumer's Scope-insensitivity

Although we think that our results are best explained by consumers' scope-insensitivity, there are two alternatives that may contribute to the observed patterns of the results. The first alternative is that the observed scope-insensitivity is partially a result of a ceiling effect. Namely, coffee, doughnuts, and fabric shopping bags are all relatively low retail value goods, so payments hit an implicit ceiling at a relatively low magnitude. Certainly it might be the case that consumers will show more scope-sensitivity when buying a higher retail value product (e.g., a blender). It is also entirely possible that the actual retail value of a product is irrelevant to the level of consumers' generosity.

The second possibility is that the pattern of consumers' behaviors in our studies reflects the behaviors of a particular population in the Bay Area where we collected our data. Although our studies show consistent patterns of consumers' generous behavior in the elective price settings using three different products in four separate locations, more research could investigate whether our findings are generalizable for products with high retail values and in different socio-economic and cultural environments. Although these alternatives offer plausible explanations for some of the results, no single explanation other than our primary account would seem to account for all of them.

We observed the data are somewhat inconsistent; the *pure altruist customer* seemed to drive payments in Studies 1 and 3, whereas the *impure altruist* or *scope-insensitive consumer* seem to have guided behavior in Studies 2 and 4. That is, people are sometimes sensitive to the size of their charitable contribution but sometimes they are not. The two alternative possibilities described above could explain these inconsistencies. First, it is possible that consumers in Studies 1and 3 represent particular segments of the Bay Area consisting of those who are sensitive to the charitable cause and the size of charitable giving. Alternatively, the customers in Studies 2 and 4 might represent relatively lower income segments with lower ceilings for payments. These are only speculations. More research could contribute to better understand what influences consumers' quantitative sensitivity to prosocial signals in different domains.

Consumer Behavior and Changes in Legislation

Although the bag law did not change the direction or significance of our main results, it substantially influenced the magnitude of the payment amount and purchase rate. Consumers were much more likely to buy a bag but paid much less for it. When paper bags became no longer free, the bag law increased the demand for our reusable bags. This boost in purchase rate, however, came mainly from the non-charity condition and the 99% charity condition (see Figure 2.4). It seems that the bag law did not alter the level of discomfort consumers felt in purchasing a bag when their payment was linked to a charitable cause.

Why did customers pay less for a bag after the legislation? One possibility is that a $0.10 increase in the price of a paper bag could ostensibly produce an increase in the (now relative) perceived value of our fabric bags, and would therefore increase payments (consistent with Kim et al. 2009). An alternative explanation is that the bag law functionally shifted people from pricing in a social market (i.e., "what is the right price to pay for an environmentally conscious bag?") to pricing in a money market (i.e., "what is the right price to pay to get a good deal on bags in this market?"(see Fiske 1992, Gneezy and Rustichini 2000, Heyman and Ariely 2004). Despite of the environmentally-friendly purpose of the law, requiring an explicit fee for a substitute product might have shifted how people estimate the price of an environmentally-friendly product. Another possibility is that,the $0.10 alternative cost could have provided a stable reference price that guided customers' payment decisions. Indeed, as can be seen in Tables 2.1 and 2.2, mean payments were lower after the bag law, and consistent with the reference price interpretations. Those payments also showed substantially less variance.

We would be remiss to suggest that the bag law was the only difference between 2012 and 2013. Most notably, the 2012 sales occurred during the holiday season, which might have contributed to a more general charitable spirit and higher overall contributions, though these seasonal effects do not offer a sufficient explanation for the increase in purchase likelihood post the holiday season in 2013. It is quite possible that the effect is multiply determined by one or more relevant social and economic variables. This particular phenomenon is challenging to study in a randomized experiment in the real world setting (i.e., we opportunistically utilized a naturally occurring legislation change). Future research could separate some of the potential social forces influencing of the environmental legislation on consumers' prosocial behavior.

Selection and Image Concerns

Consumers' image concerns could explain the low purchase rates under SSR; some consumers opt-out because they think the right price is higher than what they want to pay (Gneezy et al. 2012). It is important to note, however, that such an interpretation hinges on our assessment that the charities involved are generally well liked and comfortable to publicly endorse. Imagine instead that we had chosen more polarizing charitable partners (e.g., National Rifle Association in Ariely, Bracha, and Meier 2009). Under those circumstances, it seems likely that different groups of people would choose to buy the products not so much to signal their generosity, but rather to signal their political identity. Accordingly, although the local county food bank seemed like both a likeable and innocuous cause, it is easy to imagine that a different partner might have changed behaviors. For example, if we had partnered with the Planned Parenthood Action Fund, a non-profit organization that lobbies for pro-choice legislations, the selection effect might be stronger, attracting customers who strongly supported the cause. And those who chose to buy might have been more sensitive to how much of their payment went to charity.

SSR and Corporate Social Responsibility

Despite the obvious financial risks, SSR seems promising for sustainable company profit and sustainable support of social welfare. One competitive advantage of SSR over a conventional CSR program (e.g., fixed pricing and a portion going to charity) may be the strength of its prosocial signal. Conventional CSR programs are typically administered as the firm's short-term expense or an investment the firm hopes to harvest in the long run. Although this approach is certainly not risk-free, the company typically does not bet its entire financial performance on conventional CSR strategies. Furthermore, conventional CSR programs might be perceived as a separate communication and easily overwhelmed by the firm's core business message of selling products. In contrast, SSR is tightly integrated into the practicing firms' core business strategy, essentially informing consumers that their generosity drives the firms' bottom-line success. Therefore, compared to conventional CSR strategies, SSR is likely to be more effective in communicating its prosocial commitment and elevating the public image of the practicing firm.

SSR may differ from conventional CSR in how the "fit" or consistency between the company's brand and its prosocial cause influences the strength of its prosocial signal. Prior research suggests that a high level of consistency is more effective than a low level in enhancing consumers' perception of the company's prosocial goals and their purchase intent (Albert and Whetten 1985, Fein 1996, Hoeffler and Keller 2002, Becker-Olsen, Cudmore, and Hill 2006). Patagonia, for example, a company that specializes outdoor clothing gives 1% of total sales to environmental campaigns to prevent oil drilling in Alaska Wildlife Refugee. In this case, there is a strong fit or consistency between the company's outdoorsy brand image and its non-profit partners' pro-environmental initiative. Patagonia's customers may perceive this partnership to be appropriate and think that the company is capable in achieving its prosocial goal (Wojciszke, Brycz, and Borkenau 1993). Alternatively, some research argues that a strong consistency or an overlap between the company's established brand image and its charitable cause may dilute prosocial signals of the company's CSR program. Furthermore, consumers may simply attribute the CSR programs to the company's opportunistic or self-interested motive (Drumwright 1996, Ellen, Mohr, and Webb 2006, Foreh and Grier, 2003). In concept at least, SSR would seem to be largely immune to these constraints. A food bank might be a good charitable "fit" with a grocery store, but an animal protection group is certainly less so. Consumers behaved similar in both situations. On the other hand, there is room to think that a pairing that was too close (e.g., selling reusable bags and partnered with an environmental protection group) might begin to suggest a profit-seeking manipulation from the company. This is an open–and plausibly interesting–question for firms and charities to answer as they seek to apply SSR.

A serious concern for CSR is the long-term dilution of its appeal. When a firm first introduces its partnership with a charity, it captures consumer attention and hopefully successfully conveys pro-social intentions in the firm. Over time, however, more companies offer CSR and fewer customers notice, the benefits are lost. This concern almost certainly could apply to SSR as well, since its effectiveness relies entirely on customers being sensitive to its presence. Nevertheless, SSR has held a persistent presence for a number of firms. HumbleBundle, StoryBundle, IndieGala, Vodo.net, are all digital media firms that employ pay-what-you-want pricing, and all of them employ some version of SSR. The positive interpretation of that consensus would be that all have found a way to increase profit through the support of charity. The negative interpretation might be that a company can no longer compete in this market *without* giving money to charity.

Companies must also be concerned with changes in individual customer payments over time. Perhaps customers who were generous in the past might feel morally licensed to be a bit more selfish in the future (Monin and Miller 2001). There are some indications that SSR might be resistant to this dilution effect. HumbleBundle, for example, has used an SSR strategy and publicly reports their general sales figures; after three years, they are not seeing decreases in payments despite many of the customers being repeat purchasers. Additionally, we have analyzed individual level data from an online retailer that has sold bundles of products approximately once a month for the past two years. That firm adopted a variant of SSR (i.e., customers can choose to give 10% of their payment to a charitable cause). When tracking individual customers we can observe whether there is decline over time. There is not. Customers pay very similar amounts over time, and when compared to new customers they are similar indistinguishable overall. In combination these observations suggest that SSR may offer a sustainable long-term appeal.

Although we speak optimistically about the long-term viability of SSR, we are less sanguine about the ability of a firm to switch away from that model to either fixed pricing or standard PWYW. Either change would be noticeable, less appealing for customers, and could potentially undermine the company's public persona. The company offers its products to customers for any price and gives some of its revenue to charity will always be likeable; but the company that reverses course on those offerings will likely just seem stingy.

SSR will not work for all companies. The present work has emphasized how people respond to even a mild charitable signal, but a little bit more does not have any additional influence. We would therefore predict that a company that already was publicly contributing to charity might see less of a boost from the SSR strategy. Alternatively, consider a company firmly on the other side of the prosocial spectrum. Although SSR might alleviate consumers' distrust toward the company's motive, SSR programs may be ineffective or even backfire when adopted by companies that are perceived to be unethical (e.g., oil or cigarette companies) or purely profit-maximizing (i.e., investment banking companies). Consumers may perceive these company's SSR motives to be insincere or self-interested (Osterhus 1997, Strahilevitz 2003, Webb and Mohr 1998), and contribute little to its SSR program to punish the company.

In our studies SSR was often profitable especially when a lower percentage of customers' payments went to charity. Despite the obvious implication of these results for the practicing companies' bottom-line, we do not claim that adopting a trivial size is a suitable strategy. While a small charitable portion is more helpful than none, a company' SSR program with a trivial charitable presence could have the unintended effect of signaling an unethical or selfish motive in the long-run, potentially threatening the company's sustainability.

In this chapter, we considered charitable giving as a social stimulus to test how emotion operates to influence consumers' decisions under elective pricing setting. In Chapter 3, we take this approach to the next level to test how consumers' concerns toward others operate with a new stimulus that is more symbolic and less heavy-handed than charitable giving. Chapter 3 includes a series of field and lab studies that investigated how a mere consideration of "others" influences consumers' perception of social norms and their subsequent economic decision-making.

CHAPTER 3: PAYING MORE FOR OTHERS

Abstract

Social behavior is heavily influenced by the perception of the behaviors of others. We consider how perceptions (and misperceptions) of kindness can increase generosity in economic transactions. We investigate how these perceptions alter behavior in a novel a real-life situation which pits kindness against selfishness. That situation, *consumer elective pricing*, is defined by an economic transaction allowing people to purchase goods or services for any price (including zero). Field and lab experiments compared how people behave in two financially identical circumstances: pay-what-you-want (in which people are ostensibly paying for themselves) and pay-it-forward (in which people are ostensibly paying on behalf of someone else). In four field experiments people paid more under pay-it-forward than pay-what-you-want (Studies 1-4). Four subsequent lab studies assessed whether the salience of others explains the increased payments (Study 5), whether ability to justify lowered payments (Study 6), and whether the manipulation was operating through changing the perceptions of others' (Studies 7 and 8). When people rely on ambiguous perceptions, pay-it-forward leads to overestimating the kindness of others and a corresponding increase in personal payment. When those perceptions are replaced with explicit descriptive norms (i.e., others' payment amounts), that effect is eliminated. Finally, subsequent studies confirmed that the effects were not driven by participant confusion (Studies 9A and 9B) and not limited by the specificity of the referent other in the PIF framing (Study 9C).

People are self-interested, but they can also be surprisingly generous toward others. As economists and psychologists note, even in environments that promote material self-interest, people are frequently kinder than purely self-interested, conforming to the norms of fairness and reciprocity (Andreoni & Miller, 2002; Charness & Rabin, 2005; Falk & Fischbacher, 2006; Fehr & Schmidt, 1999; Rabin 1993). While people look to social norms to guide their behavior, social norms can be often ambiguous in unfamiliar or uncertain social contexts. In these situations individuals' beliefs about the behavior of others could be influential (Cialdini, Kallgren, & Reno, 1991; Cialdini & Trost, 1998; Sherif, 1936). Present research considers how social forces can increase kindness by influencing the perceptions of others. In particular, we investigate how implicit information about others' generous behavior influences the level of generosity. Our findings operate in a relatively narrow domain (consumer elective pricing), but as we identify throughout the paper, this domain offers opportunities to capture changes in kindness despite meaningful financial incentives pushing in the opposite direction.

People look to the behavior of others to decide how to behave themselves. They use others to learn about the relevant social norms (Cialdini et al., 2006; Cialdini, Reno, & Kallgren, 1990; Goldstein, Cialdini, & Griskevicius, 2008). Social norms are defined as the "rules and standards that are understood by members of a group, and that guide and/or constrain social behaviors without the force of laws" (Cialdini & Trost, 1998). Although a single situation may be guided by different norms, people follow those focal in their attention (Reno, Cialdini, & Kallgren, 1993). The simplest of norms are descriptive norms, which contain information about the behavior of others (Cialdini & Trost, 1998). Descriptive norms simplify social decision making by indicating justifiable course of action even under powerful uncertainty (Deutsch & Gerard, 1955; Sherif, 1936; Tesser, Campbell, & Mickler, 1983). Accordingly, when descriptive norms are made explicit, they can be very influential. People litter more in a littered environment presumably because they have learned that other people are littering (Cialdini et al., 1990; Keizer, Lindenberg, & Steg, 2008), and they hang-up their towels more often when they find that other guests are hanging up their towels too (Goldstein et al., 2008). Beliefs about others influence how we decide to behave ourselves.

People are also influenced by their interpretations of others' behavior, but that means that they may be vulnerable to their own misinterpretations. Pluralistic ignorance is one way in which people misunderstand social norms. People see a common or consensus behavior in others and infer an underlying social norm, but sometimes that inference is inaccurate. Under pluralistic ignorance, everyone might behave similarly, but an observer believes that the observed are doing so out of adherence to a different norm (Miller &McFarland, 1987). Accordingly, people will follow the behaviors of others even when their own attitudes and judgments (falsely) feel quite different (Allport, 1924; Prentice & Miller, 1993).

With such a potential misperception in mind, we ask, how does social perception influence personal prosocial behavior? Do people correctly estimate the level of others' generosity (or selfishness) in economic exchanges? A small number of studies have investigated people's perceptions of others' prosocial behaviors in contexts where their pure self-interest motive is relevant and salient. Some studies suggest that people overestimate underlying self-interest, for example, inferring that others donate blood for financial incentives (Miller and Ratner, 1998). Other studies start with a similar bias (other people are more selfish than me), but find error in self-perception; the observers are just as selfish as the people they observe (Epley & Dunning, 2000). When it comes to prosocial spending, people think that others are less likely to spend and choose to spend less.

When the prosocial trappings are stripped away, people spontaneously judge others to be much more willing to part ways with their money. When estimating how much others would be willing to pay for goods, people consistently believe that others would pay more (Frederick, 2012). In combination with the above observations, we can see that there are substantial misperceptions of others, but the direction of that misperception is substantially guided by context. This paper considers behavior in a context that is somewhat prosocial and somewhat commercial. As we will demonstrate, that combination allows for a mixture of misperception and social influence that prompts more generous behavior.

Consumer Elective Pricing: Pay-What-You-Want and Pay-it-Forward

The context for our investigation is consumer elective pricing. Consumer elective pricing, as we define it here, is any commercial transaction in which the buyer can pay any price for a good or service. Consumer elective pricing offers an opportunity to test the extent to which people deviate from pure self-interest in transactions that are both commercial and social. Maximizing immediate self-interest would move people to pay zero, whereas considering others' welfare (e.g., sellers), would push them to pay more. Furthermore, consumer elective pricing provides a conservative setting for evaluating generous behaviors because they necessarily come at directly measurable personal costs (i.e., they are incentive compatible).

The most well documented form of consumer elective pricing is pay-what-you-want (PWYW) pricing. PWYW has received much popular and academic attention. A precipitating event came in the release of (British popular music artist) Radiohead's album, "In Rainbows", in 2007. The band released the album as a pay-what-you-want download, and was rewarded with one of their most frequently purchased albums. Since Radiohead's PWYW album made headlines, more independent musicians (e.g., Girl Talk, Amanda Palmer) are adopting PWYW. But the applicability of PWYW goes beyond the marginal or "creative" market territory. For instance, Panera, a large restaurant chain, also opened PWYW cafes for soups and sandwiches. As mentioned earlier, HumbleBundle attracts millions of customers, sells many categories of products (e.g., video games, digital music, and ebooks), and is hugely profitable with a PWYW pricing set-up. Even if most customers are self-interested, there is enough social kindness left over to sustain company profitability.

Some studies have investigated this surprisingly generous behavior (e.g., Mak, Zwick, & Rao, 2010; Kim, Natter, & Spann, 2009; Regner & Barria, 2009; Regner & Riener, 2012; Schmidt, Spann, & Zeithammer, 2012). Such generous behavior is particularly surprising given that it involves explicitly financial exchanges which otherwise increase selfishness (Vohs, Mead, & Goode, 2006). PWYW transactions invoke concerns of reciprocity and fairness (e.g., Kim et al., 2009), suggesting that people feel obligated to pay, even when offered an opportunity not to. An important motive for buyers is maintaining or burnishing their self-image and self-identity in making decisions about whether to buy, and how much to pay (Gneezy, Gneezy, Nelson, & Brown, 2010; Gneezy, Gneezy, Riener, & Nelson, 2012). The latter set of studies show that self-image concerns can push people to pay more, but it may also make them less likely to buy so as to avoid the scrutiny of the situation altogether.

In these situations, how are people influenced by their beliefs about the behaviors of others? PWYW transactions are direct exchanges between a buyer and a seller, invoking concerns of reciprocity and fairness. Because the "appropriate" price is often inherently ambiguous, a useful point of reference might be to know what others pay. When that information

is not available (as it most often is not), people make guesses. Will those guesses change when different norms are highlighted, as in the case we will investigate, when payments are implicitly linked with the behavior of other customers, not just with the seller? Because descriptive norms can influence prosocial behaviors (e.g., Goldstein et al., 2008), if people judge others to be generous, that descriptive norm may induce more generosity.

We find such a framing manipulation in a close cousin of PWYW: Pay-it-forward. Under pay-it-forward (PIF) pricing, people are still given the opportunity to electively choose any price they want (including zero), but the payment is treated differently. Customers are told that their product has been paid for by a previous customer, and that their payment will be on behalf of someone else who comes later. PIF is less common than PWYW, but it does exist. To give one example, diners at Seva Café, a restaurant in Ahmedabad, India, are told that a previous guest paid for their meal as a gift, and they have a chance to make a similar gift for a future guest. Another related example is a movement of "suspended coffee" cafes in various European cities. At these coffee shops people can choose to pay for two (or more) cups of coffee, one for themselves and the other, "suspended" cup for anyone who wants it (Poggioli, 2013).

From the perspective of the seller, PIF pricing is financially identical to PWYW pricing: all customers receive a good and choose the price they want to pay, and all of that payment goes to the seller. However, a PIF framing transforms the direct reciprocal relationship between the buyer and seller under pay-what-you-want pricing to a symbolically social relationship with other customers; the receiver and giver of a gift. In this way the direct exchange with the seller also takes on a symbolic social exchange.

In this research we compare behaviors under the two forms of consumer elective pricing: PWYW and PIF. We expect that the forces influencing PWYW behavior will also influence PIF behavior. People will seek to be fair to sellers and reciprocate their efforts, and to maximize their self-image and social-image. Under PIF, despite facing an identical financial bottom line, those same forces might change qualitatively or quantitatively.

How Will Others' Behavior Influence Payments in Consumer Elective Pricing?

Although, as we describe below, our research was guided by an effort to consider (and discard) alternative explanations, our core prediction relies on the findings described above. PIF, by dint of its definition and implementation, encourages people to think about the payments of others. Furthermore, because part of consumer elective pricing is *commercial*, people think that others are paying a lot (Frederick, 2012). However, because part of consumer elective pricing is *social*, people are especially guided by the norms of others. In combination, these forces could potentially combine to lead people to increase payments under PIF relative to PWYW. Our first four studies demonstrate that relationship in field experiments.

We then present and test four possible explanations for why people might pay more under PIF than PWYW: *salience of others, differential weights for reciprocity and generosity, payment as justification,* and *generosity matching*. Each of these accounts is further detailed in the section after our report of the four field studies (Studies 1-4) but here we briefly describe these mechanisms.

Our first account, *salience of others,* predicts that people are more generous under PIF than PWYW because the PIF framing makes others more salient and present than the PWYW framing. This explicit reference to other customers may increase the pressure to reciprocate or look generous. *Differential weights for reciprocity and generosity* predicts that people pay more under PIF than PWYW because they might be basing their payment on the opportunity to be generous than social pressures of having to reciprocate the seller. Our third account, *payment as*

justification, is that people pay because they feel that they need to justify their payments more under PIF. Lastly, *generosity matching* predicts that people perceive a higher level of generosity in others under PIF than PWYW and pay more to match their perception of others. We test whether one or more of these accounts explain why people pay more under PIF than PWYW in Studies 5-8.

Consumer elective pricing provides a rich setting for evaluating generous behaviors because they can be tested both in the lab and the field. While lab experimentation is useful for testing the psychological variables in a controlled setting, it provides an abstract environment, for which behavior may systematically differ from behaviors in the field (Cialdini, 2009). Accordingly, we will use field experiments to identify the phenomenon we are investigating, and follow up with lab experiments to better understand the psychology of the phenomenon.

THE OVERALL PLAN OF THE STUDIES

The paper is roughly divided into two portions. The first goes to lengths to establish the reliability and generalizability of our primary prediction that people will pay more under PIF than under PWYW. We conducted an initial field experiment with a museum, randomly assigning visitors to receive either a PWYW or a PIF message (Study 1). We then conducted nearly exact replications at the same field setting, confirming customers' understanding of the pricing (Study 2) and the financial interpretation of the results (Study 3). We then moved to a different field setting (a gourmet coffee vendor) to test whether the effect could generalize to a very different setting (Study 4).

The second part of the paper reports our efforts to understand why people pay more under PIF. We develop a laboratory paradigm in which actual goods are exchanged for actual payments, and that also allows us to manipulate psychologically meaningful moderators, and observe their influence. We test how participants' payments are influenced by the identifiability of the giver and receiver (Study 5), the ability to communicate with the next participant (Study 6), and knowledge about the behavior of the previous participant (Studies 7 and 8). A final study includes three additional experiments in which people predicted their own behavior or the behavior of others under different articulations of the central manipulations (Experiments 9a-9c). These experiments serve to rule out some plausible confounds and provide some additional insight into the generality of the basic relationship.

STUDY 1: A Field Experiment in Museum Admission Payment

Method

We conducted a field experiment at the Cartoon Art Museum (CAM) in San Francisco. The CAM showcases 6,000 pieces of original cartoon and animation art, is located in a central part of the city, and charges approximately $7 on regular admission days. It has been hosting a Pay-What-You-Wish Tuesday on the first Tuesday of every month for more than ten years. Groups of participants (N=151 groups) in this study were the individuals who visited the museum on the first Tuesdays of September and October in 2011 from 11am to 4:30pm. We predetermined the two month window, and did not analyze any data until the completion of the experiment. For this study, and all that follow, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study (Simmons, Nelson, & Simonsohn, 2011).

During the experiment our team of experimenters functionally took over the staffing of the museum. As detailed below, they handled the random assignment, presented the manipulated payment request, administered all transactions, and recorded responses.

All visitors were assigned to one of two conditions. Participants in the PWYW condition were told "Today is a Pay-What-You-Wish Day. You can pay what you want for your admission. How much would you like to pay?" We used "Pay-What-You-Wish Day" instead of "Pay-What-You-Want Day" because that is the term the museum has typically been using in its promotional materials. Participants in the PIF conditions were told, "Today is a Pay-What-You-Wish Day. A visitor who came earlier paid for your admission. Since you are paid for, you now have a chance to pay forward the admission for another person who will come later today. How much would you like to pay forward for another person's admission?" The research assistant at the reception desk greeted visitors as they entered the museum and delivered the manipulation. When people arrived in groups (58% of the time), everyone in the group was assigned to the same condition. Conditional assignment was determined as follows. The first group of the day was randomly assigned to one of the two conditions. We then alternated after every 10 groups of visitors. This was occasionally extended to 11 or 12 consecutive groups if the groups were immediately adjacent in line because we did not want to risk contamination across conditions. Visitors trickled in throughout the day, so this was a rare occurrence. We had three large groups that consisted of more than 10 people per group. They were 36 children from a local school, 11 people from a non-profit organization for the mentally disabled, and 17 people from a halfway house. We did not communicate with each individual in these groups but gave our pricing manipulation to the group leaders whom they did not share with their group members. The analysis we report here included these people, but excluding these groups changes neither the direction nor the statistical significance of the effect[15].

Data Analysis Strategy. We recorded each participant's payment amount, the time he or she entered the museum, the number of people in a group, and immediately obvious demographic information (i.e., the gender composition of the group, and their approximate ages and ethnic backgrounds)[16]. We predicted that people would pay more under PIF pricing than under PWYW pricing.

The correct specification of the analysis was not immediately obvious. People frequently come in groups, so group-level analysis seemed appropriate. Of course, not all groups are of the same size, and group size could (and does) have a substantial influence on payment amounts. Basically, bigger groups paid more than smaller groups since more people were gaining admission. Group size also could (and does) have a perverse effect akin to social loafing (Freeman, Walker, Borden, & Latané, 1975; Lynn & Grassman, 1990; Seiter & Weger, 2010); as groups get larger there is a bit more anonymity and a little more selfishness. These concerns suggest that the size of the group should be accounted for in the analysis. However, some groups were composed of individuals who each chose to pay separately (and whose payments were independent of each other), and some groups submitted one payment, but only after each person made an individual contribution to the person directly paying the receptionist. These suggest that an individual-level analysis might be reasonable.

We decided to focus on the following specification: payment-per-person, with the group as the unit of analysis. Accordingly, that is the analysis we report in text for this study and the subsequent field studies. Nevertheless, because that decision is at least somewhat arbitrary, Table 3.1.2 in Appendix reports the results of the alternative specifications. As that table reveals, the

specification influences both the size and the statistical significance of the effects. Nevertheless, regardless of the specification, the findings remain supportive of the central hypothesis.

Results and Discussion

Groups paid more under PIF than they did under PWYW (M = $2.67 per person vs. $1.82 per person; $F(1, 149) = 4.74$, $p < .031$). For alternative specifications of the analysis, see Table 1.2 in Supplemental Materials.

Despite extremely similar wording and identical financial implications, people paid more under PIF than under PWYW. Although consistent with our prediction, we wanted to see if the pattern of results would replicate. Furthermore, when discussing this finding with colleagues some months later, someone raised the possibility that our participants might have simply misunderstood the payment description[17]. Specifically, it may have been the case that people paid more under PIF simply because they inferred that previous visitors had paid a full regular admission price of $7. We therefore rephrased our manipulation to eliminate this ambiguity and replicated the experiment in Study 2.

STUDY 2: A Second Field Experiment in Museum Admission Payment

Study 2 sought to be a nearly perfect replication of Study 1 with a small adjustment in the wording of the pricing manipulation.

Method

We conducted an experiment on one PWYW day on June 5, 2012 at CAM. The museum receives considerably higher traffic in the summer months. Accordingly, we estimated that a single day would be enough for data collection. The procedure is the same as in Study 1 except for the following: To make sure participants understood that *all* visitors (N=152 groups consisting of 372 individual visitors) [18] paid what they wanted for admission, we specifically explained that to everyone. We told participants in the PWYW condition that, "Today is a Pay-What-You-Wish Day so all visitors will be admitted regardless of how much they pay. Today, all visitors, including you, can pay any price they want for their own admission. How much do you want to pay?" Participants in the PIF condition were told, "Today is a Pay-What-You-Wish Day so all visitors will be admitted regardless of how much they pay. Today all visitors, including you, can pay any price they want for the admission of someone who comes later today. But your admission has already been paid for by someone who came earlier, and you have a chance to pay for someone else who will come later. How much do you want to pay?" We recorded the same information as in Study 1.

Results and Discussion

Consistent with the results of Study 1, participants paid significantly more under PIF. That is, even when people were specifically told that all visitors paid what they wanted for admission, they still paid more under PIF than PWYW (M = $3.07 per person vs. $2.19 per person; $F(1, 150) = 5.33$, $p < .022$)[19]. Tables 3.1.2 and 3.1.3 in Appendix reports alternative specifications of the analysis.

STUDY 3: A Third Field Experiment in Museum Admission Payment (and a test of cannibalization from other purchasing)

We had two goals for Study 3. First, we sought an additional replication of the results from Studies 1 and 2. Second, we wanted to examine, and possibly rule out, a potential

downstream consequence of our effects. In simplified form, the museum has two sources of revenue: admission payments and purchases from the museum store. If people are paying more for admission, will they pay less when they visit the gift shop later?

Method

We conducted Study 3 on four different PWYW days at CAM from November 2012 to February 2013. We gathered this larger sample so that we might have at least tolerable power for looking at variation in store purchases. The procedure is the same as in Study 1 except that this time all visitors (N = 304) were given a colored sticker and asked to put it on their shirt before they entered the museum. Visitors in the PWYW condition received a green sticker and those in the PIF condition received a blue sticker. As in the previous studies, we recorded their payments for admission (as well as group size) at the front desk. Additionally, in the museum store, a research assistant who was blind to the conditional assignment recorded the purchased items, the total receipt, and the sticker color of all customers.

Results

Replicating the previous two experiments, people paid more under PIF than PWYW (M = \$3.59 per person vs. \$2.64 per person; $F(1, 302) = 10.33$, $p = .001$). There was no evidence that our manipulation cannibalized from overall revenue however. The PIF visitors made similarly sized purchases as those in the PWYW condition (M_{PIF} = \$20.64 vs. M_{PWYW} = \$ 14.83, $t(50) = 1.14$, $p = .253$). The number of people purchasing in the PWYW relative to the PIF condition did not differ ($\chi^2 = 1.45$, $p = .23$). The combination of these factors (a slight increase in purchase likelihood from one group and a slight increase in purchase amount from the other) meant that total revenue was very similar in the two conditions.

Three field experiments at CAM provide robust evidence that people pay more under PIF than PWYW. Of course, replicable does not mean generalizable. We cannot speak to every possible application of PIF pricing, but in our next study we wanted to at least investigate that it could work in a fairly different setting. One particular concern related to generalizability is that museum admissions might be peculiar. Museums are non-profits, so payments might already feel more like donations—a feeling that might be intensified under PIF. Furthermore, museums fairly clearly have almost zero variable cost; if the museum admits 4 people or 400 people, their costs are almost unchanged. Accordingly, a single individual could reasonably infer that they personally are costless to serve, thereby plausibly becoming unusually sensitive to the manipulation. We conducted a fourth field experiment to test whether our findings would replicate with a for-profit company.

STUDY 4: A Field Experiment with a Gourmet Coffee Vendor

We sought a business collaborator with all of the following features. It sold a product that had clear unit costs. It was clearly a for-profit company. It served a different population than that served by the first three studies (whom were primarily residents of and visitors to the city of San Francisco). We also wanted to find a company for which elective pricing might be plausibly profitable because of small costs and relatively frequent purchases (e.g., we did not approach any Maserati dealerships).

Method

"Ola's Corner," a gourmet coffee vendor in the Bay Area, specializes in rare African coffees and, amongst a few other locations, operates coffee stands at local farmers markets. We sold cups of Ola's coffee at a farmers' market in Jack London Square in Oakland, California. As

with Studies 1 and 2, we used the existing infrastructure (i.e., the tent, signs, and carafes) that Dr. Ola regularly used, but completely replaced the staff with research assistants. Participating groups (N = 132 groups) in this study were the people who bought coffee on two subsequent Sundays in April 2012 approximately from 8:30 am to 2:30pm. We predetermined the two Sundays and did not analyze any data until the completion of the experiment. We asked customers who approached the counter for a cup of coffee to either pay what they wanted for a cup of coffee or pay-it-forward to another customer. Participants in the PWYW condition were told, "Today, you can pay what you want for a cup of coffee. How much would you like to pay?" In the PIF condition, they were told, "Today, you can pay what you want for a cup of coffee. A person who came earlier has paid for your coffee. Now that your coffee's been paid for, you have a chance to pay it forward to a person who will come later. How much would you like to pay forward?" Similar to the previous studies, we recorded each customer's payment, the time of the transaction, the group size, the number of cups purchased (which differ in the event that a group approaches the counter, but only some people buy coffee), and immediately obvious demographic information (i.e., the gender composition of the group, and their approximate ages and ethnic backgrounds)[20].

Results and Discussion

Consistent with the results of Studies 1-3, people paid more for coffee under PIF than PWYW (M = \$2.33 vs. \$1.93 per cup), $F(1, 130) = 6.50$, $p = .012$). Table 3.1.2 in Appendix reports alternative specifications of the analysis.

Thus far, four studies conducted in two different field settings (one non-profit and one for-profit) show that people pay more under PIF than under PWYW. These studies establish that the manipulation is consequential, but they do little to explain why this difference emerges. We try to answer this question in Studies 5-8.

Table 3.1

Study 1, 2, 3, and 4: Payment Per Person In Group (sample size, standard deviation).

	Pay-What-You-Want	Pay-it-Forward	t-tests, p-value
Study 1	\$1.89 (n = 74, 2.03)	\$2.67 (n = 77, 2.69)	$t(149) = 2.18$, $p = .031$
Study 2	\$2.19 (n = 77, 2.05)	\$3.07 (n = 75, 2.63)	$t(150) = 2.31$, $p = .022$
Study 3	\$2.64 (n = 163, 2.20)	\$3.58 (n = 141, 2.91)	$t(304) = 3.21$, $p = .001$
Study 4	\$1.93 (n = 67, 0.93)	\$2.33 (n = 65, 0.90)	$t(130) = 2.55$, $p = .012$

Why People Pay More Under PIF? Four Possible Accounts

Previous research has documented strong social norms for reciprocation (Fehr, Gächter, & Kirchsteiger, 1997; Gouldner, 1960; Goranson & Berkowitz, 1966; Trivers, 1971; Cialdini, 1993). Reciprocation implies a relationship in which one party responds in kind for a deed (good or bad) by the other party (Fehr & Gächter, 2000). Under PWYW customers might be reciprocating the seller for the product or services they received. People also reciprocate *indirectly*, returning one party's kind or unkind behavior to *another* party (Alexander, 1987; Gray, Ward, & Norton, 2012; Nowak & Sigmund, 1998). Accordingly, people might be indirectly reciprocating a customer's kind behavior to another customer under PIF. People contribute as much if reciprocity is indirect as if it is direct (Dufwenberg, Gneezy, Güth, & van Damme, 2001). Reciprocation in either direct or indirect form might explain why people pay more than \$0 overall but there is no evidence suggesting that people are more pressured to

reciprocate indirectly than directly. Furthermore, it is clear that the goods and services are available regardless of the previous parties' payments. There is little formal indebtedness. Then why do people pay more when paying forward than paying what they want?

We speculate four possible explanations that may explain the phenomenon. The first possibility, we term *salience of others,* is that people pay more under PIF than PWYW because other customers in the exchanges are made more salient in the PIF framing than in the PWYW framing. Perhaps, people are thinking of PIF as a gift exchange rather than as a financial exchange. Predictably, social exchanges engage a very different psychology than financial exchanges, and it may be the case that PIF invokes the norms and pressures of social exchanges. Although people would want to reciprocate in both PWYW and PIF exchanges, the pressure to reciprocate might be stronger when the presence of others becomes salient as in the PIF framing. PIF is merely a symbolic gift exchange in which there is no definitive "other" but people may nevertheless feel indebted and pressured to reciprocate[21].

If PIF payments are primarily driven by the salience of others, then a variable that intensifies the salience should *increase* the difference. One way to increase the salience of others in a PIF exchange is to shift the giver/recipient from ambiguous and anonymous to specific and identified. People are more engaged when processing information about a specific target (Chaiken, 1980; Petty & Cacioppo, 1986) and are more willing to help an identified victim more than a statistical victim (Schelling, 1968; Small & Loewenstein, 2003). Accordingly, if salience of others lies at the explanation for PIF effectiveness, then people will pay more under PIF when the anonymity and ambiguous identity of others in their exchange relationships is removed. We consider this possibility in Study 5.

The second account, *differential weights for reciprocity and generosity,* predicts that people pay more under PIF because of an opportunity to be generous to others. People would want to reciprocate under both PWYW and PIF but people are kinder when an exchange involves not only pure reciprocation of the kindness from one person, but it also offers an opportunity to be generous to another person as in PIF. People may also want to be generous toward the seller under PWYW. But thinking about how much they want to pay, people may think more about a fair price, rather than a price that they would pay to feel generous. Or they might feel pressured to reciprocate and pay the minimum amount that is considered appropriate to avoid social pressure (DellaVigna, List, & Malmendier, 2012). But when paying forward, they might feel less tied to the need to reciprocate but more driven by an opportunity to be generous toward the recipient of a gift and pay more than the amount that pure reciprocation would induce them to pay.

As with the *salience of others* account, the *differential weights for reciprocity and generosity* account predicts that increasing the presence of others would increase payments under PWYW and PIF. But *differential weights for reciprocity and generosity* further separates the influences of the salience of the giver and that of the recipient of a gift and predicts that increasing the identifiability of the recipient produces a higher levels of generosity. We test this possibility in Study 5.

The third account is that perhaps people are thinking of *payment as justification*. People may feel more pressured to justify their level of kindness under PIF, and in the absence of any other means of communication, an increased payment is easier to justify. People avoid morally discrediting behavior, but if they can justify it, subsequent behavior may be more unethical (Miller & Effron, 2010; Monin & Miller, 2001). Accordingly, people make more morally questionable choices if they can effectively hide their true motive and justify their behaviors

(Snyder, Kleck, Strenta, & Mentzer, 1979). Paying zero or very little under PIF could cost customers' social image since their generosity would be judged purely based on their payment amount. If, on the other hand, people are offered a costless opportunity to save their social image through justification, they may *pay less*. We consider this possibility in Study 6.

Our fourth account, *generosity matching*, develops out of our initial argument about the influence of descriptive norms. Perhaps people believe that PIF pricing increases generosity in others. Such a systematic perception informs intuitions about existing norms that subsequently guide behavior. Although descriptive norms of others' behavior are not operating on the trade-offs of the exchange itself, they are invoked indirectly by the situation. Pay-it-forward does not permit true generosity or reciprocity between customers, but it most likely makes people *think* about those constructs when they are determining their payments. Perhaps, under PIF, people believe that others paid more, and align their own behavior accordingly. We consider that possibility in Studies 7 and 8.

To test why people consistently pay more under PIF than under PWYW, we created a laboratory setting that allowed for manipulations impractical in the field, while still allowing for actual payments.

STUDY 5: A Laboratory Experiment on Identifiability, Reciprocity, and Generosity

If people pay more under PIF under PWYW because of the salience of others in the PIF pricing frame, an increase in the identifiability of the giver/recipient should increase the salience of others and therefore increase payments. Prior research has shown that social preferences are heavily influenced by knowledge of, and experience with the givers and recipients of prosocial acts (Small & Loewenstein, 2003; Small & Simonsohn, 2008). For example, in a field experiment, people gave 26% more to a charity when the recipient was already determined (a single family from a set of four) than when it would be determined later (Small & Loewenstein, 2003). Recipients were only slightly more identifiable (i.e., people do not get to know the recipients), but the financial consequences were meaningful and reliable. If a similar process underlies PIF, then as identifiability increases so should payments.

Our experimental design manipulated whether participants had direct interaction with the participant who immediately preceded, or immediately followed, in sequence. Additionally, Study 5 tested whether the account of *differential weights for reciprocity and generosity* could explain the conditional difference between PWYW and PIF. If this account were responsible for the phenomenon, people would pay more under PIF (as opposed to PWYW) because they want to be generous to the next participant rather than feel a need to reciprocate for the previous participant. Accordingly, if we increase the identifiability of the next participant then that people should pay more. If reciprocity is more important than generosity, then we could make the reverse prediction: under PIF people would pay more when the previous participant is identifiable. Our design should be able to identify if either of these mechanisms influences the effect, and then differentiate between them.

It should be noted that we conducted the experiment without any clear prediction between the proposed accounts, as each can be quite reasonably justified by previously published findings. Instead, we conducted the study to see which account seemed to offer the best account.

Method

Undergraduates at University of California, Berkeley (N = 294) participated in a 2 (Pricing: PWYW or PIF) × 3 (Social Exchange: Previous Participant, Next Participant, or

Control/no interaction) between-participants design study. We aimed for at least 40 participants per cell and collected data until the end of the semester. We did not analyze any data until the study was concluded. In this study, before the session started, some participants briefly interacted with a confederate appearing to be the previous participant, the next participant, or had no social interaction. Studies 5, 6, 7, and 8 used similar methodologies, so it is important to note that the participant samples were entirely non-overlapping.

Participants in the Previous Participant condition were escorted from the waiting room area to the experiment room, where a confederate was seated in front of a laptop. The experimenter said, "Oh, this is Sarah. She's just finishing up. Can you wait for a minute? Let me grab something. I'll be right back." After the experimenter left the room, Sarah, the confederate, introduced herself to the participant and chatted for a minute. The conversation was scripted to be similar across all participants; the confederate re-stated her name and conversed about general topics related to being in school (e.g. year in school, academic majors, etc.). The goal was to make certain that the participant was aware of "Sarah" but not to make the interaction intensely personal, which would have created a very different set of interpersonal pressures.

The experimenter then returned (after about 60 seconds) with the payment and receipt of the show-up fee and told the confederate, "Sarah, you can leave now. Thank you for your participation." The experimenter then began the session.

Participants in the Next Participant condition met the confederate in the waiting room. The experimenter greeted the participants and said, "Hi, are you (participant's name)? Great." The experimenter then asked the confederate her name. The confederate told her that her name was Sarah. The experimenter checked her clipboard and told Sarah, "Sarah, you're here early. You're scheduled for the next session, which is after (participant's name)'s session. Could you wait for 25-30 minutes? Okay, can you give me a moment? Let me grab something. I'll be right back." The experimenter left the waiting room and the confederate introduced herself to the participant and chatted for a minute following the same script as in the other condition. As in the Previous Participant condition, the experimenter was gone for approximately 60 seconds. The experimenter returned after about a minute and took the participant to an experiment room and began the session. In the control condition, participants did not interact with a confederate, and were ushered to and from the experiment room without interacting with any confederates.

After the social interaction manipulation, each participant was seated in front of a laptop, paid $10 show-up fee in $1 bills, and asked to begin the survey that was unrelated to the main purpose of this study[22]. At the end of this survey, participants were asked how much they enjoyed their experience at school, how much they liked participating in academic research, to what extent they felt that they were a generous person, and to what extent they had more or less money than an average student in the same school.

After participants completed the survey, they received the manipulation of payment type. An experimenter gave each a coffee mug with the university logo ("Cal", written in yellow writing on a blue mug) and asked the participant to either pay what he or she wanted for the mug or pay it forward for the next participant. The participants in the PWYW condition were told, "Today you will receive a Cal mug. The mug is yours, but you also have an option to pay what you want for it. You can put your payment in this envelope and drop it in the box on the desk." The participants in the PIF condition were told, "Today you will receive a Cal mug. The mug is yours. It was paid for by the participant before you. You have a chance to pay it forward to the next participant. You can put your payment in this envelope and drop it in the box on the desk." The experimenter left the room for a few minutes, returned, and asked them to proceed to a short

end survey on the computer screen. The payments participants left in the envelopes constituted a critical dependent variable.

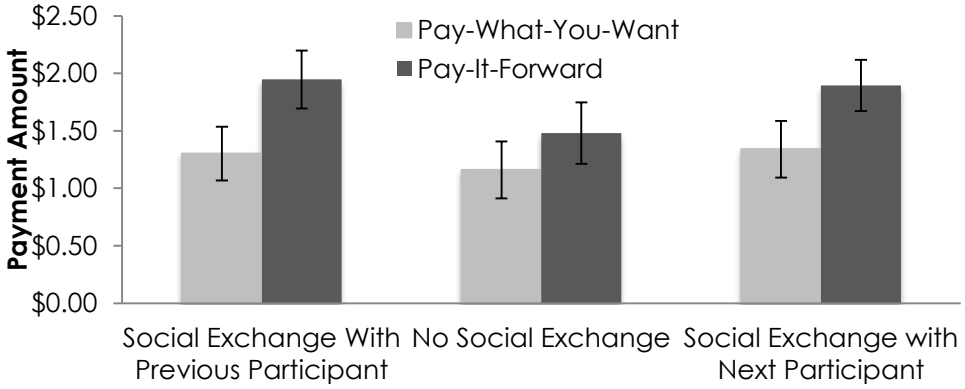
Although participants' payment amount for a mug was our main dependent variable, we assessed a number of variables that might have influenced the payment amount. After making their payment decision, participants were asked to indicate to what extent various factors influenced their payment amount for a mug. The factors, assessed with 7-point Likert-type scales, were: the $10 show-up fee, support of an academic research project, the color of the mug, novelty value of this experiment, the value of the school logo, school affiliation, an average price of a mug, and expectation by the experimenter. Furthermore, participants estimated the payment made by the previous and next participants for a mug. All participants were asked to what extent their payment was fair. They were asked to estimate the actual price of a mug at a campus bookstore. We also asked how satisfied they were with the mug and the study. Lastly, all participants were asked to provide demographic information such as age, gender, and ethnicity. All of the variables, with exact wording, are available in Supplemental Materials and are included in the available dataset [23]. They were debriefed, thanked, and excused.

Results and Discussion

Before analyzing the data we excluded six participants whom research assistants had identified as unreasonable to include: Three participants knew the confederate (and therefore could not be told that her name was "Sarah"), one participant knew about the study (and the deception) from a friend, and two participants were intoxicated with alcohol (the session was scheduled immediately after the conclusion of a college football game across the street which both participants had attended. Their inebriety was independently identified by the experimenter, the confederate, and the participants themselves). These exclusion decisions were made prior to conducting any analysis, but a post-hoc analysis including those participants changes neither the direction nor the statistical significance of any effects. See Table 3.4 in Appendix for a full reporting of these results.

The exclusion of six participants left a final sample size of 288. A 2 (Pricing: PWYW vs. PIF) x 3 (Social Exchange: Previous participant, Next participant, or Control/no interaction) ANOVA on the payment for a mug yielded only a main effect of pricing, such that PIF participants paid more for the mug than did PWYW participants (M = \$1.79 vs. \$1.27), $F(1, 282) = 6.39$, $p = .012$.

Figure 3.1. Payment Amount For A Coffee Mug In Study 5, Error bars reflect standard error of the means.



There was no evidence that knowing either the giver or receiver had any influence on payment amounts. There was no main effect of social exchange, $F(2,282) = .96$, $p = .385$, nor was there a Pricing × Social Exchange interaction, $F(2, 282) = .22$, $p = .803$. As revealed in Figure 3.1, across all Social Exchange conditions, people paid more under Pay-it-Forward than under Pay-What-You-Want, but the size of that difference was not influenced by the Social Exchange manipulation.

Predictions about the payments of the next participant and the previous participant

Generosity matching, one of our three alternative accounts for higher payments under PIF, predicts that the PIF pricing frame indirectly invokes norms of generosity and reciprocity. We reasoned that such implicit social forces operate at least partially on people's perception of others' behaviors under PIF. To test whether participants' belief about others' payments influenced their own payment, we asked all participants to estimate how much the previous and next participants would pay for the mug. Regardless of condition, actual payments were highly correlated with beliefs about the payment of the previous participant ($r = .418$) and the next participant ($r = .449$).

There was one additional effect of note: Participants thought that others—either the previous participant ($M = \$2.03$) or the next participant ($M = \$2.09$)—paid more than they did ($M = \$1.53$), $t(287) = 5.07$, $p < .001$. This effect is surprising and important. First, it is surprising in light of the many findings on egocentrism and the better-than-average effect (Alba & Hutchinson, 2000; Gilovich, Medvec, & Savitsky, 2000; Kruger, 1999; Kruger & Dunning, 1999). A simplified representation of that literature is that people think of themselves as generally better on more positive or flattering attributes, of which generosity (in the form of elective payments) would have seemed to qualify. Alternatively, if we reframe the payment as merely an expression of willingness to pay for the product, then we can appeal to a more basic phenomenon. Most people think that other people are willing to pay more for every product than they would themselves (Frederick, 2012). Why do people think that others are more generous than themselves? Furthermore, in combination with the high correlation with personal payment, how might that misperception guide behavior?

Under elective pricing, it is difficult for people to identify the "correct" payment amount; the amount that simultaneously maximizes the personal sense of propriety and frugality. One shortcut for finding that point is to ask a related question: What do other people pay? (in fact, across all study designs, participants frequently articulate this question. All research assistants were trained to say, "you can pay whatever price is right for you."). With a mental reference in mind, the payment decision is easier; people can pay an amount that is similar to what they think others must be paying. If, as we suspect, beliefs about others are critical in determining personal payments, then it is enough to believe than the specific language (PIF vs. PWYW) will influence the payments of others. If a person thinks that other people will pay more under PIF, then that inference will subsequently guide their own behavior. In the analyses below we offer some tentative tests of this hypothesis.

Testing the Mediating Role of Perceptions of Other's Payments

To test whether or not the relationship between pricing manipulation and payment for the mug was mediated by beliefs about others' payments, we conducted a mediation analysis

following the procedures recommended by Preacher and Hayes (2004). Because participants' estimations of the amounts paid by the previous and next participants were highly correlated (r=.745, p<.001), we used the average of the combined estimations as the mediating variable in our analysis.

Multiple regression analyses were conducted to assess each component of the proposed mediation model. First, the results show that our pricing manipulation (PWYW was coded 0 and PIF was coded 1) was positively associated with the perception of others' payments (B = .47, t(279) = 2.43, p = .016), and the perception of others' payments were positively associated with participants' actual payment (B = .55, t(279) = 2.58, p = .010). Lastly, the results indicate that the mediator, the perception of others' payments, was positively associated with the participants' actual payments (B = .41, t(279) = 8.43, p < .001). The 95% confidence interval of the indirect effect was obtained with 1000 bootstrap resamples, and results of the mediation analysis confirmed the mediating role of the perception of others' payments in the relation between pricing manipulation and the actual payments. (B = .226, CI = .069 to .461). Additionally, the direct effect of pricing on actual payments became non-significant (B = .24, t(279) = 1.39, p = .16) when controlling for the predicted others' payments.

Given a high correlation between the predicted others' payments and the actual payments (r= .463, p < .001), it is equally plausible that participants' actual payments mediated the perception of others' payments. This parallel account was also supported by the analysis—actual payments mediated the relationship between pricing manipulation and perception of others' payments (B = .234, CI = .054 to .488). Differentiating the two pathways would require directly manipulating the perceptions of the payments of others; a manipulation we employ in Studies 7 and 8.

The results from Study 5 replicated the results obtained in the field experiments: people paid more under PIF than under PWYW pricing. Notably, the results of Study 5 suggest that this effect cannot be entirely attributed to the *salience of others* or *differential weights for reciprocity and generosity* accounts. Furthermore, there is some support for the idea that the effect may be partially driven by participants' beliefs about the payments of others.

Before moving to a more direct test of that possibility (in Studies 7 and 8), in Study 6 we give brief consideration to a parallel account. It is possible that under PIF people feel pressured to pay more for whatever reason, but more importantly, because of the nature of the experimental design, feel unable to express an explanation for a lower payment. Accordingly, it is possible that PIF effects are artifactual. In Study 6 we manipulate the opportunity to publicly justify payments, to see if that opportunity lowers payments.

STUDY 6: Do People Pay Less If They Have A Chance To Justify Their Payment?

Study 6 examined how an opportunity to explicitly display generosity—or explicitly justify frugality—influenced payments under PIF. Specifically, we manipulated the extent to which participants could showcase their generosity/justify their frugality to the following participant (and to the experimenter). We had two primary predictions. First, because people are concerned about how they are seen by others, they will pay more when their payments could be observed. Second, if people are constrained by an inability to justify a (low) payment, then the ability to justify should reduce payments.

Method

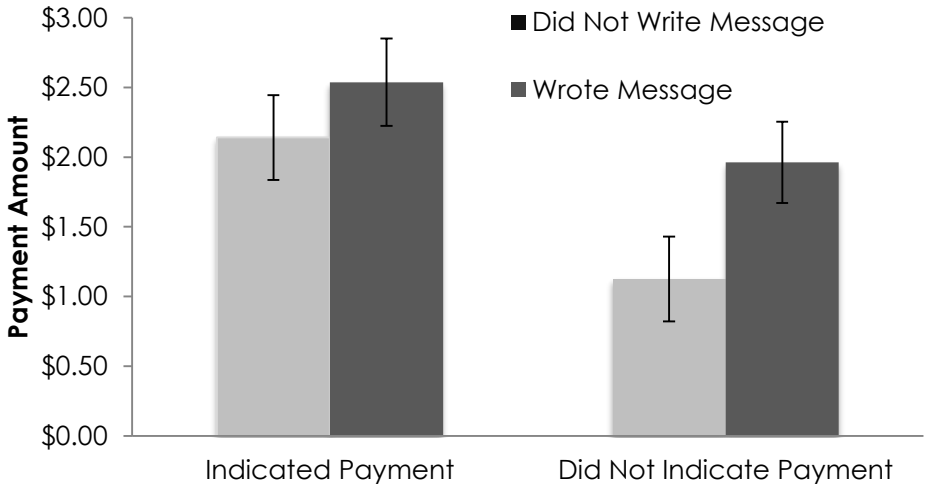
Undergraduates (N = 193) participated in a 2 (Payment Amount Information: Yes vs. No) × 2 (Message: Yes vs. No) between-participants design. All participants received the PIF instructions from Study 5 (i.e., they had been given a mug, that the next person would be given one as well, and that they could pay for the mug on behalf of the next person). Participants in three conditions were given a card to write either a message, the amount paying forward, or both. The cards were 3.5 by 5 inches, and printed on card stock (examples are shown in Figure 3.1.1-3.1.3 in Supplemental Materials)[24].

Participants were run individually. Each was greeted, seated at a laptop, paid the $10 show-up fee in $1 bills, and asked to start a survey that was unrelated to this study[25]. At the end of the survey (approximately 20 to 25 minutes) participants answered the same pre-test questions used in Study 4 (see Table 3.2 in Appendix). Participants were given the mug and received the PIF message, with an addition consistent with the conditional assignment. Participants in the Message condition, for example, were told, "Please write a message to the next participant on this card. Leave the card here (on the keyboard) when you are finished. It will be presented to the next participant." The experimenter gave each participant a Cal mug, an envelope, and a card (for all but the control condition), and left the room. A final survey asked the same questions used as in Study 4 (see Table 3.2 in Appendix).

Results and Discussion

A 2 (Payment Amount Information: Yes vs. No) × 2 (Message: Yes vs. No) between-participants ANOVA revealed that participants paid more when they had to report the amount they paid (*Ms* = $2.34 vs. $1.54), $F(1, 189) = 6.86$, $p = .010$). There was no evidence supporting the justification account. In fact, contrary to that hypothesis, payments slightly *increased* when people could justify their payment amount (*Ms* = $2.25 vs. $1.63), $F(1, 189) = 4.12$, $p = .044$. The interaction was not significant, $F(1, 189) = 0.52$, $p = .470$. One prediction was confirmed, but the other was rather soundly rejected.

Figure 3.2. Mean payment amount for a coffee mug in Study 6. Error bars reflect standard error of the means.



Regardless of these conditional differences, this study provided a second confirmation that people might choose their payments based on how they perceive the payments of others. As in Study 5, participants thought that the previous (M = \$2.47) and next (M = \$ 2.53) participants paid significantly more for the mug than they did ($M_{Predicted\ others'payments}$ = \$2.50 vs. $M_{Participant's\ payment}$= \$1.93), $t(186) = 3.60$, $p < .001$.

Across these two studies we can make a tentative claim about why PIF leads to higher payments than PWYW. The first step is predicated on pluralistic ignorance (Allport, 1924; Miller & McFarland, 1987): participants think that others are paying more. The second step is driven by social pressure: people do not want to pay much less than they think other people are paying.

Study 7 followed that logic to identify a manipulation that might operate on the pluralistic ignorance. We closely followed previous manipulations used in more complex environments (e.g., to reduce campus alcohol consumption, Schroeder & Prentice, 1998), and disabused people by giving them information about how much others have paid. Specifically, in addition to the payment wording manipulation (PWYW vs. PIF), some participants were told how much the previous participant had paid for a mug, whereas the rest were not provided this information. We predicted that, replicating Study 5, people would pay more under PIF when they did not know the payment of the previous participant, but that this effect would be eliminated when participants were told how much the previous participant had paid.

STUDY 7: Does Knowing the Payment of Others Eliminate The Influence of Pay-It-Forward?

This study examined how payments change when people know about the payments of others. We informed approximately half of the participants that the previous participant had paid \$1.50 for a mug (we adopted \$1.50 as it was roughly the average amount paid for the mug previously). In this way, although we were deceptive in telling people that the previous person had paid exactly that amount (in order to keep the conditions identical), the stated price was close to what naturalistically occurred.

Method

Undergraduates (N=198) were run individually and randomly assigned to one of four conditions in a 2 (Pricing: PWYW vs. PIF) × 2 (Previous Participant's Payment Information: Yes vs. No) between-participants design. The design was nearly identical to that employed in Studies 5 and 6. Each participant was greeted, seated at a laptop, paid a \$10 show-up fee in \$1 bills, and asked to begin an unrelated survey. At the end of this survey (which lasted approximately 20 minutes), participants answered the same questions used in Studies 5 and 6 (see Table 3.2 in Appendix). We were aiming for a minimum of 45 participants per cell, ran the study until the end of the semester, and did not analyze the data until the study was completed.

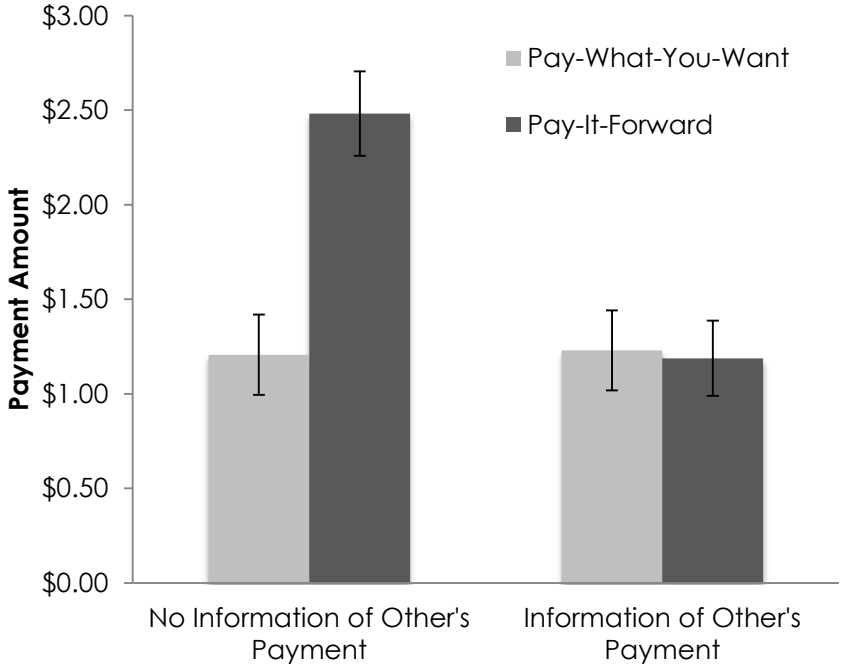
After completing the survey, participants received a university coffee mug and asked to either "pay what you want" or "pay it forward" for the next participant. At this point, approximately half of the participants were told verbally, that the previous participant had paid \$1.50 for the mug. Specifically, participants in the PWYW condition were told, "Today you will receive a Cal mug. The mug is yours, but you also have an option to pay what you want for it (the participant before you paid \$1.50). You can put your payment in this envelope and drop it in the box on the desk." Participants in the PIF condition were told, "Today you will receive a Cal mug. The mug is yours, it was paid for by the participant before you (who gave \$1.50). You have a chance to pay it forward to the next participant. You can put your payment in this

envelope and drop it in the box on the desk." The experimenter made clear that all participants would receive a mug regardless of how much they pay, and that their payment was on behalf of themselves or on behalf of the next participant, depending on condition. The experimenter left the room for a few minutes, returned, and asked participants to complete a final survey on the computer, which contained the same items as in the previous two studies. The only change from the previous studies was that participants who were told about the previous participant's payment information were additionally asked the extent to which they thought that the previous participant's payment (of $1.50) was fair.

Results and Discussion

Payment. We predicted that, in the absence of knowledge about the previous participant's payment, people would pay more under PIF than under PWYW (replicating the previous studies), but that difference would be eliminated when the previous participant's payment was revealed. We submitted participants' payments to a 2 (Pricing: PWYW vs. PIF) x 2 (Previous Participants Payment Information: Yes vs. No) ANOVA. Consistent with the results of Study 5, participants paid more under PIF than they did under PWYW, (*Ms* = $1.84 vs. $1.22, $F(1, 194) = 8.48$, $p = .004$), and paid more when they did not know about the payment of the previous participant than when they did (*Ms* = $1.85 vs. $1.21, $F(1, 194) = 8.99$, $p = .003$). Most importantly, those effects were qualified by the predicted interaction, $F(1, 194) = 9.66$, $p = .002$. When people did not know how much the previous participant had paid, they paid more under PIF than under PWYW (*Ms* = $2.48 vs. $1.21, $t(91) = 3.13$, $p = .002$), whereas people made very similar payments in those conditions when they were informed how much the previous participant had paid (*Ms* = $1.23 vs. $1.19, $t(103) = .25$, $p = .81$).

Figure 3.3. Mean payment amount for a coffee mug in Study 7. Error bars reflect standard error of the means



Predictions of the next participant's payment. Participants' estimations about the payment of the next participant closely mirrored the pattern observed with actual payments. We submitted participants' estimation to the same 2 (Pricing: PWYW vs. PIF) x 2 (Previous Participant's Payment Information: Yes vs. No) ANOVA. The main effects of both Pricing and Previous Participant's Information were significant. Estimations provided by participants in the PIF condition were significantly higher than those reported by participants in the PWYW condition (*Ms* = $2.45 vs. $1.94), $F(1, 185) = 4.47$, $p = .036$. In addition, people thought that the next participant would pay more when they were not given information about the previous participant relative to when they had learned how much the previous participant had paid (*Ms* = $2.73 vs. $1.67), $F(1, 185) = 19.38$, $p < .001$. These effects were qualified by the predicted interaction, $F(1, 185) = 12.79$, $p < .001$.

Consistent with the results presented thus far, payments for the mug were significantly higher under PIF. Also, similar to participants in Studies 5 and 6, all participants in Study 7 estimated that payments of other (previous and next) were more than they actually were ($M_{Combined\ predictions}$= $2.62 vs. $M_{Actual\ payment}$= $1.49), $t(50) = 5.14$, $p < .001$).

As in the previous studies, when people were asked to pay-it-forward, they paid more than when asked to simply pay-what-you-want. However, that effect was only observed when participants did not know how much the previous participant had paid. When participants were disabused of their belief about others' behavior, the effect was eliminated.

Participants in the PWYW conditions paid about the same whether or not they were informed of how much the previous participant had paid. As with the participants in the PIF conditions, participants who paid what they wanted were likely to be influenced by the information of the previous participant's payment amount. However, it is also possible that participants in the PWYW condition based their payment on other idiosyncratic factors such as their internal valuation of a mug or their concerns for fairness toward the experimenter, which was similar to the average payment amount of previous participants. If this were true, their payment amount might be less tied to the information of others' payment and participants would pay about the same amount (i.e., $1.50) for a mug regardless of others' payment amount. In the next study, we test this possibility by varying the previous participant's payment amount to be lower or higher than the average payment in the PWYW and PIF conditions.

Study 8: Receiving Information About A Generous (Or Stingy) Previous Participant

In this study we tested how people were influenced by information about previous participants' payments, particularly when those payments were unusually low or unusually high. To identify "low" and "high" payments, we looked at the distribution of PIF payments in Studies 5 and 7. For this study we chose $0.50 (a 31st percentile payment) and $2.50 (an 83rd percentile payment) to serve as low and high payments.

Method

Students and employees of the University of California, Berkeley (N=329) were randomly assigned to one condition in a 2 (Pricing: PWYW or PIF) × 3 (Information of the previous participant's payment: no information, low payment, or high payment) between-participant design. The procedure was identical to the one used in the previous studies.

As in the previous studies, after completing an unrelated 5 minute survey, participants received a university coffee mug and were given the payment instructions. Specifically, participants in the PWYW conditions were told, "This Cal mug is yours. All participants receive

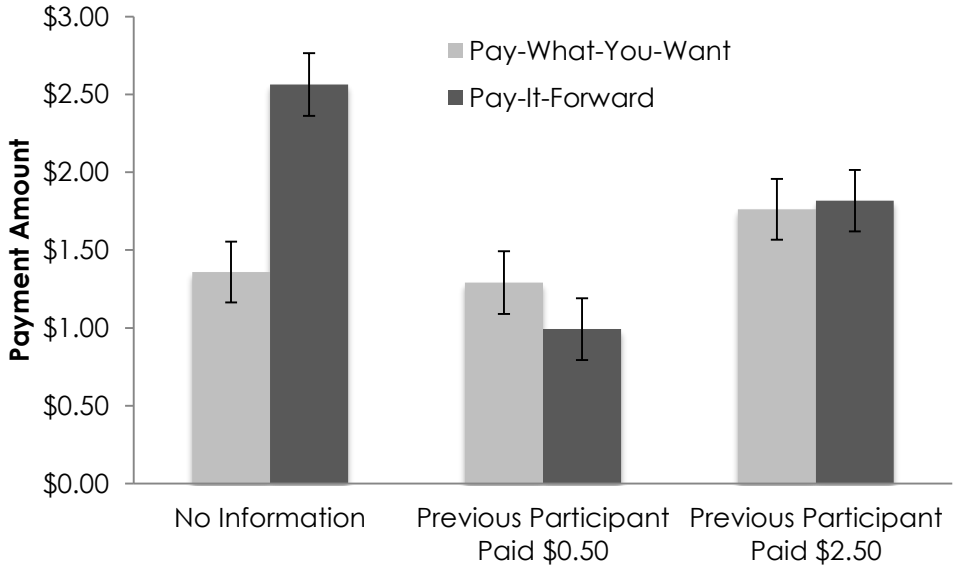
a mug regardless of how much they pay (the participant before you paid $0.50/$2.50). You can pay what you want for it. You can put your payment in this envelope and drop it in the box on the desk." Participants in the PIF conditions were told, "This Cal mug is yours. All participants receive a mug regardless of how much they pay. So this mug was paid for by the participant before you (who gave $0.50/$2.50). You have a chance to pay it forward to the next participant. You can put your payment in this envelope and drop it in the box on the desk." The experimenter left the room for a few minutes, returned, and asked them to proceed to a short survey on the computer screen which contained the same items as in the previous studies.

Results and Discussion

Payment. We excluded two participants in two different conditions because a research assistant's error meant that we couldn't be sure which participant had paid which amount (one contained $2 and the other $3. Exclusion seems overwhelmingly like the correct decision, but the results reported below change in neither direction nor statistical significance regardless of which participant paid which amount).

We predicted that, without information about others' payment, people would pay more under PIF than PWYW. Furthermore, when people were told about the payments of others, we predicted that their own payments would generally follow the information they received. We used a 2 (Pricing: PWYW or PIF) × 3 (Information about the previous participant's payment: no information, low payment, or high payment) ANOVA. As in the previous studies, overall people paid slightly more under PIF than they did under PWYW, (*Ms* = $1.47 vs. $1.79, $F(1, 321) = 4.29$, $p = .049$), and their payments varied as a function of the information they received about the payments of others, $F(2, 321) = 9.43$, $p < .001$. Most importantly, those effects were qualified by the predicted interaction, $F(2, 321) = 7.84$, $p < .001$. When participants did not know how much the previous participant had paid, they paid more under PIF than under PWYW (*Ms* = $2.57 vs. $1.37, $t(107) = 3.20$, $p = .002$). PIF and PWYW were not statistically different with either the low payment (M_{PWYW} = $1.29 vs. M_{PIF} = $0.99, $t(105) = 1.42$, $p = .158$) or the high payment (M_{PWYW} = $1.75 vs. M_{PIF} = $1.79, $t(109) = .21$, $p = .837$).

Figure 3.4. Mean payment amount for a coffee mug in Study 8. Error bars reflect standard error of the mean



Predictions of the next participant's payment. As in previous studies, all participants predicted how much the next participant would pay for a mug. Again, we used the same 2 × 3 ANOVA. In this study, there was no overall main effect of Pricing (M_{PWYW} = \$1.83 vs. M_{PIF} = \$1.86, $F(1, 315) = .04$, $p = .840$), but there was an effect of Payment Information ($F(2, 315) = 13.00$, $p < .001$.). Consistent with the previous study, those effect were qualified by an interaction, $F(2, 315) = 3.37$, $p = .036$, though in this case, the pattern was somewhat peculiar. As before, participants estimated higher payments under PIF when they did not have information about the previous participant (*Ms* = \$2.50 vs. \$1.93), but this difference was only marginally significant $t(103) = 1.518$, $p = .13$. In the low payment condition, PWYW participants predicted higher payments than did those participants in the PIF condition (M_{PWYW} = \$1.50 vs. M_{PIF} = \$1.03), $t(102) = 2.03$, $p = .045$. Predictions were very similar in the high payment condition (M_{PWYW} = \$1.70 vs. M_{PIF} = \$1.79), $t(109) = .44$, $p = .659$.

Study 9: Ruling Out Misunderstanding And Considering Near Extensions

The lab and field studies presented above compared payments under PWYW and PIF under a small handful of conditions. In combination, they present a generally robust (if still incompletely understood) phenomenon. In the process of evaluating the manuscript, our reviewers and editor identified a series of open questions or concerns about how participants interpreted the manipulations, and suggested a few possibilities for alternative phrasings that might produce similar or different results. The following experiments are our efforts to investigate those possibilities. Unlike in the previous eight studies, these studies use an online population and asked them to consider a hypothetical scenario. While giving up the realism of actual payment, they retain the core psychology of the other studies.

Experiment 9A: Are People Confused About Whether Or Not They Can Pay Any Price?

In some of these studies (i.e., Studies 2 and 8), we made an additional effort to ensure that people understood that all participants received the product regardless of payment. Nevertheless, participant confusion might still contribute to the differences between PWYW and PIF. Experiment 9A was designed to rule out that possibility.

The three experiments in Study 9 use hypothetical settings in which participants recruited from an online panel estimated other customers' as well as their own willingness to pay for a cup of coffee under PWYW or PIF. We used this coffee purchase setting because it is similar to the field setting in Study 4 in which we sold coffee at a farmers' market under PWYW and PIF. A serious concern about using this setting is that people might report significantly higher payments than would be observed in real life (i.e., it is a lot easier to part ways with a hypothetical dollar than an actual dollar). We wanted to correct for possible outliers but do so in a way that neither appreciably distorted the data nor left us open to the risk of p-hacking (Simonsohn, Nelson, & Simmons, in press). Accordingly, we set an arbitrary (but reasonable) point to winsorize the data and preregistered that specification for all studies at Open Science Framework (https://osf.io/), a public resource for documenting transparent scientific practices.

Method

Participants (N=419) recruited from the Amazon Mechanical Turk were randomly assigned to two pricing conditions, PWYW and PIF. Participants imagined that they were purchasing a cup of coffee at a regular coffee shop and estimated a typical person's and their own payment. In the PWYW condition, participants read, "The coffee shop does not use

traditional fixed prices. Instead customers are told the following: Today every customer can have coffee for any price they choose to pay. You can pay what you want for a cup of coffee." In the PIF condition, they read, "The coffee shop does not use traditional fixed prices. Instead customers are told the following: Today every customer can have coffee for any price they choose to pay. A customer who came earlier has paid for your coffee. Now that your coffee's been paid for, you have a chance to pay it forward to a customer who will come later." The exact materials are posted along with the preregistration at https://osf.io/z9q4y/.

Participants in both conditions then answered three additional questions probing whether they understood that all customers could pay *any* price they wanted for a cup of coffee. Participants saw the PWYW or PIF pricing description again and answered either yes or no to the following three questions: "Was the customer who came earlier allowed to pay any price he or she wanted for coffee?", "Are you allowed to pay any price you want for coffee?", and "Will a customer who comes later today be allowed to pay any price he or she wants for coffee?"

Results and Discussion

As we preregistered (https://osf.io/z9q4y/), we winsorized respondents' willingness-to-pay estimates at $10. We determined this number based on the field data in which the maximum payment was $5 for a cup of coffee, we wanted to allow for higher payments, but decided that any payments above $10 were unreasonable. Alternative specifications are reported in Appendix.

Replicating the results from the previous studies, respondents were willing to pay more for a cup of coffee under PIF than PWYW, M = $2.73 vs. $1.93, $t(411)=3.55$, $p < .001$. Furthermore, people also thought that a typical customer would pay more under PIF than PWYW, M = $2.70 vs. $2.16, $t(414)=2.32$, $p = .021$. 75% of the participants correctly answered all three comprehension questions and reported that the customer before and after could pay any price they wanted. If we restrict the analysis only to that 75%, the results change in neither direction nor significance. These results indicate that most people were not confused about whether or not they could pay any price under PWYW and PIF. Furthermore, the PIF effect was not driven by people who misunderstood the pricing. Those respondents were not willing to pay more under PIF than PWYW, M = $2.79 vs. $2.76, $t(102)=-.06$, $p = .952$, and also thought others would pay about the same under PWYW and PIF, M = $2.98 vs. $2.75, $t(102)=.40$, $p = .690$.

Experiment 9B: "Can" vs. "Have a Chance"

In Studies 1-8 participants were told, "you *can* pay what you want" in the PWYW condition or "you *have a chance* to pay forward" in the PIF condition. Experiment 9B was designed to rule out a possibility that the PIF effect were driven by this slight difference in the wording, "can" vs. "have a chance".

Method

Respondents (N=835) recruited from the Amazon Mechanical Turk were randomly assigned to participate in a 2 (Pricing: PWYW or PIF) x 2 (Wording: Can or Have a Chance) between-participants design study[26]. For the pricing manipulation, we used the same wordings for PWYW and PIF as in Experiment 9A with the "can" vs. "have a chance" variation. In the PWYW conditions, participants read, "Imagine there is a coffee shop that sells regular coffee. The coffee shop does not use traditional fixed prices. Instead customers are told the following: Today every customer can have coffee for any price they choose to pay. You can (have a chance) pay what you want for a cup of coffee." The PIF conditions followed the same pattern.

As in Study 9A, we asked the same three questions that probed whether or not participants understood that they and other customers could pay any price under PWYW and PIF. We also included an instructional attention check (Oppenheimer, Meyvis, & Davidenko, 2009) to identify respondents who fail to read the instruction carefully.

Results and Discussion

As with Study 9A, we pre-registered our plans for data collection and analyses (https://osf.io/a2icv/). Consistent with the pre-registered plans, our primary analysis excluded those who failed the attention check and we winsorized the willingness-to-pay estimates at $10. Alternative specifications are reported in Appendix.

We excluded 80 participants (9.6%) who failed the attention check[27]. This exclusion left a final sample size of 756. A 2 (Pricing: PWYW or PIF) x 2 (Wording: Can vs. Have a Chance) ANOVA on participants' willingness-to-pay yielded only a main effect of pricing, such that participants were willing to pay more in PIF conditions than PWYW conditions (M = $2.50 vs. $1.75), $F(1, 752) = 45.71$, $p < .001$. The "can" vs. "have a chance" wording did not influence willingness-to-pay differentially, (M = $2.12 vs. $2.13), $F(1, 752) = .02$, $p = .883$. Furthermore, the Pricing x Wording interaction was not significant, $F(1, 752) = 1.06$, $p = .304$.

A very similar pattern emerged for estimates of a typical customers' willingness-to-pay. A 2 (Pricing) x 2 (Wording) ANOVA yielded only a main effect of pricing, (M = $2.35 vs. $1.81), $F(1, 752) = 27.54$, $p < .001$. Neither the main effect of Wording, (M_{can} = $2.06 vs. M_{chance}= $2.10), $F(1, 752) = .10$, $p < .758$, nor the Pricing x Wording interaction was significant, $F(1, 752) = 3.36$, $p = .067$. The latter effect was neither hypothesized nor significant, but is directionally suggestive of a larger PIF vs. PWYW difference with "can" wording than with "have a chance" wording. These results suggest that the PIF effect observed in the previous studies was not driven by a slight wording difference we used (i.e., they can pay what they want vs. they have a chance to pay forward).

Among those who passed the attention check, 93.5% correctly believed that they, the previous, and the next customer could pay any price under PWYW or PIF. Again, if we restrict the analysis to only those people, the results change in neither direction nor significance.

Experiment 9C: Paying for, and being paid for by, a singular other or plural others

An astute reviewer asked whether or not the PIF effect was limited to the social exchange between the specific customers in the chain of PIF transactions. Is it that people feel a social connection with a specific other, or would the effect hold with for a diffuse "others"? We additionally considered a slightly stronger alternative: what if customers are told that *the company* has paid for them? Will a customer still pay forward the company's kindness to another customer? Experiment 9C tests how the information about kind behavior can be generalized beyond the specific PIF pricing framing our previous studies have used.

Method

Participants (N=1,065) were randomly assigned to one of five pricing conditions, PWYW, PIF, PIF singular, PIF plural, and PIF company. The first two conditions repeated the conditions of Experiment 9A. In the PIF singular condition participants read, "Imagine there is a coffee shop that sells regular coffee. The coffee shop does not use traditional fixed prices. Instead customers are told the following: Today every customer can have coffee for any price they choose to pay. The previous customer has paid for your coffee. Now that your coffee's been paid for, you can pay it forward to the next customer." In the PIF plural condition participants

saw, "Today every customer can have coffee for any price they choose to pay. Previous customers have paid for your coffee. You can pay it forward to future customers." In the PIF company condition, "Today every customer can have coffee for any price they choose to pay. We've paid for your coffee. Now that your coffee's been paid for, you can pay it forward to a customer who will come later."

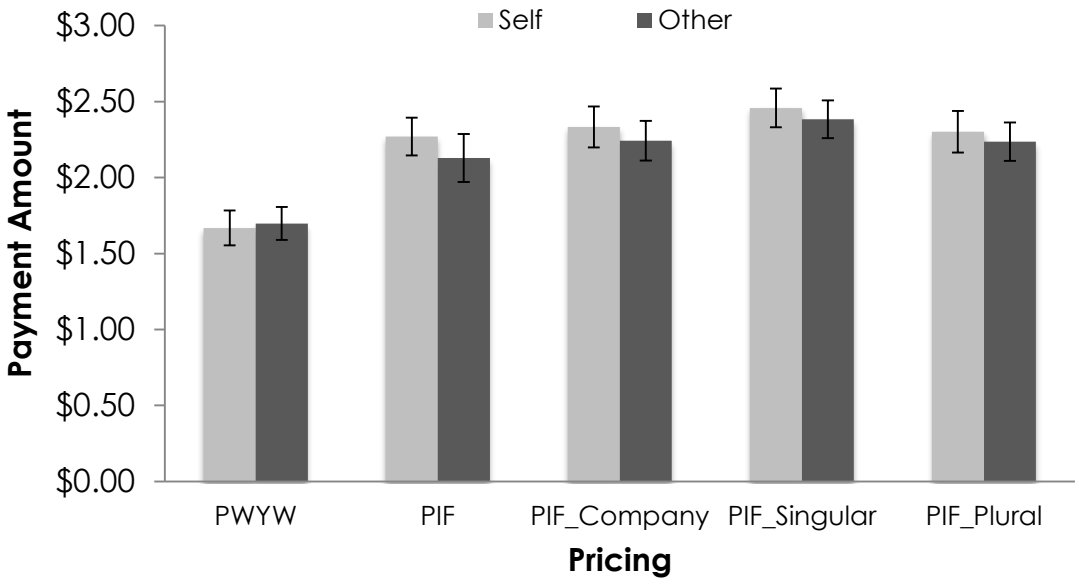
After indicating a typical person's and their own willingness-to-pay for a cup of coffee, they were asked the same three questions in Experiments 9A and 9B assessing understanding and then answered an additional attention check[28].

Results and Discussion

As in Studies 9A and 9B, we pre-registered our plans for data collection and analyses at https://osf.io/pyb6c/. In our primary analysis, we again winsorized willingness-to-pay at $10 and excluded those who failed the attention check. Alternative specifications are reported in Appendix.

169 out of 1,065 participants (16%) failed the attention check. Excluding them left us 896 participants for analysis. Relative to the PWYW condition, participants said that they would pay more for a cup of coffee in all four of the PIF conditions, $3.50 \leq ts \leq 4.61$, $ps \leq .001$. Similarly, estimates of a typical person's payments were higher in all four PIF conditions than in the PWYW condition, $3.23 \leq ts \leq 4.17$, $ps \leq .001$. There were no differences between the four PIF conditions in terms of personal payment ($.04 \leq ts \leq .93$; $.35 < ps < .97$) or estimates of others' payments ($.17 \leq ts \leq 1.06$; $.292 \leq ps < .866$).

Figure 3.5. Mean willingness-to-pay for self and other (a typical customer) for a cup of coffee in Study 9c. Error bars reflect standard error of the means



125 participants (14%) did not think that they or other customers could pay any price they wanted[29]. As in Experiments 9A and 9B, excluding these participants did not change the direction or significance of the results reported above[30].

GENERAL DISCUSSION

The present research documents a novel influence on generosity under consumer elective pricing. By merely reframing payments as being on behalf of others, despite an identical financial reality, people nevertheless pay more. Four field studies (Study1- 4) compared behaviors under PWYW to those under PIF and found robust evidence suggesting that people pay more under PIF than PWYW in both non-profit and profit settings. Subsequent lab studies investigated potential mechanisms for the reported phenomenon. Study 5 found that payments under PWYW and PIF were not influenced by the identifiability of either the giver or recipient, suggesting that the differential weighting for reciprocity and generosity could not explain the difference. Study 6 showed that the effect was not merely due to lack of opportunity to justify the payment, since even with justification opportunities payments did not decrease. Studies 7 and 8 showed that personal payments were heavily influenced by information about the payments of others. When people do not know about others' payments, PIF makes them believe that others are paying more, but knowing the payments of others entirely eliminates the effect. Experiments 9a and 9b showed that neither confusion nor wording-confound explained the observed differences. Finally, Experiment 9C tested whether or not the effect holds beyond the specific target of a giver or recipient.

Under consumer elective pricing people frequently pay for something that they can have for free. Whether or not payment is directed to a seller or another buyer, payment of any amount under consumer elective pricing indicates that people are influenced by social motives other than material self-interest. Perhaps people construe the exchange as pure charity, with the generosity directed toward a seller, an experimenter, or other buyers instead of victims of misfortune. Indeed, some of the same underlying psychological motives are likely at work in consumer elective pricing and pure giving or helping situations. For example, as the present research highlighted, people are heavily influenced by the behavior—real or perceived—of others. Famously, the same concerns have been thought to contribute to helping behaviors in an emergency situation as well (e.g., Latané & Darley, 1970).

Despite those similarities, the key determinants of the *level* of kindness (i.e., how *much* people pay or give) may differ in consumer elective pricing and charitable giving situations. In a charitable giving context, the emotional connection with the recipient substantially determines how much someone will give. Contributions increase as the emotional connection increases through features such as identifiability of the recipient, social distance from the recipient, and familiarity to the recipient's distress (Batson et al., 1983; Small & Loewenstein, 2003; Small & Simonsohn, 2008). These variables did not seem to play as big a role with consumer elective pricing in our studies. In both consumer elective pricing and pure giving situations, people may feel pressured to follow the behaviors of others in an effort to behave appropriately. On the other hand, in pure charitable giving there is a sense that any contribution might be an appreciated contribution, and contributors seek a contribution level which sufficiently relieves the distress evoked through empathy toward others. The same does not hold for consumer elective pricing, in which there is an implicit sense that payments below a certain level might be judged as too low, and instead people seek a payment level which meets the needs of looking kind, while simultaneously feeling affordable.

It should be noted that the observable generosity of others could just as easily push the opposite direction. If people are focused on the net charitable contribution, then information about generous others might prompt people to free ride instead. We are not the only ones who do not find evidence for that alternative (e.g., Andreoni & Scholz, 1998), as free riding the generosity of others is considered socially inappropriate, and therefore, psychologically costly.

Pushing even further against that alternative is the possibility that people see the contributions of others as a signal of the quality of a product or the worthiness of an act (Silverman et al., 1984; Vesterlund, 2003).

Payments in consumer elective pricing are partially driven by ambiguity over what an appropriate payment should be. Without clear standards, people need to rely on their estimates of norms. Those estimates are at risk for systematic error. People inaccurately estimate that others are willing to pay more for goods (Frederick, 2012) and try to match their payment to their perception of others' payment. We can reasonably ask then, when will people correctly estimate actual norms? Tipping is an example of an elective payment with very well understood and agreed upon norms. At restaurants, for example, people in the United States tend to use 15% as a rule of thumb and pay more depending on how they evaluate service quality and other various factors contributing to their dining experience (e.g., group size, alcohol consumption, and frequency of visits) (Bodvarsson & Gibson, 1994). Although tipping is conceptually intended to incentivize workers, the effect of service quality is minimal on the total tipping percentage (e.g., 2% in Lynn & McCall, 2000; 1.48% in Conlin et al., 2003). Instead, variation is much more influenced by the irrelevant feature of bill size (Freeman et al., 1975; Lynn & Grassman, 1990; Rogelberg, Barnes-Farrell, & Creamer, 1999).

Our research suggests that people think that others are paying more under PIF than PWYW. But we still do not know exactly *why* people estimate a higher level of generosity in others under PIF than PWYW. One possibility hinges on the well-documented (but also imperfectly understood) bias towards overestimating the willingness to pay of others (Frederick, 2012). That research shows that, across a wide array of goods and services (e.g., a can of macadamia nuts, a portable minicycle, etc.), people think that someone else who similarly likes the product would nevertheless be willing to pay more for it. Add in the much more straightforward contention that payments are influenced by what they think others are paying, and there is a plausible explanation for the effect: the PIF wording makes people think of others' payments, people overestimate the payments of others, and people adjust their own payments up to match that perception. This explanation is parsimonious, but necessarily speculative. Subsequent research could aim to isolate the role of (mis)perceptions of other payments. Any answer would likely inform not only this research, but also the research of Frederick (2012; Weaver & Frederick, 2012).

Our research indicates that people pay more when they pay it forward. We think that this phenomenon is due to indirectly (implicitly) invoked social influence. When we made those influences more explicitly relevant however, by exposing the identity of gift exchange partners (Study 5), we found no effect of the manipulation. These results suggest that the explicit identity of an exchange partner itself does not influence the level of generosity under PIF. Participants' generosity might be influenced by how *closely* participants relate to their exchange partners under PIF. Small & Simonsohn (2008) found that a closer personal relationship with victims of a misfortune increased sympathy and charitable giving. The participants in our Study 5 were functionally all at the same middling social distance of unfamiliar undergraduates at the same university. Would participants be more generous under PIF if they exchanged gifts with close friends? Since givers and receivers are not victims of misfortune, there is unlikely to be major changes in sympathy. However, the norms of generosity change with reduced social distance. Exchange relationships involve short-term interactions with strangers or acquaintances, whereas communal relationships involve long-term interactions with close friends or family members (Clark & Mills, 1979 and 1993). People in exchange relationships tend to keep record of their

exchange partners' past favor and return the favor at a comparable level. On the contrary, people in communal relationships avoided a "tit-for-tat" type of reciprocity and paid attention to their interaction partners' need and felt obligated to accommodate their need (Clark, 1984). These findings suggest that dominant norms vary depending on the nature of people's social relationships. People might follow different norms depending on how closely they relate to the person they exchange gifts with. Future research could investigate how the strength of social relationships with others influences people's identification of norms and their level of generosity under PIF.

People like to be seen as generous by others and by themselves. They also want to save their money. In concept, therefore, people will savor an opportunity to save money while also saving face. Study 6 tested this possibility. Participants could protect their self-image by justifying a low payment[31]. They did not take the opportunity; they paid slightly *more*. Perhaps participants interpreted it as an opportunity to authentically express and explain their generosity. Consistent with this possibility, when we excluded zero payments from our analysis, only the message factor significantly predicted higher payments. When people wrote a message, more people (32% vs. 19.4%) paid zero but those who paid something did not pay any less than those who indicated payments (*Ms* = $3.27 vs. $2.93).

Pay-it-forward pricing has two quite different features: receiving a gift and giving a gift. Whereas the former invokes the norms of reciprocity the latter is more closely related to generosity. Which is the more powerful influence in shaping people's behaviors under PIF? Grant and Dutton (2012) found that people behave more prosocially when reflecting on giving benefits to others than receiving benefits from others. They argue that giving to others enhances the salience and strength of a giver's identity as a capable and caring person, whereas receiving from others increases a sense of indebtedness and incompetence. Consistent with these findings, we predicted that when people were reminded of giving than receiving a gift, they pay forward a higher amount.

We conducted a pilot test of this possibility at a local Indian restaurant. *Karma Kitchen* has operated with a pay-it-forward pricing model as a Sunday lunch restaurant for many years. In this experiment, we made either the receiving or giving feature of PIF salient and recorded customers' payments. When people entered the restaurant, they were told that their meals had been paid for another customer who came earlier and they could pay it forward to another customer. At the end of their meal, customers (N=94) received one of two slight variants of the PIF language. When customers received their check, a card said either "Thanks for coming to the Karma Kitchen today. Someone who came here earlier paid for your meal as a gift. How much would you like to pay? $___.__" or a card saying "Thanks for coming to the Karma Kitchen today. Now you have a chance to pay for the meal as a gift for someone who will come later. How much would you like to pay? $___.__". Consistent with our prediction, groups of customers paid more when the card emphasized *giving* (*M* = $20.42) than when it emphasized *receiving* (*M* = $11.09; $F(1, 40)= 4.77$, $p = .035$).

We were able to conduct this experiment for only one day and had an insufficient sample size to credibly test our prediction (despite the statistical significance), and therefore, we are hesitant to conclude that the influence of generosity is stronger than reciprocity in the PIF context. However, these results hint that even though people are heavily influenced by norms of others' behaviors, they may be more responsive to the aspects of the normative appeals that enhance and strengthen their identity.

In Chapter 3 we considered a stimuli that evokes social norms of kindness. Although symbolic in nature, information about others' behavior contained in PIF significantly influences consumers' behavior compared to PWYW that lacks such information. Furthermore this effect is eliminated when people are explicitly informed of others' payment amount. In the next chapter, we extend our investigation to test how various explicit numeric information, payment anchors, influences consumers' behavior.

CHAPTER 4. ANCHORING IN PAYMENT: EVALUATING A JUDGMENTAL HEURISTIC IN FIELD EXPERIMENTAL SETTINGS

Abstract

Anchoring, the biasing of estimates towards a previously considered value, is both hugely influential in consumer judgment and frequently studied by consumer researchers. Given its proven robustness and ubiquity, anchoring should influence payments. However, most anchoring work has been in the lab, only occasionally studying payments hypothetically, and the results from field work have been mixed. Here, the authors use real transactions from an empirically-investigated and commercially-employed pricing scheme (pay-what-you-want) to improve our understanding of how anchors influence payments. Sixteen field studies (N=21,997) and four hypothetical studies (N=3,174) reveal substantial variation: although anchoring replicates both with and without financial consequences (Studies 1-2), the type and size of anchor gaps (Studies 3a-5) as well as those of the anchors themselves (Studies 6a-6b) significantly influence judgment in previously undocumented ways. The authors then test suggestions from the literature that should enhance anchoring effects (Studies 7-13) and find null results. Finally, the authors show these subtleties do not emerge in hypothetical settings (Studies 14a-14d), where anchoring is as big and reliable as the literature has previously suggested.

Consumer researchers study human psychology in the hopes that it will allow them to make better predictions about behavior in the marketplace. If consumer judgment can be well understood, then the consequences for real life consumption should be a direct extension. Sometimes those extensions are less direct than we hope.

This paper aims to take an incredibly robust judgment process, anchoring (Tversky & Kahneman 1974), and examine how it operates on payments in the field. Consistent with that aim, our goals are broad: the judgmental process under consideration is foundational in consumer research and our investigation employs many statistically powerful field experiments (16 field experiments with more than 22,000 total participants) rather than a few modest lab studies.

Our goals are also nuanced. No one doubts whether anchoring exists, but there is much less certainty about when and how its operation is bounded in the field. Even the extant literature, as we discuss in the following sections, alludes to a more complicated story than is typically articulated. We aim to present preliminary evidence on a number of factors that influence the expression of anchoring in the field. We conclude by discussing the hidden consequence of stimuli selection in anchoring studies.

A Brief Background on Anchoring

People who first answer whether an adult giraffe weighs more or less than 2100 pounds give a larger subsequent estimate for its weight than those who first answer whether it weighs more or less than 800 pounds (Frederick and Mochon 2012). This difference is due to anchoring, first articulated by Tversky and Kahneman in 1974, wherein irrelevant numbers can substantially influence numerical judgments. Since then, anchoring has become one of the most well-studied phenomena in judgment and decision making.

The giraffe example demonstrates three critical features of an anchoring paradigm. First, is the anchor, which can be arbitrary or devised by the participant. Second, is the deliberate consideration of the anchor before the final estimate. Though deliberation is widely agreed upon to yield the largest anchoring effects (e.g., Brewer and Chapman 2002; Mochon and Frederick 2013), deliberation-free anchors can seemingly also be influential (e.g., Wilson, Houston, Etling, and Brekke 1996; Critcher and Gilovich 2008). Finally, the value of the target judgment should be uncertain (e.g., anchoring is unlikely to influence estimates of the number of hours in a day[32].) Overall, when evaluating an uncertain numeric entity, higher anchors should produce higher estimates.

Researchers can all observe the effect, but they broadly disagree about the cause. Tversky and Kahneman (1974) first attributed the effect to insufficient adjustment from the anchor. That is, people mentally move away from the anchor until they reach a plausible value and then stop. Because they do not continue moving past this first plausible value, the adjustments are insufficient. Alternatively, Strack and Mussweiler (1997) proposed a selective accessibility account, in which an anchor leads people to call to mind consistent information, which supports final estimates closer to the anchor. Complicating things further is the question of whether motivation for accuracy influences the effect. Many researchers fail to find any effect of incentivized accuracy (Tversky and Kahneman, 1974; Strack and Mussweiler 1997; Chapman and Johnson 2002; Epley and Gilovich 2005), whereas others have suggested that incentives alter adjustment conditional on simply knowing which direction to adjust (Simmons, LeBoeuf, and Nelson 2010). With a different perspective, Frederick and Mochon (2012) have proposed that anchors distort the response scale rather than the judgments themselves. All judges maintain the same sense of giraffe weight, but the anchor changes how people think about the weight of a pound: an 800 pound anchor makes pounds seem like a larger unit than does a 2,100 pound

anchor; hence fewer pounds are required for the giraffe's weight. The research community has made strides in understanding anchoring, but they have hardly reached an agreement.

Our goal, however, is not to differentiate between these accounts. Instead, it is to better understand how this long-standing lab phenomenon holds up when taken into the field. In doing so, we aim to use manipulations that all theories (and theoreticians) would agree will operate on anchoring. We value this consensus, because we want to be able to interpret the presence and magnitude of anchoring effects in the field.

To do so, we use a form of consumer-elective pricing (pay-what-you-want, PWYW) to consider the operation of anchoring. Under PWYW, customers can choose what price to pay, yet consistently pay positive amounts (Kim, Natter, and Spann 2009; Armstrong-Soule and Madrigal 2014; Riener and Traxler 2012; Regner, Tobias, and Barria 2009; although see León and Noguera 2012 for a boundary condition at high retail prices). PWYW paradigms allow anchors to be organically presented as default or suggested prices. Returning to the three aspects of anchoring in this context: anchors can be presented clearly without justification (i.e., are arbitrary), must be considered and rejected (or accepted) by participants before their final payment, and the value of the goods sold in our studies are not readily known. Moreover, the combined uncertainty of personal valuation (Bettman, Luce, and Payne 1998) and socially appropriate payment (Gneezy, Gneezy, Riener, and Nelson 2012) should make customers especially susceptible to anchors.

Field Anchoring

Given how often consumers are called upon to make numeric judgments, anchoring could be important across many payment contexts. In hypothetical scenarios, anchoring effects have been shown with credit card payments (Stewart, 2009), negotiation outcomes (Mason, Lee, Wiley, and Ames 2013), and buying and selling prices (Simonson and Drolet 2004). However robust, there is still uncertainty about how these effects are expressed outside of hypothetical situations.

A smaller body of work considers anchoring effects with incentive-compatible designs. Work by Ariely, Loewenstein, and Prelec (2003) as well as Maniadis, Tufano, and List (2014) employed designs with real money and goods at stake. Both of these papers show data consistent with classic anchoring effects[33]. Both retained some contrived characteristics from the lab: the consideration and rejection of an arbitrary price derived from the last three digits of a social security number before eliciting the actual willingness to pay through a Becker, DeGroot, Marschak (BDM) procedure. Nunes and Boatwright (2004) used more naturalistic anchors (featured prices of nearby products at a local concert), but also utilized a BDM procedure. Given that the real world frequently lacks such contrivances and participants frequently misunderstand the BDM procedure's premise (Carson and Plott 2012), these findings offer an incomplete account of how anchoring operates in the field.

Finally, there is an even smaller literature documenting anchoring effects outside of the lab. One such context is auctions: several researchers have found that varying key references prices (e.g., minimum bids—Kamins, Dreze, and Folkes 2004; buy-it-now prices—Hardesty and Suter 2013; and bid options—Spann, Häubl, Skiera, and Bernhardt 2012) can elicit changes in payments in line with anchoring. However, since not every bid turns into a payment, auctions offer more than a pure hypothetical but less than a certain payment.

Charitable donations provide a closer analog, as payment is both elective in amount and guaranteed. A variety of papers examine the influence of both seller-produced (Alpizar, Carlsson, and Johansson-Stenman 2008; Martin and Randal 2008; Desmet and Feinberg 2003;

Smith and Berger 1996) and donor-derived reference prices (Croson and Shang 2008; Shang and Croson 2009) on donation behavior. These papers, too, frequently produce significant anchoring effects. Finally, there have even been some investigations of anchoring under PWYW, in which the presence (e.g., Kim, Kaufmann, and Stegemann 2013; Gneezy et al. 2012) of an anchor is varied, also frequently producing significant effects.

Those findings are largely unassailable, but they have limitations in how they can be generalized beyond the original context. As mentioned above, many papers are hypothetical and lack financial consequence, take place in an incongruent setting, or retain the artificiality of traditional lab paradigms. Of greater concern, and importance for the present paper, is that the overwhelming majority of these investigations have been isolated occurrences: sometimes comparing only one anchor to the absence of one, and often tested in only one or two studies. Such an approach is highly valuable for answering isolated precise questions, but is incomplete for making overall statements about a general process like anchoring. The present paper, with many more experiments, intends to have sufficient breadth to offer a more nuanced rendering.

This need for nuance becomes already apparent when looking more closely at the aforementioned charitable donations literature. Some studies were successful, but some were not. Given low, medium, and high anchors (suggested donations), some find that only the highest anchor generates significantly higher donations (e.g., Shang and Croson 2009), whereas others find that only lowest anchor generates significantly lower donations (e.g., Alpizar, et al 2008; Martin and Randal 2008). An even more unnerving set of studies fails to show any effects of anchors on donation amounts (e.g., Croson and Shang 2008; Desmet and Feinberg 2003; Smith and Berger 1996). These failures are hard to reconcile in the face of a literature that so uniformly reports large and pervasive effects. We will try to offer an account at the end of this paper.

Using this PWYW context, we explore how and why anchoring effects change between the lab and the field. Sixteen studies tested the size and presence of anchoring effects, and four additional studies consider similar manipulations in a hypothetical domain. The persevering reader will see our general inferences: First, although anchoring researchers agree that larger anchor gaps generate larger effects, the objective gap (the difference in absolute magnitude) is much less important than the size of the subjective gap (operationalized as the difference in the anchors' percentile ranks). Second, with extreme anchors, low anchors pull down payments more than high anchors inflate them. Finally, when these exact paradigms are taken back into the lab (where no money leaves the participant's wallet), the subjective range of payments widens, and previously inert extremely high anchors appear subjectively reasonable, and become influential.

In this research we first[34] conceptually replicate past work that has been done in both the lab and field, with both financial and non-financial consequences on (Studies 1 and 2). In Study 3, we more closely examine objective (absolute magnitude) and subjective (distributional) anchor gaps, an apparently critical distinction which has been generally neglected in the literature. Studies 4-6b reveal another asymmetry, this time relating to the anchors themselves. Using reasonable anchor sets, we then implement a variety of modifications to the basic anchoring paradigm suggested by the literature in Studies 7-13, largely revealing null results. However, in the final section, we present four hypothetical studies to demonstrate the disparity between hypothetical and real settings, as we show anchoring effects are much easier to find in hypothetical judgments. Descriptive statistics for each study are in Table 4.1. With all of these findings in mind, we present our best effort at detailing what we have learned about the robustness of anchoring in the field. Certainly, it is far less reliable than in the lab, but it is hardly absent either.

Table 4.1. Basic Info for Each Study

	Product	N	n	Anchor Type	Anchor	Anchor Percentile	Mean (SD)	Outcome Percentile
Study 1	Allocations	1,328	648	Default	50%	16.9	61.64% (15.36%)[a]	30.9
			680		90%	80.0	89.20% (6.60%)[b]	80.0
Study 2	Doughnuts	1,009	490	Suggestion	$0.25	20.3	$0.44 ($0.38)[a]	90.1
			181		$1.75	92.8	$1.04 ($0.71)[b]	50.4
			338		Control	---	$0.66 ($0.54)[c]	29.9
Study 3a	Media Bundle	303	110	Default	$3	4.6	$6.59 ($2.38)[a]	55.1
			91		$9	69.6	$7.79 ($2.44)[b]	66.3
			102		$20	96.7	$8.29 ($4.57)[b]	69.6
Study 3b	Media bundle	3,978	1,443	Default	$8	21.5	$8.88 ($4.03)[a]	23.5
			1,291		$20	95.3	$9.89 ($4.96)[b]	23.5
			1,244		$50	100.0	$9.88 ($5.63)[b]	23.5
Study 4	Media bundle	3,214	1,602	Default	$12	75.9	$8.80 ($4.64)[a]	23.2
			1,612		$15	87.1	$8.88 ($4.88)[a]	23.2
Study 5	Vodo	1,603	376	Default	$2	29.2	$11.48 ($7.87)[a]	42.7
			425		$5	30.7	$11.24 ($7.65)[a]	42.1
			419		$9	32.1	$11.05 ($7.68)[a]	41.8
			383		$12	45.6	$11.29 ($7.72)[a]	42.3
Study 6a	Doughnuts	395	113	Suggestion	$1	27.9	$0.91 ($0.38)[a]	27.9
			78		$3	94.8	$0.84($0.75)[a,b]	27.9
			106		Control	--	$0.72 ($0.45)[b]	24.9
Study 6b	Museum tickets	431	97	Suggestion	Nothing	0.0	$2.55 ($2.73)[a,c]	57.1
			117		$0.01	16.2	$2.47 ($2.36)[a,c]	52.0
			92		$5	67.3	$3.47 ($2.96)[b]	62.9
			102		Control	---	$3.24 ($2.77)[b,c]	61.3
Study 8	Museum tickets	957	211	Suggestion	$0.50	56.4	$0.93 ($2.20)[a]	64.8
			273		$1	64.9	$0.77 ($1.29)[a]	64.6
			253		$2	83.4	$0.91 ($1.36)[a]	64.8
			220		Control	---	$1.08 ($6.80)[a]	74.8
Study 9	Media bundle	1,175	378	Previous	$8	7.0	$11.10 ($3.97)[a]	68.1
			393	Payment	$12	68.2	$11.17 ($3.50)[a]	68.2
Study 10	Media bundle	2,190	1,122	Default	$14	88.1	$10.78 ($5.40)[a]	62.2
			1,068		$28.88	98.5	$11.17 ($6.19)[a]	67.5
Study 11	Museum Tickets	429	121	Suggestion	$5	93.0	$1.06($1.55)[a]	70.2
			88		$11	99.8	$0.77($1.48)[a]	68.1
			126		Control	--	$0.90($1.52)[a]	68.5
Study 12a	Media bundle	714	138	Default	$19.91	84.7	$11.67 ($6.13)[a,b]	25.6
			140		$19.99	85.9		66.1
			135		$20.00	88.1	$12.31 ($6.39)[a] $11.53 ($6.63)[a,b]	25.6
			158		$20.01	93.8	$10.52 ($7.65)[b]	25.4
			143		$20.09	95.2	$11.60 ($6.89)[a,b]	25.6
Study 12b	Media bundle	4,110	790	Default	$19.91	91.4	$10.56 ($4.81)[a]	75.9
			827		$19.99	92.2	$10.39 ($4.48)[a]	75.4
			861		$20.00	93.5	$10.32 ($4.45)[a]	75.4
			838		$20.01	97.1	$10.60 ($5.67)[a]	75.9
			794		$20.09	97.7	$10.55 ($4.71)[a]	75.9
Study	Museum	590	231	Maximum	$10	99.0	$2.72 ($3.25)[a,b]	58.7

13	tickets**		155	Price	$50	100	$2.25 ($2.27)[a]	57.3
			73		$100	100	$2.90 ($2.54)[b]	58.7
			147		Control	--	$2.01($2.94)[a,b]	57.3
Study 14a	Media bundle*	386	205	Default	$14	56.2	$13.01 ($9.82)[a]	56.2
			181		$28.88	94.6	$15.78 ($10.34)[b]	73.1
Study 14b	Media bundle*	581	274	Default	$3	21.7	$5.10 ($3.43)[a]	47.0
			307		$9	85.7	$5.99 ($3.12)[b]	47.0
Study 14c	Doughnuts*	169	61	Suggestion	$1	36.7	$0.82 ($0.82)[a]	36.1
			56		$3	89.3	$1.47 ($0.94)[b]	73.4
Study 14d	Doughnuts*	2,038	313	Suggestion	$0.00	0.0	$0.92 ($0.78)[a]	47.9
			359		$0.01	3.8	$0.50 ($0.52)[b]	29.6
			382		$0.25	11.9	$0.39 ($0.52)[c]	29.0
			304		$1.75	87.6	$1.09 ($0.61)[d]	83.3
			359		$50	100	$1.52 ($1.26)[e]	87.6
			321		Control	--	$1.04 ($0.73)[f]	73.4

Within each study, means with different subscripts differ significantly ($p < .05$) *Studies 14a-d are all hypothetical ** The means were controlled for month-to-month variation.

For all experiments reported in this paper, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures. In no case did we analyze the data before reaching our pre-determined sample size. All data and materials are posted online.

Examining Different Types of Anchor Gaps

Study 1: Field Anchoring With Non-Payments

Before diving into our investigation of reference prices and payments in anchoring, we wanted to first ensure that anchoring was attainable in a field setting more generally. To retain ecological validity, however, we sought out a naturally-occurring context where people made numeric decisions that lacked financial consequences for themselves.

A media retailer (with whom we collaborate in many subsequent studies) allows customers to choose the percentage of their payments which goes to the products' developers or to the retailer (on a bi-polar scale, increasing in one percentage point units). The situation is low-stakes for the customer, since she doesn't have to pay more depending on the allocation, but it still retains the ecological validity of the field setting. In line with the anchoring literature, we predicted that higher allocations would be made under higher allocation defaults. In this and all subsequent studies with this retailer, the sample size was determined to be every customer who completed and did not cancel their purchase during the promotion.

Method

Customers (N=1,328) were randomly assigned to one of six default allocations (note that we will refer only to the authors' allocation for ease of explanation): 49%, 50%, 51%, 89%, 90%, and 91%. These numbers were chosen for two reasons: First, the company required that our conditions average out to the previous default of 70% to the authors. Second, precise anchors have been shown to more strongly influence anchoring effects and could be readily adapted (e.g., Janiszewski and Uy 2008) to field studies.

Results

Customers allocated more of their payment to developers when they saw a higher anchor. A one-way ANOVA revealed differences among the conditions, $F(5, 1322) = 366.98$, $p < .001$. Further analysis showed that this effect was driven primarily by the large differences between the three low ($M = 61.64\%$) and three high anchors ($M = 89.20\%$), $t(1326) = 42.82$, $p < .001$, $d = 2.35$, in line with our general prediction. The three low anchors were not different from each other, $F(2, 645) = 0.08$, $p = .922$, but the three high anchors showed some differences, $F(2,677) = 4.36$, $p = .013$. The lowest high anchor ($M_{89\%} = 88.23\%$), generated a lower allocation that the medium high allocation ($M_{90\%}=89.27\%$), which in turn generated a lower allocation than the highest anchor than the highest anchor ($M_{91\%} = 90.02\%$).

Perhaps the result was due to laziness or inattention rather than anchoring. To confirm the operation of anchoring, we reanalyzed the data excluding customers who had simply accepted the defaults, leaving 449 customers (33.81% of the original sample). First, note that this is (excessively) conservative; the strongest possible anchoring effect would be to predict no adjustment at all, and we are systematically eliminating those responses. Even in this restricted sample the effect was still large and significant between the three high ($M = 85.71\%$) and three low defaults ($M = 73.66\%$), $t(447) = 8.20$, $p < .001$, $d = 0.85$.

Study 2: Field Anchoring With Real Payments

After this conceptual anchoring replication in a low-stakes context, the remaining studies consider the more consequential context of actual payments. In the lab, researchers can choose anchors independent of concerns for profitability. One result is that an overwhelming majority of laboratory studies use extreme anchors, i.e., those which would represent uncommonly small or uncommonly large responses (e.g., Jacowitz and Kahneman 1995; Ariely, et al 2003; Nunes and Boatwright 2004). We mimicked this flexibility by creating our own retailer (a campus doughnut stand) which could be more sensitive to researcher whim than profit-orientation.

Although we examine actual payments in Study 2, we otherwise stayed very close to lab paradigms (e.g., Brewer and Chapman 2002): Customers were forced to consider (and accept or reject) an initial anchor and adjust their payment. We accomplished this by forcing customers to choose between a fixed default and an additional option to specify their own price. Because this modification more closely matches the classic, highly successful anchoring paradigm, we predicted that higher anchors would be associated with higher payments.

Additionally, we guarded against another alternative explanation. Previous research suggests that anchors might only be effective for participants who did not know that they were entering a PWYW transaction (Gautier and van der Klaauw 2012). Study 2 manipulated whether customers had this foreknowledge.

Method

The experiment contained two manipulations: foreknowledge and anchor value. The first addressed the selection concern by randomly assigning people to condition either before they chose to buy or after they chose to buy. The second manipulation was the anchor: some people were simply offered PWYW (a no-anchor control), whereas others were presented a very low or very high anchor in addition to PWYW.

We sold glazed doughnuts at an outdoor plaza at the University of California, Berkeley for 27 days in March and April of 2014. People (N = 70,091) passing by the doughnut stand saw our shop's sign, "Dream Fluff Doughnuts!" In the selection conditions, the sign also read "Pay What You Want," "$0.25 or Pay What You Want," or "$1.75 or Pay What You Want" (see Figure 4.1.2.1 in Appendix for an example of the signs). We chose these anchors because our previous experience in this domain suggested those amounts were subjectively far apart for customers.

The sign changed in a randomized order for every 200 people passing by our doughnut stand[35]. In the no-selection condition, the sign simply said "Dream Fluff Doughnuts." Once customers approached the shop, they were told their price would be determined by a random draw. Customers reached into an opaque box and drew the price information. We recorded date/ time of transactions, number of passersby, payment, customer group size, gender, and age. We predetermined to collect data until we had at least 100 observations per condition.

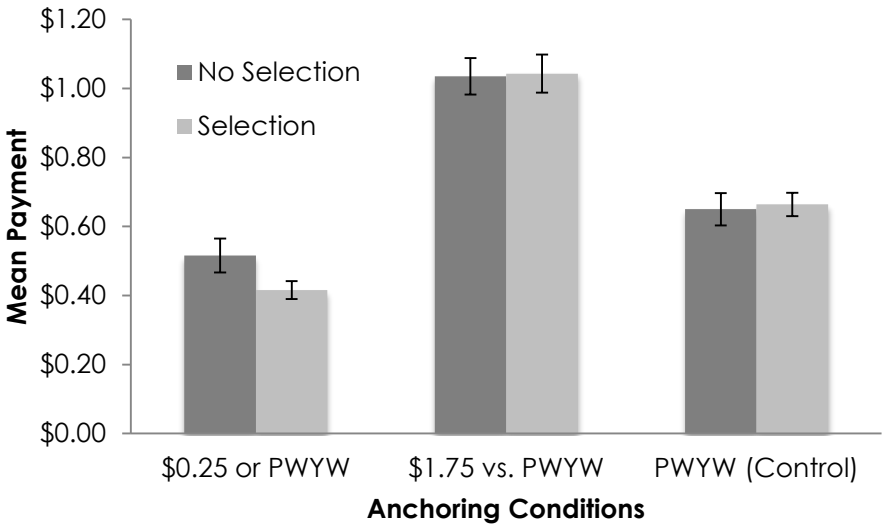
Results

Customers (N = 892 groups; N=1,038 individuals) bought 1,054 doughnuts[36]. We used individual doughnut purchase as unit of analysis with the average payment per doughnut as the dependent measure. We excluded 29 transactions by the experimenters' friends[37], which left us 1,009 purchases for analysis.

Consistent with basic self-selection, people were strongly influenced by posted prices on the signs: People were more likely to buy a doughnut when the sign said "$0.25 or PWYW" than when it said "$1.75 or PWYW" (382 out of 14,631 vs. 87 out of 14,548), $\chi^2(1) = 186.89$, $p < .001$. Furthermore, people seeing the former were more likely to purchase than those seeing either the PWYW sign (220 out of 13,639), $\chi^2(N=28,270) = 33.73$, $p < .001$ or the more

ambiguous "Dream Fluff" sign (320 out of 28,073), $\chi^2(1) = 128.72$, $p < .001$. Even though all participants could pay what they wanted, the invitation to "pay $0.25 or pay what you want" was more motivating than any other sign.

Figure 4.1. Means for Study 2



However, a 3 (anchor: $0.25 vs. PWYW, $1.75 vs. PWYW, or PWYW) x 2 (selection: absent or present) ANOVA on payments revealed only a main effect of anchor, $F(2, 1003) = 74.29$, $p < .001$. Despite the substantial differences in purchase rate, there were no effects of selection on average payments, ($M_{selection}$= $0.71 vs. $M_{noselection}$ = $0.74), $F(1, 1003) = .55$, $p = .46$. Critically, the interaction between selection and anchor was not significant, $F(2, 1003) = 1.15$, $p = .316$.

Customers paid more for a doughnut under "$1.75 or PWYW" than PWYW (*Ms* = $1.04 vs. $0.66), $t(517) = 6.77$ $p < .001$. But they paid more under PWYW than "$0.25 or PWYW", (*Ms* = $0.66 vs. $0.44), $t(826) = 6.92$, $p < .001$. Most importantly, people paid more under $1.75 or PWYW than under $0.25 or PWYW, (*Ms* = $1.04 vs. $0.44), $t(669) = 14.02$, $p < .001$, $d = 1.08$, a very large anchoring effect.

Study 3a: Considering Different Types Of Gaps

Although we replicated a standard anchoring effect with a standard, large anchor gap, Study 2—and most anchoring studies—did not differentiate the types of anchor gaps, e.g., in absolute terms, standardized terms, or in percentile[38] rank. As we discovered across many of the next studies, although anchor gaps can be defined in many ways (perhaps too many), not all anchor gaps lead to anchoring effects. Studies 3a and 3b examine anchors that have a large absolute gap, but are subjectively similar (i.e., similar in percentile rank of payment), and compare those results to anchors both absolutely and subjectively far apart on these measures.

Method

We again collaborated with the PWYW media retailer from Study 1, but now focused on payments rather than allocations. This company typically offers a two or three-week promotion in which customers can pay what they want for a collection of thematically organized media

goods. There are a few critical features in each promotion: First, there is a minimum price (e.g., $1 for this bundle of six goods). Second, there is a "bonus" price, above which any elective payment also buys additional 2-4 goods (e.g., $9, for ten in total). Third, prices are identified by either typing in a box or by selecting on a sliding scale, which moves in $1 increments and ranges from the minimum to $100. Customers also may donate 10% of their payment to charity.

During Study 3a's (N=303) promotion, the minimum price was $2 and the bonus price was $6. We randomly assigned customers to see a $3, $9, or $20 default. Our lowest and two highest defaults were separated by large gaps in both absolute value ($6 and $17, respectively) and percentile rank (69.6 and 96.7, respectively). Critically, however, our two highest defaults had a relatively small gap in percentile rank (27.1), but a large gap in absolute value ($11). Although in this and many subsequent studies participants are not directly prompted to accept or reject the anchor as they were in Study 2, we believe having to move the slider away from the default involves a similar cognitive process. (Also see Wilson, et al (1996) and Critcher and Gilovich (2008) for anchoring with even less direct anchor consideration.)

Results

Payments differed between the three conditions, $F(2, 300) = 7.51$, $p = .001$. Payments were lower for the $3 anchor ($M$ = $6.59) than for either the $9 ($M$ = $7.79) or the $20 anchor $20 ($M$ = $8.29), $t(199) = 3.53$, $p = .001$; $t(210) = 3.44$, $p = .001$, respectively. However, payments did not differ between the two higher anchors, $t(191) = .94$, $p = .350$.

Study 3b: Replicating Study 3a

Study 3a demonstrates that the type of anchor gaps matter in the field; however, the sample is idiosyncratically small, so we include a near replication in Study 3b.

Method

In Study 3b (N = 3,978) the minimum price was $3 and the bonus price was $10. We randomly assigned visitors to see an $8, $20, or $50 default, allowing us a similar pattern of objective and subjective gaps as in Study 3a.

Results

As before, payments different across the three conditions, $F(2, 3975) = 19.62$, $p < .001$. Payments were lower under the $8 anchor ($M$ = $8.88) than under both the $20 ($M$ = $9.89) and the $50 anchors ($M$ = $9.88), $t(2732) = 5.88$, $p < .001$; $t(2685) = 5.36$, $p < .001$, respectively, but the two higher anchors produced similar mean payments, $t(2533) = .04$, $p = .971$.

Discussion

Whereas Studies 1 and 2 conceptually replicated past work on anchoring, Studies 3a and 3b were more complicated. Some anchor gaps produced large and reliable effects (e.g., $8 vs. $20) but other gaps did not (e.g., $20 vs. $50). As we will elaborate through this paper, the mix of findings suggests that even with a very large objective anchor gap (absolute value), effects are largely dependent on the subjective gap (percentile rank).

Past research has generally been indifferent to the type of anchor gap. In the process, it has also generally ignored the types of anchors involved (e.g., extremely high, moderately low). The following section gives empirical consideration to types of anchors and types of gaps.

Asymmetries in Anchor-Gaps and Anchor Perception

Study 4: Attempting Narrow Gaps

Study 4 attempts a more conservative test of anchoring. While remaining in a high-powered, ecologically valid, commercial context, we employed a smaller anchor gap. As in Studies 3a and 3b, we vary default prices and measure payments.

Method

Study 4 (N = 3,214) took place with the PWYW media retailer from Studies 1 and 3; the minimum price for this promotion was $1 and the bonus price was $10. Customers were randomly assigned to either a $12 or $15 anchor, relatively common payments. We also ran two additional manipulations involving the option to donate 10% to charity. We did not predict interactions with the default manipulation and not find them, so they are not discussed further.

Results

There were no differences in average payment between the two conditions ($M_{\$12}$ = $8.80 vs. $M_{\$15}$ = $8.88), $t(3212) = 0.47$, $p = .641$.

Both $12 and $15 were above the median payment ($10) and therefore, somewhat uncommon payments. Other work has suggested that uncommonly high anchors are less influential (Mussweiler and Strack 2001). In concept, this null effect occurs because adjustment stops at the boundary of possible values (e.g., Tversky and Kahneman 1974). When two values are above, all adjustments will stop at the boundary, eliminating any difference. However, Mussweiler and Strack (2001) used anchors well beyond the 100th percentile judgment (e.g., 214 as Gandhi's age at his death); whereas our anchors in Study 4, while high, are not nearly so extreme. Magnitude is important, but anchor gap seems necessary to explain the effect.

Study 5: Attempting Wider Gaps

If the $3 anchor gap had translated into a $3 payment difference it would have been financially substantial, but a $3 gap may have seemed subjectively similar to customers, as those were only 10 percentiles apart. Perhaps customers called to mind the same types of supporting information (according to the selective accessibility account) or similarly changed their perceptions of the scale (in line with the scale distortion account). Study 5 titrates and widens the anchor gap to consider these possibilities. We use four levels of anchors to investigate this similarity account, predicting higher anchors would be associated with higher average payments. The sample size was determined to be all purchases made during the promotion.

Method

We conducted a field experiment (N = 1,603 customers) with Vodo (www.vodo.net), a retailer of independently published media. Vodo periodically offers a three-week PWYW promotion for a bundle of several products (e.g., movies, games, etc.). Customers paying more than the current average payment receive four additional products. If they beat the current average payment by more than $7.50, customers receive three additional products. Customers choose their price by either typing in a box or using a sliding scale, which moves in $0.10 increments. The minimum payment is $1.

Site visitors were randomly assigned to anchors of $2, $5, $9, or $12. We intended these amounts to fall equally on either side of the average payment, in the hope that this would match objective anchor gaps to subjective gaps (past promotions had averages between $4 and $6).

Results

The promotion did better than expected, resulting in average payments between $9 and $12 across the three-week period. Therefore, the three low anchors ($2, $5, and $9) represented

similar percentile payments: 30.0, 31.4, and 32.3, respectively, with the highest anchor, $12, generally remaining above average and resulting in a 60.7 percentile rank (using inclusive percentiles). The four means were similar, $F(3, 1599) = 0.21$, $p = .890$. The largest gap of the set, $2 vs. $12, actually showed a non-significant trend opposite the anchors ($M_{\$2} = \$11.48$, $M_{\$12} = \$11.29$), $t(757) = .34$, $p = .736$. We interpret this null effect as suggesting that the anchor gap needed to be even larger.

Study 6a: Considering High Anchors

Study 6a attempts to widen the subjective anchor gap once again, however, a second interest was in considering how to handle unusually common payments. Consider a $1 payment for a doughnut purchase in Study 2: a $0.99 payment is at the 66th percentile, but because 25% of customers pay $1, a $1 payment might alternatively be interpreted as a 91th percentile payment, a 66th percentile payment, or something in between. In Study 6a, we use the $1 payment as the lower anchor and compare it to the unambiguously high anchor of $3 (95th percentile). Three additional conditions considered a true control (PWYW without any anchor) and two fixed price conditions (equivalent to the anchors) allow for assessment of overall demand.

Method

The experiment environment was identical to that of Study 2. We sold glazed doughnuts on the UC Berkeley campus for 17 days from October 2013 to December 2013. People (N = 44,483) passing our doughnut stand saw one of our five shop signs: "Dream Fluff Doughnuts! ($1, $3, Pay What You Want, $1 or Pay What You Want, or $3 or Pay What You Want)." See Figure 4.1.7. in Appendix for a signage example. The sign changed in a randomized order after every 200 people passed by. We recorded the date and time of transactions, number of passersby, purchase price, customer group size, gender, and approximate age of customers. We predetermined to collect data until we had at least 100 observations in the largest of the PWYW conditions.

Results

The total of 393 groups of customers (N = 501 individuals) bought 440 doughnuts. In line with Study 2, we used individual doughnut purchase as unit of analysis with the average payment per doughnut as a dependent measure. We excluded 37 purchases by the experimenters' friends, seven purchases in which customers were not assigned to a randomized pricing condition (e.g., because they arrived before the signs could be switched), and one purchase in which the payment information was missing from our analysis, which left us 395 purchases for analysis.

Average payments did not differ between $1 or PWYW and $3 or PWYW conditions, ($M_{\$1 \text{ or PWYW}} = \$0.91$ vs. $M_{\$3 \text{ or PWYW}} = \$0.84$), $t(189) = .79$, $p = .436$, $d = .11$. Based on these results, it appears that customers saw this common (i.e., chosen by 62%) anchor as a higher one, making our subjective anchor gap much smaller than we had wanted. Payments under both anchor conditions were also higher than under the anchor-free condition ($M_{\$1 \text{ or pwyw}}$ vs. $M_{\text{pwyw}} = \$0.72$), $t(217) = 3.28$, $p = .001$, and ($M_{\$3 \text{ or pwyw}}$ vs. M_{pwyw}), $t(182) = 1.34$, $p = .182$.

Study 6b: Considering Low Anchors

In order to further probe the potential differential in anchor consideration in Study 6a (i.e., the null effect between a $3 anchor and a PWYW-only control, but the significant difference between the $1 anchor and control), we implemented a somewhat similar design in Study 6b. To better test this and to ensure we would find anchoring effects, Study 6b employs

anchor gaps closer to lab-levels of extremity (16^{th} vs. 67^{th} percentiles), while retaining our control condition. This also allows us to test whether participants would reject *all* extreme anchors—both high and low—or whether we would find another asymmetry, in magnitude.

Method

Visitors (N= 431 groups of visitors or 909 individuals) to the Cartoon Art Museum on the Pay-What-You-Wish Days in June, July, August, and September 2014 were randomly assigned to one of the four conditions: PWYW, pay nothing or PWYW, pay $0.01 or PWYW, pay $5 or PWYW. We randomized the experimental condition by changing it every 10 groups of visitors. This admission pricing manipulation was verbally delivered to the visitors by the museum staff (i.e., our research assistants) at the reception desk. For example, visitors in the second condition were told, "Thanks for coming to the museum today. You can pay $0.01 or pay-what-you-want for your admission. How much would you like to pay?" We predetermined to collect data until we had at least 100 observations per condition.

Results

We analyzed group of visitors as unit of analysis with the average admission payment per person per group as the main dependent variable, a specification that makes sense and that we have used previously (Jung, Nelson, Gneezy, and Gneezy 2014). Due to a miscommunication, only three of our four conditions were run in June. We excluded these data from the following analyses but including them does not change the direction or significance of the results. For the remaining three months, the month variable did not influence the payment amount significantly and is not discussed any further.

The average payment amount differed significantly across the four conditions, $F(3, 427) = 3.63$, $p = .013$. The average payments in the two low anchor (nothing vs. $0.01) conditions did not differ, (*Ms* =$2.55 vs. $2.47), $t(212)=.24$, $p = .815$. Visitors paid more when they paid what they wanted and were not provided any anchor than when they were given a choice between paying nothing and paying what they wanted, but this payment difference was only marginally significant, (*Ms* =$3.24 vs. $2.55), $t(210) = 1.81$, $p = .072$. They paid significantly more when they were not provided with any anchor than when they could pay $0.01 or pay what they wanted, (*Ms* =$3.24 vs. $2.47), $t(230) = 2.28$, $p = .024$.

In line with Study 6a, visitors did not pay more when they were provided with a very high anchor value ($5) than when they were not provided with any anchor, ($M_{\$5}$ = $3.47 vs. $M_{control}$=$3.24), $t(215) = .60$, $p = .549$. But they paid more in the high anchor condition than in the nothing vs. pay-what-you-want, (*Ms* =$3.47 vs. $2.55), $t(197) = 2.28$, $p = .024$, or in the $0.01 vs. pay-what-you-want conditions, (*Ms* =$3.47 vs. $2.47), $t(217) = 2.79$ $p = .006$.

Discussion

Study 6 showed that more explicit consideration of the anchor, as in most lab studies, does not facilitate anchoring effects with smaller anchor gaps. In fact, in Studies 4-6, we found overall that the subjective gap between the anchors must be quite wide to elicit significant effects, suggesting that the anchor gaps chosen by convention in the anchoring literature may actually be the minimum requirement.

Study 6a also revealed a previously undiscovered nuance: if the gap is too wide, leaving anchors to be too extreme, customers may reject them from consideration and behave as if they saw an only slightly high anchor. Moreover, this aversion to extremeness appears to be asymmetric in our data: while customers may be put off by an implication that they should make large payments, they embrace the tacit permission to pay a small amount (see Croson and Shang

(2008) for a similar effect of low anchors through social information in donations). We investigate this finding further in a hypothetical domain in Study 14d.

Testing Insights from the Literature in the Field

Having identified and at least partially addressed empirical gaps in the literature regarding perception of anchor gaps and magnitudes, we move away from these issues and focus more broadly on other potential insights from the anchoring literature. This large body of work rightly suggests that anchor size and gaps are not the sole factors that drive anchoring effects. These additional proposed factors have been hypothesized to replicate outside the lab, but have not actually been tested. Because, in real commercial settings, instantiating the sufficiently wide gap may be challenging for companies, we look to these insights to help facilitate anchoring effects with smaller, more reasonable gaps. In Studies 7-13, we complicate our previously simple designs to implement various manipulations that are believed to enhance anchoring effects.

Study 7: Anchors That Inform About The Behavior Of Others

In Study 7, we manipulated whether customers were informed about the average payments of others. Using the average payment as the anchor has a couple of advantages. We could be confident that the anchors were plausible numbers and therefore plausibly more influential. Armstrong-Soule and Madrigal (2014) demonstrated in a hypothetical context that anchors that set injunctive social norms would be more influential than those that set descriptive norms (i.e., from the company, as in previous studies) and Smith, Windmeijer, and Wright (2014) found similar results within charitable donations. Hence, we predicted that higher payments would be associated with higher average payments shown.

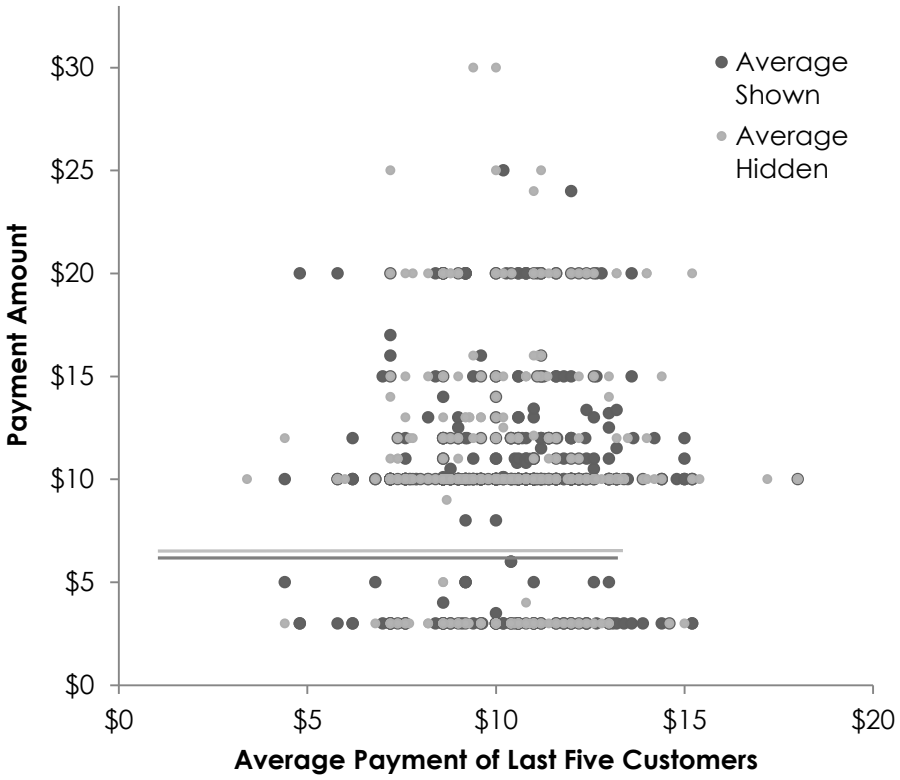
Method

Study 7 (N = 1,074 customers) took place on the same media retailer as in previous studies. This promotion had a minimum price of $3, a bonus price of $10, and a default price of $15. Half of the site's visitors saw the average of the last five payments, labeled "Current average purchase price," above the payment slider. We chose the average of the last five payments over the cumulative average to ensure substantial variability in the anchors shown. The average was updated every seven minutes. The other half saw no average price information.

Results

The average price ranged between $4.40 and $18.00 with a median of $10.27. We exclude the first 115 customers from our analyses, as they were not shown an average price. Contrary to our hypothesis, there was no correlation between the payments of those who saw the average price and the price they saw, $r = .005$, $p = .909$. Even if the relationship had been reliable, we would be very concerned about a confound (e.g., people who buy at high payment times are also more likely to see high average payments). To address this concern, in the average hidden condition, we tracked value of the anchor a customer would have seen. Since that anchor was unseen, if it were related to payments then we would suspect the confound. Accordingly, our critical analysis was to regress average price (whether real or placebo), condition, and their interaction on actual payments. This interaction was not significant, $t(962) = 0.42$, $p = .676$.

Figure 4.2. Payments and Regressions from Study 7



Study 8: Average Payment Information And Price Defaults

Study 8 tests a combination of manipulations from the previous studies in the hopes of identifying a mix that influenced payment (e.g., Croson and Shang, 2008). Participants saw one of three suggested prices ($0.50, $1, or $2) and approximately half of the participants were told that the average payment was $1[39]. For this study, we manipulated the price information in the context of elective admission donations at a Bay Area children's museum. The museum hosts "Free Wednesday" on the first Wednesday of each month. We collected our data on two Free Wednesdays in November 2013 and February 2014.

Method

Groups of visitors (N = 957 groups, 2,761 individual visitors) were randomly assigned to one of eight conditions: 4 (suggested donation amount: No information, $0.50, $1, or $2) x 2 (average donation amount: No information or $1) between-participants study[40]. We selected these numeric values from the distribution of donation amount in the previous month. On average, visitors gave about $0.94 (including 66% of visitors who did not donate).

Each group of visitors was asked to fill out a card that contained our main manipulations. Visitors read, "Today you can donate any amount to support the museum. (The suggested donation amount is [$0.50/$1/$2] per person). (Each visitor to the museum donates $1 on average.) How much would you like to donate?" Visitors also indicated their group size and

home zip code on the card. (Examples of the cards can be found in Figure 4.1.9 in Appendix.) We predetermined to collect data until we had at least 100 groups per condition.

Results

We submitted the average payment per person per group to a 4 (suggested donation: No information, $0.50, $1, or $2) x 2 (average donation: No information or $ 1) ANOVA. The main effects of both the suggested donation amount, $F(3, 949) = .56$, $p = .641$, and the average donation amount, $F(1, 949) = 3.05$, $p = .081$, were not significant. Nor was the interaction between the two variables, $F(3, 949) = .95$, $p = .418$.

Study 9: Anchoring On The Payment Of An Identifiable Other

Knowing about a higher average payment did not lead people to pay more money than knowing about a lower average payment. One possibility for the absence of this effect was that a statistical representation is simply less notable to a customer (as in Small and Loewenstein (2003)). Therefore, in Study 9, instead of presenting participants with an average price, we told them what the previous single customer had paid. We presented customers with a plausible, though not actual payment of the previous customer. We predicted, then, that customers who learned of higher previous payments would pay more than those who learned of lower payments.

Method

For Study 9, we returned to the online PWYW media retailer in their (N=1,175), which had a minimum price of $3, a bonus price of $10, and a default price of $15 for this promotion. Visitors to the site were randomly assigned to see the text "The previous customer paid: $8.00," "The previous customer paid: $12.00," or no additional text above the payment slider.

Results

There was no evidence of anchoring; people paid very similar amounts after learning that the previous customer had paid $8 ($M$ = $11.10) as they did after learning that the previous customer had paid $12 ($M$ = $11.17), $t(769) = 0.27$, $p = .785$.

Study 10: Anchoring With Real Retail Prices

One possible reason for difficulty in Studies 7-9 is that the anchors were too "pushy", and had the effect of partially encouraging participants to discount them. Accordingly, in Study 10, we based one anchor on an explicitly seller-derived fact: the good's true retail price, sometimes labeling it as such. As it happens, this manipulation also made the anchor a precise number, instead of a round one, which, as previously discussed, has been shown to increase the weight on an anchor (e.g., Janiszewski and Uy 2008). We predicted that higher average payments would be associated with higher default prices and inclusion of the retail price.

Method

Study 10 (N=2,190) took place with the same PWYW retailer with the minimum price of $1 and the bonus price of $10. Visitors to the site were randomly assigned to see either a $14 default or a $28.88 default. In addition, half of the visitors saw the message "Full Retail Value: $28.88!" above the payment slider. (Due to a programming error, this manipulation began on the second day of the promotion. For the first day, participants were assigned to only a default condition, without the retail price information.)

Results

Because of the programming error, we divided our analyses into two parts: payments on the first day, without a retail price manipulation and subsequent payments with it enacted. Regardless, there were no significant effects. On the first day, anchors did not affect payments, $t(683) = .02$, $p = .991$. Although this null persisted for the rest of the promotion, payments did trend marginally in the direction of anchoring, as customers in the $14 condition paid less ($M$ = $10.68) than those in the $28.88 condition ($M$ = $11.26), $t(1503) = 1.91$, $p = .057$. The presence of the $28.88 retail price neither produced a main effect by itself, $F(1, 1501) = .72$, $p = .395$, nor did it interact with the default manipulation, $F(1, 1501) = .01$, $p = .92$. Even employing a high, precise, and justified anchor was not sufficient to elicit significant anchoring effects relative to a lower, imprecise, and unjustified one.

Study 11: Anchoring With Suggested Payments

Although justifying the anchor appeared not to work in Study 10, maybe that effect could emerge when paired with a hint of persuasion. In Study 11, we presented people with suggested payments in addition to (and below) justified reference prices.

Method

For Study 11 (N = 429 groups, 1,234 individuals), we returned to the children's museum from Study 8 conducted a 2 (regular admission fee ($11): present or absent) x 2 (suggested donation amount ($5): present or absent) between-participants study. For example, participants in the regular-fee or suggested-fee-present condition saw, "Your admission has been sponsored by ScholarShare[41]. On most days, the general admission is $11 per person. Today you can give a gift of any amount to support the museum. The suggested donation amount is $5 per person. How much would like to give?" All visitors indicated their group size and home zip code.

Results

We analyzed the average donation amount per person as our main dependent variable with groups as unit of analysis. A 2 (regular admission fee ($11): present or absent) x 2 (suggested donation amount ($5): present or absent) ANOVA showed that neither variables influenced payments significantly. Visitors donated similar amounts regardless of whether they saw the regular admission fee, ($M_{\$5\text{\& regular fee present}}$ = $0.96 vs. $M_{\$5\text{\& regular fee absent}}$ = $1.06), $F(1, 425) = 0.56$, $p = .456$. Their payments did not differ whether or not they were provided with a suggested donation amount, ($M_{\$11\text{ \& suggestion present}}$ = $0.96 vs. $M_{\$11\text{\& suggestion absent}}$ = $0.77), $F(1, 425) = 1.26$, $p = .263$. Furthermore, the interaction was not significant, $F(1, 425)=.01$, $p= .909$.

Studies 12a And B: Anchoring With Precise Values

Studies 10 and 11 used anchoring manipulations grounded in justification and legitimate suggestion, but neither showed reliable effects. However, Study 10, in particular, had a nearly marginally significant effect. One possibility is that was due to the use of a precise anchor, a feature shown to increase anchoring effects in other studies by shrinking the unit of adjustment (Janiszewski and Uy 2008; Mason et al. 2013). Studies 12a aimed to test this possibility more thoroughly by seeing whether or not precise anchors produced smaller adjustments than round anchors. Because the anchors in this study were quite high, we reasoned that higher average payments would be associated with precise, versus round anchors.

Method

We again worked with the PWYW media retailer for our experiment, during a promotion (N=714 customers) with a minimum price of $1 and bonus price of $10. Visitors to the site were randomly assigned to see one of five default prices: $19.91, $19.99, $20.00, $20.01, or $20.09.

Results

A one-way ANOVA did not suggest any significant differences in average payment across the five conditions, $F(4, 709)=1.35$, $p = .439$. However, further analysis revealed a significant difference between two conditions: payments under the $19.99 default ($M$ = $12.31) were actually significantly *higher* than those under the $20.01 default ($M$ = $10.52), $t(296) = 2.18$, $p = .030$. This reverse anchoring effect directly contradicted the literature on anchoring and precision. Nevertheless, because this effect was not predicted and was from a relatively small sample, we decided to run a highly direct replication of Study 12a.

This replication (Study 12b) was identical to 12a in domain and design, but had a much larger sample: 4,110 bundles were sold during this promotion. This time, however, no mean differences were significant, $F(4, 4105) = .53$, $p = .731$, including between $19.99 ($M$ = $10.39) and $20.01 ($M$ = $10.60), $t(1663) = 0.86$, $p = .39$.

Study 13: Anchoring With Maximum Possible Payments

In the preceding studies, to different degrees, all of our anchors could be construed as suggestions. Although some manipulations could be seen as especially heavy handed in this respect, past research has suggested that even arbitrary, unexplained defaults are frequently seen this way as well (McKenzie et al, 2006). Perhaps in a real market setting, in which the anchor presenter has incentives for higher payments, nearly any anchor would be entirely ignored. Study 10 sought to counteract this potential problem. Perhaps, we reasoned, we needed an anchoring manipulation that reduced any risk for reactance by in fact encouraging customers to adjust their payments. Rather than manipulating defaults, we manipulated the stated "maximum acceptable payment." In accordance with the anchoring literature, we predicted that as the maximum payment went up, so would average payments.

Method

We conducted a series of experiment at the Cartoon Art Museum (as described in Study 6b) on the museum's pay-what-you-wish day in January, April, May, and July 2012. In each of those months, we tested a set of maximum prices on how much visitors pay for their admission.

In those four periods, groups of visitors (N= 606 groups, 1,349 individuals) were assigned to a set of maximum prices. The receptionist (our research assistant) told the visitors: "Thanks for coming to the Cartoon Art Museum. Today is Pay What You Wish Day. You can pay what you want for your admission. But, the maximum we can accept is [amount in dollars] per person. How much would you like to pay?" In this experiment, we recorded payment amount, group size, time of transaction, and readily identifiable demographic information such as gender and ethnicity. Each month featured a slightly different set of maximum payment anchors: $10, $50, and $100.

Results

Although anchoring effects appeared sporadically over the four month-long periods (see Item 4.3.4 in Appendix for full discussion of the results), they were conflicting and unreliable: a combined analysis dampened any seeming anchoring effect. We analyzed all of the anchor differences while controlling for month-to-month variation (which was substantial). There were no differences in payments between the different anchor conditions except the two highest

anchor conditions, $50 and $100 (see Table 4.1). We are disinclined to read too much into the unpredicted pattern between the $50 and $100 anchors as it has not been observed in any other studies (including other studies conducted at the same museum). Our analysis controlled for month-to-month variation, but there is room for an unidentified systematic influence due to the lack of perfect random assignment across four months of data collection. Though we predict that this pattern would not replicate under better conditions, future research could investigate it.

Discussion

Without the perfectly sized gap discussed in the previous sections, payments seemed generally insensitive to the size of the anchor (low in Study 8, high in Studies 12a and 12b), if and how it was justified (Studies 7-11), its level of precision (Studies 12a and 12b), and its framing as something other than a suggestion (Study 13). It seems these findings from the literature do not translate nearly as well as the effect upon which they build when taken into the domain of anchoring in payment.

Returning to the Lab

Rather than attribute the null effects from the previous to the financial, field component, it is prudent that we consider that the studies could simply have been poorly designed or in confusing contexts (e.g., Vodo in Study 5). To eliminate this concern and affirm the influence of field settings, we replicated four of the above studies as hypothetical designs. If our paradigms are truly at fault, then we should continue to see mostly null results. If, however, it is the inclusion of real money, then we should see significant effects when using hypothetical versions of the experiment. Therefore, in the following section, we chose a representative study from each of the preceding three sections—examining different types of gaps, asymmetries in magnitude, and returning to the literature—to replicate in a hypothetical domain, as well as a conceptual replication that tests all our primary findings in one setting.

Study 14a: Lab Replication Of Study 10

In Study 14a, we attempted to replicate Study 10, which took place on the PWYW media bundle retailer's website with a high default of $28.88 and a low default of $14.00. In addition, because customers use a sliding scale to choose their price whereas most anchoring studies simply provide a text box for participants' estimates, we wanted to make sure this difference was not behind our null effects. Therefore, we also manipulated the response format in Study 14a. We predicted no effect of response format, but, contrary to the results of Study 10 and consistent with the literature, we predicted higher payments under higher defaults.

Method

We aimed to recruit 100 participants per condition, and so recruited 400 participants on Amazon Mechanical Turk (MTurk). All participants were shown a screenshot from the original promotion that displayed all available goods. Participants were asked to imagine they were interested in purchasing the bundle and informed of the minimum and bonus prices. They were then asked how much they would pay for the bundle. Participants in the slider condition answered using a sliding scale; those in the box condition typed their answers into a text box. The default settings of the box or slider depended on condition: $14.00 or $28.88.

Results

Fourteen participants gave an answer less than the minimum price or did not complete the survey and were excluded from analyses, leaving 386 participants. We conducted a 2 (default price: $14 or $28.88) x 2 (response format: slider vs. text box) ANOVA on payment. People paid

slightly more when using the slider than when using the text box (M_{slider} = \$15.28 vs. $M_{textbox}$ = \$13.35), $F(1, 382) = 3.21$, $p = .074$, but, critically, did not interact with the anchoring effect, $F(1, 382) = .49$, $p = .483$. Most important, participants who saw a \$14 anchor paid significantly less (M=\$13.01) than those who saw a \$28.88 anchor (M = \$15.78), $t(384) = 2.68$, $p = .008$, d=0.28. Although this distribution resembled our field sample (e.g., the minimum, bonus, and default prices were all relatively popular), the upper bound of payments was much higher than what we usually see. Therefore, we winsorized our results at the 95th percentile (\$28.88) and repeated our analysis: we still found no significant interaction, $p = .157$, and the anchoring effect remained highly significant, $F(1,382) = 16.01$, $p < .001$, $d = 0.41$. Importantly, though, response format, was even further from statistically significant, $p = .543$.

Study 14b:Lab Replication Of Study 3a

Having eliminated response format as a potential suppressor of anchoring effects, in Study 14b, we varied the presence of the bonus price, another idiosyncrasy in some of our paradigms. Here, we sought to replicate our significant results from Study 3a. We predicted a significant effect of anchor amount, but no effect of bonus presence.

Method

We aimed to recruit 150 participants per condition, and so recruited 613 participants from MTurk. As in Study 14a, participants were shown an image of the goods, informed of the minimum and default prices, and asked how much they would pay. We manipulated the default price: \$3 or \$9. Participants in the bonus-present condition were also informed of the bonus price and goods. Participants in the bonus-absent condition saw the same set of goods, but were not informed of a bonus price (Goods that would have been marked as a bonus had those labels digitally removed).

Results

Thirty-two participants either provided a price below the minimum or failed to complete the survey, leaving us with 581 participants in our analysis. We conducted a 2 (default price: \$3 or \$9) x 2 (bonus presence: yes or no) ANOVA on payments, which yielded null effects for both bonus presence, $F(1, 577) = .00$, $p = .999$ and the interaction with default price, $F(1, 577) = .27$, p=.602. Replicating our results in Study 3a, participants would pay less after seeing a \$3 default (M = \$5.10) than after seeing a \$9 default (M = \$5.99), $F(1, 577) = 10.70$, $p = .001$, $d = 0.28$.

Study 14c: Lab Replication Of Study 6a

Study 14c was a direct, but hypothetical, replication of Study 6a, in which participants had the opportunity to buy a doughnut at a fixed price (the anchor) or PWYW. We predicted that now, as in Study 14a, we would find significant effects in the absence of actual payments.

Method

We aimed to recruit approximately 50 participants per condition in Study 14c, and recruited 169 participants on MTurk. Participants were shown images of the signs used in Study 6a, which read "Dream Fluff Doughnuts! \$1 or PWYW," "…\$3 or PWYW," or simply "…PWYW," depending on condition. They were then asked how much they would pay for a doughnut. We also included exploratory, secondary manipulations of predicting the purchase price for another customer or a retail price. This manipulation was not significant and did not interact with the anchor manipulation, and is not discussed further.

Results

Because we nearly never encountered payments greater than $5 in our field data, we winsorized payments at $5. A one-way ANOVA on payments showed significant differences across conditions, $F(2, 176) = 6.88$, $p = .001$. In line with our predictions, participants in the $1 condition paid significantly less (M = $0.82) than did participants in the $3 condition ($M$ = $1.49), $t(115)=4.08$, $p < .001$. Similar to Study 6a, payments in the $3 condition did not differ from payments in the control condition (M = $1.30), $t(106) = 0.88$, $p = .382$; however, payments in this $1 condition were significantly lower, $t(111) = 2.46$, $p = .015$.

Study 14d: Lab Replication Of Study 2

Having cast doubt on the possibility that the setting of our anchoring studies or the way we elicited estimates caused our null results in Studies 14a-c, in Study 14d, we wanted to investigate further into the role of the anchor gap. That is, we had seen small gaps fail (e.g., Study 4) and large gaps succeed (e.g., Study 1), but knew little about the boundaries of this effect. For Study 14d, we sought to replicate Study 2 while adding additional conditions to look at other, theoretically interesting anchors.

Method

We recruited 2,100 participants from MTurk, as we believed this to be the largest sample we could recruit in a reasonable timeframe. We used images of the signs used in Study 2, which read "Dream Fluff Doughnuts! $X or PWYW." What replaced the $X with "Pay nothing, $0.01, $0.25, $1.75, or $50", depending on the condition. We also included a control condition that did not have an anchor. The $0.25, $1.75, and control condition were taken from Study 2. After providing their estimates, participants received an attention check.

Including $50 and $0.01 allowed us to see how participants used anchors that were beyond or at the boundaries of payments and compare that to how they used anchors that were still extreme, but to a lesser extent. Having a "Pay nothing or PWYW" condition allowed us to examine how participants think of PWYW pricing. One potential concern with using PWYW pricing to study anchoring is PWYW may essentially generate an anchor of $0, which could mute anchoring effects.

Results

We pre-registered the sample size, materials, exclusions, and analyses with the Open Science Framework, available here: https://osf.io/qfjht/. By that pre-registered protocol, we first excluded participants who failed the attention check (n = 68) and then winsorized payments at $5 (which was a 99th percentile payment). Most conditions differed from most other conditions. A one-way ANOVA on payments revealed significant differences across the conditions, $F(5, 2032) = 102.25$, $p < .001$. We replicated Study 2's anchoring effect: participants who saw the $0.25 anchor paid less ($M$ = $0.39) than participants who saw the $1.75 anchor ($M$=$1.09), $t(684) = 16.30$, $p < .001$. Anchoring effects were not capped at the very high value of $1.75, as $50 was even higher ($M_{\$50}$ = $1.52), $t(661) = 5.44$, $p < .001$.

There were some interesting differences at the lowest anchors, as the $0.25 anchor produced a somewhat *lower* mean payment than the $.01 anchor ($Ms$= $0.39 vs. $0.50), $t(739) = 2.93$, $p = .004$, and a *substantially* lower payment than observed in the "pay nothing or PWYW" condition ($M_{\$0}$ = $0.92), $t(693) = 10.56$, $p < .001$.

Although we pre-registered the winsorized specification, it is worth noting that all but one difference was significant when analyzing the untransformed data. That alternative

specification is reported in Table 4.3.3 in Appendix, along with winsorized and natural log means and comparisons for every study.

Discussion

Studies 14a-d demonstrated that our paradigms are likely not at fault for our null results in previous studies. Two studies (14a and b) also showed us that the peculiarities of one of our study sites likely did not suppress anchoring effects. In addition, in Studies 14a and c, we showed significant effects with the same designs that produced null effects in the field. Finally, Study 14d showed that anchors in a hypothetical context can be fairly close together—both objectively and subjectively—and still elicit significant differences in payments.

GENERAL DISCUSSION

Our aims for this paper were simple: we sought to conduct a detailed investigation into the operation of a core judgment process (anchoring) in a meaningful real-life setting (payment). The simplicity of the goals is incommensurate with the complexity of the findings. The sixteen field experiments (and four hypothetical experiments) give cumulative insight into when and why anchoring will influence payments.

Overall, if we were to offer a single summary insight, it would be that the published literature makes anchoring effects look unrealistically large and easy to find. For example, when the anchor gaps were imperfectly titrated, but still looked reasonable, we did not observe anchoring with relatively straightforward defaults (Studies 4, 5, 10, 12a, 12b), with more elaborate references to the payments of others (Studies 7 and 9), or reference to retail prices (Study 10), with direct donations (Studies 6b, 8, and 13), physical purchases (Study 6a), or online payments (Studies 4, 9, 10, 11, and 12). Those null findings tell one (disappointing) story, but there is more than one story in these findings.

We often found anchoring effects and they were occasionally quite large. In order to reconcile the potent and pervasive null findings with the sporadic but emphatic significant results we peered closer at the stimuli. The selection of stimuli, then, offers a second (and more intriguing) story. Journal articles present results with pristine clarity. The unique operationalization of the study – meant to stand in for a larger construct – feels reasonable or even obvious. That is especially true for anchoring, in which researchers can (mostly quite reasonably) decide to choose any arbitrary numerical judgment (e.g., clowns per million residents) and any pair of numbers (7 vs. 70), and anticipate a reliable effect. However, even a modest specification in the context (e.g., payment) renders those assumptions cavalier and ineffective. Despite conscientious efforts to find a reliable anchoring paradigm, it still took many attempts (and tens of thousands of participants) before we had some sense of the variables that mattered. Below we detail what we believe those variables to be.

The lab versus the field. The most obvious of our three speculative explanations is the impact of taking a paradigm out of the lab and into the field and vice versa. Most lab researchers think that their effects will look different (and smaller) in the field. Relative to tidy lab studies, field settings introduce noise in measurement and in manipulation. Consistent with that thinking, when we took some of our null findings out of the field and back into the lab (e.g., Studies 14a and c), we show highly significant results. Nevertheless, we think that the inherent noisiness of the field is an unlikely explanation for the observed effects. In our case, "lab vs. field" stands in for more than just variation in noise and signal. Perhaps it is just that anchoring effects are small when real money is at stake? An unrealistically high anchor was quite influential in

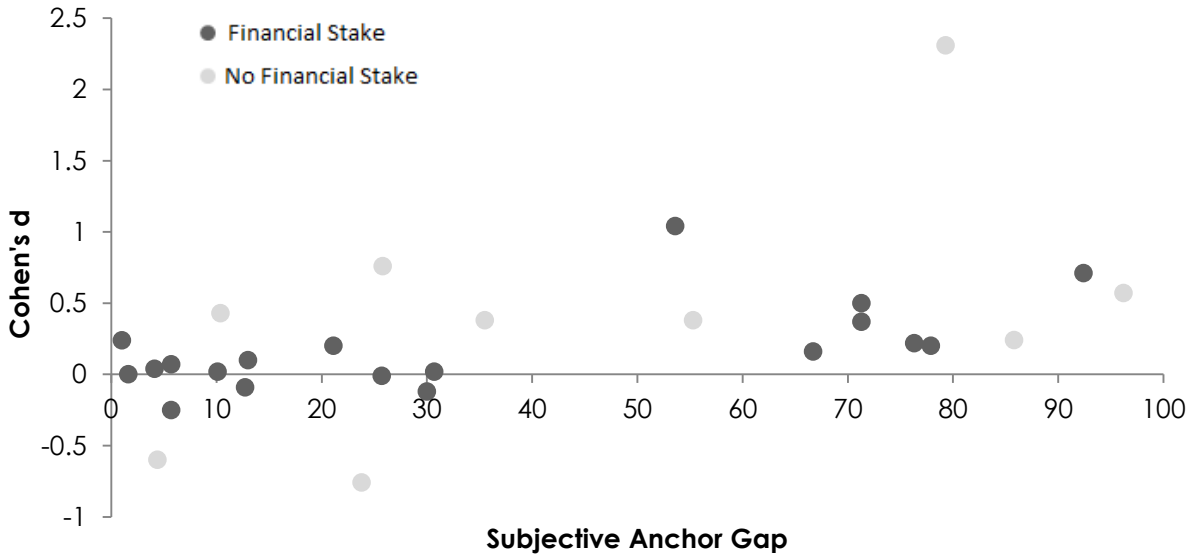
hypotheticality but not in reality. When buying hypothetical confections with hypothetical currency, very large anchors are quite influential.

Perhaps the inclusion of real payments is driving this effect? Although we do not present any lab studies with real payments, we do present the opposite: a field experiment without a payment (i.e., Study 1). That study showed a very large anchoring effect. Our (likely uncontested) guess is real payments are less sensitive to anchors than are hypothetical payments.

Asymmetry of magnitude perception. In a PWYW context, low anchors license low payments. Because customers want to avoid acting stingy without actually paying too much (Gneezy et al. 2012), a low anchor does more than simply distort the scale; it licenses a low payment. High anchors, on the other hand, lack such influence, and could even motivate some reactance. Therefore, in the field, very high anchors may operate as merely slightly high.

Size of anchor gap. It may be the case that extremely high anchors are ignored, but it is certainly the case that the high anchor has to be substantially higher than the low anchor. Moreover, that anchor gap needs to be considered in terms of subjective size – which we conceptualize as how that anchor would appear as a percentile rank across all payments – rather than objective size. In a number of our studies, anchors that were financially far apart (e.g., $20 vs. $50 in Study 3b) yielded no significant differences. We attribute these null effects to the absence of a similarly large gap in their percentile ranks (99th vs. 100th). Consider the dark points (representing field experiments with payments) in Figure 4.3. For every anchor gap smaller than 50 percentile points, we observed non-significant effects and very small effect sizes. For six of the seven gaps larger than 50 percentile points we observed statistically significant effects (and generally larger effect sizes).

Figure 4.3. Subjective Anchor Gap versus Cohen's d Across All Twenty Studies



Though a focus on something as simple as the subjective anchor-gap might seem obvious, it has been missed by previous researchers testing anchoring effects in the field. For example, successful instances of anchoring in the field (e.g., Croson and Shang 2008) derived their anchors from each customer's past donation, placing them somewhat above or below that

amount—a strategy likely to maximize effect sizes by ensuring a large subjective gap with anchors still within the range of consideration. Other attempts, such as Smith and Berger (1996) also based anchors off of prior donations, but in the same direction with a much smaller gap (10% vs. 50% of the prior amount), which may have contributed to their null result. Others still (e.g., Alpizar, et al 2008) failed to account for the distribution entirely, choosing anchors on an objective scale alone (e.g., $2, $5, $10, anchors in a distribution with 48% $0 payments and mean of $2.39).

These three explanations (real vs. hypothetical payments, asymmetries of influence for low and high anchors, and the subjective anchor gap) don't rule out the operation of a more mundane concern with simple noisiness of measurement in the field. Given that we observed some large effects in the field, and generally find evidence that participants are attentive to the anchors (i.e., because the specific anchor values are popular payments), this account is neither parsimonious nor insightful.

Another alternative explanation could be how we provided participants with the anchors. Defaults, while seemingly arbitrary (i.e., lacking explicit justification), are often perceived as recommendations (McKenzie 2006). Anchors are frequently defined as the exact opposite, with several paradigms even informing participants that they are not useful.

First of all, there is every reason to predict that, if an anchor is judged as an intentional and informed suggestion that it would be *more* influential, not less influential (e.g., Armstrong-Soule and Madrigal 2014). Second, there is some reason to believe that even the most explicitly arbitrary anchors are seen as recommendations (Danilowitz, Frederick, and Mochon 2014). For the present studies, we cannot judge whether or not our anchors are perceived as particularly non-arbitrary, nor what the consequence of that judgment would be. Importantly, however, across our studies we do use a mix of anchors—some intended to be more meaningful and others more arbitrary—without a closely corresponding mix of results.

Anchoring effects are extraordinarily robust and replicable when studied in the lab, but can become more subtle and fragile when taken into the field, especially into a monetary domain. The gap between anchors needs to be wide enough (wider than previously believed) to elicit a difference, but perhaps not so wide that the high anchors become extreme and less influential. There is anchoring in payment, but, despite the large literature from the lab, more work is still needed to fully understand it.

CHAPTER 5. CONCLUSION

This research documents consumers' decisions when they are allowed to be both maximally selfish and kind toward others. Our results suggest that consumers' level of selfishness or kindness toward others is quite malleable depending on the salient social forces in their environment. Consumers pay more in response to their emotion when a portion of their payment has a social value of benefitting a charity; they are sensitive to the presence or absence of charity but not to the size of charitable giving (Chapter 2). When told that others are being kind, consumers believed that others are being kinder than they are and pay more for others than for themselves (Chapter 3). When paying for things, consumers do not base their payment heavily on payment anchors, unless they are paying for things hypothetically or an anchor value meant that the price is sufficiently low (Chapter 4). These studies show nuances in consumers' decision-making in achieving their economic goal of maximizing material self-interest and their social goal of being appropriate and kind toward others.

REFERERENCES

Alba, J. W., & Hutchinson, J. W. (2000). Knowledge calibration: What consumers know and what they think they know. *Journal of Consumer Research*, *27*(2), 123-156.

Albert, S., & Whetten, D. A. (1985). Organizational identity. *Research in organizational behavior*.

Alexander, R. D. (1987). *The biology of moral systems*. Transaction Books.

Allport, F. H. *Social psychology*. Boston: Houghton Mifflin, 1924.

Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2005). How much do we care about absolute versus relative income and consumption?. *Journal of Economic Behavior & Organization*, *56*(3), 405-421.

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *The economic journal*, 464-477.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737-753.

Andreoni, J., & Scholz, J. K. (1998). An econometric analysis of charitable giving with interdependent preferences. *Economic Inquiry*, *36*(3), 410-428.

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, 544-555.

Ariely, D., Loewenstein, G., & Prelec, D. (2006). Tom Sawyer and the construction of value. *Journal of Economic Behavior & Organization*, *60*(1), 1-10.

Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, *45*(3), 706.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin*, *117*(3), 497.

Becker-Olsen, K. L., Cudmore, B. A., & Hill, R. P. (2006). The impact of perceived corporate social responsibility on consumer behavior. *Journal of Business Research*, *59*(1), 46-53.

Bem, D. J. (1972). Self-perception theory. *Advances in experimental social psychology*, *6*, 1-62.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of consumer research*, *25*(3), 187-217.

Bodner, R., & Prelec, D. (2003). Self-signaling in a neo-Calvinist model of everyday decision making. *Psychology of economic decisions*, *1*, 105-126.

Bodvarsson, Ö. B., & Gibson, W. A. (1994). Gratuities and customer appraisal of service: Evidence from Minesota restaurants. *Journal of Socio-Economics*,*23*(3), 287-302.

Brewer, N. T., & Chapman, G. B. (2002). The fragile basic anchoring effect.*Journal of Behavioral Decision Making*, *15*(1), 65-77.

Cason, T. N., & Plott, C. R. (2014). Misconceptions and game form recognition of the BDM method: challenges to theories of revealed preference and framing.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, *39*(5), 752.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. *Heuristics and biases: The psychology of intuitive judgment*, 120-138.

Charness, G., & Rabin, M. (2005). Expressed preferences and behavior in experimental games. *Games and Economic Behavior*, *53*(2), 151-169.

Cialdini, R. B. (1993). *Influence*. HarperCollins.
Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*(1), 5-6.
Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance.
Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. *Social Influence*, *1*(1), 3-15.
Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in experimental social psychology*, *24*(20), 1-243.
Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, *58*(6), 1015-1026.
Clark, M. S. (1984). A distinction between two types of relationships and its implications for development. *Boundary areas in social and developmental psychology*, 241-270.
Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of personality and social psychology*, *37*(1), 12.
Clark, M. S., & Mils, J. (1993). The difference between communal and exchange relationships: What it is and is not. *Personality and Social Psychology Bulletin*, *19*(6), 684-691.
Conlin, M., Lynn, M., & O'Donoghue, T. (2003). The norm of restaurant tipping.*Journal of Economic Behavior & Organization*, *52*(3), 297-321.
Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, *21*(3), 241-251.
Croson, R., & Shang, J. Y. (2008). The impact of downward social information on contribution decisions. *Experimental Economics*, *11*(3), 221-233.
Danilowitz, J. S., Frederick, S., & Mochon, D. Anchoring Unveiled. *Psychology*,*83*, 60-74.
David, P., Kline, S., & Dai, Y. (2005). Corporate social responsibility practices, corporate identity, and purchase intention: A dual-process model. *Journal of Public Relations Research*, *17*(3), 291-313.
DellaVigna, S., List, J. A., & Malmendier, U. (2009). *Testing for altruism and social pressure in charitable giving* (No. w15629). National Bureau of Economic Research.
DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics*, *127*(1), 1-56.
Desmet, P., & Feinberg, F. M. (2003). Ask and ye shall receive: The effect of the appeals scale on consumers' donation behavior. *Journal of Economic Psychology*, *24*(3), 349-376.
Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology*, *51*(3), 629.
Drumwright, M. E. (1996). Company advertising with a social dimension: The role of noneconomic criteria. *The Journal of Marketing*, 71-87.
Dufwenberg, M., Gneezy, U., Güth, W., & Van Damme, E. (2001). Direct vs indirect reciprocity: an experiment. *Homo Oecon*, *18*, 19-30.
Dunning, D. (2007). Self-image motives and consumer behavior: How sacrosanct self-beliefs sway preferences in the marketplace. *Journal of Consumer Psychology*, *17*(4), 237-249.
Ellen, P. S., Webb, D. J., & Mohr, L. A. (2006). Building corporate associations: Consumer attributions for corporate socially responsible programs. *Journal of the Academy of Marketing Science*, *34*(2), 147-157.

Epley, N., & Dunning, D. (2000). Feeling" holier than thou": are self-serving assessments produced by errors in self-or social prediction?. *Journal of personality and social psychology*, *79*(6), 861.

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self- generated and externally provided anchors. *Journal of Behavioral Decision Making*, *18*(3), 199-212.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, *54*(2), 293-315.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, *14*(3), 159-181.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817-868.

Fehr, E., Brown, M., & Zehnder, C. (2009). On reputation: A microfoundation of contract enforcement and price rigidity*. *The Economic Journal*, *119*(536), 333-353.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica: journal of the Econometric Society*, 833-860.

Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology*, *70*(6), 1164.

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological review*, *99*(4), 689.

Foreh, M. R., & Grier, S. (2003). When is honesty the best policy? The effect of stated company intent on consumer skepticism. *Journal of Consumer Psychology*, *13*(3), 349-356.

Frederick, S. (2012). Overestimating Others' Willingness to Pay. *Journal of Consumer Research*, *39*(1), 1-21.

Frederick, S. W., & Mochon, D. (2012). A scale distortion theory of anchoring.*Journal of Experimental Psychology: General*, *141*(1), 124.

Frederick, S., & Fischhoff, B. (1998). Scope (in) sensitivity in elicited valuations.*Risk Decision and Policy*, *3*(2), 109-123.

Freeman, S., Walker, M. R., Borden, R., & Latané,, B. (1975). Diffusion of responsibility and restaurant tipping: Cheaper by the bunch. *Personality and Social Psychology Bulletin*, *1*(4), 584-587.

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of consumer research*, 1-31.

Gautier, P. A., & Klaauw, B. V. D. (2012). Selection in a field experiment with voluntary participation. *Journal of Applied Econometrics*, *27*(1), 63-84.

Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of personality and social psychology*, *78*(2), 211-222.

Glazer, A., & Konrad, K. A. (1996). A signaling explanation for charity. *The American Economic Review*, 1019-1028.

Gneezy, A., Gneezy, U., Nelson, L. D., & Brown, A. (2010). Shared social responsibility: A field experiment in pay-what-you-want pricing and charitable giving. *Science*, *329*(5989), 325-327.

Gneezy, A., Gneezy, U., Riener, G., & Nelson, L. D. (2012). Pay-what-you-want, identity, and self-signaling in markets. *Proceedings of the National Academy of Sciences*, *109*(19), 7236-7240.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly journal of economics*, 791-810.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472-482.

Goranson, R. E., & Berkowitz, L. (1966). Reciprocity and responsibility reactions to prior help. *Journal of Personality and Social Psychology*, *3*(2), 227.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American sociological review*, 161-178.

Grant, A., & Dutton, J. (2012). Beneficiary or Benefactor Are People More Prosocial When They Reflect on Receiving or Giving?. *Psychological science*,*23*(9), 1033-1039.

Gray, K., Ward, A.F., & Norton, M.I. (2012). Paying it forward: Generalized Reciprocity and the Limits of Generosity. *Journal of Experimental Psychology*: General.

Greenwald, A. G., & Breckler, S. J. (1985). To whom is the self presented. *The self and social life*, *126*, 145.

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers.*American Economic Review*, 277-282.

Hardesty, D. M., & Suter, T. A. (2013). Maximizing willingness to bid within "Buy It Now" auctions. *Journal of Business Research*, *66*(4), 554-558.

Heyman, J., & Ariely, D. (2004). Effort for payment a tale of two markets.*Psychological science*, *15*(11), 787-793.

Hoeffler, S., & Keller, K. L. (2002). Building brand equity through corporate societal marketing. *Journal of Public Policy & Marketing*, *21*(1), 78-89.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: on the affective psychology of value. *Journal of Experimental Psychology: General*,*133*(1), 23.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, *21*, 1161-1166.

Jung, M. H., Nelson, L. D., Gneezy, A., & Gneezy, U. (2014). Paying more when paying for others. *Journal of personality and social psychology*, *107*(3), 414.

Kamins, M. A., Dreze, X., & Folkes, V. S. (2004). Effects of seller- supplied prices on buyers' product evaluations: Reference prices in an internet auction context. *Journal of Consumer Research*, *30*(4), 622-628.

Karlan, D., List, J. A., & Shafir, E. (2011). Small matches and charitable giving: Evidence from a natural field experiment. *Journal of Public Economics*, *95*(5), 344-350.

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, *322*(5908), 1681-1685.

Kim, J. Y., Natter, M., & Spann, M. (2009). Pay what you want: A new participative pricing mechanism. *Journal of Marketing*, *73*(1), 44-58.

Kim, J. Y., Kaufmann, K., & Stegemann, M. (2014). The impact of buyer–seller relationships and reference prices on the effectiveness of the pay what you want pricing mechanism. *Marketing Letters*, *25*(4), 409-423.

Kruger, J. (1999). Lake Wobegon be gone! The" below-average effect" and the egocentric nature of comparative ability judgments. *Journal of personality and social psychology*, *77*(2), 221.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, *77*(6), 1121-1134.
Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?*. New York: Appleton-Century Crofts.
León, F. J., Noguera, J. A., & Tena-Sánchez, J. (2012). How much would you like to pay? Trust, reciprocity and prosocial motivations in El trato. *Social Science Information*, *51*(3), 389-417.
Lynn, M., & Grassman, A. (1990). Restaurant tipping: an examination of three 'rational'explanations. *Journal of Economic Psychology*, *11*(2), 169-181.
Lynn, M., & McCall, M. (2000). Gratitude and gratuity: a meta-analysis of research on the service-tipping relationship. *Journal of Socio-Economics*, *29*(2), 203-214.
Mak, V., Zwick, R., & Rao, A. R. (2010, May). "Pay what you want" as a profitable pricing strategy: Theory and experimental evidence. In *Proceedings of the Management Science and Operations Seminar Series*.
Mak, V., Zwick, R., & Rao, A. R. (2010, May). Pay what you want" as a profitable pricing strategy: Theory and experimental evidence. In *Proceedings of the Management Science and Operations Seminar Series*.
Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *The American Economic Review*,*104*(1), 277-290.
Martin, R., & Randal, J. (2008). How is donation behaviour affected by the donations of others?. *Journal of Economic Behavior & Organization*, *67*(1), 228-238.
Mason, M. F., Lee, A. J., Wiley, E. A., & Ames, D. R. (2013). Precise offers are potent anchors: Conciliatory counteroffers and attributions of knowledge in negotiations. *Journal of Experimental Social Psychology*, *49*(4), 759-763.
Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, *415*(6870), 424-426.
Miller, D. T., & Effron, D. A. (2010). Chapter Three-Psychological License: When it is Needed and How it Functions. *Advances in experimental social psychology*, *43*, 115-155.
Miller, D. T., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and social Psychology*,*53*(2), 298.
Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of personality and social psychology*,*74*(1), 53.
Mohamud, A. (2007) Nearly Half of all U.K. Downloaders Pay for New Radiohead Album, Press Release, retrived from: http://www.comscore.com/Insights/Press_Releases/2007/11/UK_Radiohead_Downloaders
Mohin, T. J. (2012). Changing Busine Corporations (pp. 191-192). *UK, Greenleaf*.
Mohr, L. A., & Webb, D. J. (2005). The effects of corporate social responsibility and price on consumer responses. *Journal of Consumer Affairs*, *39*(1), 121-147.
Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of personality and social psychology*, *81*(1), 33.
Mussweiler, T., & Strack, F. (2001). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition*, *19*(2), 145-160.
Nowak, M. A., & Sigmund, K. (1998). The dynamics of indirect reciprocity.*Journal of theoretical Biology*, *194*(4), 561-574.

Nunes, J. C., & Boatwright, P. (2004). Incidental prices and their effect on willingness to pay. *Journal of Marketing Research*, *41*(4), 457-466.
Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867-872.
Osterhus, T. L. (1997). Pro-social consumer influence strategies: when and how do they work?. *The Journal of Marketing*, 16-29.
Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and Persuasion* (pp. 1-24). Springer New York.
Poggioli, S. (2013) EU Embraces 'Suspended Coffee': Pay It Forward With A Cup of Joe, *National Public Radio* Retrieved from http://www.npr.org/blogs/thesalt/2013/04/24/178829301/eu-embraces-suspended-coffee-pay-it-forward-with-a-cup-of-joe
Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717-731.
Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of personality and social psychology*, *64*(2), 243.
Rabin, M. (1993). Incorporating fairness into game theory and economics.*Advances in Behavioral Economics*, 297.
Regner, T., & Barria, J. A. (2009). Do consumers pay voluntarily? The case of online music. *Journal of Economic Behavior & Organization*, *71*(2), 395-406.
Regner, T., & Barria, J. A. (2009). Do consumers pay voluntarily? The case of online music. *Journal of Economic Behavior & Organization*, *71*(2), 395-406.
Regner, T., & Riener, G. (2012). *Voluntary payments, privacy and social pressure on the internet: A natural field experiment* (No. 82). DICE Discussion Paper.
Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The transsituational influence of social norms. *Journal of personality and social psychology*, *64*(1), 104.
Reputation Institute's CSR RepTrak 2013, retrieved: http://www.reputationinstitute.com/thought-leadership/global-reptrak-100
Riener, G., & Traxler, C. (2012). Norms, moods, and free lunch: Longitudinal evidence on payments from a Pay-What-You-Want restaurant. *The Journal of Socio-Economics*, *41*(4), 476-483.
Rogelberg, S. G., Barnes-Farrell, J. L., & Creamer, V. (1999). Customer service behavior: The interaction of service predisposition and job characteristics. *Journal of Business and Psychology*, *13*(3), 421-435.
Schelling, T. C. (1968). The life you save may be your own. *Problems in public expenditure analysis*, *127*.
Schlenker, B. R. (1986). Self-identification: Toward an integration of the private and public self. In *Public self and private self* (pp. 21-62). Springer New York.
Schmidt, K. M., Spann, M., & Zeithammer, R. (2014). Pay what you want as a marketing strategy in monopolistic and competitive markets. *Management Science*.
Schmidt, K., Spann, M., & Zeithammer, R. (2012). Pay What You Want as a Marketing Strategy in Monopolistic and Competitive Markets. *Available at SSRN 2191934*.

Schroeder, C. M., & Prentice, D. A. (1998). Exposing Pluralistic Ignorance to Reduce Alcohol Use Among College Students1. *Journal of Applied Social Psychology*, *28*(23), 2150-2180.

Seiter, J. S., & Weger Jr, H. (2010). The Effect of Generalized Compliments, Sex of Server, and Size of Dining Party on Tipping Behavior in Restaurants1.*Journal of Applied Social Psychology*, *40*(1), 1-12.

Shang, J., & Croson, R. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal*, *119*(540), 1422-1439.

Sherif, M. (1936). *The psychology of social norms*. Harper & Brothers.

Silverman, W. K., Robertson, S. J., Middlebrook, J. L., & Drabman, R. S. (1984). An investigation of pledging behavior to a national charitable telethon.*Behavior Therapy*, *15*(3), 304-311.

Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors?. *Journal of personality and social psychology*, *99*(6), 917.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, *22*(11), 1359-1366.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (in press). P-Curve: A key to the file drawer. *Journal of Experimental Psychology: General.*

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2014). Anchoring is not a false-positive: Maniadis, Tufano, and List's (2014)'failure-to-replicate'is actually entirely consistent with the original. *Tufano, and List's*.

Simonson, I., & Drolet, A. (2004). Anchoring effects on consumers' Willingness- to- Pay and Willingness- to- Accept. *Journal of consumer research*, *31*(3), 681-690.

Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, *26*(1), 5-16.

Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, *26*(1), 5-16.

Small, D. A., & Simonsohn, U. (2008). Friends of victims: personal experience and prosocial behavior. *Journal of Consumer Research*, *35*(3), 532-542.

Smith, A. (1937). *The wealth of nations [1776]* (p. 421). na.

Smith, G. E., & Berger, P. D. (1996). The impact of direct marketing appeals on charitable marketing effectiveness. *Journal of the Academy of Marketing Science*, *24*(3), 219-231.

Smith, S., Windmeijer, F., & Wright, E. (2014). Peer effects in charitable giving: Evidence from the (running) field. *The Economic Journal*.

Snyder, M. L., Kleck, R. E., Strenta, A., & Mentzer, S. J. (1979). Avoidance of the handicapped: an attributional ambiguity analysis. *Journal of Personality and social Psychology*, *37*(12), 2297.

Social Investment Forum. 2009. 2009 Report on Socially Responsible Investing Trends in the United States. Washington, D.C.: Social Investment Forum.

Soule, C. A. A., & Madrigal, R. (2014). Anchors and norms in anonymous pay-what-you-want pricing contexts. *Journal of Behavioral and Experimental Economics*.

Spann, M., Häubl, G., Skiera, B., & Bernhardt, M. (2012). Bid-elicitation interfaces and bidding behavior in retail interactive pricing. *Journal of Retailing*,*88*(1), 131-144.

Stewart, N. (2009). The cost of anchoring on credit-card minimum repayments.*Psychological Science*, *20*(1), 39-41.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*(3), 437.

Strahilevitz, L. J. (2003). Charismatic code, social norms, and the emergence of cooperation on the file-swapping networks. *Virginia Law Review*, 505-595.

Tesser, A., Campbell, J., & Mickler, S. (1983). The role of social pressure, attention to the stimulus, and self- doubt in conformity. *European Journal of Social Psychology*, *13*(3), 217-233.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly review of biology*, 35-57.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.

Vesterlund, L. (2003). The informational value of sequential fundraising. *Journal of Public Economics*, *87*(3), 627-657.

Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *science*, *314*(5802), 1154-1156.

Weaver, R., & Frederick, S. (2012). A reference price theory of the endowment effect. *Journal of Marketing Research*, *49*(5), 696-707.

Webb, D. J., & Mohr, L. A. (1998). A typology of consumer responses to cause-related marketing: From skeptics to socially concerned. *Journal of Public Policy & Marketing*, 226-238.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, *125*(4), 387.

Wojciszke, B., Brycz, H., & Borkenau, P. (1993). Effects of information content and evaluative extremity on positivity and negativity biases. *Journal of Personality and Social Psychology*, *64*(3), 327.

APPENDIX

CHAPTER 2. Signaling Virtue: Charitable Behavior under Consumer Elective Pricing

Figure 2.1.1. Examples of Store Sign Used in Study 1

Figure 2.1.2. Reusable Grocery Bag with a Charity or Commercial Logo in Study 1.

Table 2.1.2
Study 1: Summary of the Results per Condition

Logo Type and Pricing	Total Number of Passerby	Number of Customers	Purchase Rate	Average Payment Per Bag (Std. Error)	Standard Deviation	Median
Commercial logo Bag at PWYW	2489	113	4.94%	$1.41 (0.12)	$1.3	$1
Commercial logo Bag at PWYW & 1% to charity	2546	86	3.65%	$2.43 (0.33)	$3.04	$1.98
Commercial logo Bag at PWYW & 50% to charity	2489	69	2.81%	$3.71 (0.55)	$4.53	$2
Charity logo Bag at PWYW & 1% to charity	2319	64	3.10%	$1.98 (0.20)	$1.61	$2
Charity logo Bag at PWYW & 50% to charity	2551	68	2.90%	$4.25 (0.58)	$4.82	$3

Table 2.1.3
Study 1: Mean Payment Comparisons (collapsed across the sign variable)

	Average Payment	t-test, p-value	Log (Average Payment+1)	t-test, p-value
0% vs. 1%	$1.40 vs. $2.24	$t(261) = 3.19$, $p = .002$	0.33 vs. 0.43	$t(261) = 3.29$, $p = .001$
0% vs. 50%	$1.40 vs. $3.98	$t(248) = 5.68$, $p < .001$	0.33 vs. 0.57	$t(248) = 6.26$, $p < .001$
1% vs. 50%	$2.24 vs. $3.98	$t(285) = 3.97$, $p < .001$	0.43 vs. 0.57	$t(285) = 3.72$, $p < .001$

Table 2.1.4
Study 1: Mean Payment Comparisons

	Purchase Price	t-test, p-value	Log_{10} (Purchase Price+1)	t-test, p-value
Commercial Logo				
0% vs.1%	\$1.40 vs. \$2.43	$t(197) = 3.22$, $p = .002$	0.33 vs. 0.43	$t(197) = 3.00$, $p = .003$
0% vs. 50%	\$1.40 vs. \$3.71	$t(180) = 5.09$, $p < .001$	0.33 vs. 0.53	$t(180) = 5.02$, $p < .001$
1% vs. 50%	\$2.43 vs. \$3.71	$t(153) = 2.11$, $p = .046$	0.43 vs. 0.53	$t(153) = 2.01$, $p = .051$
Charity Logo				
0% (with the commercial logo) vs. 1%	\$1.40 vs. \$1.98	$t(175) = 2.59$, $p = .011$	0.33 vs. 0.42	$t(175) = 2.60$, $p = .010$
0% (with the commercial logo) vs. 50%	\$1.40 vs. \$4.25	$t(179) = 5.93$, $p < .001$	0.33 vs. 0.58	$t(179) = 6.17$, $p < .001$
1% vs. 50%	\$2.43 vs. \$4.25	$t(130) = 3.58$, $p < .001$	0.43 vs. 0.58	$t(130) = 3.28$, $p = .001$
1% to Charity				
Commercial Logo vs. Charity Logo	\$2.43 vs. \$1.98	$t(148) = 1.08$, $p = .284$	0.43 vs. 0.42	$t(148) = .38$, $p = .702$
50% to Charity				
Commercial Logo vs. Charity Logo	\$3.71 vs. \$4.25	$t(135) = .67$, $p = .506$	0.53 vs. 0.58	$t(135) = .87$, $p = .388$
0% vs. Charity Percentage Average	\$1.40 vs. \$3.80	$t(398) = 4.55$, $p < .001$	0.33 vs. 0.49	$t(398) = 5.13$, $p < .001$

Table 2.1.5
Study 1: Purchase Rate Comparisons

	Purchase Rates	Chi-Square, p-value
With the Commercial Logo		
0% vs.1%	4.54% vs. 3.38%	$\chi^2 = 4.48$, $p = .034$
0% vs. 50%	4.54% vs. 2.77%	$\chi^2 = 11.04$, $p < .001$
1% vs. 50%	3.38% vs. 2.77%	$\chi^2 = 1.55$, $p = .213$
With the Charity Logo		
0% (with the commercial logo) vs. 1%	4.54% vs.2.76%	$\chi^2 = 10.73$, $p = .001$

0% (with the commercial logo) vs. 50%	4.54% vs. 2.67%	$\chi 2 = 12.78$, $p < .001$
1% vs. 50%	3.38% vs. 2.77%	$\chi^2 = 0.04$, $p = .84$
1% to Charity		
Commercial Logo vs. Charity Logo	3.38% vs. 2.76%	$\chi^2 = 1.55$, $p = .213$
50% to Charity Commercial Logo vs. Charity Logo	2.76% vs. 2.67%	$\chi^2 = 0.05$, $p = .816$
0% vs. Charity Percentage Average	4.54% vs. 2.90%	$\chi^2 = 17.18$, $p < .001$

Table 2.1.6
Study 1: Purchase Rate Comparisons

	Purchase Rates	Chi-Sqaure, p-value
0% vs. 1%	4.54% vs. 3.08%	$\chi^2 = 10.13$, $p = .001$
0% vs. 50%	4.54% vs. 2.72%	$\chi^2 = 17.22$, $p = .001$
1% vs. 50%	3.08% vs. 2.72%	$\chi^2 = 1.17$, $p = .279$

Table 2.1.7
Study 1: Purchase Price Comparison (collapsed across charitable conditions)

	Average Payment	t-test, p-value	Log(Payment+1)	t-test, p-value
Commercial Logo (excluding 0%) vs. Charity Logo	\$3.00 vs. \$3.14	$t(285) = .326$, $p = .745$	0.48 (0.30) vs. 0.50 (0.30)	$t(285) = .684$, $p = .494$

Table 2.1.8
Study 1: Purchase Rate Comparison

	Purchase Rates	Chi Square, p-value
Commercial Logo (excluding 0%) vs. Charity Logo	3.08% vs. 2.71%	$\chi^2 = 1.19$, $p = .275$

Figure 2.2.1. Store Signs Used in Study 2.

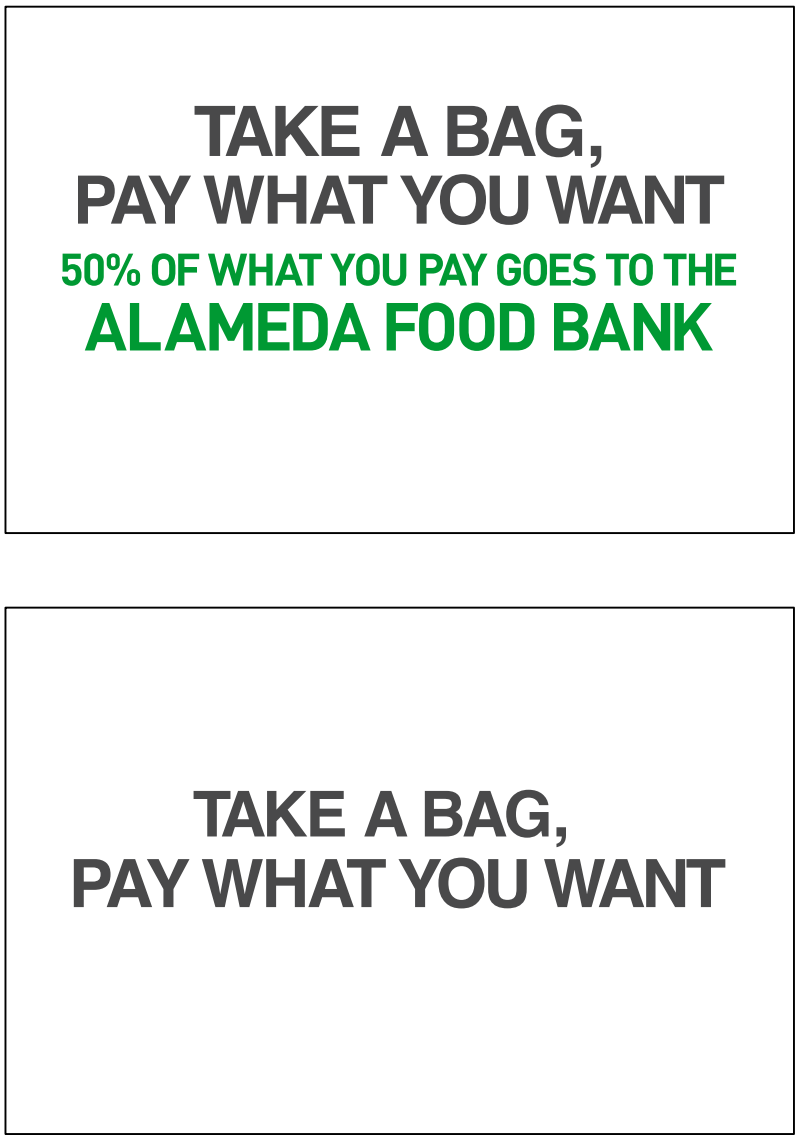


Table 2.2.1
Study 2: Descriptives

Percentage to Charity	**0%**	**1%**	**50%**	**99%**	**100%**
Number of Bags sold in 2012	56	47	40	39	45
Number of Bags sold in 2013	126	84	85	97	95
Total Transactions	182	131	125	136	140
Total Passerby in 2012	2069	1926	2169	2029	1956
Total Passerby in 2013	3295	3315	3517	3338	3447
Total Passerby	5364	5241	5686	5367	5433
Average Payment in 2012 (St. Error)	$1.26 (0.20)	$3.13 (0.50)	$4.59 (0.95)	$5.83 (1.01)	$4.30 (0.98)
Standard Deviation of Mean in 2012	$1.49	$3.43	$6.00	$6.32	$6.55
Average Payment in 2013 (St. Error)	$0.85 (0.07)	$1.76 (0.27)	$2.05 (0.24)	$1.99 (0.30)	$2.92 (0.55)
Standard Deviation of Mean in 2013	$0.77	$2.46	$2.17	$2.97	$5.37
Average Purchase Rate 2012	2.7%	2.4%	1.8%	1.9%	2.3%
Average Purchase Rate 2013	3.8%	2.5%	2.4%	2.9%	2.8%

Table 2.2.2
Study 2: Mean Payment and Log_{10} (Mean Payment +1) Comparisons

	Purchase Price	t-test, p-value	Log_{10} (Purchase Price+1)	t-test, p-value
0% vs. 1%	\$0.98 vs. \$2.24	$t(311)=5.44$, $p < .001$	0.26 vs. 0.41	$t(311)=6.15$, $p < .001$
0% vs. 1% (2012)	\$1.26 vs. \$3.13	$t(101)=3.69$, $p < .001$	0.30 vs. 0.52	$t(101)=4.49$, $p < .001$
0% vs. 1% (2013)	\$0.85 vs. \$1.76	$t(208)=3.88$, $p < .001$	0.24 vs.0.35	$t(208)=4.18$, $p < .001$
0% vs. 50%	\$0.98 vs. \$2.86	$t(305)=6.07$, $p < .001$	0.26 vs. 0.45	$t(305)=7.11$, $p < .001$
0% vs. 50% (2012)	\$1.26 vs. \$4.59	$t(94)=3.99$, $p < .001$	0.30 vs. 0.55	$t(94)=3.95$, $p < .001$
0% vs. 50% (2013)	\$9.85 vs. \$2.05	$t(209)=5.69$, $p < .001$	0.24 vs. 0.41	$t(209)=6.25$, $p < .001$
0% vs. 99%	\$0.98 vs. \$3.09	$t(316)=6.09$, $p < .001$	0.26 vs. 0.47	$t(316)=7.97$, $p < .001$
0% vs. 99% (2012)	\$1.26 vs. \$5.83	$t(93)=5.22$, $p < .001$	0.30 vs. 0.67	$t(93)=6.06$, $p < .001$
0% vs. 99% (2013)	\$0.85 vs. \$5.83	$t(221)=4.15$, $p < .001$	0.24 vs. 0.39	$t(221)=6.11$, $p < .001$
0% vs. 100%	\$0.98 vs. \$3.36	$t(320)=5.45$, $p < .001$	0.26 vs. 0.46	$t(320)=6.81$, $p < .001$
0% vs. 100% (2012)	\$1.26 vs. \$4.30	$t(99)=3.37$, $p =.001$	0.30 vs. 0.51	$t(99)=3.46$, $p =.001$
0% vs. 100% (2013)	\$0.85 vs. \$2.92	$t(219)=4.27$, $p < .001$	0.24 vs. 0.43	$t(219)=6.00$, $p < .001$
1% vs. 50%	\$2.25 vs. \$2.86	$t(254)=1.40$, $p =.163$	0.41 vs. 0.45	$t(254)=1.16$, $p =.248$
1% vs. 50% (2012)	\$3.13 vs. \$4.59	$t(85)=1.12$, $p =.160$	0.52 vs. 0.55	$t(85)=.45$, $p =.655$
1% vs. 50% (2013)	\$1.76 vs. \$2.05	$t(167)=.80$, $p =.424$	0.35 vs. 0.41	$t(167)=1.49$, $p =.140$
1% vs. 99%	\$2.25 vs. \$3.09	$t(265)=1.80$, $p =.073$	0.41vs. 0.47	$t(265)=1.80$, $p =.073$
1% vs. 99% (2012)	\$3.13 vs. \$5.83	$t(84)=2.52$, $p =.014$	0.52 vs. 0.67	$t(84)=2.10$, $p =.039$
1% vs. 99% (2013)	\$1.76 vs. \$2.00	$t(179)=.572$, $p =.572$	0.35 vs. 0.39	$t(179)=1.18$, $p =.238$
1% vs. 100%	\$2.25 vs. \$3.36	$t(269)=1.98$, $p =.049$	0.41 vs. 0.46	$t(269)=1.24$, $p =.215$
1% vs. 100% (2012)	\$3.13 vs. \$4.30	$t(90)=1.08$, $p =.284$	0.52 vs. 0.51	$t(90)=.051$, $p =.959$

1% vs. 100% (2013)	$1.76 vs. $2.92	$t(177)=1.82$, $p=.071$	0.35 vs. 0.43	$t(177)=1.87$, $p=.063$
50% vs. 99%	$2.86 vs. $3.09	$t(259)=.44$, $p=.660$	0.45 vs. 0.47	$t(259)=.51$, $p=.610$
50% vs. 99% (2012)	$4.59 vs. $5.83	$t(77)=.90$, $p=.373$	0.55 vs. 0.67	$t(77)=1.33$, $p=.19$
50% vs. 99% (2013)	$2.05 vs. $1.99	$t(180)=.14$, $p=.89$	0.41 vs. 0.39	$t(180)=.40$, $p=.69$
50% vs. 100%	$2.86 vs. $3.36	$t(263)=.81$, $p=.416$	0.45 vs. 0.46	$t(263)=.130$, $p=.897$
50% vs. 100% (2012)	$4.59 vs. $4.30	$t(83)=.21$, $p=.832$	0.55 vs. 0.52	$t(83)=.43$, $p=.672$
50% vs. 100% (2013)	$2.05 vs. $2.92	$t(178)=1.40$, $p=.163$	0.41 vs. 0.43	$t(178)=0.59$, $p=.555$
99% vs. 100%	$3.09 vs. $3.36	$t(274)=.43$, $p=.668$	0.47 vs. 0.46	$t(274)=.36$, $p=.720$
99% vs. 100% (2012)	$5.83 vs. $4.30	$t(82)=1.09$, $p=.280$	0.67 vs. 0.51	$t(82)=1.81$, $p=.074$
99% vs. 100% (2013)	$1.99 vs. $2.92	$t(190)=1.49$, $p=.139$	0.39 vs. 0.43	$t(190)=.978$, $p=.329$

Table 2.2.3
Study 2: Mean Payment (Standard Deviation) Per Condition After Excluding Outliers (Standardized Residuals < 4)

	0%	**1%**	**50%**	**99%**	**100%**
2012	$1.26 (1.49) n = 56	$2.76 (2.36) n= 46	$2.88 (3.15) n = 36	$5.83 (6.32) n = 39	$2.52 (3.00) n = 41
2013	$0.85 (0.77) n = 126	$1.76 (2.46) n = 84	$2.05 (2.17) n = 85	$1.62 (1.40) n=95	$1.95 (2.04) n=91

Table 2.3.1
Study 1-4: Skewness of Payment Amount

	Payment Skewedness (St.Error)	Log_{10}(Payment +1) Skewedness (St.Error)
Study 1	3.15(.122)	0.65 (.122)
Study 2	4.12(.091)	1.274(.091)
Study 3	2.53(.173)	0.644 (.173)
Study 4	3.77(.119)	0.665 (.119)

Table 2.3.2.
Study 3. Summary Descriptives

Pricing	Number of Customers	Average Payment Per Bag (Std. Error)	Standard Deviation	Median
PWYW	60	$2.13 (0.12)	$0.95	$2
PWYW & 10% to charity	66	$2.53 (0.13)	$1.03	$2
PWYW & 50% to charity	67	$3.49 (0.27)	$2.22	$2.50

Table 2.3.3
Study 3. One-Way ANOVA

Dependent Variable: Mean Payment Amount per Cup of Coffee sold per Person	Dependent Variable: Log_{10} (Mean Payment Amount per Cup of Coffee Sold per Person +1)
Pricing: $F(2, 190) = 13.24$, $p < .001$	Pricing: $F(2, 190) = 13.86$, $p < .001$

Figure. 2.4.1. Store Sign Used in Study 4

Table 2.4.1
Study 4: Mean Payment (Standard Deviation, Standard Error) and Sample Size.

	PWYW	SSR with 10% Charity	SSR with 50% Charity
Anonymous Payment	$0.73 ($0.57, 0.07), n=66	$1.10 ($1.09, 0.13), n=74	$0.81 ($0.86, 0.10), n=73
Direct Payment	$0.67 ($0.72, 0.09), n=68	$1.10 ($0.97, 0.12), n=61	$1.13 ($1.24, 0.15), n=70

Table 2.4.2
Study 4 : The Analysis Of Variance

Dependent Variable: Average Payment Amount per Donut per Person	Dependent Variable: Log_{10}(Average Payment Amount per Donut per Person +1)
Main Effects: Pricing: $F(1, 406) = 6.32$, $p = .002$ Anonymity: $F(1, 406) = 0.97$, $p = .325$ Interaction Effect: $F(2, 406) = 1.70$, $p = .183$	Main Effects: Pricing: $F(1, 406) = 8.51$, $p < .001$ Anonymity: $F(1, 406) = 1.30$, $p = .255$ Interaction Effect: $F(2, 406) = 2.76$, $p = .065$

CHAPTER 3: P*aying More for Others*

Table 3.1.1
Studies 1-4 : Skewness of Payment Amounts

	Average Payment Skewedness/Group Level(St.Error)	Log_{10}(Average Payment +1) Skewness/Group Level(St.Error)	Payment Skewedness/Individual Level (St.Error)	Log_{10}(Payment +1) Skewness/ Individual Level(St.Error)
Study 1	1.354(.197)	.222(.197)	.849(.138)	.087(.138)
Study 2	.933(.197)	-.057(.197)	1.264(.130)	.096(.130)
Study 3	1.081(.140)	-234(.140)	1.537(.103)	-.125(.103)
Study 4	.610(.211)	-.748(.211)	.432(.185)	-.854(.185)

Table 3.1.2
Studies 1-4:Mean Payment Amounts

	a. Average payment per person per group	b. Average payment per person per group (controlling for group size)	c. Payment per person	d. Payment per person (controlling for group size)
Study 1	M_{PIF} = \$2.67 vs. M_{PWYW} = \$1.82 $F(1, 149) = 4.74$, p=.031	M_{PIF} = \$2.63 vs. M_{PWYW} = \$1.86 $F(1, 148) = 4.06$, $p = .046$	M_{PIF} = \$3.12 vs. M_{PWYW} = \$1.89 $F(1, 309) = 18.32$, $p < .001$	M_{PIF} = \$2.84 vs. M_{PWYW} = \$2.14 $F(1, 308) = 6.65$, $p = .010$
Study 2	M_{PIF} = \$3.07 vs. M_{PWYW} = \$2.19 $F(1, 150) = 5.33$, $p = .022$	M_{PIF} = \$3.16 vs. M_{PWYW} = \$2.09 $F(1, 149) = 8.26$, $p = .005$	M_{PIF} = \$2.64 vs. M_{PWYW} = \$1.84 $F(1, 351) = 6.79$, $p = .001$	M_{PIF} = \$2.70 vs. M_{PWYW} = \$1.78 $F(1, 350) = 18.85$, $p < .001$
Study 3	M_{PIF} = \$3.59 vs. M_{PWYW} = \$2.64 $F(1, 302) = 10.33$, $p = .001$	M_{PIF} = \$3.56 vs. M_{PWYW} = \$2.66 $F(1, 301) = 9.55$, p =.002	M_{PIF}=\$3.20 vs. M_{PWYW}=\$2.50 $F(1, 556) = 11.45$, p=.001	M_{PIF}=\$3.15 vs. M_{PWYW} = \$2.54 $F(1, 555) = 8.77$, p = 0.003
Study 4	M_{PIF} = \$2.33 vs. M_{PWYW} = \$1.93 $F(1, 130)=6.50$, p = .012	M_{PIF} = \$2.33 vs. M_{PWYW} = \$1.93 $F(1, 129) = 6.45$, $p = .012$	M_{PIF} = \$2.36 vs. M_{PWYW} = \$1.90 $F(1, 171) = 12.13$, $p = .001$	M_{PIF}=\$2.36 vs. M_{PWYW}=\$1.90 $F(1, 170) = 11.70$, $p = .001$

Table 3.1.3
Studies 1-4: Log Transformed Payment Amounts

	a. Log_{10}(Average payment amount per person per group +1)	b. Log_{10}(Average payment amount per person per group (controlling for group size) +1)	c. Log_{10}(Payment amount per person +1)	d. Log_{10}(Payment amount per person(controlling for group size)+1)
Study 1	M_{PIF} = 0.45 vs. M_{PWYW} = 0.35 $F(1, 149) = 4.21$, $p = .042$	M_{PIF} = 0.45 vs. M_{PWYW} = 0.36 $F(1, 148) = 3.50$, $p = .064$	M_{PIF} = 0.49 vs. M_{PWYW} = 0.34 $F(1, 309)=14.60$, $p < .001$	M_{PIF} = 0.44 vs. M_{PWYW} = 0.38 $F(1, 308) = 3.05$, $p = .082$
Study 2	M_{PIF} = 0.51 vs. M_{PWYW} = 0.413 $F(1, 150) = 4.36$ $p= .039$	M_{PIF} = 0.53 vs. M_{PWYW} = 0.40 $F(1, 149) = 7.24$, $p = .008$	M_{PIF} = 0.47 vs. M_{PWYW} = 0.37 $F(1, 351) = 7.25$, $p =. 001$	M_{PIF} = 0.48 vs. M_{PWYW} = 0.36 $F(1, 350) = 8.98$, $p < .001$
Study 3	M_{PIF} = 0.57 vs. M_{PWYW} = 0.48 $F(1, 302)=6.86$, $p=.009$	M_{PIF} = 0.56 vs. M_{PWYW} = 0.48 $F(1, 301) = 6.31$, $p = .013$	M_{PIF} = 0.47 vs. M_{PWYW} = 0.53 $F(1, 556) = 7.72$, $p = .006$	M_{PIF} = 0.47 vs. M_{PWYW} = 0.53 $F(1, 555) = 5.54$, $p = .019$
Study 4	M_{PIF} = 0.51 vs. M_{PWYW} = 0.44 $F(1, 130) = 7.88$, $p = .006$	M_{PIF} = 0.51 vs. M_{PWYW} = 0.43 $F(1, 129)=7.82$, $p=.006$	M_{PIF} = 0.51 vs. M_{PWYW} = 0.44 $F(1, 171) = 13.63$, $p < .001$	M_{PIF} = 0.51 vs. M_{PWYW} = 0.44 $F(1, 170) = 13.19$, $p < .001$

Table 3.2.
Studies 5-8 : List of Collected Measures On a 1-7 Likert-scale (1: Not much at all 4: Neutral 7: Very much)

After the Unrelated Survey Before the Pricing Manipulation
How much have you enjoyed your experience at UC Berkeley?
How much do you like participating in academic research?
To what extent do you feel that you are a generous person?
To what extent do you think that you have more or less money than an average UCB student?
After the Pricing Manipulation (after paying for a coffee mug)
How much money do you predict that the next participant would pay for a mug?
How much do you think that the participant before you had paid for a mug?
How much do you think is the price of the same mug in the UCB bookstore?
To what extent did the following factors (a-k) influence your decision on the amount of your payment?
a. Support of an academic research project
b. Color of the mug
c. Novelty of this experiment
d. $10 payment
e. Average price of a mug
f. Expectation by the experimenter
g. Value of the Cal Logo/UCB Affiliation
h. Cost of manufacturing a mug (only in Study 7)
i. The amount of money the previous participant paid for a mug (only in Study 7. Participants who were not informed about the previous participant's payment were asked)

j. The participant who will participate after me (only in Study 7)

k. The participant who participated before me (only in Study 7)

How many mugs do you own?

How often do you use a mug?

How satisfied are you with the mug? (only in Study 5, 6, and 8)

How satisfied are you with this consumer behavior study experience? (only in Study 5, 6, and 8)

To what extent do you feel that your payment was fair?

To what extent do you feel that the previous participant's payment was fair? (In Study 7 only. Participants who were told about the previous participant's payment were asked.)

Pease list other factors, if any, that influenced your decision on the amount of your payment.

Age

Gender

Ethnicity

Table 3.3.
Studies 5-8: Skewness of the Average Payment Amounts

	Payment Skewedness (St.Error)	Log_{10}(Payment +1) Skewedness (St.Error)
Study 5	1.870(.144)	.249 (.144)
Study 6	1.587(.175)	.227(.175)
Study 7	2.399(.173)	.171(.173)
Study 8	2.396(.135)	.119(.135)

Table 3.4
Study 5: Exclusions of Participants and Analysis of Variance

	Dependent Variable: Payment	Dependent Variable: Log_{10} (Payment +1)
Excluding none	Main Effects: Pricing: $F(1, 288) = 5.47$, $p = .020$ Social Exchange: $F(2, 288) = 0.76$, $p = .469$ Interaction Effect: $F(2, 288) = 0.13$, $p = .882$	Main Effects: Pricing: $F(1, 288) = 2.80$, $p = .096$ Social Exchange: $F(2, 288) = 0.15$, $p = .862$ Interaction Effect: $F(2, 288) = 0.04$, $p = .963$
Excluding 3 who knew the confederate	Main Effects: Pricing: $F(1, 285) = 5.97$, $p = .015$ Social Exchange: $F(2, 285) = 0.82$, $p = .441$ Interaction Effect: $F(2, 285) = 0.17$, $p = .846$	Main Effects: Pricing: $F(1, 285) = 3.30$, $p = .070$ Social Exchange: $F(2, 285) = 0.15$, $p = .860$ Interaction Effect: $F(2, 285) = 0.01$, $p = .990$
Excluding 3 who knew the confederate and 2 who were inebriated	Main Effects: Pricing: $F(1, 283) = 6.06$, $p = .014$ Social Exchange: $F(2, 283) = 0.90$, $p = .410$ Interaction Effect: $F(2, 283) = 0.18$, $p = .836$	Main Effects: Pricing: $F(1, 283) = 3.36$, $p = .068$ Social Exchange: $F(2, 283) = 0.19$, $p = .824$ Interaction Effect: $F(2, 283) = 0.01$, $p = .990$

Excluding 3 who knew the confederate, 2 who were inebriated, and 1 who knew about the study from a friend	Main Effects: Pricing: $F(1, 282) = 6.39$, $p = .012$ Social Exchange: $F(2, 282) = 0.96$, $p = .385$ Interaction Effect: $F(2, 282) = 0.22$, $p = .803$	Main Effects: Pricing: $F(1, 283) = 3.67$, $p = .056$ Social Exchange: $F(2, 283) = 0.24$, $p = .780$ Interaction Effect: $F(2, 283) = 0.02$, $p = .981$
Excluding 3 who knew the confederate, 2 who were inebriated, 1 who knew about the study, and 6 who knew the experimenter	Main Effects: Pricing: $F(1, 276) = 6.03$, $p = .015$ Social Exchange: $F(2, 276) = 0.78$, $p = .459$ Interaction Effect: $F(2, 276) = 0.21$, $p = .810$	Main Effects: Pricing: $F(1, 276) = 3.50$, $p = .063$ Social Exchange: $F(2, 276) = 0.19$, $p = .829$ Interaction Effect: $F(2, 276) = 0.03$, $p = .973$

Table 3.5.1
Study 6: Mean Payment Amounts (sample size, standard deviation)

	Payment Not Shown	Payment Shown
No Message	\$1.13 (n = 48, 1.33)	\$2.14 (n = 48, 1.98)
Message	\$1.96 (n = 52, 2.62)	\$2.54 (n = 45, 2.24)

Table 3.5.2
Study 6: Analysis of Variance

Dependent Variable: Payment	Dependent Variable: Log_{10}(Payment +1)
Main Effects: Payment Reporting: $F(1, 189) = 6.86$, $p = .010$ Message: $F(1, 189) = 4.12$, $p = .044$ Interaction Effect: $F(1, 189) = 0.52$, $p = .470$	Main Effects: Payment Reporting: $F(1, 189) = 11.04$, $p = .001$ Message: $F(1, 189) = 1.83$, $p = .178$ Interaction Effect: $F(1, 189) = 0.42$, $p = .518$

Table 3.6.1.
Study 7: Mean Payment Amounts (sample size, standard deviation)

	No Information of Other's Payment	Information of Other's Payment
Pay-What-You-Want	$1.21 (n = 49, 1.47)	$1.23(n = 49, 0.93)
Pay-It-Forward	$2.48(n = 44, 2.40)	$1.19(n = 56, 0.81)

Table 3.6.2
Study 7: Analysis of Variance

Dependent Variable: Payment Amount	Dependent Variable: Log_{10}(Payment Amount+1)
Main Effects: Pricing: $F(1, 194) = 8.48$, $p = .004$ Information about others' payment: $F(1, 194) = 8.99$, $p = .003$	Main Effects: Pricing: $F(1, 194) = 7.90$, $p = .005$ Information about others' payment: $F(1, 194) = 2.54$, $p = .113$
Interaction Effect: $F(1, 194) = 9.66$, $p = .002$	Interaction Effect: $F(1, 194) = 8.11$, $p = .005$

Table 3.7.1
Study 8: Mean Payment Amounts (sample size, standard deviation)

	No Information of Other's Payment	Previous Participant Paid $0.50	Previous Participant Paid $2.50
Pay-What-You-Want	$1.36 (n = 56, 1.20)	$1.29 (n = 53, 1.29)	$1.76 (n = 56, 1.25)
Pay-It-Forward	$2.56 (n = 53, 2.54)	$0.99 (n = 54, 0.84)	$1.82 (n = 55, 1.07)

Table 3.7.2.
Study 8: Analysis of Variance

Dependent Variable: Payment Amount	Dependent Variable: Log_{10}(Payment Amount+1)
Main Effects: Pricing: $F(1, 321) = 3.91$, $p = .049$ Information about others' payment: $F(2, 321) = 9.43$, $p < .001$	Main Effects: Pricing: $F(1, 321) = 3.87$, $p = .050$ Information about others' payment: $F(2, 321) = 9.04$, $p < .001$
Interaction Effect: $F(2, 321) = 7.84$, $p < .001$	Interaction Effect: $F(2, 321) = 5.19$, $p = .006$

Table 3.8.1
Study 9A: Means and Standard Deviations of Willingness-to-Pay (WTP) for Self and Other

		Pay What You Want (n=200)	Pay It Forward (n=213)
WTP Winsorized at $10	Self	$1.96 (2.14)	$2.73 (2.38)
	Other	$2.14 (2.30)	$2.67 (2.33)
Raw WTP	Self	$2.04 (2.58)	$3.34 (5.65)
	Other	$2.22 (2.70)	$3.28 (5.65)
Natural Log (WTP Plus One)	Self	0.92 (0.55)	1.18 (0.64)
	Other	0.97 (0.57)	1.18 (0.62)
WTP Winsorize at $5	Self	$1.69 (1.32)	$2.39 (1.55)
	Other	$1.82 (1.42)	$2.33 (1.45)

Table 3.8.2
Study 9A: Number and Percent of Participants Who Believed That They and Others Could Pay Any Price

	Customer Before You	You	Customer After You
PWYW Correct	168/200	182/200	170/200
% Correct Answer in PWYW	84%	91%	85%
PIF Correct	182/211	186/211	177/211
% Correct Answer In PIF	86.26%	88.15%	83.89%
Total Correct Answer	350/411	368/411	347/411
% Correct Answer	85.16%	89.53%	84.42%

Table 3.8.3
Study 9B: Means and Standard Deviations of Willingness-to-Pay (WTP) for Self and Other

		PWYW Can (n=196)	PWYW Chance (n=186)	PIF Can (n=185)	PIF Chance (n=189)
WTP Winsorized at $10	Self	$1.68(1.36)	$1.81 (1.52)	$2.55 (1.68)	$2.12 (1.59)
	Other	$1.70 (1.37)	$1.92 (1.50)	$2.43 (1.29)	$2.45 (1.59)
Raw WTP	Self	$1.68 (1.36)	$3.10 (18.26)	$57.63(749.74)	$2.53 (2.22)
	Other	$1.69(1.37)	$3.48 (21.90)	$3.18 (10.94)	$2.27 (1.28)
Natural Log (WTP Plus One)	Self	0.88 (0.43)	0.94 (0.56)	1.21 (0.74)	1.15 (0.45)
	Other	0.90 (0.42)	0.98 (0.56)	1.16 (0.48)	1.11 (0.38)
WTP Winsorized at $5	Self	$1.63 (1.19)	$1.74 (1.28)	$2.45 (1.39)	$2.37 (1.36)
	Other	$1.65 (1.17)	$1.87 (1.32)	$2.36 (1.27)	$2.26 (1.25)

Table 3.8.4
Study 9B: Number and Percent of Participants Who Believed That They and Others Could Pay Any Price

	Customer Before You	You	Customer After You
PWYW Can Correct	190/196	194/196	194/196
% Correct Answer in PWYW Can	96.94%	98.98%	98.98%
PWYW Chance Correct	182/186	181/186	183/186
% Correct Answer in PWYW Chance	97.85%	97.31%	98.39%
PIF Can Correct	176/185	176/185	177/185
% Correct Answer In PIF Can	95.14%	95.14%	95.68%
PIF Chance Correct	177/189	185/189	186/189
% Correct Answer In PIF Chance	93.65%	97.88%	98.41%
Total Correct Answer	725/756	736/756	740/756
% Correct Answer	95.90%	97.35%	97.88%

Table 3.8.5
Study 9C: Means and Standard Deviations of Willingness-to-Pay (WTP) for Self and Other

		PWYW (n=174)	**PIF (n=206)**	**PIF Company (n=159)**	**PIF Singular (n=176)**	**PIF Plural (n=181)**
WTP Winsorized at $10	Self	$1.67 (1.51)	$2.27 (1.78)	$2.33(1.70)	$2.46(1.69)	$2.30 (1.84)
	Other	$1.70 (1.43)	$2.13 (2.27)	$2.24 (1.64)	$2.38 (1.64)	$2.24 (1.70)
Raw WTP	Self	$1.67 (1.51)	$2.58 (4.44)	$2.54 (3.41)	$2.77 (4.43)	$2.55 (3.78)
	Other	$1.70 (1.47)	$2.50 (5.19)	$2.27 (1.83)	$2.69(4.22)	$2.71(5.28)
Natural Log (WTP Plus One)	Self	0.88 (0.43)	1.08 (0.51)	1.11 (0.49)	1.15 (0.50)	1.08 (0.52)
	Other	0.90 (0.42)	1.10 (0.48)	1.08 (0.45)	1.14 (0.49)	1.09 (0.51)
WTP Winsorize at $5	Self	$1.58 (1.15)	$2.12 (1.29)	$2.22 (1.32)	$2.34 (1.31)	$2.16 (1.42)
	Other	$1.64 (1.17)	$2.00 (1.93)	$2.15 (1.33)	$2.28 (1.29)	$2.10 (1.23)

Table 3.8.6
Study 9C: Number and Percent of Participants Who Believed That They and Others Could Pay Any Price

	Customer Before You	You	Customer After You
PWYW Correct	163/174	170/174	168/174
% Correct Answer in PWYW	93.68%	97.70%	97.00%
PIF Correct	192/206	198/206	193/206
% Correct Answer In PIF	93.20%	96.12%	93.69%
PIF Singular Correct	148/159	146/153	148/159
% Correct Answer In PIF	93.08%	95.42%	93.08%
PIF Plural Correct	166/176	164/176	165/176
% Correct Answer In PIF	94.32%	93.18%	93.75%
PIF Company Correct	159/181	169/181	159/181
% Correct Answer In PIF	87.85%	93.37%	87.85%
Total Correct Answer	838/896	847/896	833/896
% Correct Answer	93.53%	94.53%	92.97%

Figure 3.1.1. Card Used in the Message Condition in Study 6

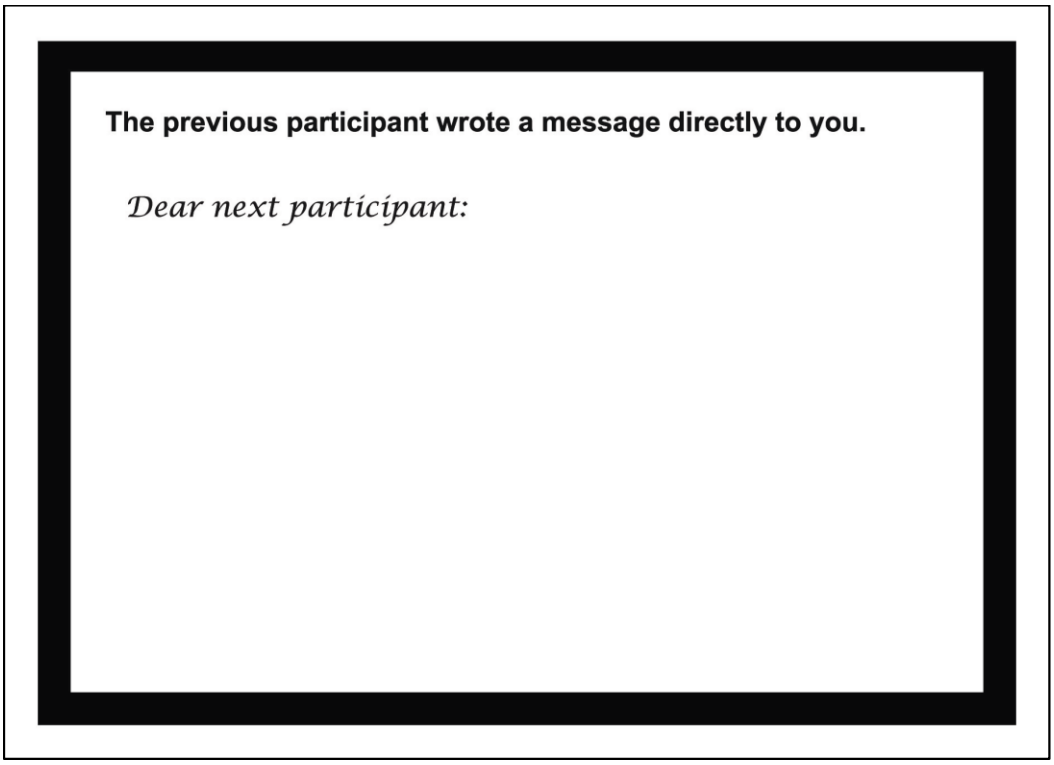


Figure 3.1.2. Card Used in the Payment Reporting Condition in Study 6

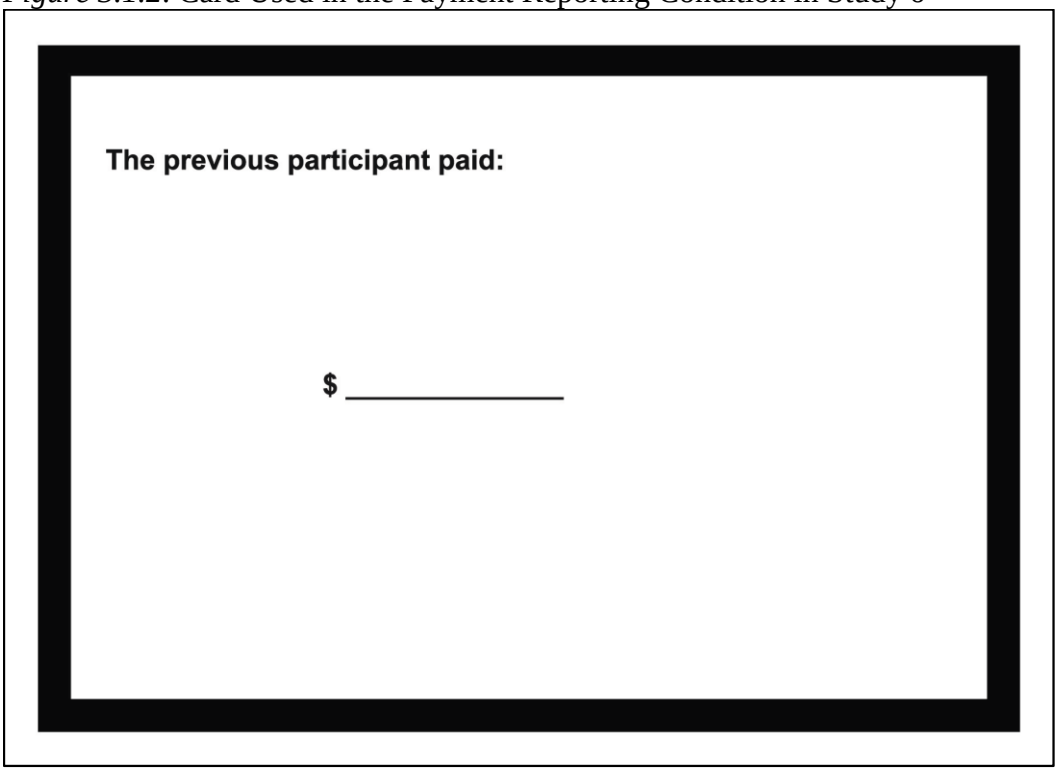
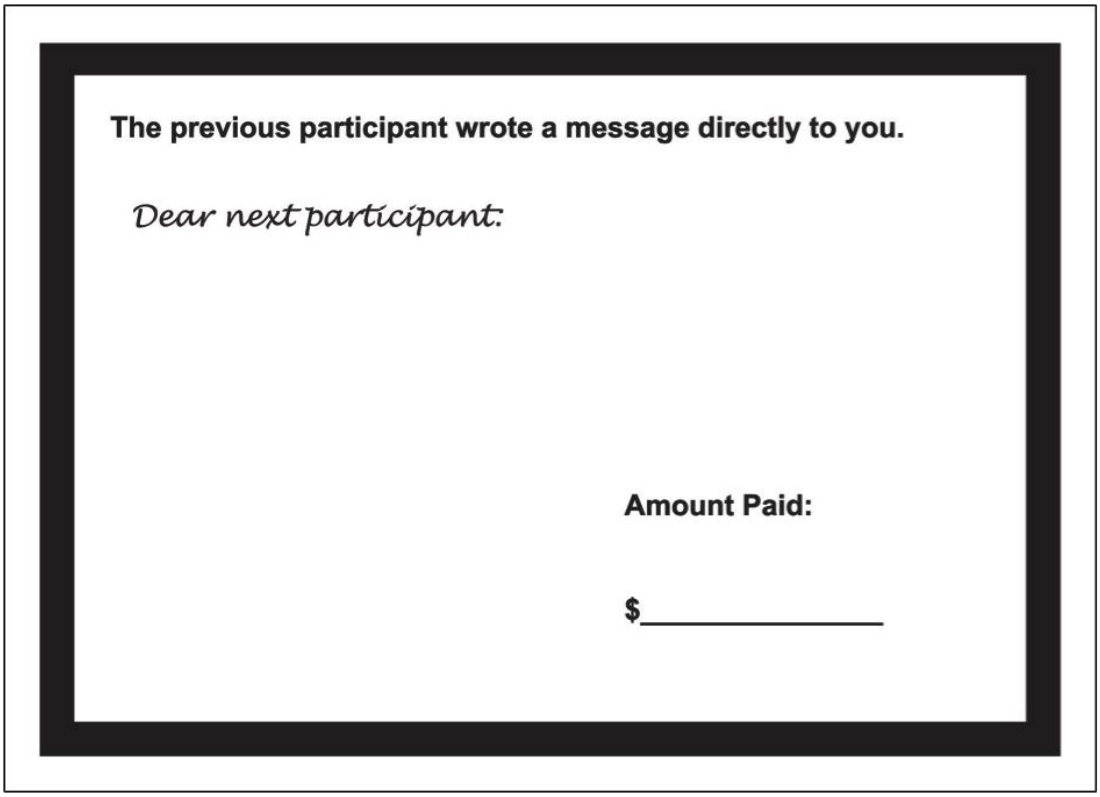


Figure 3.1.3. Card Used in the Message and Payment Reporting Condition in Study 6



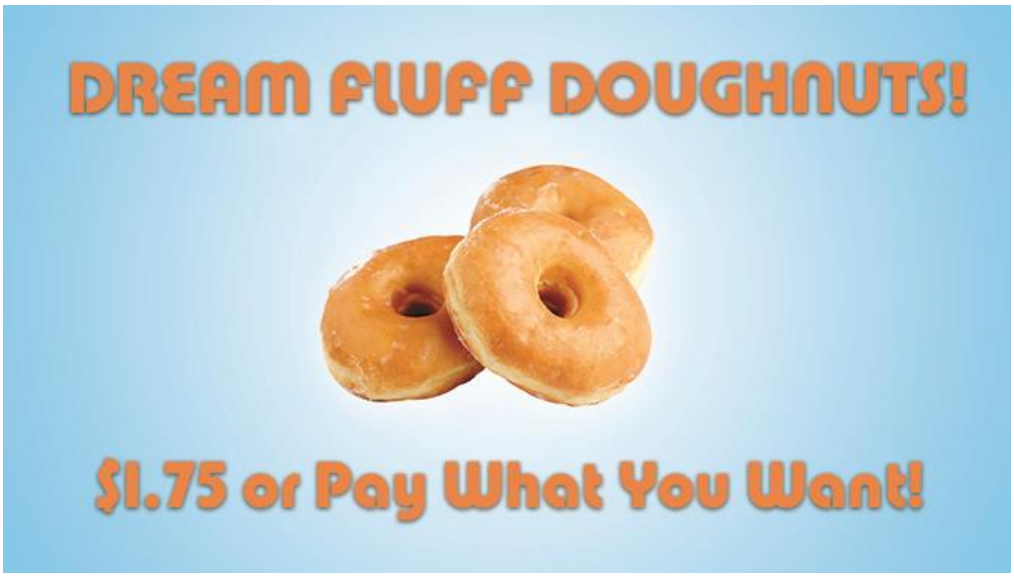
CHAPTER 4: ANCHORING IN PAYMENT: EVALUATING A JUDGMENTAL HEURISTIC IN FIELD EXPERIMENTAL SETTINGS

4.1. STUDY MATERIALS

4.1.1.1. Screenshot from Study 1. 49% default allocation condition.

4.1.1.2. Screenshot from Study 1. 89% default allocation condition.

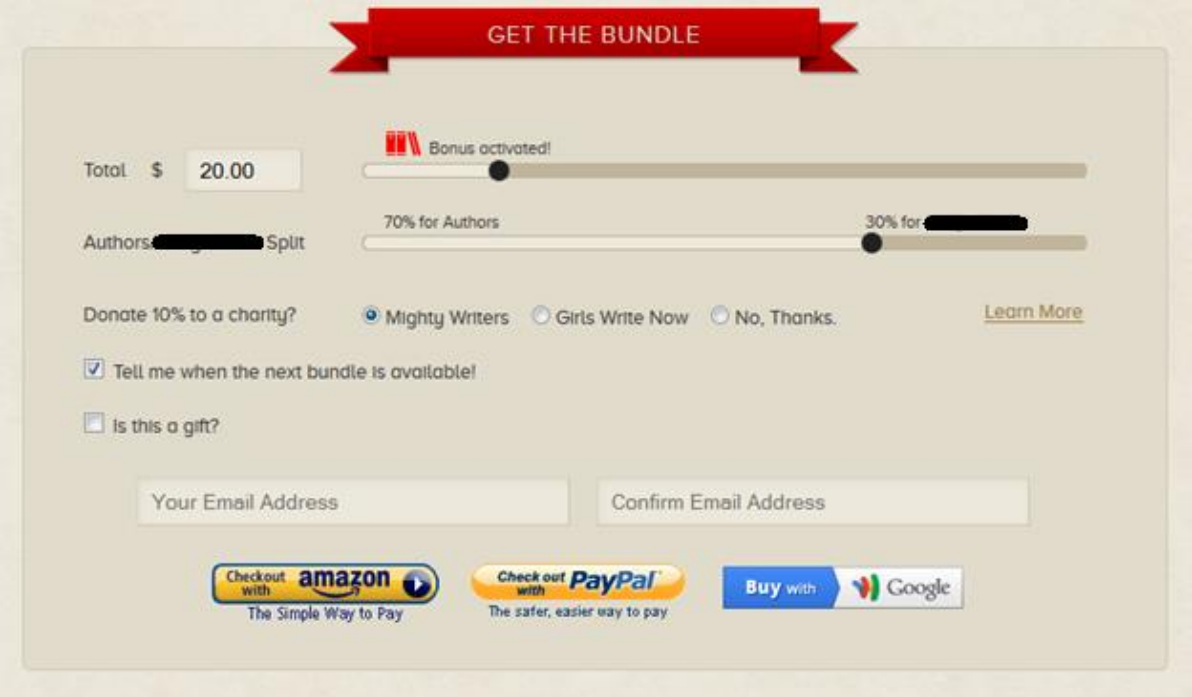
4.1.2.1. Sign Used in Study 2, $1.75 or Pay What You Want condition



4.1.2.2 Doughnut Stand at UC Berkeley in Study 2

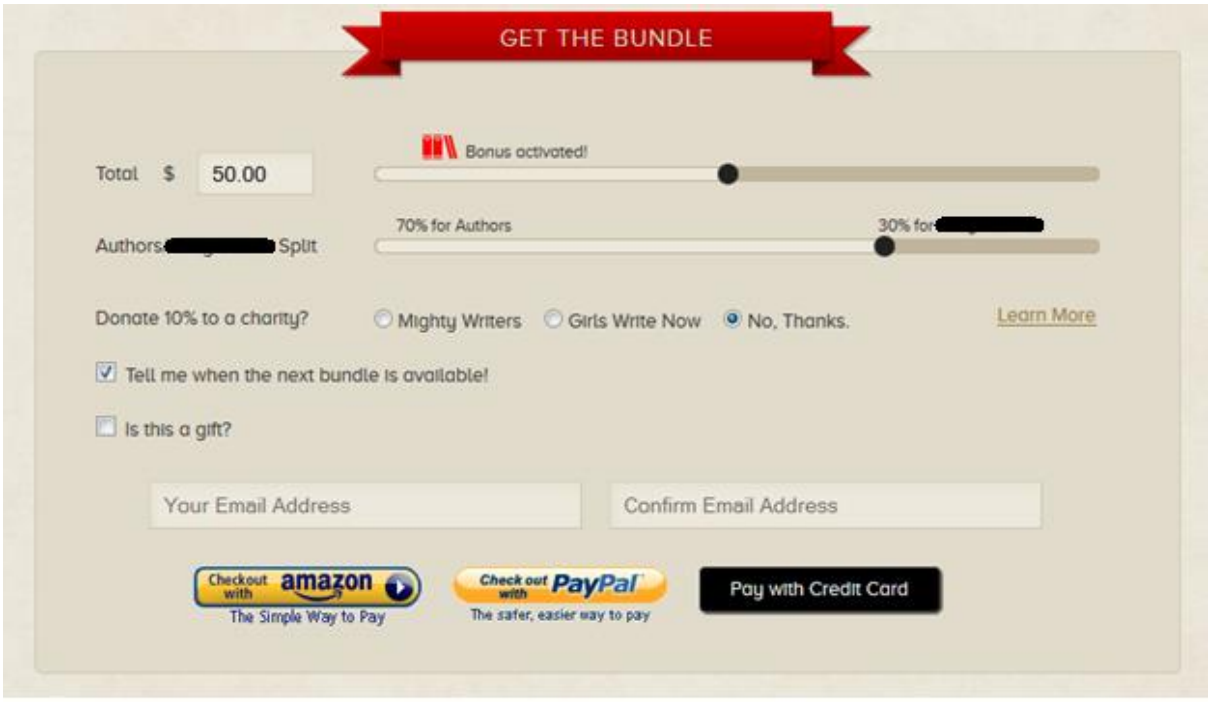
4.1.3.1. Screenshot from Study 3a. $3 condition

4.1.3.2. Screenshot from Study 3a. $20 condition

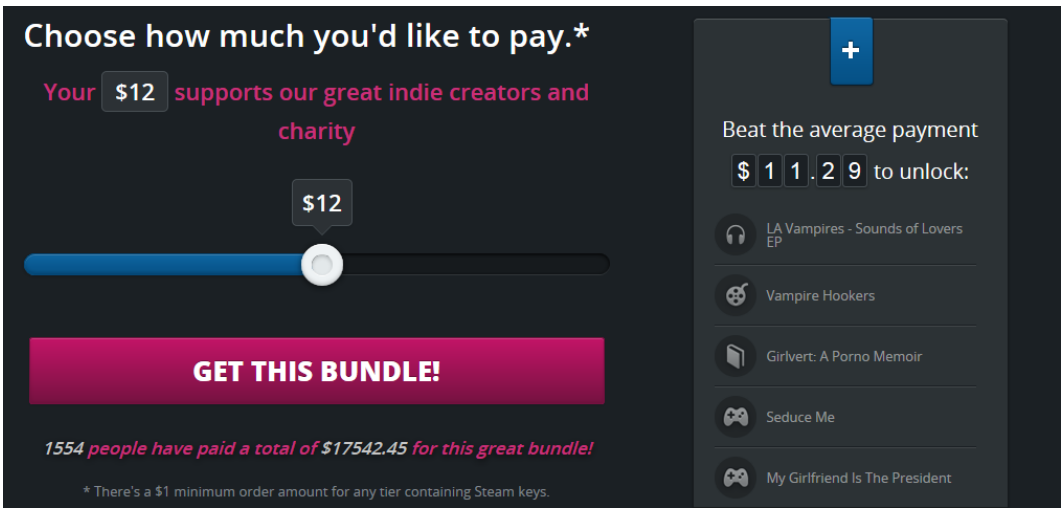


4.1.4.1. Screenshot from Study 3b. $8 condition

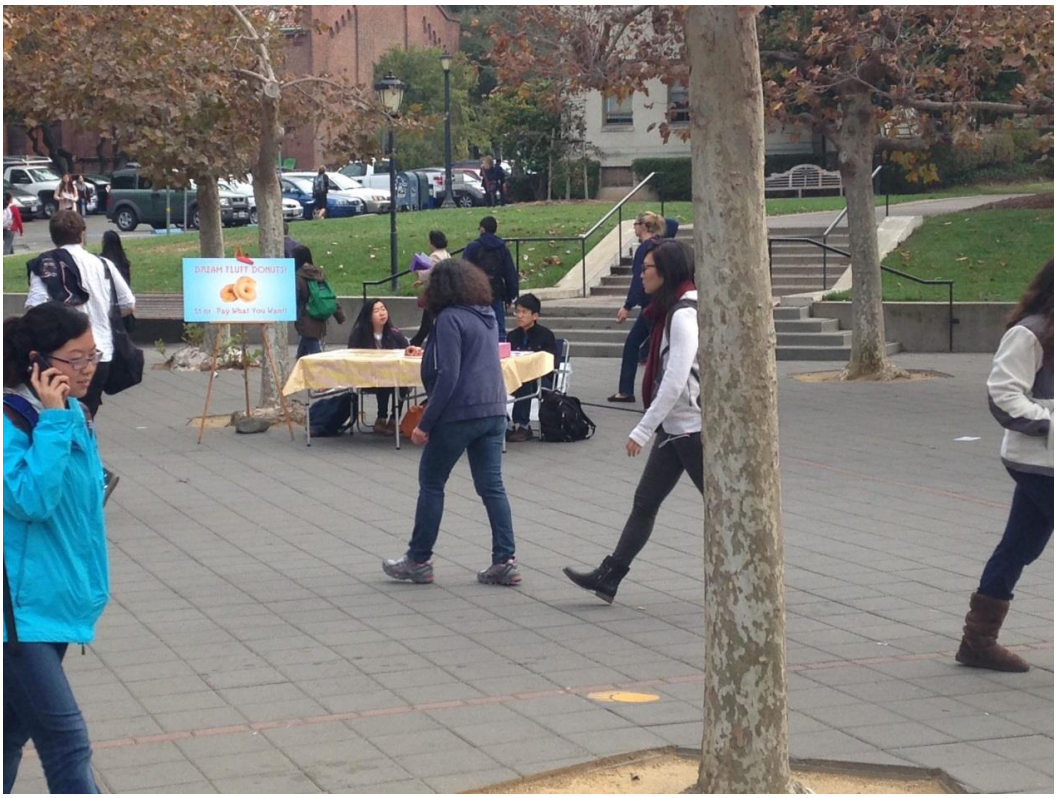
4.1.4.2. Screenshot from Study 3b. $50 condition



4.1.5. Screenshot from Study 2 (Vodo.net), $12 condition

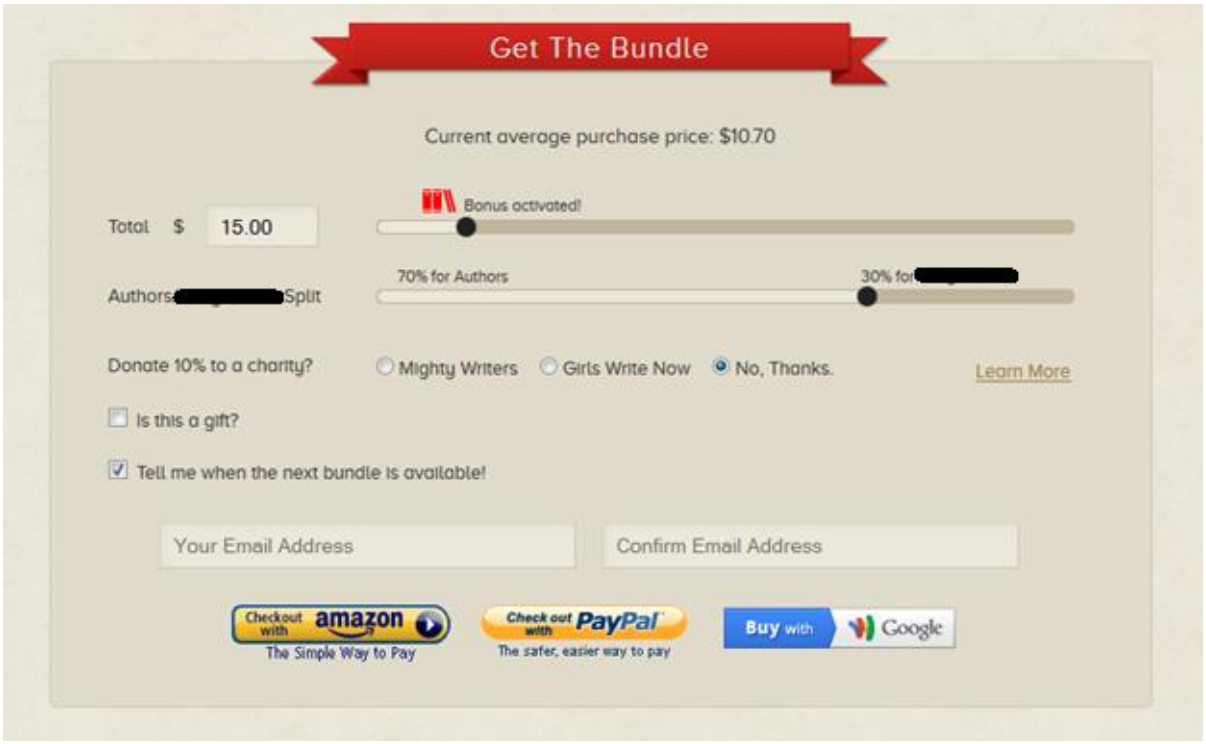


4.1.6. Doughnut Stand at UC Berkeley in Study 6a

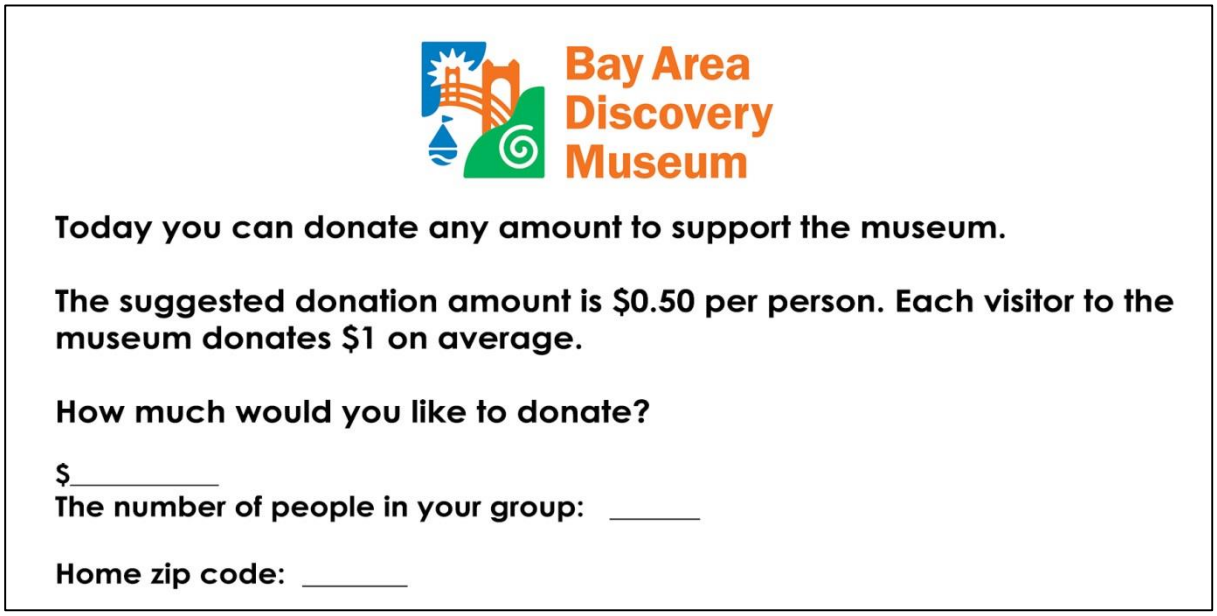


4.1.7. Sign used in Study 6a. $3 or pay-what-you-want condition.

4.1.8. Screenshot from Study 7. The average price shown condition.



4.1.9. Card used in Study 1. $0.50 suggested amount and $1 average donation condition.

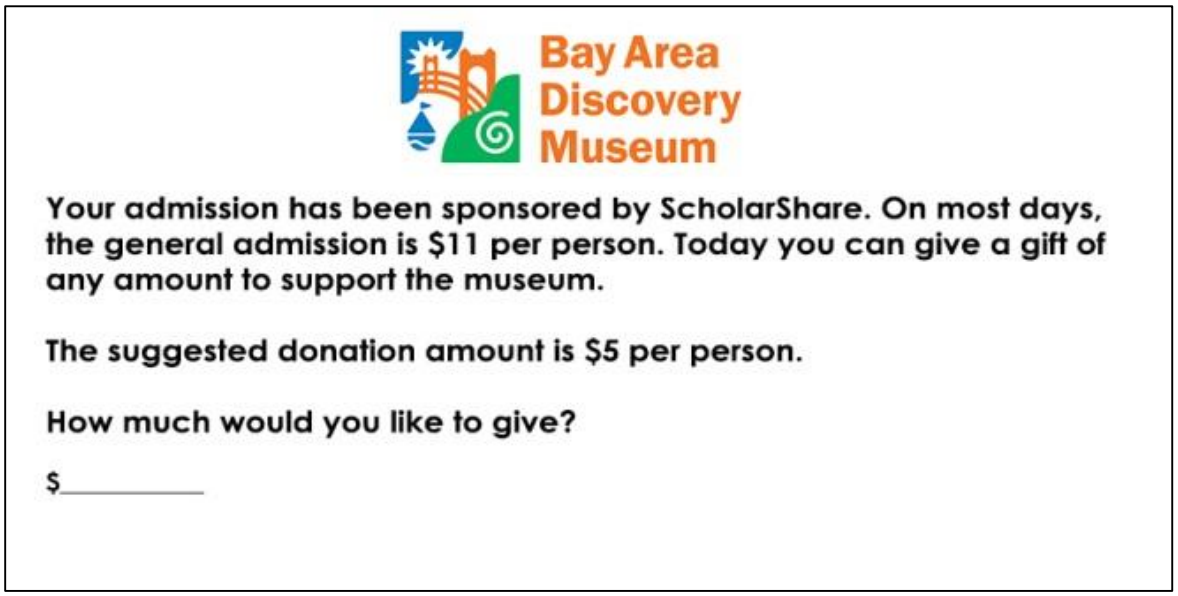


4.1.10. Screenshot from Study 9. $12 previous customer payment condition.

4.1.11. Screenshot from Study 10. Retail Price Salient and $28.88 condition

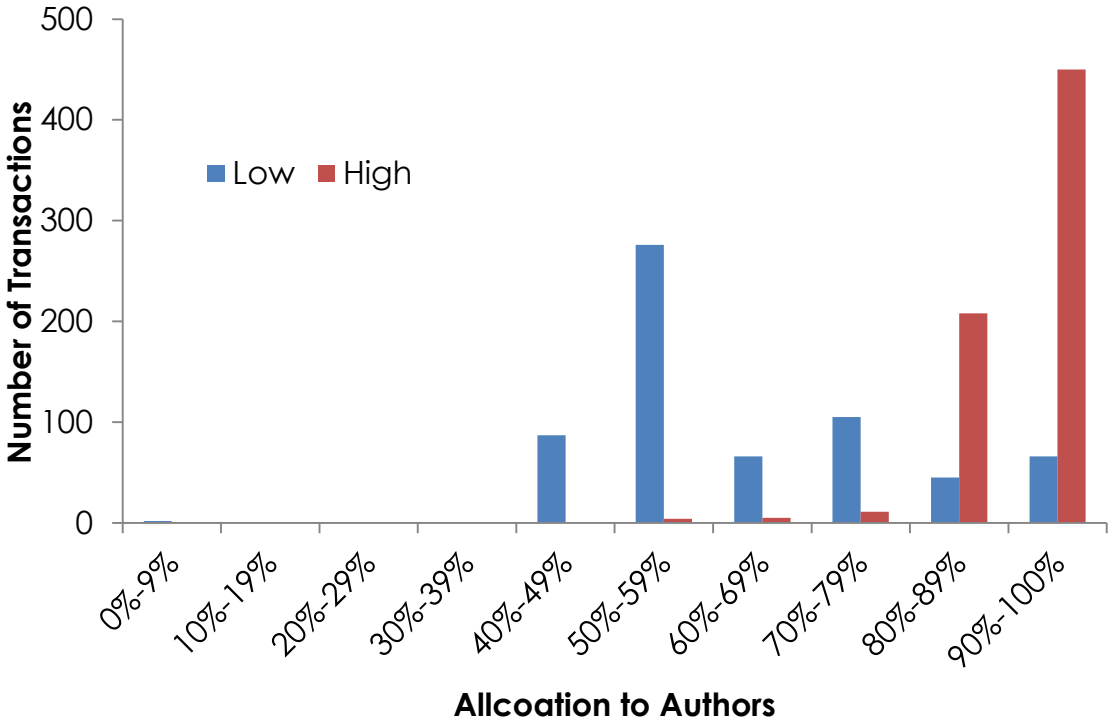
4.1.12. Screenshot from Study 10. Retail Price Not Salient and $14 condition

4.1.13. Card used in Study 11. $11 admission and $5 suggestion condition.



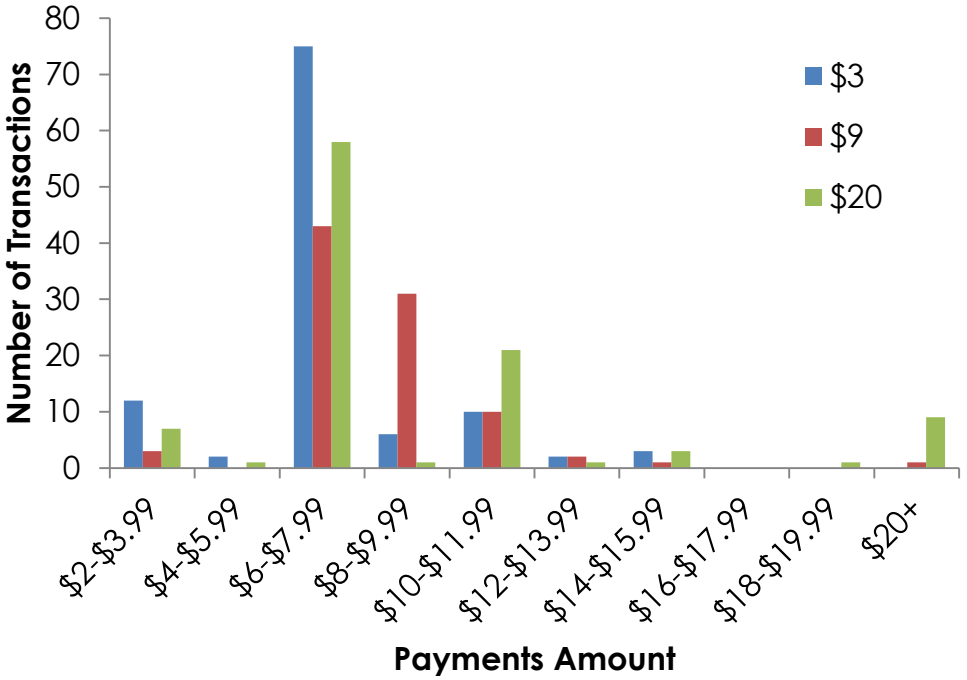
4.2. HISTOGRAMS OF PAYMENT

4.2.1. Payment Distribution in Study 1



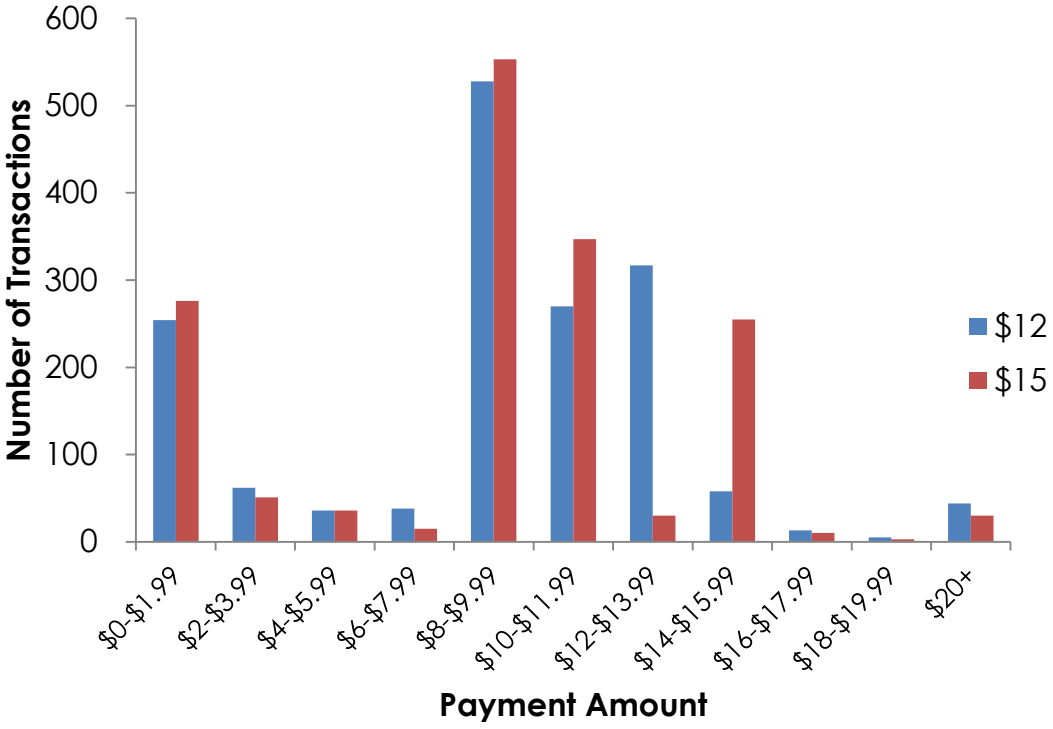
4.2.2. Payment Distribution in Study 2

4.2.3.1 Payment Distribution in Study 3a

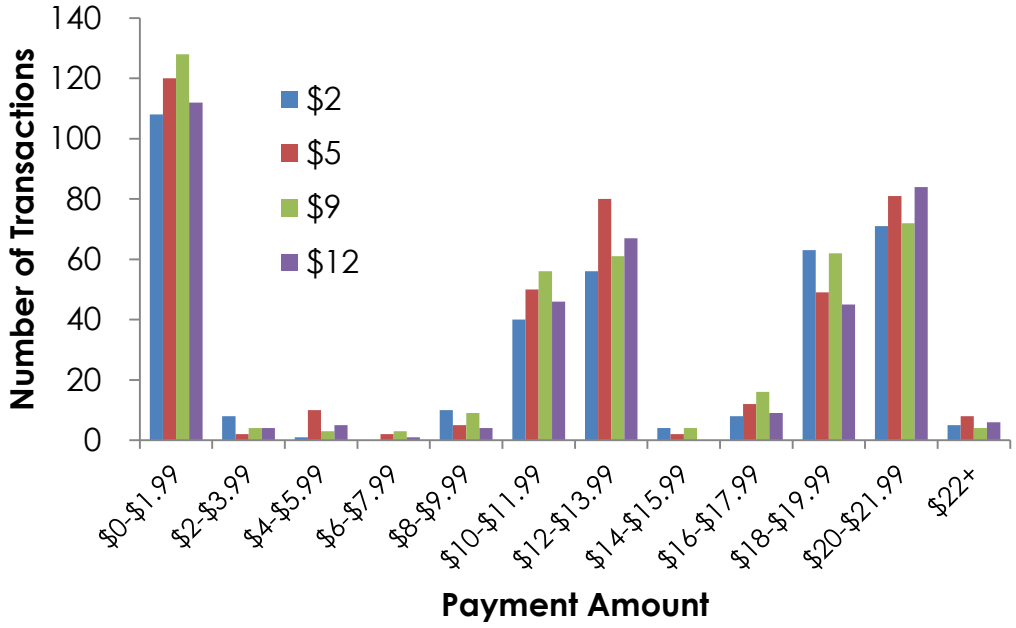


4.2.3.2. Payment Distribution in Study 3b

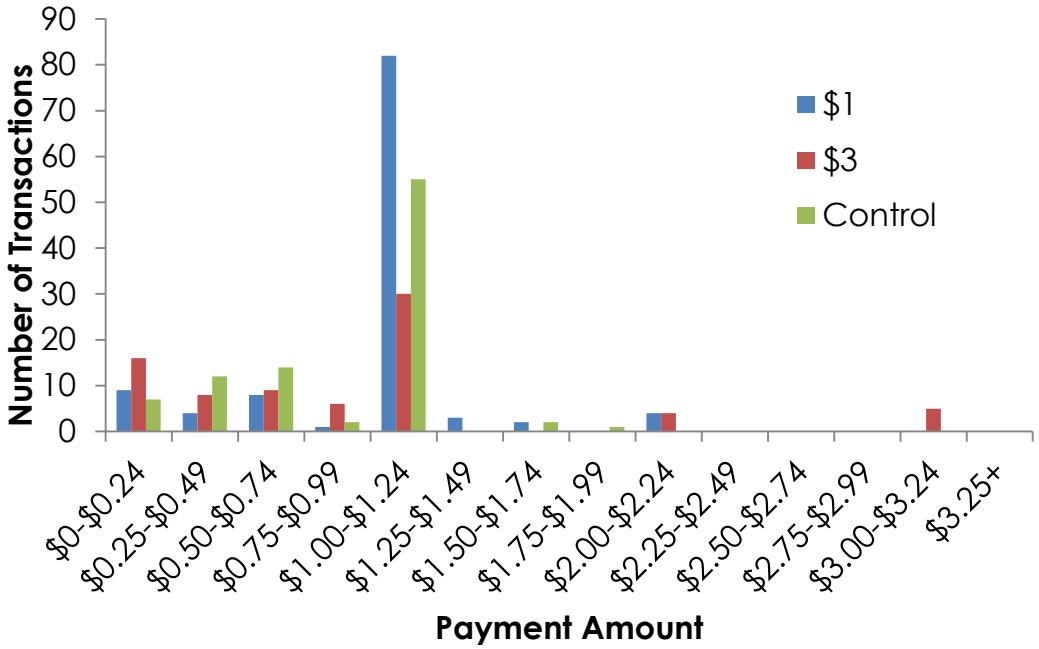
4.2.4. Payment Distribution in Study 4



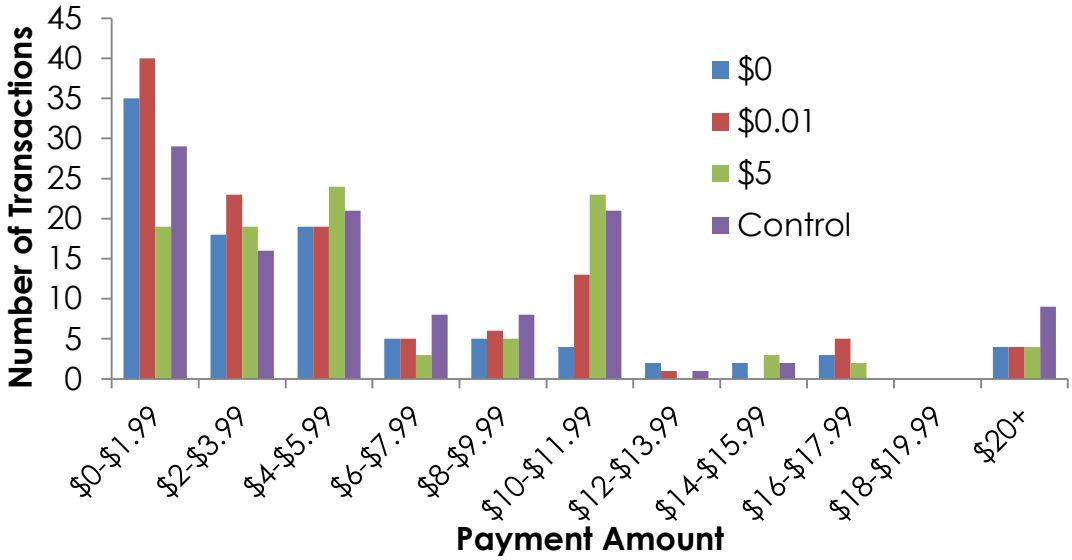
4.2.5. Payment Distribution in Study 5



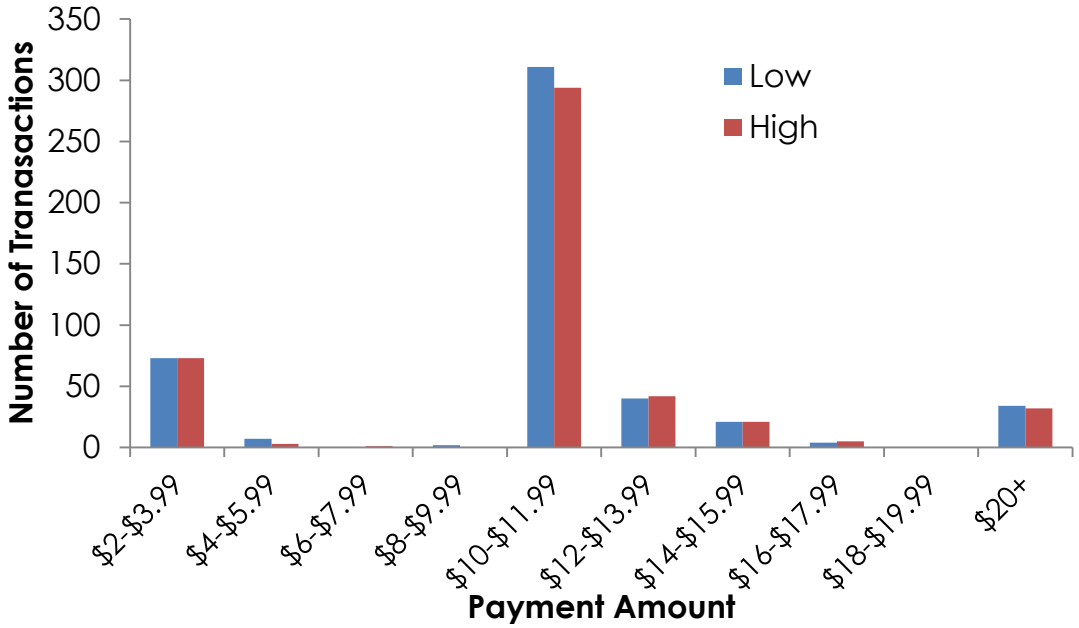
4.2.6.1. Payment Distribution in Study 6a



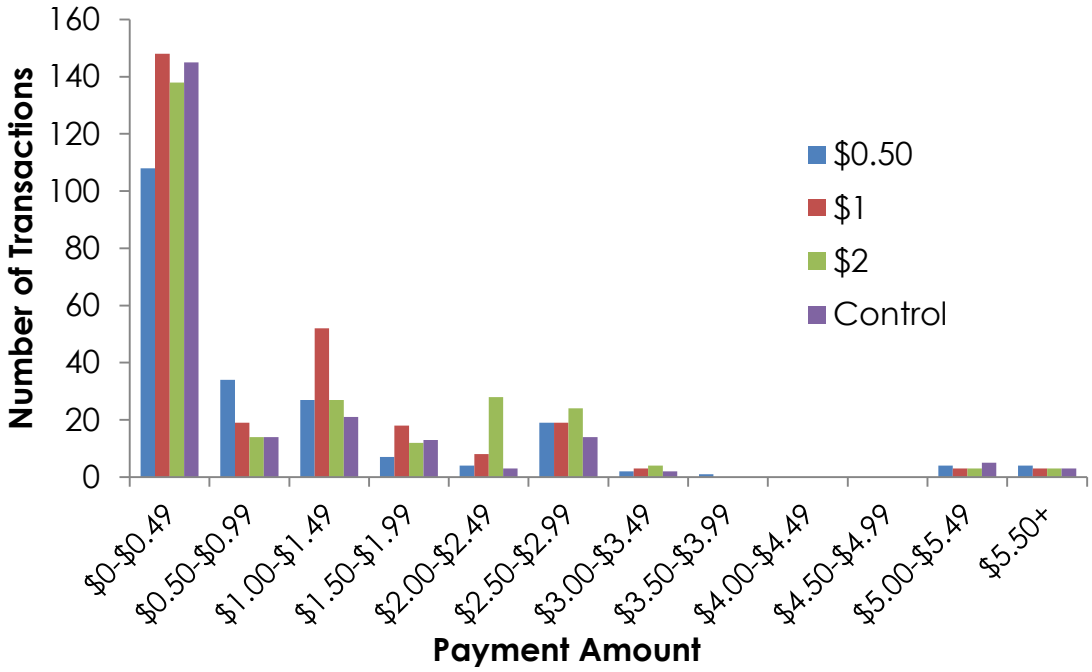
4.2.6.2. Payment Distribution in Study 6b



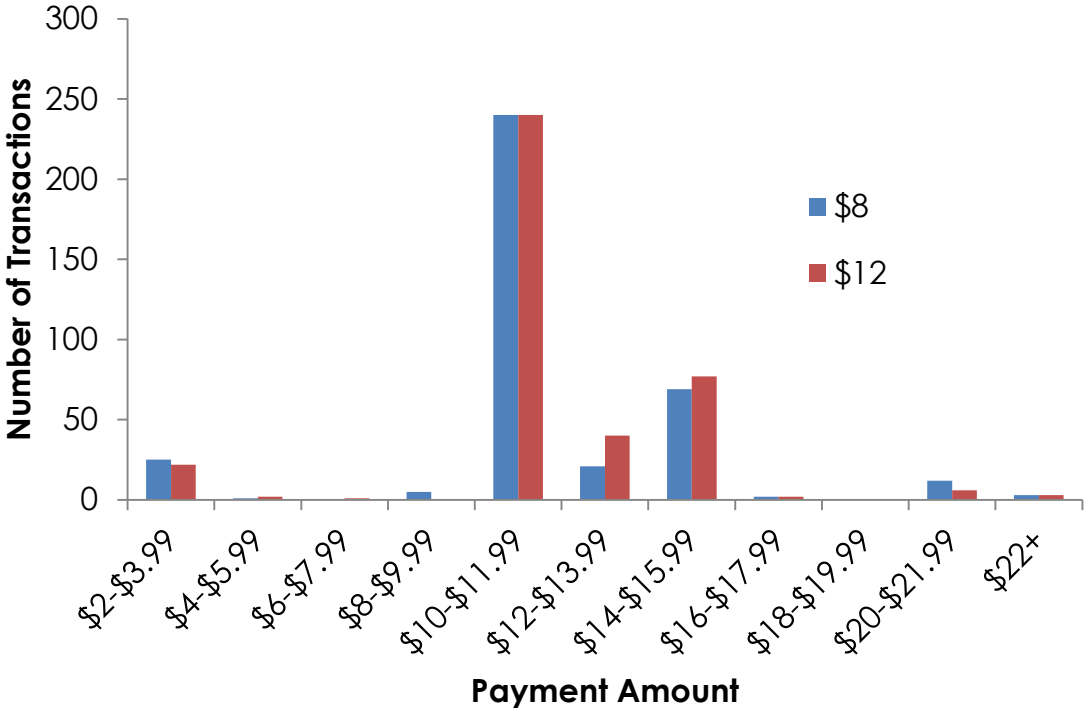
4.2.7. Payment Distribution in Study 7



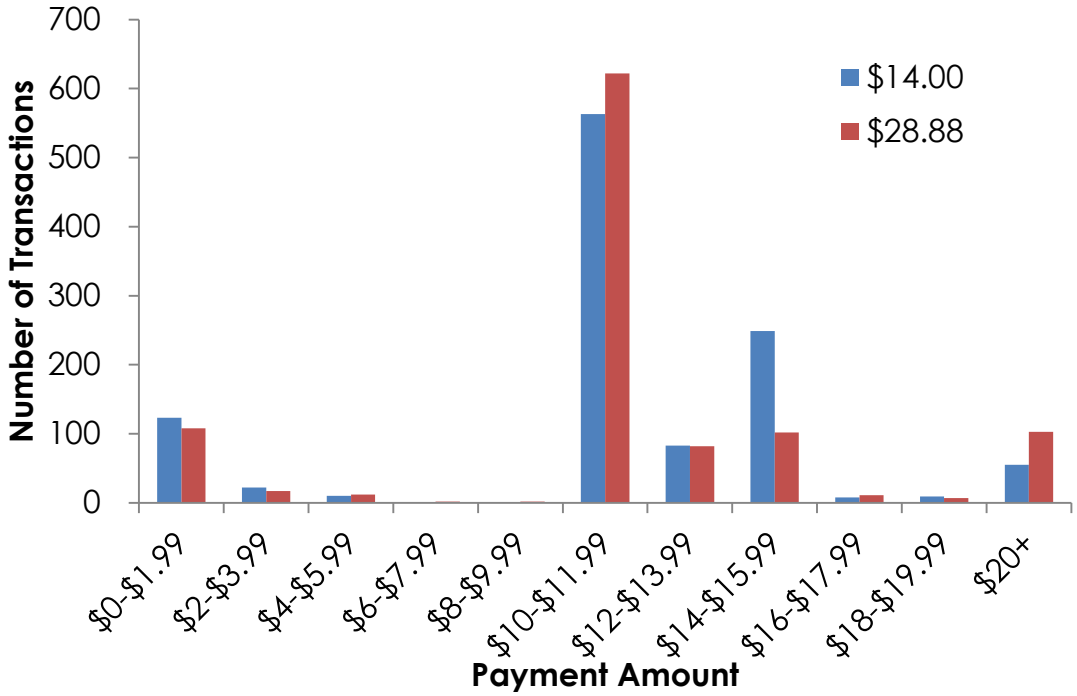
4.2.8. Payment Distribution in Study 8



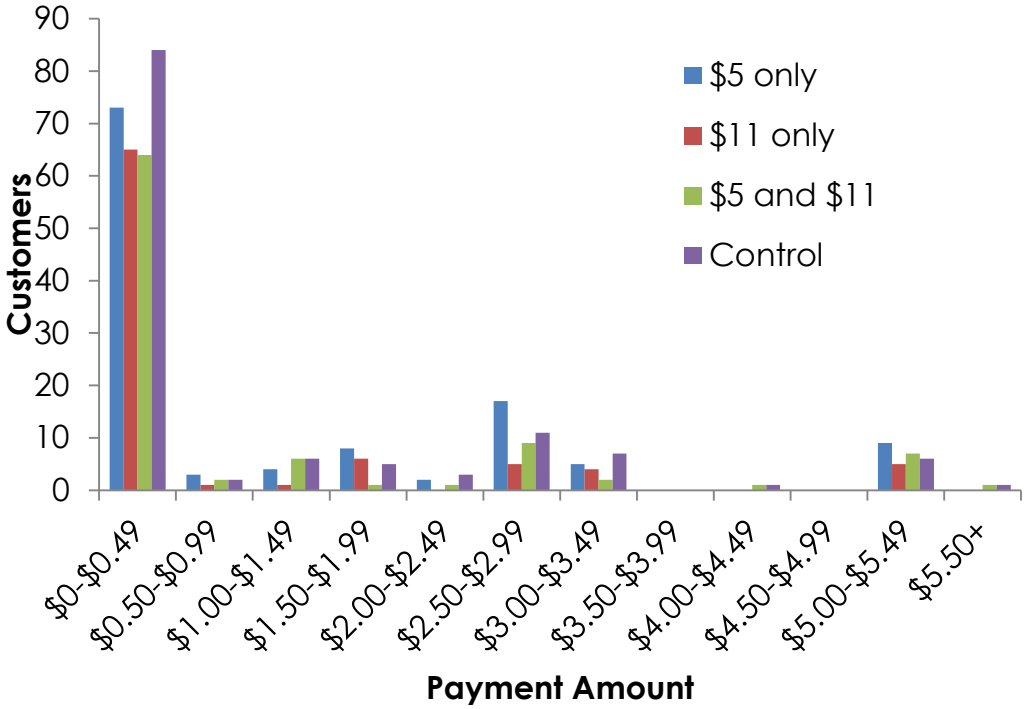
4.2.9. Payment Distribution in Study 9



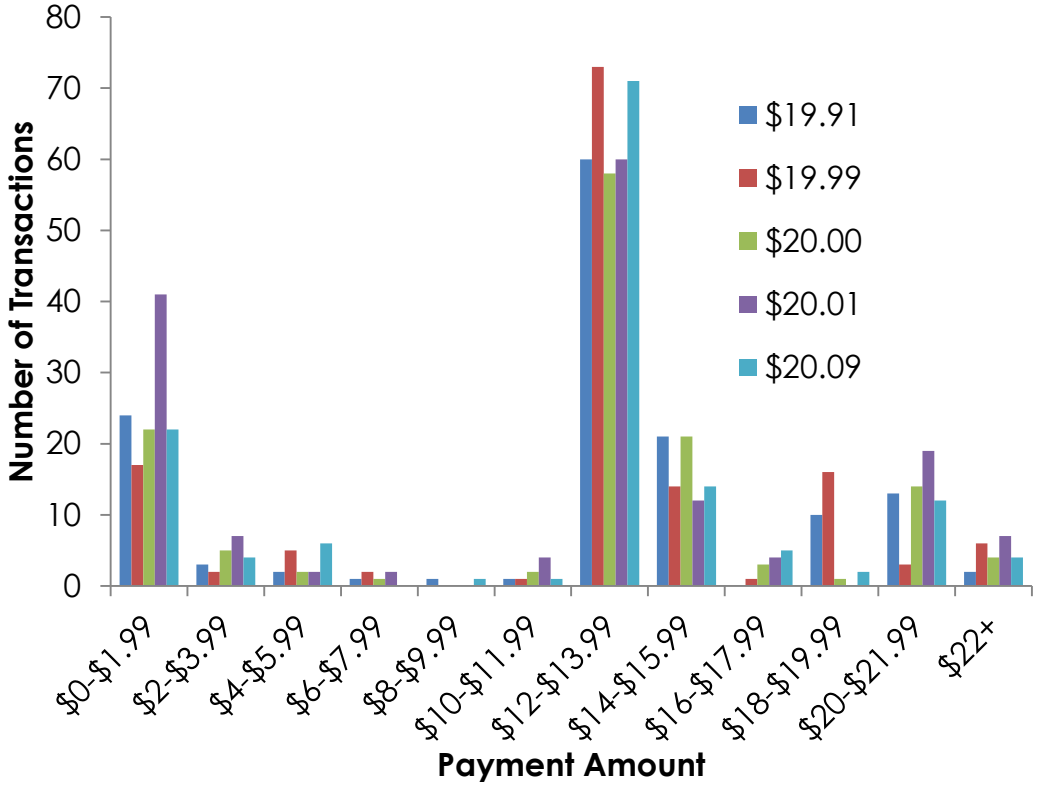
4.2.10. Payment Distribution in Study 10



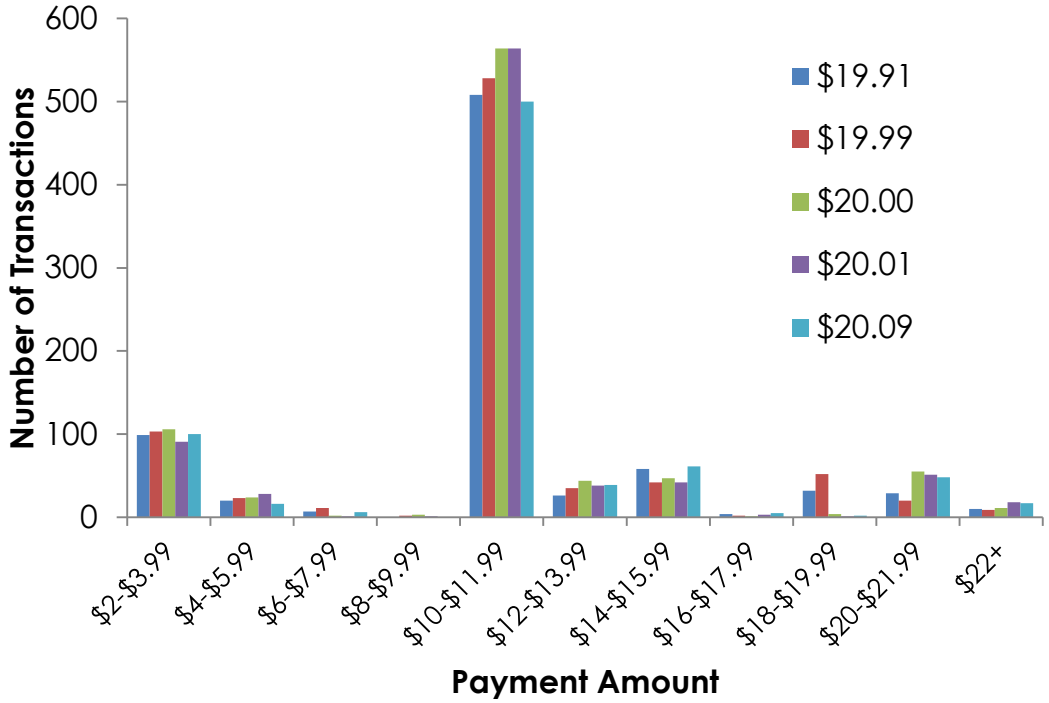
4.2.11. Payment Distribution in Study 11



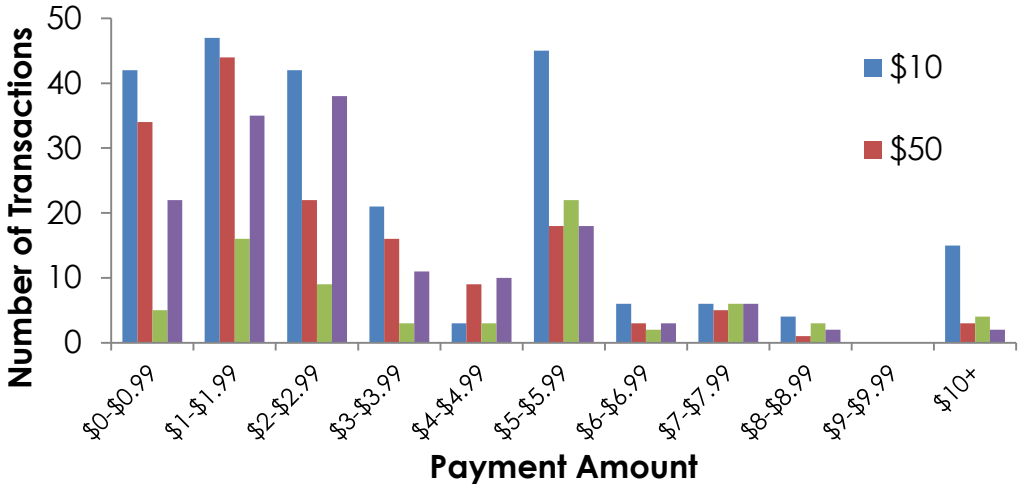
4.2.12.1. Payment Distribution in Study 12a



4.2.12.2. Payment Distribution in Study 12b

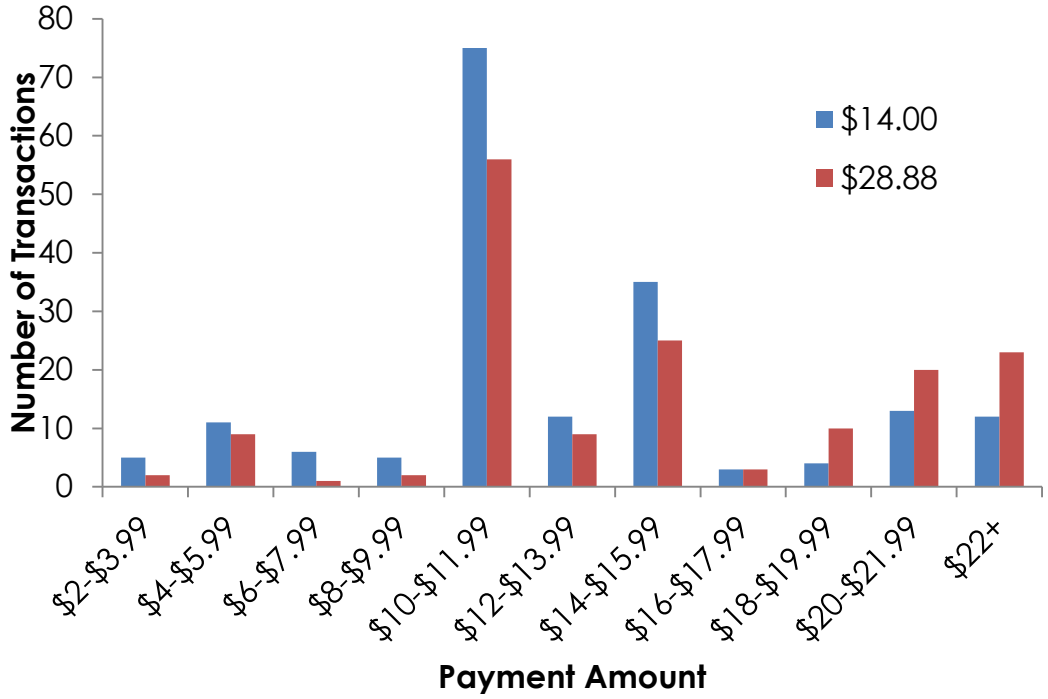


4.2.13. Payment Distribution in Study 13

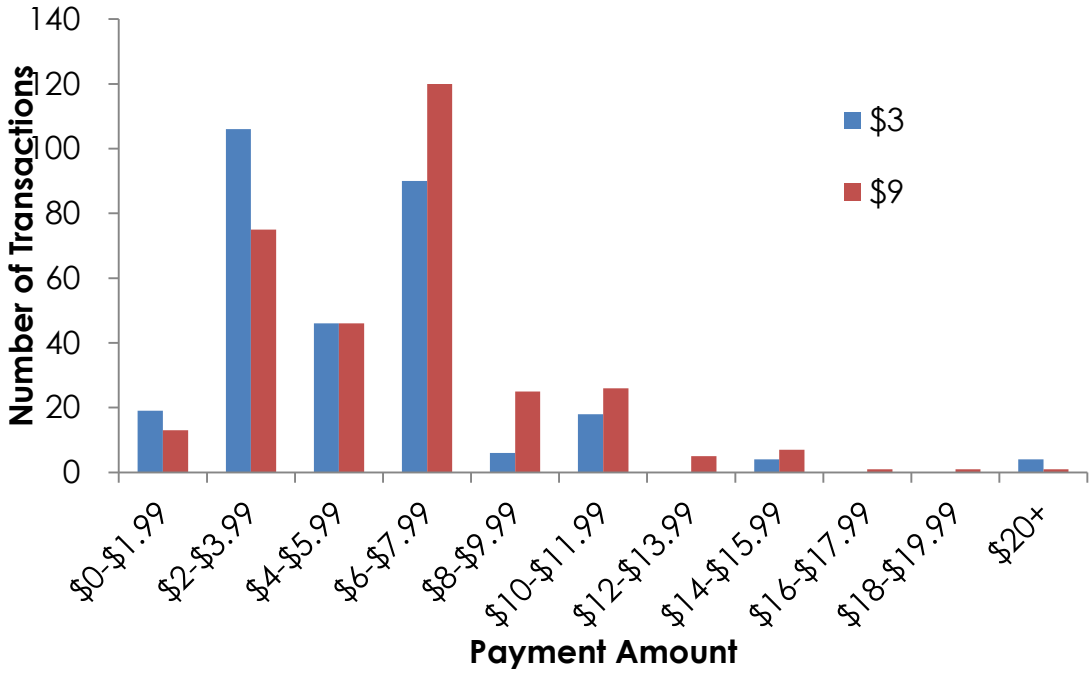


Note: As discussed in the main text, different sets of conditions were run in different months, hence not all conditions have equal sample sizes (e.g., many more overall were in the $10 vs. $100 condition).

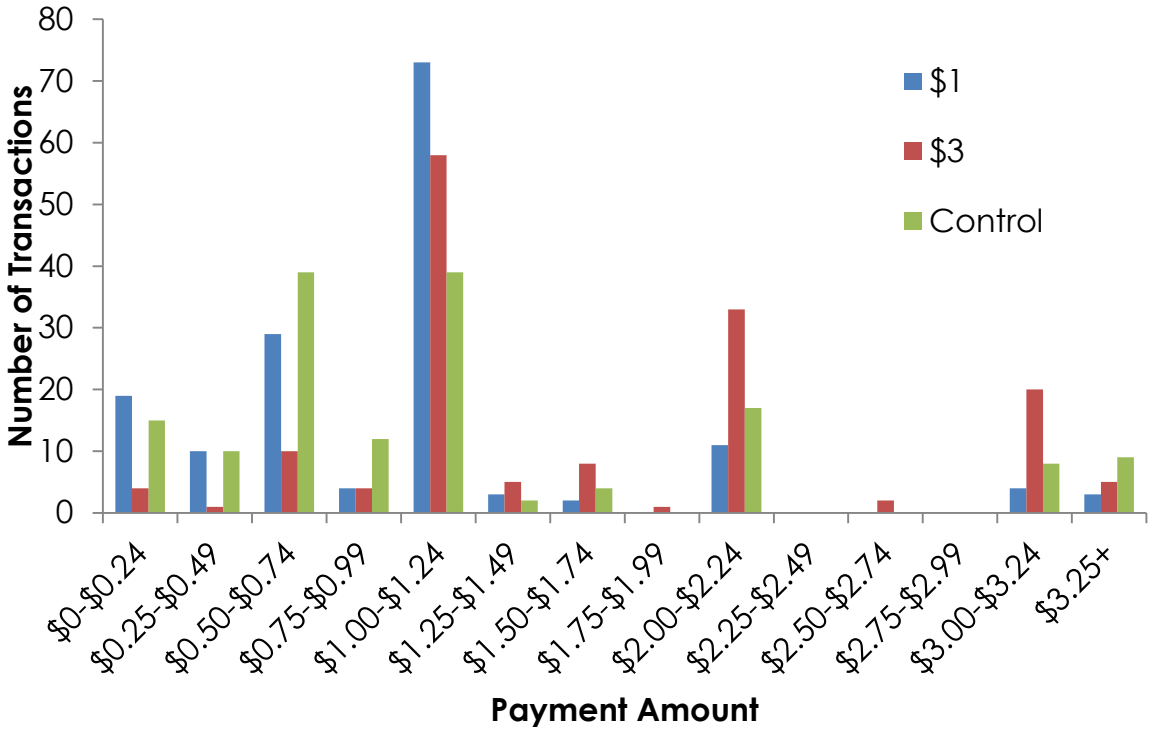
4.2.14.1. Payment Distribution in Study 14a



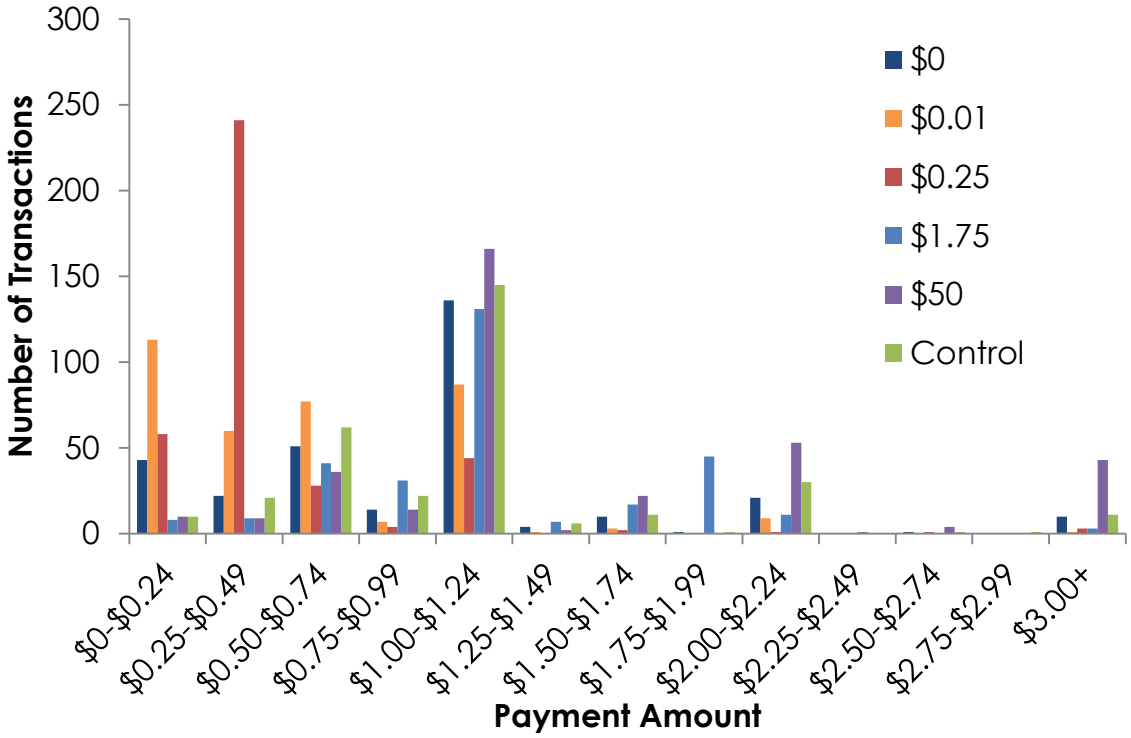
4.2.14.2. Payment Distribution in Study 14b



4.2.14.3. Payment Distribution in Study 14c



4.2.14.4. Payment Distribution in Study 14d



4.3. ADDITIONAL ANALYSES

4.3.1 Purchase Rate Information

	Product	Anchor	Customers	Visitors	Purchase Rate	Chi-squared test
Study 2	Doughnuts	$0.25	490	14,631	3.35%[a]	χ^2=646.66, p<.001
		$1.75	181	14,548	1.24%[b]	
		No anchor	338	28,073	1.20%[b]	
Study 3b	Media	$8	1,443	19,663	6.8%[a]	χ^2=16.47, p<.001
	bundle	$20	1,291	19,739	6.1%[b]	
		$50	1,244	19,737	5.9%[b]	
Study 4	Media	$12	1,602	18,634	7.9%[a]	χ^2=.01, p=.937
	bundle	$15	1,612	18,805	7.9%[a]	
Study 6a	Doughnuts	$1	113	9,128	1.24%[a]	χ^2=7.67, p=.022
		$3	78	8,794	0.89%[b,c]	
		No anchor	106	9,477	1.12%[a,c]	
Study 9	Media	$8	378	4,112	8.4%[a]	χ^2=.50, p=.778
	bundle	$12	393	4,130	8.7%[a]	
		No anchor	407	4,202	8.8%[a]	
Study 10	Media	$14	790	9,371	7.8%[a]	χ^2=3.73, p=.053
	bundle	$28.88	715	9,407	7.1%[a]	
Study 12a	Media	$19.99	140	5,795	2.4%[a]	χ^2=.93, p=.334
	bundle	$20.01	158	5,838	2.6%[a]	
Study 12b	Media	$19.99	827	12,998	6.0%[a]	χ^2=.00, p=.998
	bundle	$20.01	838	13,169	6.0%[a]	

For studies not included here, we were either unable to collect purchase rate information or the purchase rate was 100% (MTurk and museum studies; Studies 1 and 3a are excluded due to a mechanical error in data collection).

Purchase rates with different subscripts within studies differ significantly ($p < .05$).

4.3.2. Alternate Specifications for Mean Comparisons

Study	Product	Per-Cell N	Anchor	Raw Mean (SD)	Winsorized Mean (SD)	Ln Mean (SD)	Per-Cell N	Excl. Anchor Payments (SD)*	Per-Cell N	Excl. Bonus Payments (SD)
1	Allocations	648	50%	61.64%[a] (15.36%)	61.66%[a] (14.63%)	4.10[a] (0.29)	256	72.55%[a] (12.73%)		
		680	90%	89.20%[b] (6.60%)	89.03%[b] (5.17%)	4.50[b] (0.14)	98	85.86%[b] (14.91%)		
2	Dough-nuts	490	$0.25	$0.44[a] ($0.38)	$0.43[a] ($0.33)[a]	0.34[a] (0.22)	277	$0.58[a] ($0.46)		
		181	$1.75	$1.04[b] ($0.71)	$1.03[b] ($0.70)	0.65[b] (0.37)	141	$1.00[b] ($0.70)		
		338	No anchor	$0.66[c] ($0.54)	$0.65 [c] ($0.48)	0.46[c] (0.30)	290	$0.73[c] ($0.56)		
3a	Media bundle	110	$3	$6.59[a] ($2.38)	$6.62[a] ($2.32)	1.98[a] (0.32)	101	$6.85[a] ($2.25)	49	$7.32[a] ($3.45)
		91	$9	$7.79[b] ($2.44)	$7.77[b] ($2.12)	2.14[b] (0.29)	63	$7.08[a] ($2.25)	59	$8.76[b] ($2.55)
		102	$20	$8.29[b] ($4.57)	$7.90[b] ($3.33)	2.12[b] (0.46)	92	$7.21[a] ($2.86)	54	$10.33[b] ($5.56)
3b	Media bundle	1,443	$8	$8.88[a] ($4.03)	$8.70 ($3.24)[a]	2.20[a] (0.46)	1,349	$8.79[a] ($3.84)	622	$7.40[a] ($5.82)
		1,291	$20	$9.89[b] ($4.96)	$9.39[b] ($3.35)	2.29[b] (0.46)	1,207	$9.17[b] ($4.18)	491	$9.71[b] ($8.04)
		1,244	$50	$9.88[b] ($5.63)	$9.23[b] ($3.36)	2.27[b] (0.48)	1,208	$9.38[b] ($4.40)	470	$9.69[b] ($9.17)
4	Media bundle	1,602	$12	$8.80[a] ($4.64)	$8.58[a] ($4.10)	2.11[a] (0.69)	1,231	$7.70[a] ($4.74)	1,089	$8.70[a] ($5.62)
		1,612	$15	$8.88[a] ($4.88)	$8.71[a] ($4.45)	2.10[a] (0.70)	1,331	$7.65[a] ($4.47)	1,076	$8.81[a] ($5.97)
5	VOD O Bundle	376	$2	$11.48[a] ($7.87)	$11.35[a] ($7.65)	2.16[a] (0.99)	315	$11.62[a] ($8.48)		
		425	$5	$11.24[a] ($7.65)	$11.12[a] ($7.46)	2.16[a] (0.97)	349	$11.24[a] ($8.38)		
		419	$9	$11.05[a] ($7.68)	$10.99[a] ($7.57)	2.13[a] (0.99)	359	$10.96[a] ($8.27)		
		383	$12	$11.29[a] ($7.72)	$11.17[a] ($7.57)	2.15[a] (0.98)	315	$11.26[a] ($8.46)		
6a	Dough-nuts	113	$1	$0.91[a] ($0.38)	$0.91[a] ($0.38)	0.62[a] (0.22)	31	$0.66[a] ($0.67)		
		78	$3	$0.84 [a,b] ($0.75)	$0.84[a,b] ($0.75)	0.54[a,b] (0.37)	44	$0.49[a] ($0.57)		

		106	No anchor	$0.72[b] ($0.45)	$0.72[b] ($0.45)	0.51[b] (0.28)	51	$0.43[a] ($0.50)		
6b	Museum tickets	97	$0	$2.55[a,c] ($2.73)	$2.51[a,c] ($2.63)	0.96[a] (0.81)	50	$3.44[a] ($2.78)		
		117	$0.01	$2.47[a,c] ($2.36)	$2.45[a,c] ($2.31)	1.03[a,c] (0.66)	90	$2.32[b] ($2.31)		
		102	$5	$3.47[b] ($2.96)	$3.30[b] ($2.37)	1.29[b] (1.68)	61	$3.18[a,b] ($3.40)		
		115	No anchor	$3.24[b,c] ($2.77)	$3.16[b,c] ($2.60)	1.19[b,c] (0.76)	73	$3.66[a] ($2.83)		
8	Museum tickets	211	$0.50	$0.93[a] ($2.20)	$0.86[a] ($1.57)	0.43[a] (0.57)	159	$0.97[a] ($2.52)		
		273	$1	$0.77[a] ($1.29)	$0.77[a] ($1.29)	0.42[a] (0.51)	208	$0.69[a] ($1.45)		
		253	$2	$0.91[a] ($1.36)	$0.90[a] ($1.34)	0.47[a] (0.56)	206	$0.77[a] ($1.45)		
9	Media bundle	378	$8	$11.10[a] ($3.97)	$10.81[a] ($2.94)	2.44[a] (0.36)	357	$11.08[a] ($4.07)	165	$12.52[a] ($5.71)
		393	$12	$11.17[a] ($3.50)	$10.97[a] ($2.84)	2.45[a] (0.32)	357	$11.09[a] ($3.66)	185	$12.49[a] ($4.78)
		407	No anchor	$11.45[a] ($4.61)	$11.09[a] ($3.19)	2.46[a] (0.37)	391	$11.44[a] ($4.70)	193	$13.06[a] ($6.33)
10	Media bundle	790	$14	$10.68[a] ($5.57)	$10.46[a] ($4.48)	2.31[a] (.65)	665	$10.06[a] ($5.87)	432	$11.24[a] ($7.49)
		715	$28.88	$11.26[a] ($6.22)	$10.76[a] ($4.62)	2.36[a] (.62)	697	$10.89[b] ($5.77)	335	$12.69[b] ($8.88)
11										
	Museum Tickets	121	$5	$1.06[a] ($1.55)	$1.06[a] ($1.55)	0.49 (0.65)	112	$0.75[a] ($1.11)		
		88	$11	$0.77[a] ($1.48)	$0.77[a] ($1.48)	0.34 (0.60)	82	$0.46[a] ($0.97)		
		94	Both	$0.96[a] ($1.85)	$0.90[a] ($1.58)	0.41 (0.66)	86	$0.52[a] ($1.00)		
		126	Control	$0.90a ($1.52)	$0.89[a] ($1.47)	0.42 (0.63)	120	$0.62[a] (1.10)		
12a										
	Media bundle	140	$19.99	$12.31[a] ($6.39)	$10.95[a] ($5.58)	2.40[a] (0.73)	122	$11.18[a] ($6.07)	75	$12.59[a] ($8.75)
		158	$20.01	$10.52[b] ($7.65)	$10.17[b] ($6.72)	2.21[b] (0.93)	139	$9.23[b] ($7.24)	105	$9.78[b] ($9.31)
12b	Media bundle	827	$19.99	$10.39[a] ($4.48)	$10.32[a] ($4.27)	2.35[a] (0.44)	756	$9.48[a] ($3.52)	352	$10.91[a] ($6.83)

		838	$20.01	$10.60[a] ($5.67)	$10.32[a] ($4.01)	2.36[a] (0.43)	789	$10.02[b] ($5.32)	351	$11.44[a] ($8.70)
13	Museum tickets	231	$10	$2.72[a,b] ($3.25)	$2.58[a,b] ($2.47)	1.01[a,b] (0.73)	218	$2.32[a,b] ($2.87)		
		155	$50	$2.25[a] ($2.27)	$2.24[a] ($2.15)	0.89[a] (0.68)	152	$2.04[a] ($2.01)		
		73	$100	$2.90[b] ($2.94)	$2.93[b] ($2.51)	1.12[b] (0.67)	70	$2.77[b] ($2.72)		
		147	No anchor	$2.01[a,b] ($2.54)	$2.05[a,b] ($2.10)	0.87[a,b] (0.63)	146	$1.97[a,b] ($2.48)		
14a (lab)	Media bundle	205	$14	$13.01[a] ($9.82)	$12.33[a] ($5.70)	2.48[a] (0.58)	191	$12.94[a] ($10.17)	137	$14.51 ($11.74)[a]
		181	$28.88	15.78[b] ($10.34)	$14.91[b] ($6.87)	2.67[b] (0.57)	176	$15.49[b] ($10.29)	126	$18.30[b] ($11.53)
14b (lab)	Media bundle	274	$3	$5.10[a] ($3.43)	$4.84[a] ($2.44)	1.69[a] (0.47)	245	$5.32[a] ($3.54)	197	$4.75[a] ($3.99)
		307	$9	$5.99[b] ($3.12)	$5.76[b] ($2.54)	1.85[b] (0.45)	271	$6.06[b] ($3.14)	223	$5.99[b] ($3.67)
14c (lab)	Dough-nuts	61	$1	$0.91[a] ($1.33)	$0.79[a] ($0.67)	0.54[a] (0.40)	81	$0.71[a] ($1.71)		
		56	$3	$1.84[b] ($3.25)	$1.45[b] ($0.84)	0.87[b] (0.47)	73	$2.19[a] ($4.53)		
		52	No anchor	$2.26[a,b] ($6.93)	$1.17[b] ($0.90)	0.78[b] (0.64)	108	$2.71[a] ($8.43)		
14d (lab)	Dough-nuts	313	$0	$0.92[a] ($0.80)	$0.85[a] ($0.55)	0.59[a] (0.35)	257	$1.10[a] ($0.77)		
		359	$0.01	$0.56[b] ($1.37)	$0.50[b] ($0.47)	0.36[b] (0.33)	221	$0.85[b] ($1.68)		
		382	$0.25	$0.48[b,c] ($1.66)	$0.37[c] ($0.34)	0.30[c] (0.29)	129	$0.98[a,b,c] ($2.79)		
		304	$1.75	$1.24[a,d] ($2.86)	$1.06[d] ($0.47)	0.71[d] (0.31)	244	$1.02[a,b,d] ($0.51)		
		359	$50	$2.03[e] ($4.12)	$1.21[e] ($0.56)	0.87[e] (0.53)	344	$1.96[e] ($3.29)		
		321	No anchor	$1.04[d] ($0.73)	$0.97[f] ($0.58)	0.66[d] (0.30)	301	$1.10[a,c,d] ($0.71)		

Means with different subscripts within studies differ significantly ($p < .05$)
We controlled for the month variable in Study 13.
We winsorize at the 5th and 95th percentile of the grand distribution.
*For the means excluding anchor payments, we exclude any payments that match any condition's anchors. For example, in Study 4 ($12 vs. $15 anchors), any payments of $12 or $15 are excluded, regardless of condition.

4.3.3 Statistical Tests for Mean Comparisons

Study	Product	Anchor Comparisons	Raw Means	Winsorized Means	Ln Means	Excl. Anchor Payments*	Excl. Bonus Payments
1	Allocations	90% v. 50%	$t(1326)=42.82$ $p<.001$	$t(1326)=45.88$ $p<.001$	$t(1326)=31.95$ $p<.001$	$t(352)=8.38$, $p<.001$	
2	Doughnuts	$1.75 v. $0.25	$t(669)=14.02$, $p<.001$	$t(669)=15.06$, $p<.001$	$t(669)=13.24$, $p<.001$	$t(416)=7.25$, $p<.001$	
		No anchor v. $0.25	$t(826)=6.92$, $p<.001$	$t(826)=7.69$, $p<.001$	$t(826)=6.98$, $p<.001$	$t(565)=3.39$, $p=.001$	
		$1.75 v. No anchor	$t(517)=6.77$, $p<.001$	$t(517)=7.33$, $p<.001$	$t(517)=6.18$, $p<.001$	$t(429)=4.30$, $p<.001$	
3a	Media bundle	$9 v. $3	$t(199)=3.53$, $p=.001$	$t(199)=3.62$, $p<.001$	$t(199)=3.64$, $p<.001$	$t(162)=.65$, $p=.520$	$t(106)=2.50$, $p=.014$
		$20 v. $3	$t(210)=3.44$, $p=.001$	$t(210)=3.26$, $p=.001$	$t(210)=2.67$ $p=.008$	$t(191)=.98$, $p=.330$	$t(101)=3.27$, $p=.001$
		$20 v. $9	$t(191)=.94$, $p=.350$	$t(191)=.33$, $p=.745$	$t(191)=.21$, $p=.831$	$t(153)=.30$, $p=.768$	$t(111)=1.96$ $p=.053$
3b	Media bundle	$20 v. $8	$t(2732)=5.88$, $p<.001$	$t(2732)=5.47$, $p<.001$	$t(2732)=5.36$, $p<.001$	$t(2554)=2.44$, $p=.015$	$t(1111)=5.56$ $p<.001$
		$50 v. $8	$t(2685)=5.36$, $p<.001$	$t(2685)=4.15$, $p<.001$	$t(2685)=4.43$, $p<.001$	$t(2555)=3.67$, $p<.001$	$t(1090)=5.03p<.001$
		$50 v. $20	$t(2533)=.04$, $p=.971$	$t(2533)=1.20$, $p=.230$	$t(2533)=.76$, $p=.449$	$t(2413)=1.21$, $p=.228$	$t(959)=.04$, $p=.968$
4	Media bundle	$15 v. $12	$t(3212)=.47$, $p=.641$	$t(3212)=.82$, $p=.410$	$t(3212)=.17$, $p=.862$	$t(2560)=.29$, $p=.774$	$t(2163)=.45$, $p=.654$
5	NSFW bundle	-See tables below-					
6a	Doughnuts	$1 v. $3	$t(189)=.79$, $p=.436$	$t(189)=.79$, $p=.436$	$t(189)=1.92$, $p=.056$	$t(73)=1.17$, $p=.247$	
		$1 v. No anchor	$t(217)=3.28$, $p=.001$	$t(217)=3.28$, $p=.001$	$t(217)=3.39$, $p=.001$	$t(80)=1.83$, $p=.072$	
		$3 v. No anchor	$t(182)=1.34$, $p=.182$	$t(182)=1.34$, $p=.182$	$t(182)=.67$, $p=.505$	$t(93)=.62$, $p=.535$	
6b	Museum tickets	See tables below					
8	Museum tickets	$0.50 v. $1	$t(482)=.99$, $p=.321$	$t(482)=.09$, $p=.926$	$t(482)=.37$, $p=.710$	$t(365)=1.34$, $p=.182$	
		$0.50 v. $2	$t(462)=.14$, $p=.886$	$t(462)=1.39$, $p=.165$	$t(462)=.65$, $p=.514$	$t(363)=.99$, $p=.325$	
		$2 v. $1	$t(524)=1.17$, $p=.244$	$t(524)=1.62$, $p=.106$	$t(524)=1.13$, $p=.261$	$t(412)=.52$, $p=.608$	
9	Media bundle	$12 v. $8	$t(769)=.27$, $p=.785$	$t(769)=.76$, $p=.448$	$t(769)=.69$, $p=.488$	$t(712)=.03$, $p=.974$	$t(348)=.05$, $p=.962$
		No anchor v. $8	$t(783)=1.15$, $p=.253$	$t(783)=1.29$, $p=.198$	$t(783)=.89$, $p=.373$	$t(746)=1.12$, $p=.265$	$t(356)=.85$, $p=.397$
		No anchor v. $12	$t(798)=.25$, $p=.801$	$t(798)=.58$, $p=.561$	$t(798)=.25$, $p=.801$	$t(746)=1.13$, $p=.259$	$t(376)=.99$, $p=.324$

10	Media bundle	$28.88 v. $14	$t(1503)=1.91$ $p=.057$	$t(1503)=1.28$, $p=.202$	$t(1503)=1.59$, $p=.112$	$t(1360)=2.65$, $p=.008$	$t(765)=2.44$, $p = .015$
11	Museum tickets	$5 vs. $11	$t(207)=1.37$ $p=.17$	$t(207)=1.37$ $p=.17$	$t(207)=1.68$ $p=.10$	$t(192)=1.86$ $p=.065$	
		$5 vs. Control	$t(245)=0.81$ $p=.42$	$t(245)=0.89$ $p=.38$	$t(245)=0.95$ $p=.34$	$t(230)=0.29$ $p=.769$	
		$5 vs. Both	$t(213)=0.43$ $p=.67$	$t(213)=0.76$ $p=.45$	$t(213)=0.94$ $p=.35$	$t(196)=1.86$ $p=.065$	
		$11 vs. Control	$t(212)=0.64$ $p=.52$	$t(212)=0.59$ $p=.56$	$t(212)=0.83$ $p=.41$	$t(200)=1.47$ $p=.144$	
		$11 vs. Both	$t(180)=0.77$ $p=.44$	$t(180)=0.57$ $p=.57$	$t(180)=0.68$ $p=.50$	$t(166)=.376$ $p=.707$	
		Control vs. Both	$t(218)=0.26$ $p=.80$	$t(218)=0.04$ $p=.97$	$t(218)=0.09$ $p=.80$	$t(204)=1.12$ $p=.929$	
12a	Media bundle	$19.99 v. $20.01	$t(296)=2.18$, $p=.030$	$t(296)=2.47$, $p=.014$	$t(296)=2.95$, $p=.003$	$t(259)=2.35$, $p=.020$	$t(178)=2.05$, $p=.042$
12b	Media bundle	$20.01 v. $19.99	$t(1663)=.86$, $p=.390$	$t(1663)=.05$, $p=.958$	$t(1663)=.78$ $p=.434$	$t(1543)=2.31$, $p=.021$	$t(701)=.90$, $p=.370$
13	Museum tickets**	$0 vs. $10	$t(373)=1.81$, $p=.072$	$t(373)=1.60$, $p=.111$	$t(373)=1.34$, $p=.182$	$t(359)=.96$, $p=.34$	
		$0 vs. $50	$t(297)=1.51$, $p=.132$	$t(297)=1.56$, $p=.120$	$t(297)=1.41$, $p=.160$	$t(293)=1.09$, $p=.278$	
		$0 vs. $100	$t(216)=1.80$, $p=.073$	$t(216)=2.14$, $p=.033$	$t(216)=1.99$, $p=.048$	$t(212)=1.44$, $p=.151$	
		$10 vs. $50	$t(381)=1.81$, $p=.072$	$t(381)=1.457$, $p=.146$	$t(381)=1.68$, $p=.094$	$t(365)=1.31$, $p=.190$	
		$10 vs. $100	$t(299)=1.35$, $p=.175$	$t(299)=1.25$, $p=.213$	$t(299)=1.33$, $p=.185$	$t(283)=1.42$, $p=.157$	
		$50 vs. $100	$t(224)=4.68$, $p<.001$	$t(224)=4.75$, $p<.001$	$t(224)=5.35$, $p<.001$	$t(218)=4.69$, $p<.001$	
14a (lab)	Media bundle	$28.88 v. $14	$t(384)=2.69$, $p=.007$	$t(384)=4.03$, $p<.001$	$t(384)=3.25$, $p=.001$	$t(265)=2.39$, $p=.018$	$t(261)=2.64$, $p=.009$
14b (lab)	Media bundle	$9 v. $3	$t(579)=3.27$, $p=.001$	$t(579)=4.46$, $p<.001$	$t(579)=4.07$, $p<.001$	$t(514)=2.48$, $p=.014$	$t(418)=3.30$, $p= .001$
14c (lab)	Dough-nuts	$3 v. $1	$t(115)=2.07$, $p=.040$	$t(115)=4.72$, $p<.001$	$t(115=4.12$, $p<.001$	$t(60)=1.76$, $p=.084$	
		No anchor v. $1	$t(111)=1.50$, $p=.136$	$t(111)=2.57$, $p=.011$	$t(111)=2.41$, $p=.018$	$t(67)=1.35$, $p=.181$	
		$3 v. No anchor	$t(106)=.41$, $p=.684$	$t(106)=1.667$, $p=.098$	$t(106)=.87$, $p=.385$	$t(61)=.29$, $p=.771$	
14d (lab)	Dough-nuts	-See tables below-					

The larger mean is presented first.

We winsorize at 95th percentile for all field studies. For hypothetical studies, we winsorize at the highest readily found field value.
*For the means excluding anchor payments, we exclude any payments that match any condition's anchors. For example, in Study 4 ($12 vs. $15 anchors) any payments of $12 and $15 are excluded, regardless of condition.
** The tests were conducted while controlling for month-to-month variation.

4.3.3.1. Statistical Tests for Mean Comparisons – Study 5

Raw Means	$2	$5	$9	$12
$2				
$5	*t*(799)=.45, *p*=.656			
$9	*t*(793)=.78, *p*=.434	*t*(842)=.27, *p*=.723		
$12	*t*(757)=.34, *p*=.736	*t*(806)=.10, *p*=.921	*t*(800)=.44, *p*=.658	
Winsorized	$2	$5	$9	$12
$2				
$5	*t*(799)=.37, *p*=.664			
$9	*t*(793)=.69, *p*=.493	*t*(842)=.27, *p*=.789		
$12	*t*(757)=.24, *p*=.809	*t*(806)=.19, *p*=.853	*t*(800)=.44, *p*=.659	
LN Means	$2	$5	$9	$12
$2				
$5	*t*(799)=.11, *p*=.916			
$9	*t*(793)=.56, *p*=.579	*t*(842)=.47, *p*=.639		
$12	*t*(757)=.15, *p*=.885	*t*(806)=.04, *p*=.965	*t*(800)=.41, *p*=.681	
Excluding Anchors	$2	$5	$9	$12
$2				
$5	*t*(662)=.61, *p*=.569			
$9	*t*(672)=1.01, *p*=.311	*t*(706)=.45, *p*=.652		
$12	*t*(628)=.54, *p*=.586	*t*(662)=.01, *p*=.993	*t*(672)=.45, *p*=.656	

4.3.3.2. Statistical Tests for Mean Comparisons – Study 6b

Raw Means	$0	$0.01	$5	Control
$0				
$0.01	*t*(212)=.24, *p*=.815			
$5	*t*(197)=2.28, *p*=.024	*t*(217)=2.79, *p*=.006		
Control	*t*(210)=1.81, *p*=.072	*t*(230)=2.28, *p*=.024	*t*(215)=.60, *p*=.549	
Winsorized	$0	$0.01	$5	Control
$0				
$0.01	*t*(212)=.17, *p*=.865			
$5	*t*(197)=2.24, *p*=.022	*t*(217)=2.69, *p*=.008		
Control	*t*(210)=1.81, *p*=.073	*t*(230)=2.20, *p*=.029	*t*(215)=.43, *p*=.669	
LN Means	$0	$0.01	$5	Control
$0				
$0.01	*t*(212)=0.47, *p*=.721			
$5	*t*(197)=3.12, *p*=.002	*t*(217)=2.85, *p*=.005		
Control	*t*(210)=2.21, *p*=.028	*t*(230)=1.78, *p*=.077	*t*(215)=.95, *p*=.344	
Excluding Anchors	$0	$0.01	$5	Control
$0				
$0.01	*t*(138)=2.57, *p*=.011			
$5	*t*(109)=.45, *p*=.657	*t*(149)=1.86, *p*=.065		
Control	*t*(121)=.42, *p*=.679	*t*(161)=3.34, *p*=.001	*t*(132)=.89, *p*=.374	

4.3.3.3. Statistical Tests for Mean Comparisons – Study 14d

Raw Means	$0	$0.01	$0.25	$1.75	$50	No anchor
$0						
$0.01	*t*(670)=4.07, *p*<.001					
$0.25	*t*(693)=4.25, *p*<.001	*t*(739)=.67, *p*=.502				
$1.75	*t*(615)=1.94, *p*=.056	*t*(661)=4.00, *p*<.001	*t*(684)=4.34, *p*<.001			
$50	*t*(670)=4.68, *p*<.001	*t*(716)=6.40 *p*<.001	*t*(739)=6.76, *p*<.001	*t*(661)=2.80 , *p*=.005		
No anchor	*t*(632)=1.98, *p*=.048	*t*(678)=5.60, *p*<.001	*t*(701)=5.57, *p*<.001	*t*(623)=1.22 , *p*=.223	*t*(678)=4.23 , *p*<.001	

Winsorized	$0	$0.01	$0.25	$1.75	$50	No anchor
$0						
$0.01	*t*(670)=9.06, *p*<.001					
$0.25	*t*(693)=14.22 , *p*<.001	*t*(739)=4.29, *p*<.001				
$1.75	*t*(615)=5.10, *p*<.001	*t*(661)=15.46 , *p*<.001	*t*(684)=22.39 , *p*<.001			
$50	*t*(670)=8.33, *p*<.001	*t*(716)=18.52 , *p*<.001	*t*(739)=24.91 , *p*<.001	*t*(661)=3.61 , *p*<.001		
No anchor	*t*(632)=2.93, *p*=.004	*t*(678)=12.79 , *p*<.001	*t*(701)=18.90 , *p*<.001	*t*(623)=2.24 , *p*=.025	*t*(678)=5.71 , *p*<.001	

LN Means	$0	$0.01	$0.25	$1.75	$50	No anchor
$0						
$0.01	*t*(670)=8.46, *p*<.001					
$0.25	*t*(693)=11.59 , *p*<.001	*t*(739)=2.68, *p*=.007				
$1.75	*t*(615)=4.65, *p*<.001	*t*(661)=13.81 , *p*<.001	*t*(684)=17.54 , *p*<.001			
$50	*t*(670)=8.09, *p*<.001	*t*(716)=15.36 , *p*<.001	*t*(739)=18.14 , *p*<.001	*t*(661)=4.66 , *p*<.001		
No anchor	*t*(632)=3.05, *p*=.002	*t*(678)=12.46 , *p*<.001	*t*(701)=16.20 , *p*<.001	*t*(623)=1.88 , *p*=.060	*t*(678)=6.18 , *p*<.001	

Excluding	$0	$0.01	$0.25	$1.75	$50	No

Anchors						anchor
$0						
$0.01	*t*(476)=2.14, *p*=.033					
$0.25	*t*(384)=.65, *p*=.514	*t*(348)=.54, *p*=.590				
$1.75	*t*(499)=1.44, *p*=.151	*t*(463)=1.47, *p*=.142	*t*(371)=.20, *p*=.839			
$50	*t*(599)=4.12, *p*<.001	*t*(563)=4.65, *p*<.001	*t*(471)=3.02, *p*=.003	*t*(586)=4.46, *p*<.001		
No anchor	*t*(556)=.09, *p*=.929	*t*(520)=2.62, *p*=.024	*t*(428)=.67, *p*=.501	*t*(543)=1.45, *p*=.149	*t*(643)=4.49, *p*<.001	

4.3.4 By-Month Method and Analyses for Study 13

In January, visitors (N = 149 groups, 315 individuals) were randomly assigned to either the $10 or $50 maximum payment condition. People paid more when they were provided with a low maximum payment amount, $10, than a high one, $50, $M_{\$50}$ = $2.08 vs. $M_{\$10}$ = $2.99, $t(147) = 2.13$, $p = .035$, $d = 0.35$. Inspired by this (reversed) success, we sought replication.

In April (N = 140 groups, 330 individuals), we replicated the same anchors in addition to a control condition in which visitors were not provided with any maximum payment amount information. Average payments did not differ across the three conditions, $F(2, 139) = 1.62$, $p = .20$, and in fact there was a non-significant reversal of the effect observed in January ($M_{\$10}$ = $3.39 vs. $M_{\$50}$ = $2.97), $t(91) = .644$, $p = .52$. Completing the comparisons, people paid similarly in the $50 and control conditions ($M_{control}$ = $2.28 vs. $M_{\$50}$ = $2.97), $t(85) = 1.39$, $p = .17$, and in the $10 and control conditions ($M_{control}$ = $2.28 vs. $M_{\$10}$ = $3.39), $t(91) = 1.70$, $p = .094$.

In May, partially inspired by the surprising, but non-significant results, we tried to replicate the results from April. Groups of visitors (N= 84 groups, 149 individuals) were randomly assigned to the $10 maximum, $50 maximum, or control condition. The results were consistent with what we found in April; none of the comparisons of the average payment amounts significantly differed across the three conditions.

Perhaps, we reasoned, there was still potential for this paradigm to demonstrate anchoring effects, so in July we increased the high anchor to $100 while keeping the lower anchor (at $10). Visitors (N= 223 groups, 551 individuals) were randomly assigned to one of the following conditions: no maximum payment information, $10, or $100 condition. The average payment amounts did not differ across the three conditions. Visitors paid similar amounts in the $10 and $100 conditions, ($M_{\$10}$ = $4.00 vs. $M_{\$50}$ = $3.74), $t(147) = .49$, $p = .622$.

FOOT NOTES

[1] For this study and others studies in this paper there was no deception and we donated exactly the advertised amounts to the charities identified across the four studies.

[2] There were three ways to enter the store; two main entrances for pedestrians and one pathway that connected a parking lot to the store. We covered the two main entrances but not the pathway because the pathway was too narrow to set up a table without obstructing the traffic to the store. Although imperfect, we made efforts to cover the most traffic by setting up tables at the two main entrances to minimize the number of customers avoiding our tables (Androni, Rao, and Trachtman, 2011).

[3] Of the thirty five people who donated but did not take a bag, 1 were in PWYW, 16 were in SSR and 1% going to charity, 18 were in SSR and 50% going to charity. We included these pure donations in our analysis of charitable surplus and purchase rates.

[4] It is plausible that shoppers were more familiar with the grocery store than the non-profit organization and had a particular preference for the store because it offered a wide range of organic vegan products that could not be found in typical grocery stores. To test their preference for a bag with a commercial or charitable logo, we conducted a subsequent study in which customers were presented with both bag types and chose one they preferred. Our sign read, "Please take a bag. It's free." Among those who took a bag (N = 142), 58.5% chose a bag with the store logo, 32.4% chose a bag with the charity logo, and 7% equally preferred either bag type. Although people preferred a bag with the store logo to one with the charity logo (X^2 = 19.45, $p < .001$) when bags were free, they were equally likely to purchase a bag or paid about the same for either bag type when they could pay any price they wanted. It is possible that they were choosing more frequently and paying more for a bag with a public signal of charitable behavior than what they would have if they could have chosen a bag type. However, our study design and results do not offer conclusive insights to this possibility.

[5] We gave away a few bags to research assistants and kept some as samples.

[6] When customers wanted to purchase more than one bag, research assistants politely informed that they could allow only one bag per person. There are 19 transactions in which customers insisted on purchasing multiple bags or asked other shoppers to buy additional bags for them. In our analysis, they are considered as one transaction with the average payment per bag as a dependent variable.

[7] There were 7, 17, 15, 25, and 21 of pure donation in the PWYW, SSR with 1% to charity, 50% to charity, 99% to charity, and 100% to charity conditions, respectively.

[8] If customers did not read it out loud to the cashier, the cashier read it to the customers to make sure the price information was clearly delivered. Neither tea nor coffee price was displayed. If customers wanted tea, they were told that it was $3 for a cup.

[9] Customers who wanted both tea and coffee were not exposed to our manipulation because we were concerned that the customers who wanted coffee would be influenced by the price for the tea.

[10] Coffee was sold only in large plain white cups.

[11] No customers refused to purchase coffee when they were told to draw a piece of paper that contained pricing information.

[12] Forty individuals decided not to purchase a doughnut when they were told to select a price from the box.

[13] Some people immediately gave money to the cashier and took a doughnut before the cashier had an opportunity to ask them to draw a piece of paper from the box for the price. But this rarely happened.
[14] 10 of these people were in the PWYW, 11 to the SSR-10% going to charity, and 16 to the SSR-50% going to charity conditions.
[15] The results after excluding these cases: M_{PIF} = \$2.73 vs. M_{PWYW} = \$1.84; $F(1, 146) = 5.73$, $p = .025$.
[16] Although we collected this demographic information, we never seriously considered it. In general, women pay slightly more than men (in Studies 1, 2, 4, and 6) and older people pay more than younger people (in Studies 1, 2 and 3). Because those factors were not systematically linked to conditional assignment, none influenced the operation of the manipulation. Those variables are visible in the available data sets.
[17] Thanks, Clayton Critcher.
[18] One student group (n = 18) came with their teacher as part of a class trip. Another group (n = 5) had a pre-paid "Go Card" which guaranteed free admission to all museums. Neither group received either version of the manipulation and hence neither could be included in the analysis.
[19] People paid more, in both conditions, in Study 2 than in Study 1. Though it is possible that our clarified wording had this effect, it seems just as likely that it is a seasonality effect. Perhaps people simply pay more in the springtime?
[20] As with Study 1-3, we did not have any specific predictions for these measures and so we do not report those results here. Nevertheless, those variables and those data are available with the complete data sets.
[21] Customers' concern for other buyers' is necessarily not literal; the next customer gets the same offer regardless. But they may still feel direct guilt for paying zero or very little.
[22] The unrelated experiment concerned evaluating a food item, its randomization was independent of this study, and the data were never analyzed together.
[23] We included these items primarily as potential exploratory variables if, in fact, we did not observe any significant effects on the critical dependent variable (i.e., payment). Accordingly, with the exception of the analysis of the items about the predicted payments of others, we do not elaborate on their analysis. We included the exact same items in the subsequent studies so as to replicate as closely as possible across studies and to leave consistent data were anyone to subsequently go back and reanalyze the experimental data. Accordingly, all of the measures are reported in Table 3.2 in Appendix, and all of the data are available for download.
[24] To eliminate potential need for exclusions, experimenters were scheduled so that they never knew the participants. Also, there were no obviously drunk participants in this or the remaining studies.
[25] The unrelated survey was about how people evaluate political campaign advertising. We analyzed the two studies separately.
[26] We excluded those who participated in Study 9A.
[27] Including those who failed the attention check did not change the direction or significance of the results.
[28] In this study, because of a software cap at 1000, we could not systematically block all respondents who had previously participated in a similar study (e.g., Study 9A). Instead, we simply asked participants whether or not they had completed surveys using a similar coffee

purchasing setting before and told them that their payment will not depend on how they answered this question.

[29] 9.2% of the participants in PWYW, 13.6 % in PIF, 11.4% in PIF singular, 19.9% in PIF plural, 15.7% in PIF company.

[30] 86 participants (9.6%) indicated that they had completed a similar survey before. Excluding these repeat survey takers did not change the direction or significance of our results reported above.

[31] An even more severe form would be to lie, by making a small donation but claiming a large donation on the card. Seven participants showed a mismatch between actual and reported payments, six of whom reported a higher number than they had actually paid. There were no conditional differences in these payments, but the behavior is still intriguing.

[32] Rather than make this assertion, we simply asked 201 MTurkers the number of hours in a day. Half first indicated whether the answer was greater or less than four hours, half greater or less than 44 hours. There was no anchoring effect (M = 24.0 hours, SD = 0.0).

[33] Although only the former claims such a finding (Simonsohn, Simmons, and Nelson 2014).

[34] By editorial request, we report the studies in an order that maximizes logical progression, rather than the chronology in which they were conducted.

[35] This experiment was conducted at a high-traffic location where many students pass by on their way to their classes. For logistical reasons, we counted passerby walking toward only one direction, into campus.

[36] Seven individuals bought more than one doughnut. The rest bought just one. We used the average payment per doughnut for those seven people.

[37] So as not to risk revealing the true nature of our doughnut stand to others, friends of research assistants were permitted to make purchases.

[38] Although there are several, very similar ways to define a percentile, we define ours as the percentage of payments below the payment of interest, unless stated otherwise.

[39] This was a mild deception. The stated average was similar, but not identical, to other averages.

[40] We added the control condition in which no anchor was provided in the second month of the experiment. The results were not influenced by month $F(1,948)=.215$, $p=.643$, and our analysis does not include the month as a covariate.

[41] ScholarShare was one of the museum's actual Free Wednesday sponsors at the time.