MINAH H. JUNG, HANNAH PERFECTO, and LEIF D. NELSON*

Anchoring, the biasing of estimates toward a previously considered value, is a long-standing and oft-studied phenomenon in consumer research. However, most anchoring work has been in the lab, 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 better understand how anchors influence payments. Sixteen field studies (N = 21,997) and four hypothetical studies (N = 3,174) reveal four main points: (1) Although anchoring replicates both with and without financial consequences (Studies 1–2), the percentile rank gap between anchors in the distribution of payments is a much stronger predictor of anchoring emerging than merely the absolute gap between the anchors on a number line (Studies 3–5). (2) Low anchors influence payments more than high anchors (Studies 6a–b). (3) Findings from the literature that should enhance anchoring effects—anchor precision, descriptive and injunctive norms, nonsuggestions—yield null results in payment (Studies 7–13). (4) The above patterns do not emerge in hypothetical settings (Studies 14a–d), in which anchoring is as big and reliable as the literature has previously suggested.

Keywords: anchoring and adjustment, experimental pricing, numeric judgment, field experiment, heuristic

Online Supplement: http://dx.doi.org/10.1509/jmr.14.0238

Anchoring in Payment: Evaluating a Judgmental Heuristic in Field Experimental Settings

Consumer researchers study human psychology in the hopes that it will enable 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 article aims to take an incredibly robust judgment process, anchoring (Tversky and 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 that 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

When people are asked to guess whether an adult giraffe weighs more or less than 2,100 pounds and then to estimate the animal's weight, they give a larger estimate of weight than people who are first asked to guess whether an adult giraffe weighs more or less than 800 pounds (Frederick and Mochon 2012). This difference is due to anchoring, first articulated by Tversky and Kahneman (1974), wherein the presence of irrelevant numbers can substantially influence numerical judgments. Since then, anchoring has become one of the most-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, who is prompted to provide an estimate of a certain value. Second is the deliberate consideration of the anchor before making a final estimate. Although it is widely agreed that deliberation yields the largest anchoring effects (e.g., Brewer and Chapman 2002; Mochon and Frederick 2013), research has shown that deliberation-free anchors can also be influential (e.g., Critcher and Gilovich 2008; Wilson et al. 1996). 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).[1] Overall, when an uncertain numeric entity is evaluated, higher anchors should produce higher estimates.

Researchers can all observe the effect, but they broadly disagree about the cause. Tversky and Kahneman (1974) attribute the effect to insufficient adjustment from an 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) propose a selective accessibility account, in which an anchor leads people to call to mind information that supports final estimates closer to the anchor. Complicating things further is the question of whether motivation for accuracy influences the effect. Many researchers have failed to find any effect of incentivized accuracy (Chapman and Johnson 2002; Epley and Gilovich 2005; Strack and Mussweiler 1997; Tversky and Kahneman 1974), whereas others have suggested that incentives alter adjustment of estimates conditional on simply knowing in 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; thus, 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 they consistently pay more than zero (Armstrong, Soule, and Madrigal 2015; Kim, Natter, and Spann 2009; Regner and Barria 2009; Riener and Traxler 2012; but see León, Noguera, and Tena-Sánchez 2012 for a boundary condition at high retail prices). The PWYW paradigm allows anchors to be organically presented as default or suggested prices. To return to the three aspects of anchoring in this context: in our studies, anchors are presented clearly without justification (i.e., arbitrarily), people must and reject (or accept) the anchor before their final payment, and the values of the goods sold are not readily known. Moreover, the combined uncertainty of personal valuation (Bettman, Luce, and Payne 1998) and socially appropriate payment (Gneezy et al. 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 et al. 2013), and buying and selling prices (Simonson and Drolet 2004). However robust this evidence, there is still uncertainty about how these effects are expressed outside of hypothetical situations.

A smaller body of work has considered anchoring effects with incentive-compatible designs. Work by Ariely, Loewenstein, and Prelec (2003), as well as Maniadis, Tufano, and List (2014) employs designs with real money and goods at stake. Both of these articles show data consistent with classic anchoring effects.[2] Both also retain 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 the actual willingness to pay was elicited through a Becker–DeGroot–Marschak (BDM) procedure. Nunes and Boatwright (2004) use more naturalistic anchors (featured prices of nearby products at a local concert) but also utilize a BDM procedure. Given that the real world frequently lacks such contrivances and that 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 reference prices (e.g., minimum bids [Kamins, Dreze, and Folkes 2004], buy-it-now prices [Hardesty and Suter 2013], bid options [Spann et al. 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, in that payment is both elective in amount and guaranteed. A variety of

---

[1]Rather than make this assertion, we simply asked 201 MTurk members 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).

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

studies have examined 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 (Croson and Shang 2008; Shang and Croson 2009) reference prices on donation behavior. These studies, too, frequently produce significant anchoring effects. Finally, there have even been some investigations of anchoring under PWYW, in which the presence of an anchor is varied (e.g., Gneezy et al. 2012; Kim, Kaufmann, and Stegemann 2013), 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 previously, many studies 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 study, 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 it is incomplete for making overall statements about a general process like anchoring. The present study, with many more experiments, intends to have sufficient breadth to offer a more nuanced rendering.

This need for nuance becomes already apparent when we look more closely at the aforementioned literature on charitable donations. Some studies have been successful, but some have 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, Carlsson, and Johansson-Stenman 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 article.

Using this PWYW context, we explore how and why anchoring effects change between the lab and the field. Sixteen studies test 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, we consider these gaps in terms of both the absolute gap (the numeric difference between the anchor values) and the distributional gap (the difference in the anchors' percentile ranks in the distribution of payments). We find that the latter is a much better predictor of anchoring effects than the former. 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 range of payments widens in the distribution: as a result, previously inert extremely high anchors now appear reasonable and become influential.

In this research, we first[3] conceptually replicate past work that has been done in both the lab and the field, with

both financial and nonfinancial consequences, in Studies 1 and 2. In Study 3, we more closely examine absolute and distributional anchor gaps, an apparently critical distinction that has been generally neglected in the literature. Studies 4–6 reveal another asymmetry: low anchors influence payments more than high anchors. 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 (Studies 14a–d), we present four hypothetical studies to demonstrate the disparity between hypothetical and real settings; we show anchoring effects are much easier to find in hypothetical judgments. Descriptive statistics for each study are in Table 1, and histograms for each study, along with robustness checks for each study (e.g., log-transforming), are available in the Web Appendix. 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.

For all experiments reported in this article, 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 predetermined sample size. All data and materials are available in the online supplement.

EXAMINING DIFFERENT TYPES OF ANCHOR GAPS

Study 1: Field Anchoring with Nonpayments

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 in which people make numeric decisions that lack financial consequences for themselves.

An online media retailer (with whom we collaborate in many subsequent studies in this article) allows customers to choose the percentage of their payments that goes to the products' developers vs. to the retailer (on a bipolar scale, increasing in one-percentage-point units). The situation has low stakes for the customer because he/she pays the same amount regardless of 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 when there were higher default allocations. 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 allocation to the e-book authors, for ease of explanation): 49%, 50%, 51%, 89%, 90%, and 91%. We chose these numbers for two reasons: First, the company required that our conditions average out to their previous default of 70% going to the developers. Second, precise anchors have been shown to more strongly influence anchoring effects than round-number anchors 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 analysis of variance (ANOVA) revealed differences among

---

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

Table 1
BASIC INFO FOR EACH STUDY

	Product	N	Anchor Type	Anchor	n	Percentile of Anchor	M (SD)Outcome	Percentile of Mean Outcome
Study 1	Allocation of payment	1,328	Default	50%	648	16.9	61.64% (15.36%)[a]	30.9
				90%	680	80.0	89.20% (6.60%)[b]	80.0
Study 2	Doughnuts	1,009	Suggestion	$.25	490	20.3	$.44 ($.38)[a]	90.1
				$1.75	181	92.8	$1.04 ($.71)[b]	50.4
				Control	338	—	$.66 ($.54)[c]	29.9
Study 3a	Media bundle	303	Default	$3	110	4.6	$6.59 ($2.38)[a]	55.1
				$9	91	69.6	$7.79 ($2.44)[b]	66.3
				$20	102	96.7	$8.29 ($4.57)[b]	69.6
Study 3b	Media bundle	3,978	Default	$8	1,443	21.5	$8.88 ($4.03)[a]	23.5
				$20	1,291	95.3	$9.89 ($4.96)[b]	23.5
				$50	1,244	100.0	$9.88 ($5.63)[b]	23.5
Study 4	Media bundle	3,214	Default	$12	1,602	75.9	$8.80 ($4.64)[a]	23.2
				$15	1,612	87.1	$8.88 ($4.88)[a]	23.2
Study 5	Vodo	1,603	Default	$2	376	29.2	$11.48 ($7.87)[a]	42.7
				$5	425	30.7	$11.24 ($7.65)[a]	42.1
				$9	419	32.1	$11.05 ($7.68)[a]	41.8
				$12	383	45.6	$11.29 ($7.72)[a]	42.3
Study 6a	Doughnuts	395	Suggestion	$1	113	27.9	$.91 ($.38)[a]	27.9
				$3	78	94.8	$.84 ($.75)[a,b]	27.9
				Control	106	—	$.72 ($.45)[b]	24.9
Study 6b	Museum tickets	431	Suggestion	Nothing	97	.0	$2.55 ($2.73)[a,c]	57.1
				$.01	117	16.2	$2.47 ($2.36)[a,c]	52.0
				$5	92	67.3	$3.47 ($2.96)[b]	62.9
				Control	102	—	$3.24 ($2.77)[b,c]	61.3
Study 8	Museum tickets	957	Suggestion	$.50	211	56.4	$.93 ($2.20)[a]	64.8
				$1	273	64.9	$.77 ($1.29)[a]	64.6
				$2	253	83.4	$.91 ($1.36)[a]	64.8
				Control	220	—	$1.08 ($6.80)[a]	74.8
Study 9	Media bundle	1,175	Previous payment	$8	378	7.0	$11.10 ($3.97)[a]	68.1
				$12	393	68.2	$11.17 ($3.50)[a]	68.2
Study 10	Media bundle	2,190	Default	$14	1,122	88.1	$10.78 ($5.40)[a]	62.2
				$28.88	1,068	98.5	$11.17 ($6.19)[a]	67.5
Study 11	Museum tickets	429	Suggestion	$5	121	93.0	$1.06 ($1.55)[a]	70.2
				$11	88	99.8	$.77 ($1.48)[a]	68.1
				Control	126	—	$.90 ($1.52)[a]	68.5
Study 12a	Media bundle	714	Default	$19.91	138	84.7	$11.67 ($6.13)[a,b]	25.6
				$19.99	140	85.9	$12.31 ($6.39)[a]	66.1
				$20.00	135	88.1	$11.53 ($6.63)[a,b]	25.6
				$20.01	158	93.8	$10.52 ($7.65)[b]	25.4
				$20.09	143	95.2	$11.60 ($6.89)[a,b]	25.6
Study 12b	Media bundle	4,110	Default	$19.91	790	91.4	$10.56 ($4.81)[a]	75.9
				$19.99	827	92.2	$10.39 ($4.48)[a]	75.4
				$20.00	861	93.5	$10.32 ($4.45)[a]	75.4
				$20.01	838	97.1	$10.60 ($5.67)[a]	75.9
				$20.09	794	97.7	$10.55 ($4.71)[a]	75.9
Study 13	Museum tickets	590	Maximum price	$10	231	99.0	$2.72 ($3.25)[a,b]	58.7
				$50	155	100	$2.25 ($2.27)[a]	57.3
				$100	73	100	$2.90 ($2.54)[b]	58.7
				Control	147	—	$2.01 ($2.94)[a,b]	57.3
Study 14a	Media bundle	386	Default	$14	205	56.2	$13.01 ($9.82)[a]	56.2
				$28.88	181	94.6	$15.78 ($10.34)[b]	73.1
Study 14b	Media bundle	581	Default	$3	274	21.7	$5.10 ($3.43)[a]	47.0
				$9	307	85.7	$5.99 ($3.12)[b]	47.0
Study 14c	Doughnuts	169	Suggestion	$1	61	36.7	$0.82 ($.82)[a]	36.1
				$3	56	89.3	$1.47 ($.94)[b]	73.4
Study 14d	Doughnuts	2,038	Suggestion	$.00	313	.0	$.92 ($.78)[a]	47.9
				$.01	359	3.8	$.50 ($.52)[b]	29.6
				$.25	382	11.9	$.39 ($.52)[c]	29.0
				$1.75	304	87.6	$1.09 ($.61)[d]	83.3
				$50	359	100	$1.52 ($1.26)[e]	87.6
				Control	321	—	$1.04 ($.73)[f]	73.4

Notes: Within each study, means with different subscripts differ significantly ($p < .05$). In Study 13, the means were controlled for month-to-month variation. Studies 14a–d are all hypothetical.

the conditions (F(5, 1,322) = 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 the three high (M = 89.20%) anchors (t(1, 326) = 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) = .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 middle high anchor ($M_{90\%} = 89.27\%$), which in turn generated a lower allocation than the highest anchor ($M_{91\%} = 90.02\%$).

Perhaps the result was due to customers' laziness or inattention rather than anchoring. To confirm the operation of anchoring, we reanalyzed the data excluding customers who had simply accepted the default allocation, leaving 449 customers (33.81% of the original sample). First, note that this is (excessively) conservative; the strongest possible anchoring effect would be no adjustment in allocations from the default at all, and we are systematically eliminating the responses that fit that pattern. Even in this restricted sample, the effect was still large and significant between the three high (M = 85.71%) and the three low (M = 73.66%) defaults (t(447) = 8.20, $p < .001$; d = .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, that is, those that would represent uncommonly small or uncommonly large responses (e.g., Ariely, Loewenstein, and Prelec 2003; Jacowitz and Kahneman 1995; 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.

In addition, we guarded against another alternative explanation. Previous research has suggested 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 manipulates whether customers have this foreknowledge.

Method. The experiment contained two manipulations: foreknowledge and anchor value. The first addressed the selection concern by randomly assigning people to a 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 the Web Appendix for images of the signs). We chose these anchors because our previous experience in this domain suggested those amounts were far apart in terms of their percentile ranks in the distribution of customers' payments in previous data sets.

We changed the sign in a randomized order for every 200 people passing by our doughnut stand.[4] 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 out a price. For each transaction, we recorded date, time, payment, customer group size, customer gender, and customer age. We also recorded the total number of passersby in each condition. We predetermined to collect data until we had at least 100 observations per condition.

Results. Customers (N = 892 groups, 1,038 individuals) bought 1,054 doughnuts.[5] We used individual doughnut purchase as unit of analysis, with the average payment per doughnut as the dependent measure. We excluded 29 transactions in which the customers were the experimenters' friends,[6] which left us 1,009 purchases for analysis (see Figure 1).

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 Pay What You Want" than when it said "$1.75 or Pay What You Want" (382 customers out of 14,631 passersby vs. 87 customers out of 14,548 passersby; $\chi^2(1) = 186.89$, $p < .001$). Furthermore, people seeing the $.25 or PWYW sign were more likely to purchase than those seeing either the PWYW sign (220 customers out of 13,639 passersby; $\chi^2(N = 28,270) = 33.73$, $p < .001$) or the more ambiguous "Dream Fluff" sign (320 customers out of 28,073 passersby), $\chi^2(1) = 128.72$, $p < .001$. Even though all participants could pay what they wanted, the invitation to pay $.25 or pay what you want was more motivating than any other sign.

However, a 3 (anchor: $.25 or PWYW vs. $1.75 or PWYW vs. PWYW) × 2 (selection: absent vs. present) ANOVA on payments revealed only a main effect of anchor (F(2, 1,003) = 74.29, $p < .001$). Despite the substantial differences in purchase rate, there were no effects of selection on average payments ($M_{selection}$ = $.71 vs. $M_{no\ selection}$ = $.74; F(1, 1,003) = .55, $p = .46$). Critically, the interaction between selection and anchor was not significant (F(2, 1,003) = 1.15, $p = .316$).

Customers paid more for a doughnut under the $1.75 or PWYW sign than under the PWYW sign ($M_{\$1.75}$ = $1.04 vs. M_{PWYW} = $0.66; t(517) = 6.77, $p < .001$). But they paid
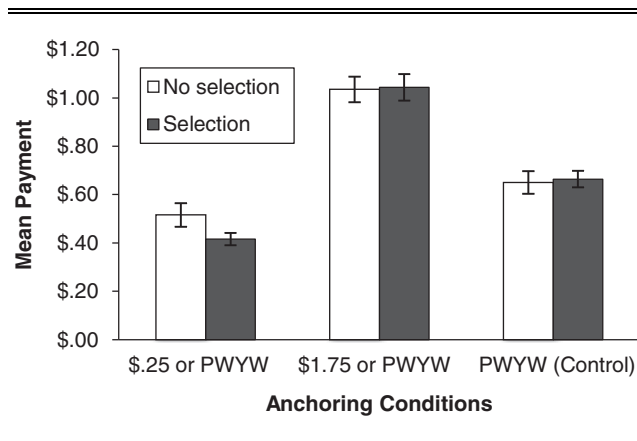
---

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

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

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

more under the PWYW sign than under the $.25 or PWYW sign ($M_{PWYW} = \$.66$ vs. $M_{\$.25} = \$.44$; t(826) = 6.92, $p <$ .001). Most importantly, people paid more under the $1.75 or PWYW sign than under the $.25 or the PWYW sign ($M_{\$1.75} = \$1.04$ vs. $M_{\$.25} = \$.44$; 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—like most anchoring studies—does not differentiate the types of anchor gaps, for example, in absolute terms, standardized terms, or percentile[7] 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 similar in the distribution (i.e., similar in percentile rank of payment) and compare those results with anchors that are both absolutely and distributionally far apart on these measures.

Method. We again collaborated with the PWYW media retailer from Study 1, but we now focused on payment amounts rather than allocation proportions. This company occasionally 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 items). Second, there is a "bonus" price, above which any elective payment also buys an additional two to four items (e.g., $9, for ten items in total). Third, customers identify their chosen prices either by typing a number in a box or by selecting a number on a sliding scale, which moves in $1 increments and ranges from the minimum price to $100. Customers also may elect that 10% of their payment will go to charity.

[7]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.

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. Thus, the lowest default was separated from the middle and highest defaults by large gaps in both absolute value ($6 and $17, respectively) and percentile rank in the distribution of payments (69.6 and 96.7, respectively). Critically, however, the gap between the two higher defaults was relatively small in terms of percentile rank (27.1) but large in terms of absolute value ($11). Although in this and many subsequent studies, participants were not directly prompted to accept or reject the anchor as they were in Study 2, we believe that having to move the slider away from the default involves a similar cognitive process (see also Wilson et al. [1996] and Critcher and Gilovich [2008] for anchoring with even less direct anchor consideration).

Results. Payments differed between the three anchor 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 (M = $8.29) anchor ($3 vs. $9: t(199) = 3.53, p = .001; $3 vs. $20: t(210) = 3.44, p = .001). However, payments did not differ between the $9 and $20 anchors (t(191) = .94, p = .350).

Study 3b: Replicating Study 3a

Study 3a demonstrates that the types of anchor gaps matter in the field. However, the study's 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 a default of $8, $20, or $50, allowing us a similar pattern of absolute and distributional anchor gaps as in Study 3a.

Results. As before, payments differed across the three anchor conditions (F(2, 3,975) = 19.62, p < .001). Payments were lower for the $8 anchor (M = $8.88) than for either the $20 (M = $9.89) or the $50 (M = $9.88) anchors ($8 vs. $20: t(2,732) = 5.88, p < .001; $8 vs. $50: t(2,685) = 5.36, p < .001), but the $20 and $50 anchors produced similar mean payments (t(2,533) = .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 throughout this article, the mix of findings suggests that even with a very large absolute anchor gap (based on absolute values of anchors), effects are largely dependent on the distributional gap (based on percentile rank in payments).

Past research on anchoring 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 GAP 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 varied default prices and measured payments.

Method. Study 4 (N = 3,214) involved 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 see either a $12 or a $15 default, both of which are relatively common payments at the retailer under these conditions. We also ran two additional manipulations involving the option to direct 10% of the payment to charity. We did not predict interactions between the additional manipulations and the default manipulation, and we did not find them, so they are not discussed further.

Results. There was no difference in average payment between the two conditions ($M_{\$12}$ = $8.80 vs. $M_{\$15}$ = $8.88; t(3,212) = .47, p = .641).

Both $12 and $15 are above the median payment ($10) and were therefore somewhat uncommon payments in this study. 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 anchors are uncommonly high, all adjustments will stop at the boundary, eliminating any difference between the anchors. However, Mussweiler and Strack (2001) use anchors well beyond the highest estimations observed in their study (e.g., offering 214 years as a possibility for Gandhi's age at his death), whereas our anchors in Study 4, although 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 in Study 4 had translated into a $3 payment difference, it would have been financially substantial, but anchors that were $3 apart might have seemed too similar to customers: they were only 10 percentile ranks apart. Perhaps customers called to mind the same types of supporting information for the two anchors (according to the selective accessibility account) or similarly changed their perceptions of the scale for each (in line with the scale distortion account). Study 5 titrates and widens the anchor gap to consider these possibilities. We use four levels of anchor to investigate this similarity account, predicting that higher anchors are associated with higher average payments. We determined the sample size 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 (movies, games, etc.). Customers who pay more than the current average payment receive four additional products. If they beat the current average payment by more than $7.50, they receive three further 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.

In our study, site visitors were randomly assigned to see a default 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 structure would match absolute anchor gaps

to distributional gaps (in similar past promotions, average payments had been between $4 and $6).

Results. The promotion did better than expected, resulting in average payments between $9 and $12 across the three-week period. The values of three lowest anchors ($2, $5, and $9) were close in percentile rank in the distribution of payments: 30.0, 31.4, and 32.3, respectively, with the value of 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, 1,599) = .21, p = .890). The largest gap in the set, from $2 to $12, actually showed a nonsignificant trend opposite the anchors ($M_{\$2}$ = $11.48 vs. $M_{\$12}$ = $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 distributional 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 $.99 payment was at the 66th percentile, but because 25% of customers paid $1, a $1 payment might alternatively be interpreted as a 91st-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 with the unambiguously high anchor of $3 (95th percentile in Study 2). Three additional conditions were a true control (PWYW without any anchor) and two fixed-price conditions (prices equivalent to the two anchors) to allow for assessment of overall demand.

Method. The experiment environment was identical to that of Study 2. We sold glazed doughnuts on the University of California, Berkeley, campus on 17 days between October 2013 and December 2013. People (N = 44,483) passing our doughnut stand saw one of our five shop signs: "Dream Fluff Doughnuts!" followed by "$1," "$3," "Pay What You Want," "$1 or Pay What You Want," or "$3 or Pay What You Want." For signage examples, see the Web Appendix. We changed the sign 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 PWYW condition with the greatest number of purchases.

Results. A 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}}$ = $.91 vs. $M_{\$3 \text{ or PWYW}}$ = $.84; t(189) = .79, p = .436; d = .11). From these results, it appears that customers saw the common $1 payment (i.e., chosen by 62%) as a higher anchor, making our distributional anchor gap much smaller than we had

wanted. Payments in the $1 or PWYW condition were higher than in the anchor-free condition ($M_{\$1 \text{ or PWYW}}$ = $.91 vs. M_{PWYW} = $.72; t(217) = 3.28, p = .001), but payments in the $3 or PWYW condition did not differ from those in the anchor-free condition ($M_{\$3 \text{ or PWYW}}$ = $.84 vs. M_{PWYW} = $.72; t(182) = 1.34, p = .182).

Study 6b: Considering Low Anchors

In order to further probe the potential differential 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 the strong anchor effect on payments, Study 6b employs anchor gaps closer to lab levels of extremity (16th vs. 67th percentiles), while retaining our control condition. This structure 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, 909 individuals) to the Cartoon Art Museum on its Pay-What-You-Wish Days in June, July, August, and September 2014 were randomly assigned to one of four conditions: PWYW; pay nothing or PWYW; pay $.01 or PWYW; or 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 pay $.01 or PWYW condition were told, "Thanks for coming to the museum today. You can pay $.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 et al. 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 and $.01) conditions did not differ ($M_{\text{nothing or PWYW}}$ = $2.55 vs. $M_{\$.01 \text{ or PWYW}}$ = $2.47; t(212) = .24, p = .815). Visitors paid more when they 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 (M_{PWYW} = $3.24 vs. $M_{\text{nothing or PWYW}}$ = $2.55; t(210) = 1.81, p = .072). They paid significantly more when they were not provided with any anchor than when they were given the choice between paying $.01 and paying what they wanted (M_{PWYW} = $3.24 vs. $M_{\$.01 \text{ or PWYW}}$ = $2.47; t(230) = 2.28, p = .024).

In line with the results of Study 6a, visitors did not pay more when they were provided with a very high anchor ($5) than when they were not provided with any anchor ($M_{\$5}$ = $3.47 vs. M_{PWYW} = $3.24; t(215) = .60, p = .549). But they paid more in the high anchor condition than in either the nothing or PWYW condition ($M_{\$5}$ = $3.47 vs. $M_{\text{nothing or PWYW}}$ = $2.55; t(197) = 2.28, p = .024) or the $.01 or PWYW condition ($M_{\$5}$ = $3.47 vs. $M_{\$.01 \text{ or PWYW}}$ = $2.47; t(217) = 2.79 p = .006).

Discussion

Studies 6a–b show 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 distributional gap between the anchors must be quite wide to elicit significant effects, suggesting that the anchor gaps chosen by convention in the anchoring literature could actually be the minimum requirement.

Study 6a also reveals a previously undiscovered nuance: if the gap is too wide, leaving anchors to be too extreme, customers might 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: whereas customers can be put off by an implication that they should make a large payment, they embrace the tacit permission to pay a small amount (for a similar effect of low anchors through social information in donations, see Croson and Shang 2008). 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 magnitudes and gaps are not the sole factors that drive anchoring effects. 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 can 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. We could be confident that the anchors were plausible numbers and therefore plausibly more influential. Armstrong Soule and Madrigal (2015) demonstrate in a hypothetical context that anchors that set injunctive social norms are more influential than those that set descriptive norms (i.e., from the company, as in previous studies), and Smith, Windmeijer, and Wright (2014) find similar results in the context of charitable donations. Therefore, we predicted that higher payments would be associated with higher average payments shown.

Method. Study 7 (N = 1,074 customers) took place at 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 past five payments, labeled "Current average purchase price," above the payment slider (the other half of the visitors saw no information about average payment).

We chose the average of the past five payments, rather than the cumulative average, to ensure substantial variability in the anchors shown. The site updated the displayed average payment every seven minutes.

Results. The average payment ranged between $4.40 and $18.00, with a median of $10.27. We exclude the first 115 customers from our analyses because they were not shown an average payment. Contrary to our hypothesis, there was no correlation between the payments of those who saw an average price and the prices 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 times of day when the average payment tends to be higher are more likely to see high average payments). To address this concern, in the condition with no average payment shown, we tracked value of the anchor a customer would have seen had it been displayed (i.e., the average of the past five payments). Because that anchor was unseen, if it were related to payments, then we would suspect a 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) = .42, *p* = .676; see Figure 2).

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). 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. Participants saw one of three suggested donations ($.50, $1, and $2), and approximately half of the participants were told that the average payment was $1.[8]

Method. Groups of visitors (N = 957 groups, 2,761 individual visitors) were randomly assigned to one of eight conditions in a 4 (suggested donation amount: no information, $.50, $1, or $2) × 2 (average donation amount: no information or $1) between-participants study.[9] We selected these numeric values from the distribution of donation amounts in the previous month. On average, visitors gave approximately $.94 (including the 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?" (Text in parentheses was shown only in the relevant conditions.) Visitors also indicated their group size and home zip code on the card. Examples of the cards can be found in the Web Appendix. We predetermined to collect data until we had at least 100 groups per condition.

---

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

[9]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.

Figure 2
STUDY 7: WE FIND THAT PAYMENTS ARE NOT INFLUENCED BY THE CUSTOMER KNOWING THE AVERAGE PRICE PAID BY PREVIOUS CUSTOMERS, AS SHOWN BY THE REGRESSION LINES FOR EACH GROUP
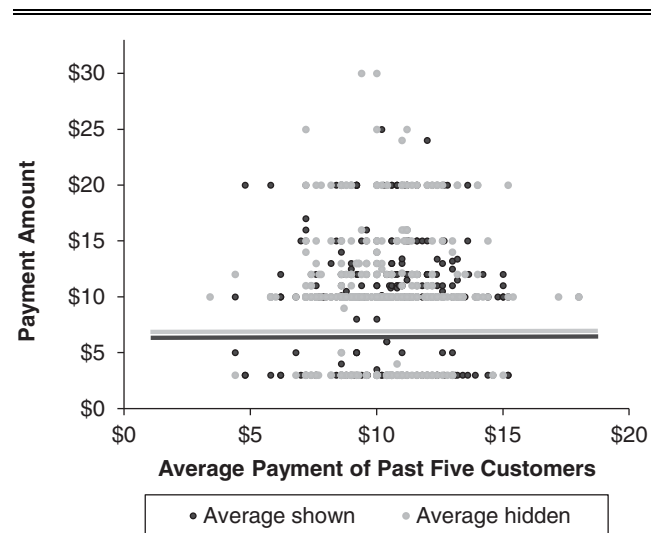
Results. We submitted the average payment per person per group to a 4 (suggested donation amount: no information, $.50, $1, or $2) × 2 (average donation amount: no information or $1) ANOVA. The main effects of both the suggested donation amount (F(3, 949) = .56, *p* = .641) and the average donation amount (F(1, 949) = 3.05, *p* = .081) were not significant. Neither 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) than a real amount a customer paid. Therefore, in Study 9, instead of presenting participants with an average price, we showed them a plausible amount that we claimed the previous single customer had paid. We predicted, then, that customers who saw higher previous payments would pay more than those who saw lower payments.

Method. For Study 9, we returned to the online PWYW media retailer during a promotion for a book bundle that had a minimum price of $3, a bonus price of $10, and a default price of $15. Visitors to the site (N = 1,175) were randomly assigned to see "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) and after learning that the previous customer had paid $12 (M = $11.17; t(769) = .27, *p* = .785).

Study 10: Anchoring with Real Retail Prices

One possible reason for the difficulty in Studies 7–9 is that the anchors were too "pushy" and thus had the effect of

encouraging participants to discount them. Accordingly, in Study 10, we based one anchor on an explicitly seller-derived fact: the product'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 media retailer. The minimum price for the promotion was $1, and the bonus price was $10. Visitors to the site were randomly assigned to see either a $14 or a $28.88 default price. In addition, half of the visitors saw the message "Full Retail Value: $28.88!" above the payment slider. (Due to a programming error, this second manipulation began on the second day of the promotion. For the first day, participants were assigned to only one of two default conditions, 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 the manipulation. Regardless, there were no significant effects of anchoring. On the first day, anchors did not affect payments ($t(683) = .02$, $p = .991$). Although this null result persisted for the rest of the promotion, payments did trend marginally in the direction of anchoring; after the first day, customers in the $14 condition paid less (M = $10.68) than those in the $28.88 condition (M = $11.26; $t(1,503) = 1.91$, $p = .057$). The presence of the $28.88 retail price neither produced a main effect by itself ($F(1, 1,501) = .72$, $p = .395$) nor interacted with the default manipulation ($F(1, 1,501) = .01$, $p = .92$). 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, that effect might emerge when the anchor is paired with a hint of persuasion. In Study 11, we presented people with justified reference prices in addition to suggested payments that were set lower than the reference prices.

Method. For Study 11 (N = 429 groups, 1,234 individuals), we returned to the children's museum from Study 8 and conducted a 2 (regular admission fee [$11]: present or absent) × 2 (suggested donation amount [$5]: present or absent) between-participants study. For example, participants in the regular fee–present and suggested fee–present condition saw, "Your admission has been sponsored by ScholarShare.[10] 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) × 2 (suggested donation amount [$5]: present or

---

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

absent) ANOVA showed that neither variable influenced payments significantly. Visitors donated similar amounts regardless of whether they saw the regular admission fee ($M_{suggestion\ present,\ regular\ fee\ present}$ = $.96 vs. $M_{suggestion\ present,\ regular\ fee\ absent}$ = $1.06; $F(1, 425) = .56$, $p = .456$). Their payments also did not differ depending on whether they were provided with a suggested donation amount, ($M_{suggestion\ present,\ regular\ fee\ present}$ = $.96 vs. $M_{suggestion\ absent,\ regular\ fee\ present}$ = $.77; $F(1, 425) = 1.26$, $p = .263$). Furthermore, the interaction between these variables was not significant ($F(1, 425) = .01$, $p = .909$).

Studies 12a–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 marginally significant effect. One possibility is that the effect 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). Study 12a aimed to test this possibility more thoroughly by investigating whether precise anchors produce 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 (vs. 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, and $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 in the $19.99 condition (M = $12.31) were actually significantly *higher* than those in the $20.01 condition (M = $10.52; $t(296) = 2.18$, $p = .030$). This reverse anchoring effect directly contradicts 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 Study 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 ($F(4, 4,105) = .53$, $p = .731$), including between the $19.99 condition (M = $10.39) and the $20.01 condition (M = $10.60; $t(1,663) = .86$, $p = .39$).

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 of our manipulations could be seen as especially heavy-handed in this respect, past research has suggested that even arbitrary, unexplained default prices are frequently seen this way as well (McKenzie, Liersch, and Finklestein 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 13 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 experiments at the Cartoon Art Museum (as described in Study 6b) on the museum's Pay-What-You-Wish Days in January, April, May, and July 2012. In each of those months, we tested a set of maximum prices that visitors could pay for their admission.

In those four periods, groups of visitors (N = 606 groups, 1,349 individuals) were assigned to a maximum price condition ($10, $50, or $100). 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 [$10/$50/$100] 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 (for details, see Web Appendix).

Results. Although anchoring effects appeared sporadically across the museum's four PWYW days (for a full discussion of the results, see Web Appendix), 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 among the different anchor conditions, except between the two highest anchor conditions, $50 and $100 (see Table 1). We are disinclined to read too much into the unpredicted pattern between the $50 and $100 anchors because 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. We predict that this pattern would not replicate under better conditions; however, future studies could investigate it.

Discussion

Without the perfectly sized gap discussed in the previous sections, payments seemed generally insensitive to the anchor's size (low anchors in Study 8, high anchors in Studies 12a–b), whether and how it was justified (Studies 7–11), its level of precision (Studies 12a–b), 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 studies to the financial, field component, it is prudent that we consider that the studies could simply have been poorly designed or conducted 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, our null results are due to the inclusion of real money, then we should see significant effects when using hypothetical versions of the experiment. Therefore, in the following section, we choose a representative study from each of the preceding three sections—examining different types of anchor gaps, asymmetries in anchor 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

In Study 14a, we attempted to replicate Study 10, which took place on the PWYW media retailer's website with a high default price of $28.88 and a low default price of $14.00. In addition, because customers at this site 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 responsible for 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 with 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 we informed them 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 price setting 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 vs. $28.88) × 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. $M_{text\ box}$ = $13.35; $F(1, 382)$ = 3.21, p = .074), but, critically, response format 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 = .28). Although this distribution resembled that of 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 saw in Study 10. 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 = .41). Importantly, though, response format was even further from statistically significant (p = .543).

Study 14b

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 price 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 (items that would have been marked as bonuses 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 $9) × 2 (bonus price: present or absent) ANOVA on payments, which yielded null effects for both bonus presence ($F(1, 577) = .00$, $p = .999$) and its interaction with default price ($F(1, 577) = .27$, $p = .602$). Replicating our results in Study 3a, participants in this study paid less after seeing a $3 default ($M = \$5.10$) than after seeing a $9 default ($M = \$5.99$; $F(1, 577) = 10.70$, $p = .001$; $d = .28$).

Study 14c

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, as in Study 14a, we would find significant effects of anchors 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!" followed by "$1 or Pay What You Want," "$3 or Pay What You Want," or simply "Pay What You Want," 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. These manipulations were not significant and did not interact with the anchor manipulation, and they are not discussed further.

Results. Because we nearly never encountered payments greater than $5 in our field data, we winsorized payments at $5. A one-way ANOVA on payments showed significant differences across anchor conditions ($F(2, 176) = 6.88$, $p = .001$). In line with our predictions, participants in the $1 condition paid significantly less ($M = \$.82$) than 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$; $t(106) = .88$, $p = .382$); however, payments in $1 condition were significantly lower than in the control condition ($t(111) = 2.46$, $p = .015$).

Study 14d

Having cast doubt, in Study 14a–c, on the possibility that the setting of our anchoring studies or the way we elicited estimates caused our previous null results, in Study 14d, we wanted to investigate further into the role of the anchor gap. That is, we had seen small gaps fail to have an effect (e.g.,

Study 4) and large gaps succeed (e.g., Study 1), but we 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 because we believed this to be the largest sample we could recruit in a reasonable time frame. We used images of the signs used in Study 2, with the varying part reading "Pay [nothing/$.01/$.25/$1.75/$50] or Pay What You Want," depending on the condition. We also included a control condition that did not have an anchor. The $.25, $1.75, and control conditions were adopted from the default manipulation in Study 2. After providing their estimates of what they would pay, 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 with 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 that PWYW could essentially generate an anchor of $0, which could mute anchoring effects.

Results. We preregistered the sample size, materials, exclusions, and analyses with the Open Science Framework, available at https://osf.io/qfjht/. By that preregistered 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, 2,032) = 102.25$, $p < .001$). We replicated Study 2's anchoring effect: participants who saw the $.25 anchor paid less ($M = \$.39$) than participants who saw the $1.75 anchor ($M = \$1.09$; $t(684) = 16.30$, $p < .001$). Furthermore, anchoring effects were not capped at the very high value of $1.75; payments in the $50 condition were even higher ($M = \$1.52$; $t(661) = 5.44$, $p < .001$).

There were some interesting differences at the lowest anchors: the $.25 anchor produced a somewhat *lower* mean payment ($M = \$.39$) than the $.01 anchor ($M = \$.50$; $t(739) = 2.93$, $p = .004$), and a *substantially* lower payment than observed in the "pay nothing or PWYW" condition ($M = \$.92$; $t(693) = 10.56$, $p < .001$).

Although we preregistered the winsorized specification, it is worth noting that all but one difference was significant when we analyzed the untransformed data. That alternative specification is reported in the Web 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 14b) also showed us that the peculiarities of one of our study sites likely did not suppress anchoring effects. In addition, in Studies 14a and 14c, we showed significant effects using 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 absolutely and

Figure 3
DISTRIBUTIONAL ANCHOR GAP VERSUS COHEN'S D ACROSS ALL STUDIES
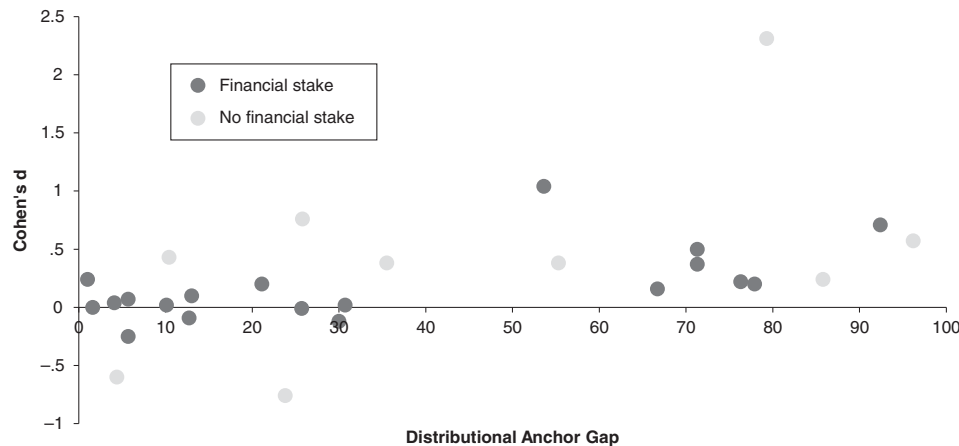
distributionally—and still elicit significant differences in payments.

GENERAL DISCUSSION

Our aims for this study 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, in the present study, 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), more elaborate references to the payments of others (Studies 7 and 9), reference to retail prices (Study 10), 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 (e.g., 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. Next, 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 14c), we found highly significant results. Nevertheless, we think that the inherent noisiness of the field is an unlikely explanation for the observed null effects. In our case, "lab versus 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 that include 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 that 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 might operate as merely slightly high.

Size of Anchor Gap

It could 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 to produce an effect. Moreover, that anchor gap needs to be considered in terms of the anchors' places in the distribution (i.e., their percentile ranks across all payments), rather than in terms of their absolute magnitudes. 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 3. For every anchor gap smaller than 50 percentile points, we observed nonsignificant 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).

Although a focus on something as simple as the distributional anchor gap might seem obvious, it has been missed by previous researchers who tested anchoring effects in the field. For example, successful instances of anchoring in the field (e.g., Croson and Shang 2008) have 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 distributional gap, with anchors still within the range of consideration. Other attempts, such as Smith and Berger (1996), have 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 could have contributed to their null result. Others (e.g., Alpizar, Carlsson, and Johansson-Stenman 2008) have still failed to account for distribution entirely, choosing anchors on absolute scale alone (e.g., $2, $5, $10—anchors in a distribution with 48% $0 payments and a $2.39 mean).

These three explanations (real vs. hypothetical payments, asymmetries of influence for low and high anchors, and the distributional 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, although seemingly arbitrary (i.e., lacking explicit justification), are often perceived as recommendations (McKenzie, Liersch, and Finklestein 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 was judged as an intentional and informed suggestion, it would be more influential, not less (e.g., Armstrong Soule and Madrigal 2015). 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 nonarbitrary, 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 they 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.

REFERENCES

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman (2008), "Anonymity, Reciprocity, and Conformity: Evidence from Voluntary Contributions to a National Park in Costa Rica," *Journal of Public Economics*, 92 (5), 1047–60.

Ariely, Dan, George Loewenstein, and Drazen Prelec (2003), "'Coherent Arbitrariness': Stable Demand Curves Without Stable Preferences," *Quarterly Journal of Economics*, 118 (1), 73–106.

Armstrong Soule, Catherine A. and Robert Madrigal (2015), "Anchors and Norms in Anonymous Pay-What-You-Want Pricing Contexts," *Journal of Behavioral and Experimental Economics*, 57 (August), 167–75.

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.

Brewer, Noel T. and Gretchen B. Chapman (2002), "The Fragile Basic Anchoring Effect," *Journal of Behavioral Decision Making*, 15 (1), 65–77.

Carson, Timothy N. and Charles R. Plott (2012), "Misconceptions and Game Form Recognition of the BDM Method: Challenges to Theories of Revealed Preference and Framing," working paper, Purdue University.

Chapman, Gretchen B. and Eric J. Johnson (2002), "Incorporating the Irrelevant: Anchors in Judgments of Belief and Value," in Heuristics and Biases: The Psychology of Intuitive Judgment, Thomas Gilovich, Dale Griffin, and Daniel Kahneman, eds. Cambridge, UK: Cambridge University Press, 120–28.

Critcher, Clayton R. and Thomas Gilovich (2008), "Incidental Environmental Anchors," *Journal of Behavioral Decision Making*, 21 (3), 241–51.

Croson, Rachel and Jen Yue Shang (2008), "The Impact of Downward Social Information on Contribution Decisions," *Experimental Economics*, 11 (3), 221–33.

Danilowitz, Jennifer, Shane W. Frederick, and Daniel Mochon, "Anchoring as Inference," working paper, Yale School of Management, Yale University.

Desmet, Pierre and Fred M. Feinberg (2003), "Ask and Ye Shall Receive: The Effect of the Appeals Scale on Consumers' Donation Behavior," *Journal of Economic Psychology*, 24 (3), 349–76.

Epley, Nicholas and Thomas Gilovich (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.

Frederick, Shane W. and Daniel Mochon (2012), "A Scale Distortion Theory of Anchoring," *Journal of Experimental Psychology: General*, 141 (1), 124–33.

Gautier, Pieter A. and Bas van der Klaauw (2012), "Selection in a Field Experiment with Voluntary Participation," *Journal of Applied Econometrics*, 27 (1), 63–84.

Gneezy, Ayelet, Uri Gneezy, Gerhard Riener, and Leif D. Nelson (2012), "Pay-What-You-Want, Identity, and Self-Signaling in Markets," *Proceedings of the National Academy of Sciences of the United States of America*, 109 (19), 7236–40.

Hardesty, David M. and Tracy A. Suter (2013), "Maximizing Willingness to Bid Within 'Buy It Now' Auctions," *Journal of Business Research*, 66 (4), 554–58.

Jacowitz, Karen E. and Daniel Kahneman (1995), "Measures of Anchoring in Estimation Tasks," *Personality and Social Psychology Bulletin*, 21 (11), 1161–66.

Janiszewski, Chris and Dan Uy (2008), "Precision of the Anchor Influences the Amount of Adjustment," *Psychological Science*, 19 (2), 121–27.

Jung, Minah H., Leif D. Nelson, Ayelet Gneezy, and Uri Gneezy (2014), "Paying More When Paying for Others: Consumer Elective Pricing with Pay-It-Forward Framing," *Journal of Personality and Social Psychology*, 107(3), 414–31.

Kamins, Michael A., Xavier Dreze, and Valerie S. Folkes (2004), "Effects of Seller-Supplied Prices on Buyers' Product Evaluations: Reference Prices in an Internet Auction Context," *Journal of Consumer Research*, 30 (4), 622–28.

Kim, Ju-Young, Katharina Kaufmann, and Manuel Stegemann (2013), "The Impact of Buyer–Seller Relationships and Reference Prices on the Effectiveness of the Pay What You Want Pricing Mechanism," *Marketing Letters*, 25 (4), 409–42.

———, Martin Natter, and Martin Spann (2009), "Pay What You Want: A New Participative Pricing Mechanism," *Journal of Marketing*, 73 (January), 44–58.

León, Francisco J., José A. Noguera, and Jordi Tena-Sánchez (2012), "How Much Would You Like to Pay? Trust, Reciprocity, and Prosocial Motivations in El Trato," *Social Sciences Information*, 5 (3), 389–417.

Maniadis, Zacharias, Fabio Tufano, and John A. List (2014), "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects," *American Economic Review*, 104 (1), 277–90.

Martin, Richard and John Randal (2008), "How Is Donation Behaviour Affected by the Donations of Others?" *Journal of Economic Behavior & Organization*, 67 (1), 228–38.

Mason, Malia F., Alice J. Lee, Elizabeth A. Wiley, and Daniel R. Ames (2013), "Precise Offers are Potent Anchors: Conciliatory Counteroffers and Attributions of Knowledge in Negotiations," *Journal of Experimental Social Psychology*, 49 (4), 759–63.

McKenzie, Craig R.M., Michael J. Liersch, and Stacey R. Finkelstein (2006), "Recommendations Implicit in Policy Defaults," *Psychological Science*, 17 (5), 414–20.

Mochon, Daniel and Shane Frederick (2013), "Anchoring in Sequential Judgments," *Organizational Behavior and Human Decision Processes*, 122 (1), 69–79.

Mussweiler, Thomas and Fritz Strack (2001), "Considering the Impossible: Explaining the Effects of Implausible Anchors," *Social Cognition*, 19 (2), 145–60.

Nunes, Joseph C. and Peter Boatwright (2004), "Incidental Prices and Their Effect on Willingness to Pay," *Journal of Marketing Research*, 41 (November), 457–66.

Regner, Tobias and Javier A. Barria (2009), "Do Consumers Pay Voluntarily? The Case of Online Music," *Journal of Economic Behavior & Organization*, 71 (2), 395–406.

Riener, Gerhard and Christian Traxler (2012), "Norms, Moods, and Free Lunch: Longitudinal Evidence on Payments from a Pay-What-You-Want Restaurant," *Journal of Socio-Economics*, 41 (4), 476–83.

Shang, Jen Yue and Rachel Croson (2009), "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Good," *Economic Journal*, 119 (540), 1422–39.

Simmons, Joseph P., Robyn A. LeBoeuf, and Leif D. Nelson (2010), "The Effect of Accuracy Motivation on Anchoring and Adjustment: Do People Adjust from Provided Anchors?" *Journal of Personality and Social Psychology*, 99 (6), 917–32.

Simonsohn, Uri and Joseph P. Simmons and Leif D. Nelson (2014), "Anchoring Is Not a False-Positive: Maniadis, Tufano, and List's (2014) 'Failure-to-Replicate' Is Actually Entirely Consistent with the Original," working paper, The Wharton School, University of Pennsylvania.

Simonson, Itamar and Aimee Drolet (2004), "Anchoring Effects on Consumers' Willingness-to-Pay and Willingness-to-Accept," *Journal of Consumer Research*, 31 (3), 681–90.

Small, Deborah A. and George Loewenstein (2003), "Helping a Victim or Helping the Victim: Altruism and Identifiability," *Journal of Risk and Uncertainty*, 26 (1), 5–16.

Smith, Gerald E. and Paul D. Berger (1996), "The Impact of Direct Marketing Appeals on Charitable Marketing Effectiveness," *Journal of the Academy of Marketing Science*, 24 (3), 219–31.

Smith, Sarah, Frank Windmeijer, and Edmund Wright (2014), "Peer Effects in Charitable Giving: Evidence from the (Running) Field," *Economic Journal*, 125 (585), 1053–71.

Spann, Martin, Gerald Häubl, Bernd Skiera, and Martin Bernhardt (2012), "Bid-Elicitation Interfaces and Bidding Behavior in Retail Interactive Pricing," *Journal of Retailing*, 88 (1), 131–44.

Stewart, Neil (2009), "The Cost of Anchoring on Credit-Card Minimum Repayments," *Psychological Science*, 20 (1), 39–41.

Strack, Fritz and Thomas Mussweiler (1997), "Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility," *Journal of Personality and Social Psychology*, 73 (3), 437–46.

Tversky, Amos and Daniel Kahneman (1974), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185 (4157), 1124–31.

Wilson, Timothy D., Christopher E. Houston, Kathryn M. Etling, and Nancy Brekke (1996), "A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents," *Journal of Experimental Psychology: General*, 125 (4), 387–402.

Web Appendix

Anchoring in Payment: Evaluating a Judgmental Heuristic in Field Experimental Settings

Minah H. Jung, Hannah Perfecto, Leif D. Nelson

TABLE OF CONTENTS

1. STUDY MATERIALS

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

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

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

1.2.2 Doughnut Stand at UC Berkeley in Study 2

1.3.1. Screenshot from Study 3a. $3 condition

1.3.2. Screenshot from Study 3a. $20 condition

1.4.1. Screenshot from Study 3b. $8 condition

1.4.2. Screenshot from Study 3b. $50 condition

1.5. Screenshot from Study 2 (Vodo.net), $12 condition
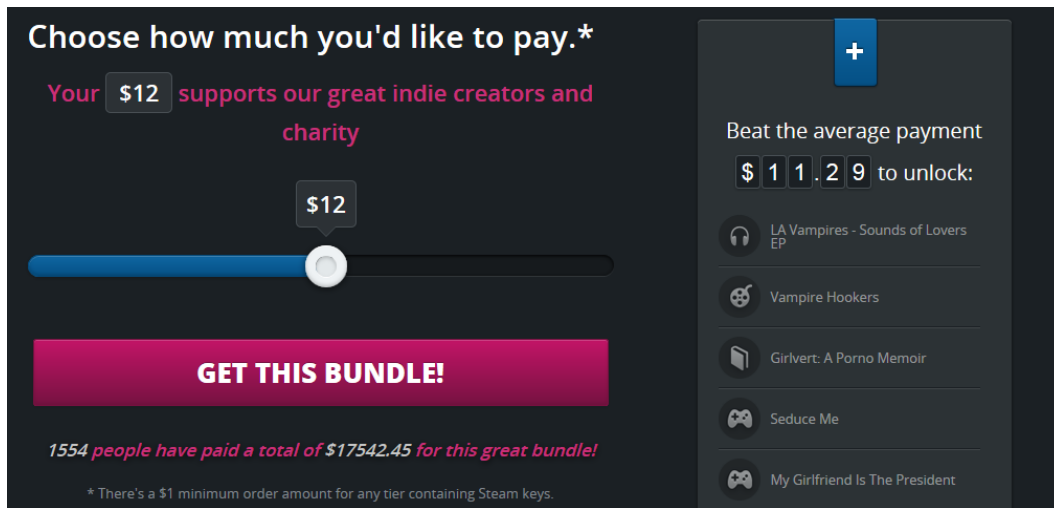
1.6. Doughnut Stand at UC Berkeley in Study 6a

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

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

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

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

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

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

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

Bay Area Discovery Museum

Your admission has been sponsored by ScholarShare. 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 you like to give?

$_____

2. HISTOGRAMS OF PAYMENT

2.1. Payment Distribution in Study 1
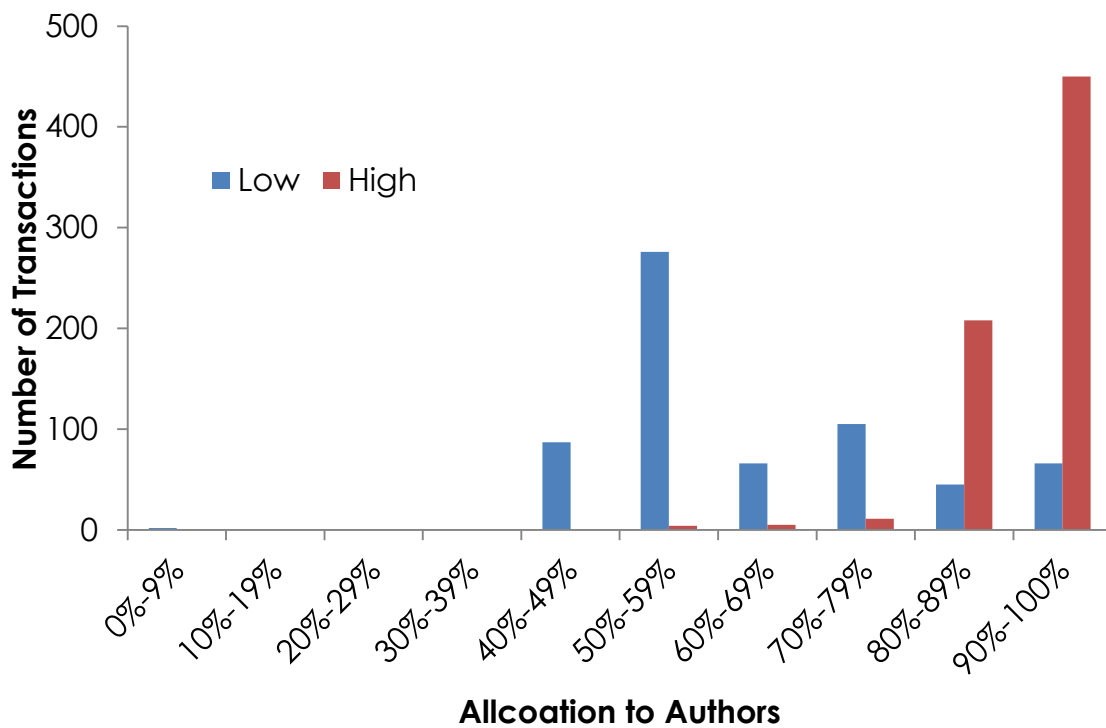
2.2. Payment Distribution in Study 2
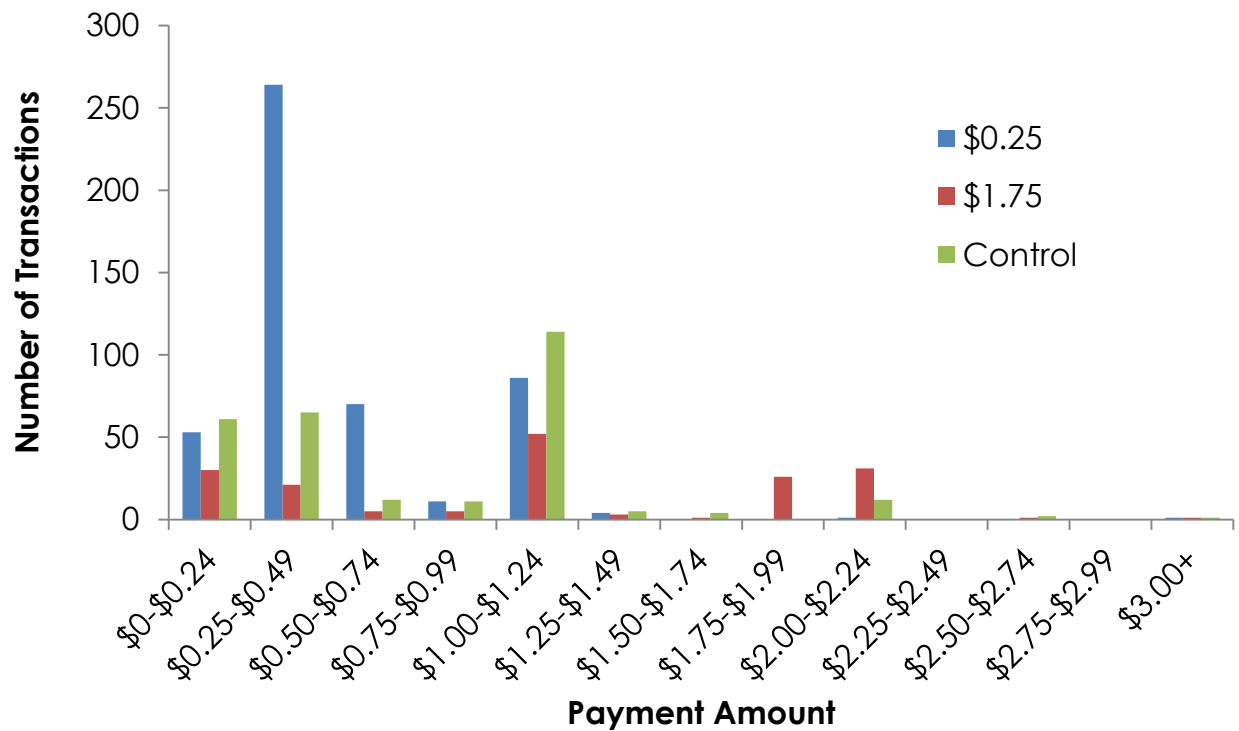
2.3.1 Payment Distribution in Study 3a
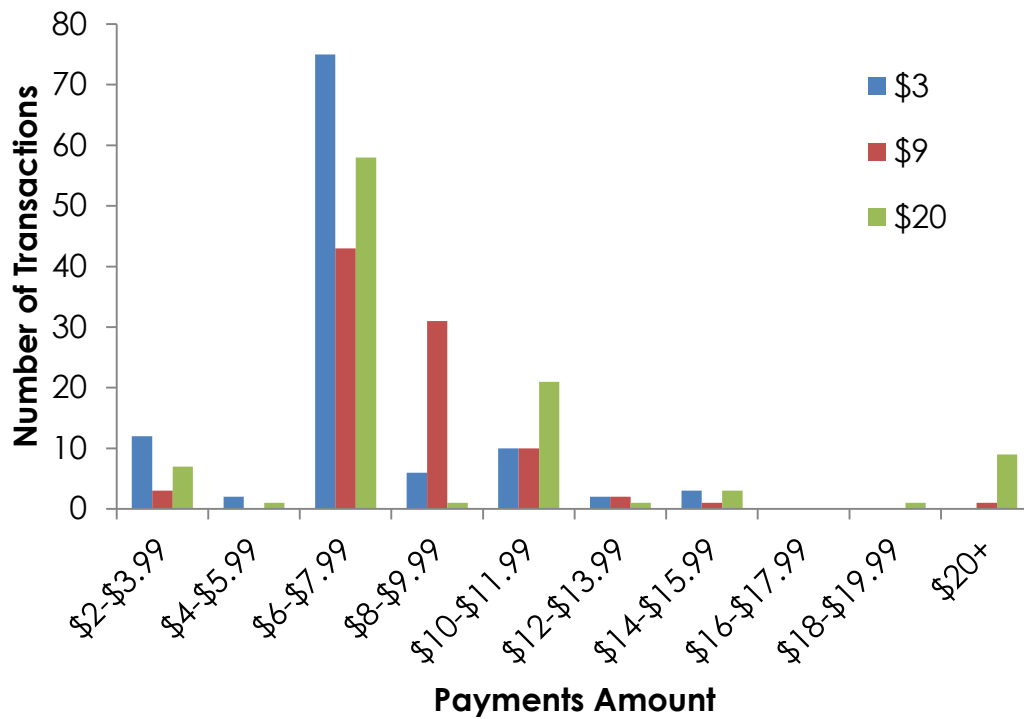
2.3.2. Payment Distribution in Study 3b
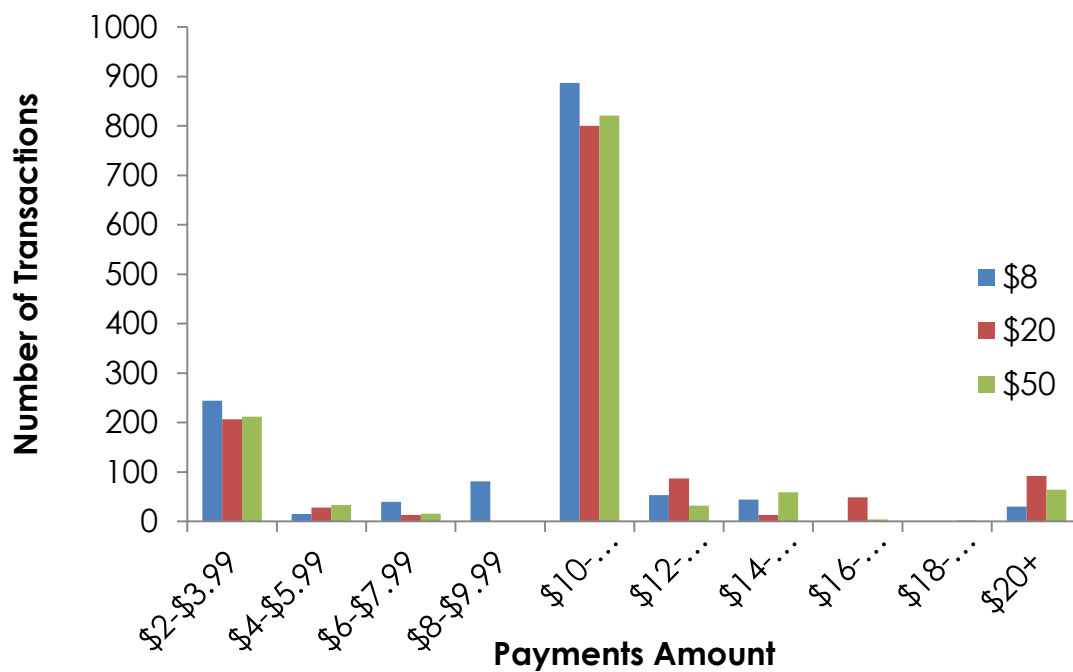
2.4. Payment Distribution in Study 4
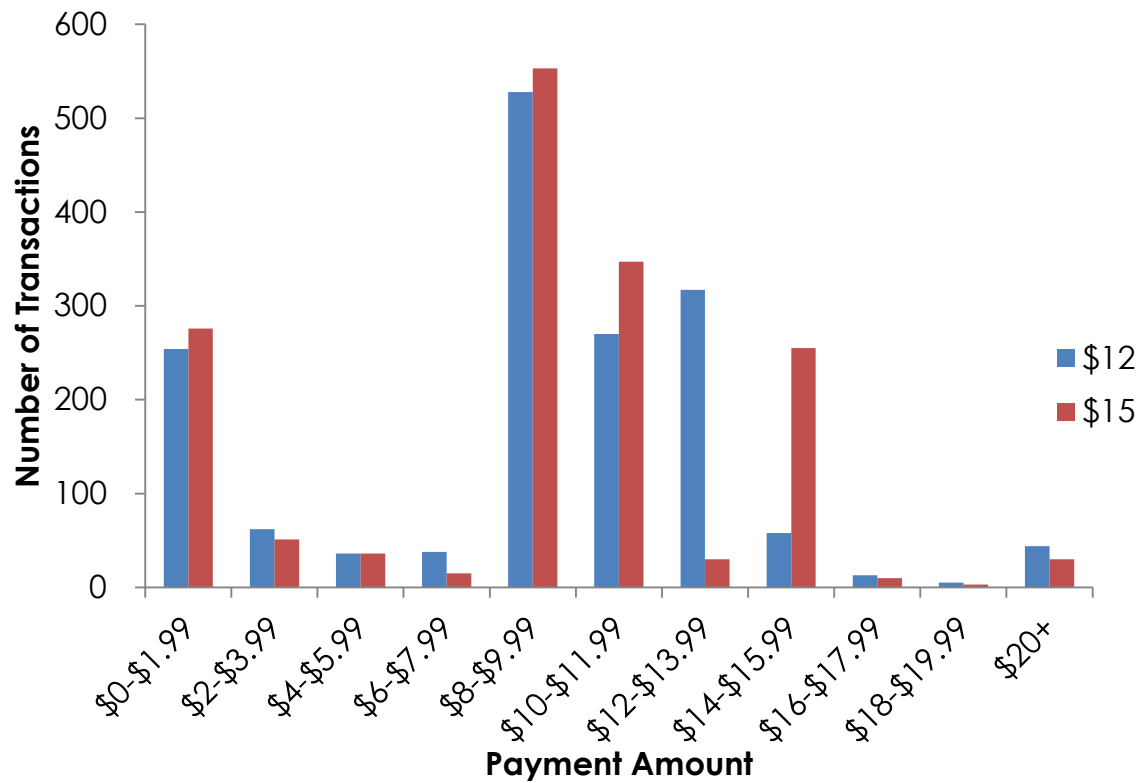
2.5. Payment Distribution in Study 5
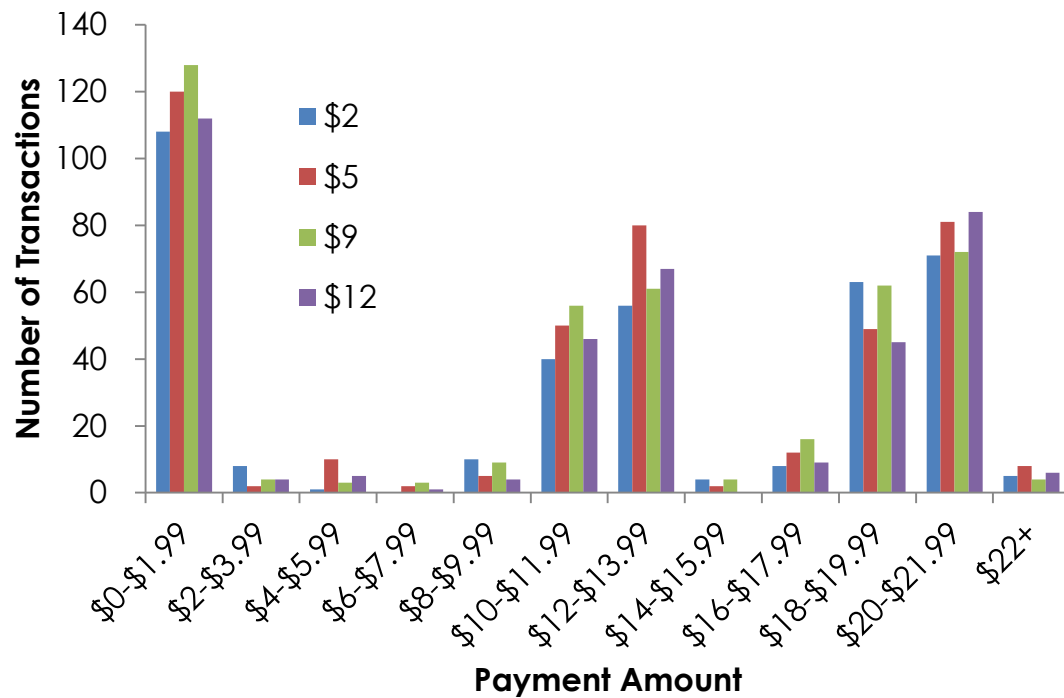
2.6.1. Payment Distribution in Study 6a
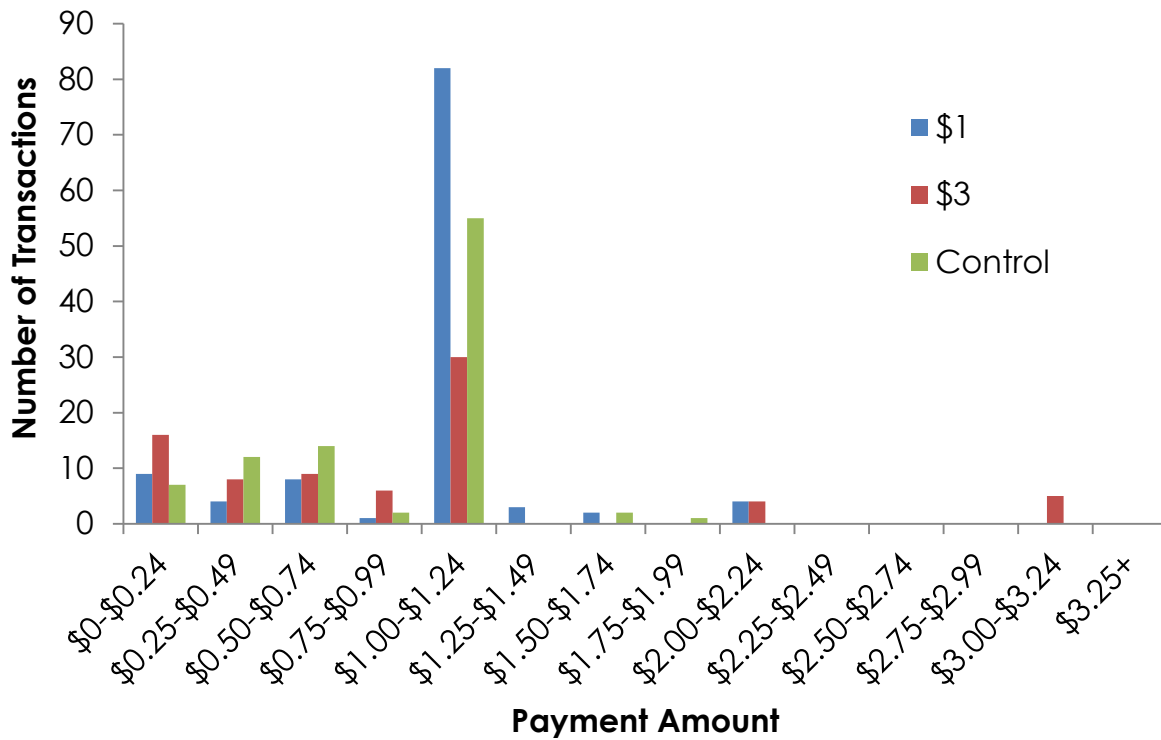
2.6.2. Payment Distribution in Study 6b
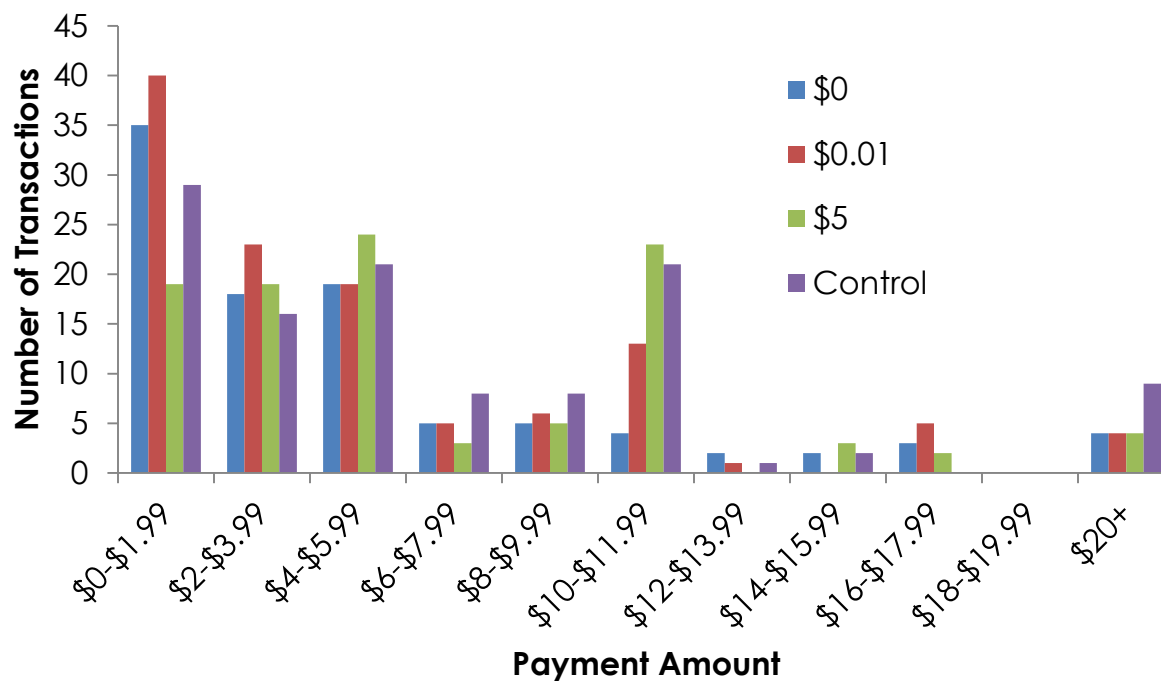
Study 2.7. Payment Distribution in Study 7
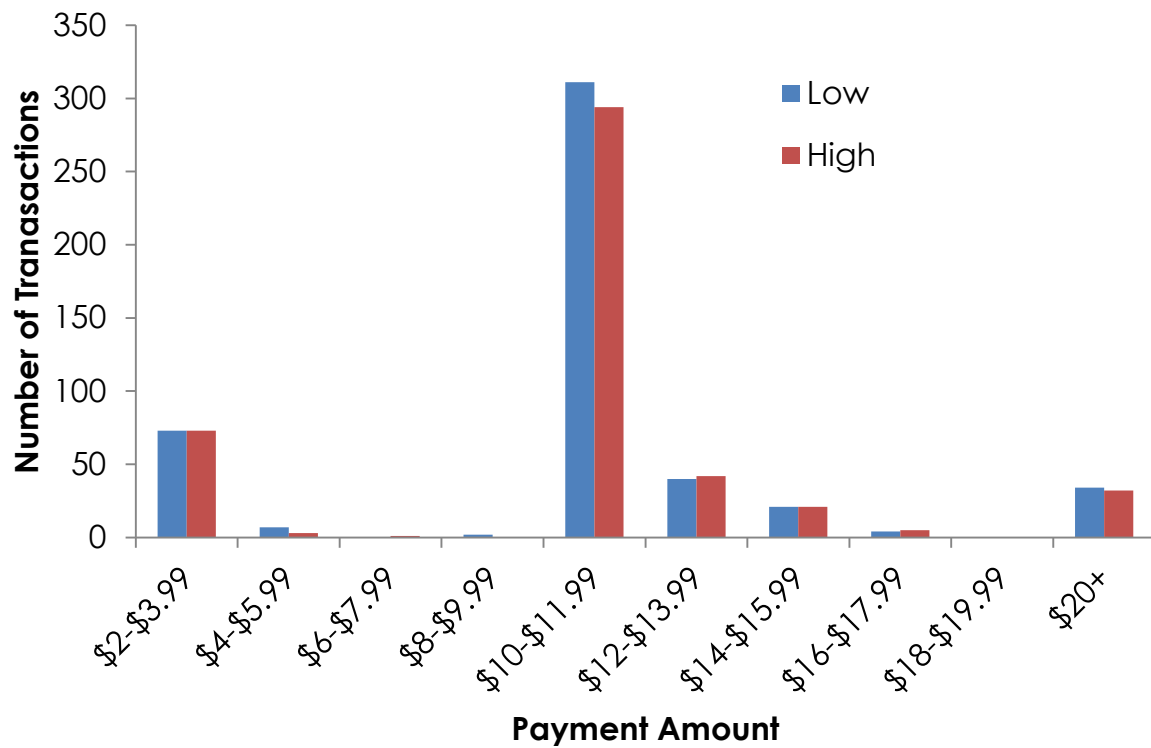
2.8. Payment Distribution in Study 8
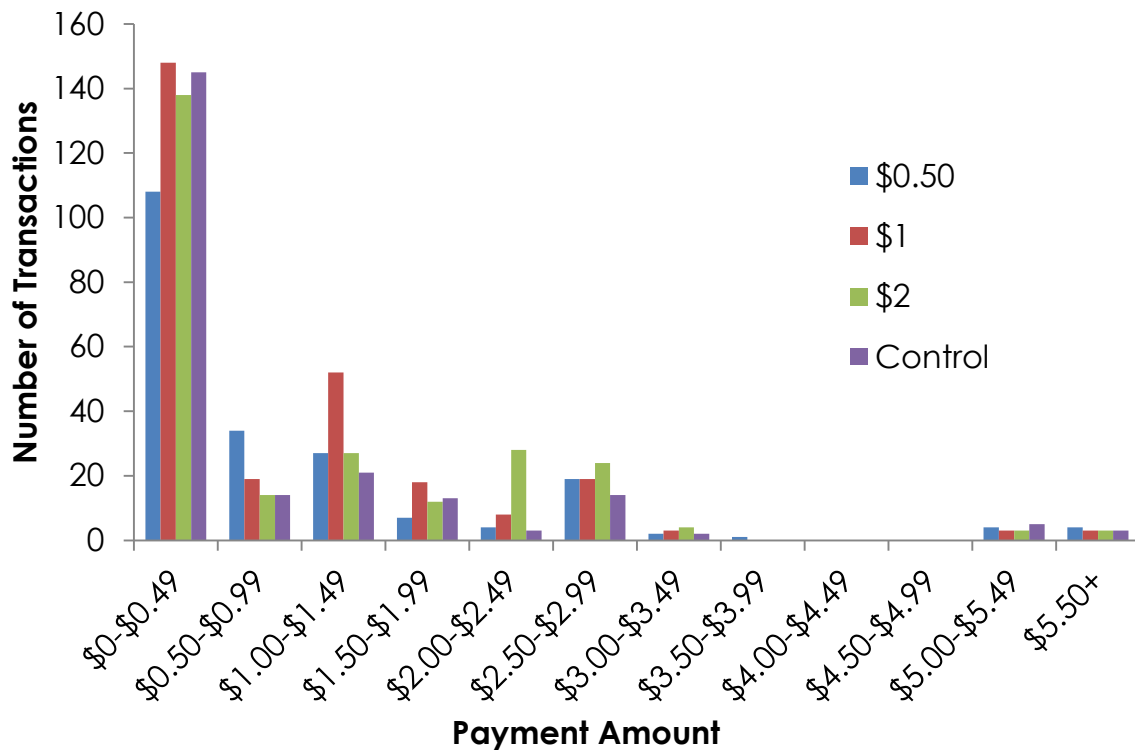
2.9. Payment Distribution in Study 9
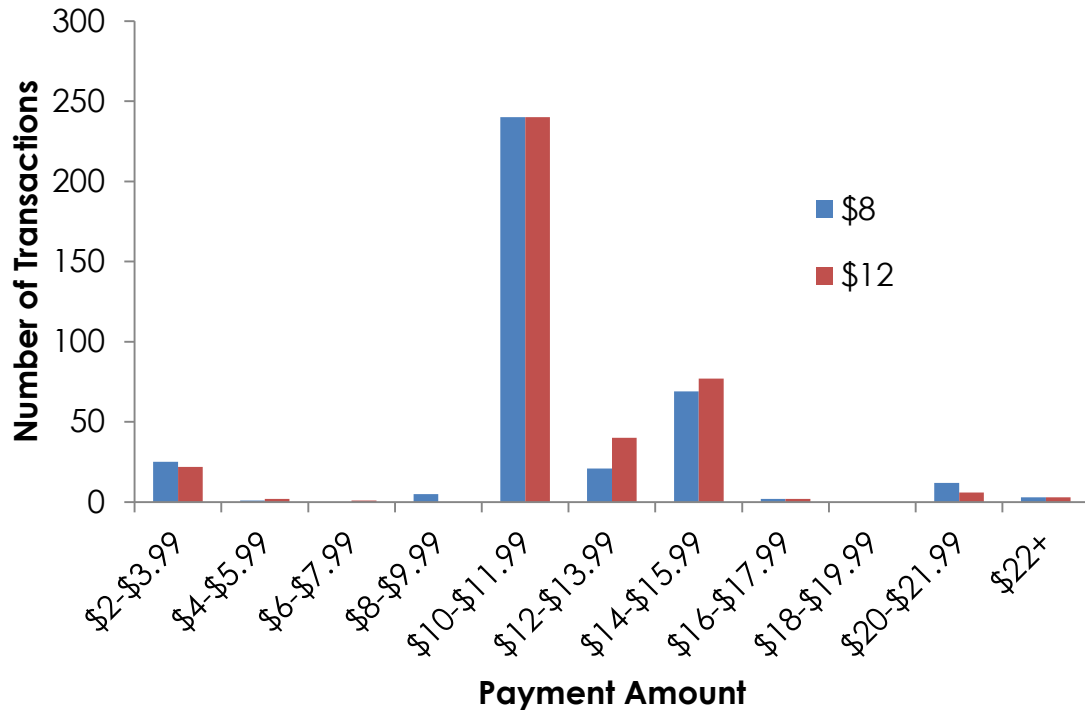
2.10. Payment Distribution in Study 10
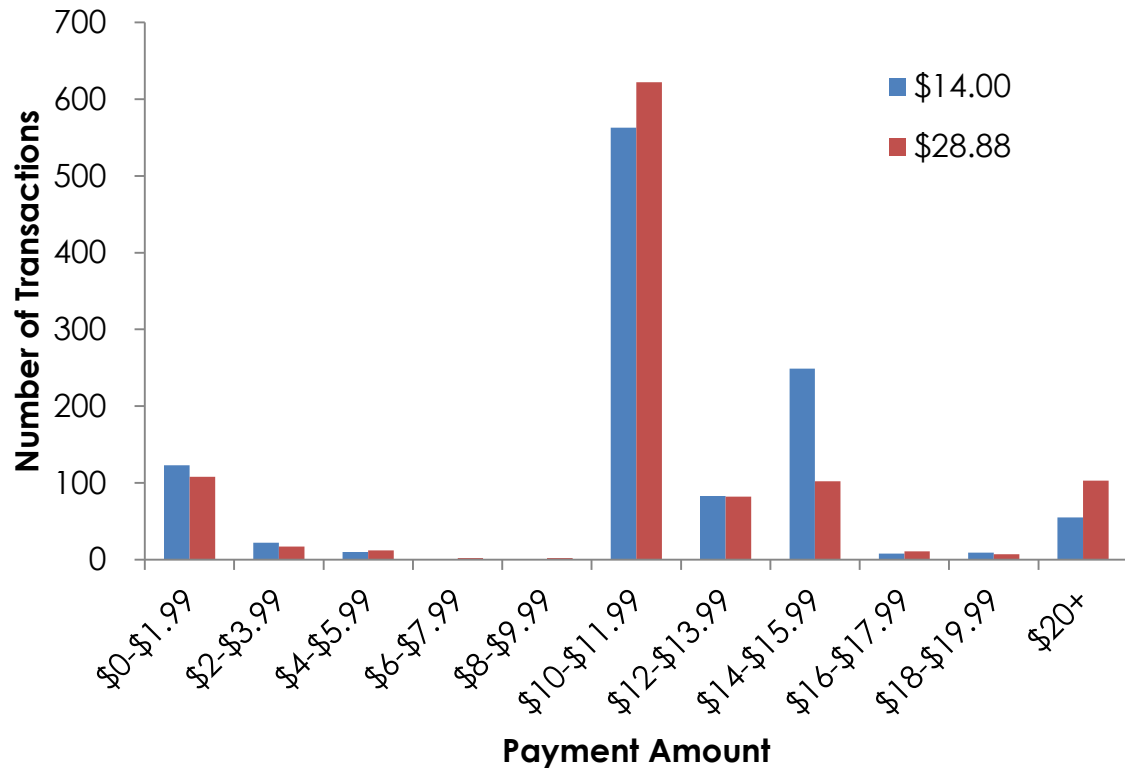
2.11. Payment Distribution in Study 11
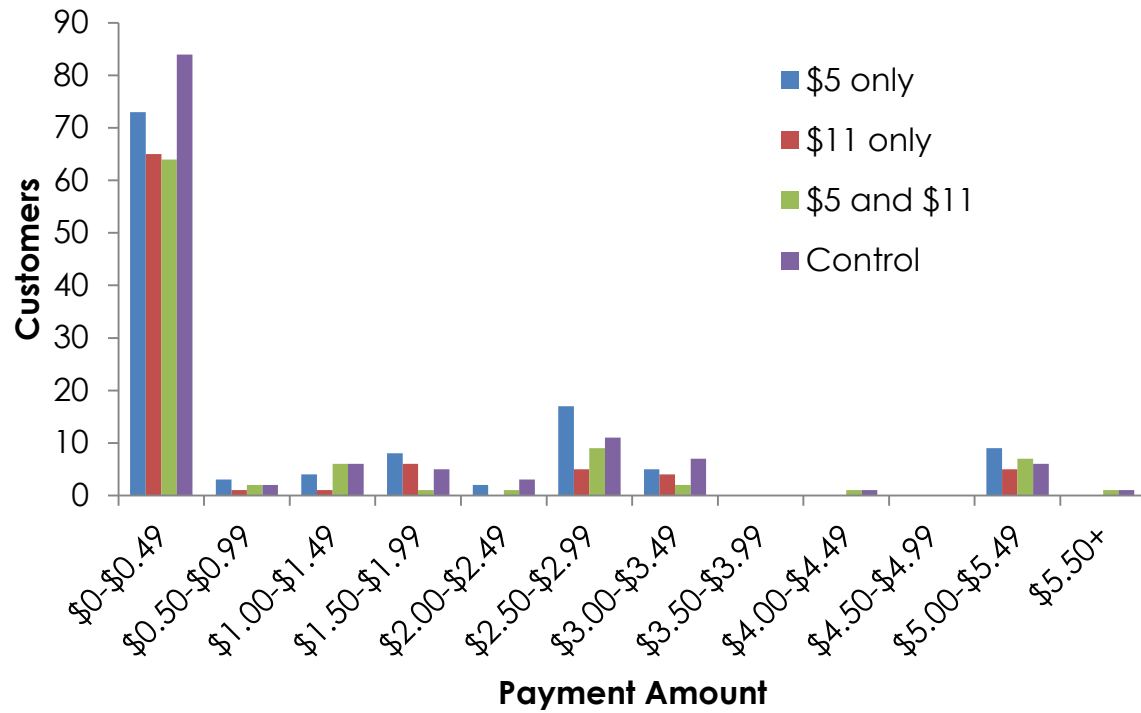
2.12.1. Payment Distribution in Study 12a
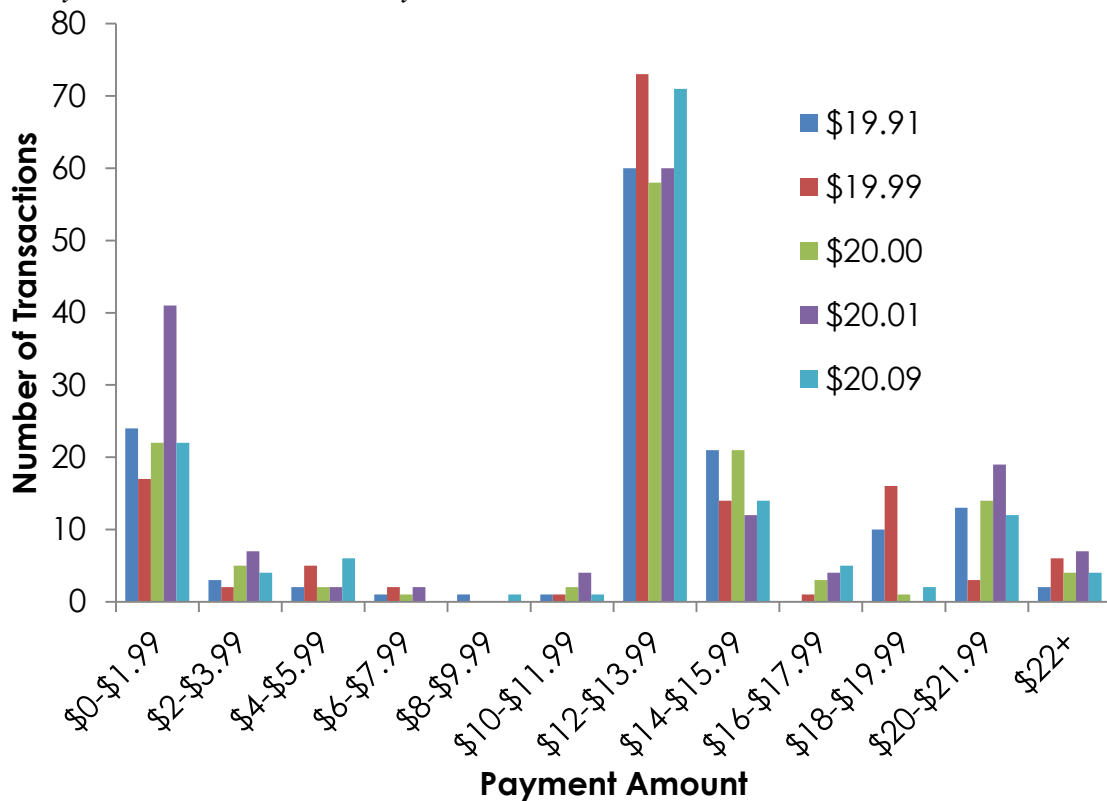
2.12.2. Payment Distribution in Study 12b
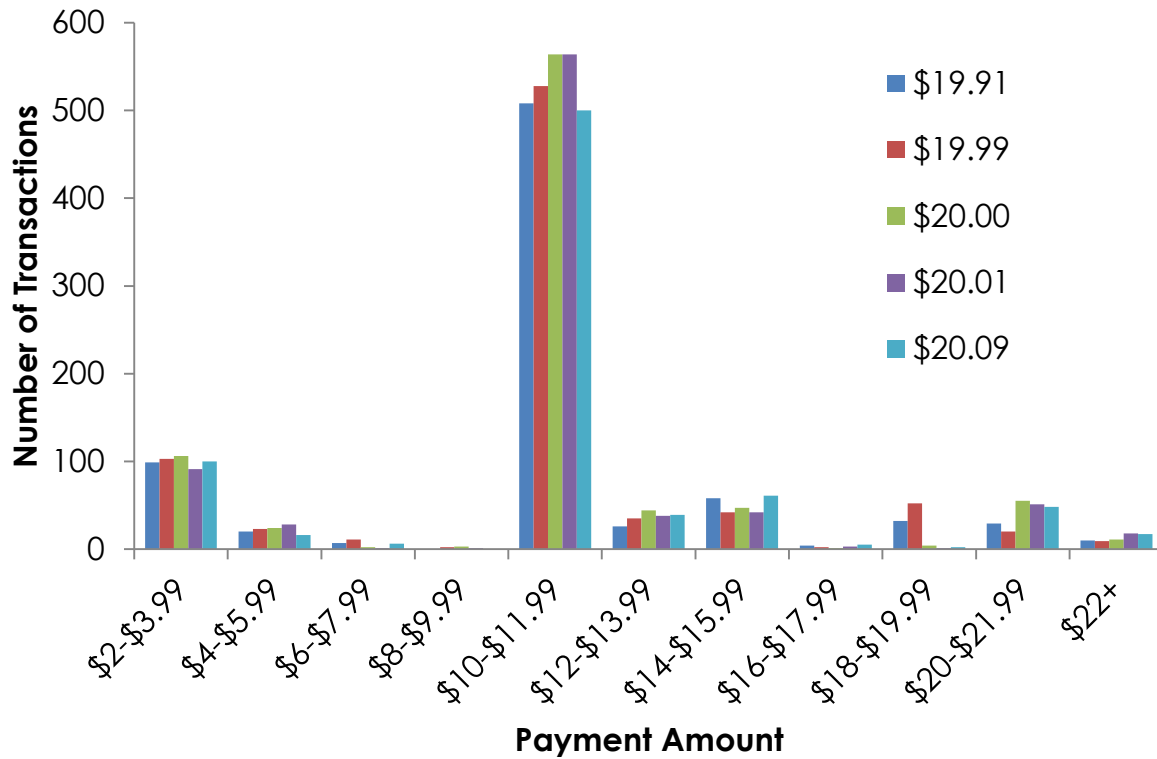
2.13. Payment Distribution in Study 13
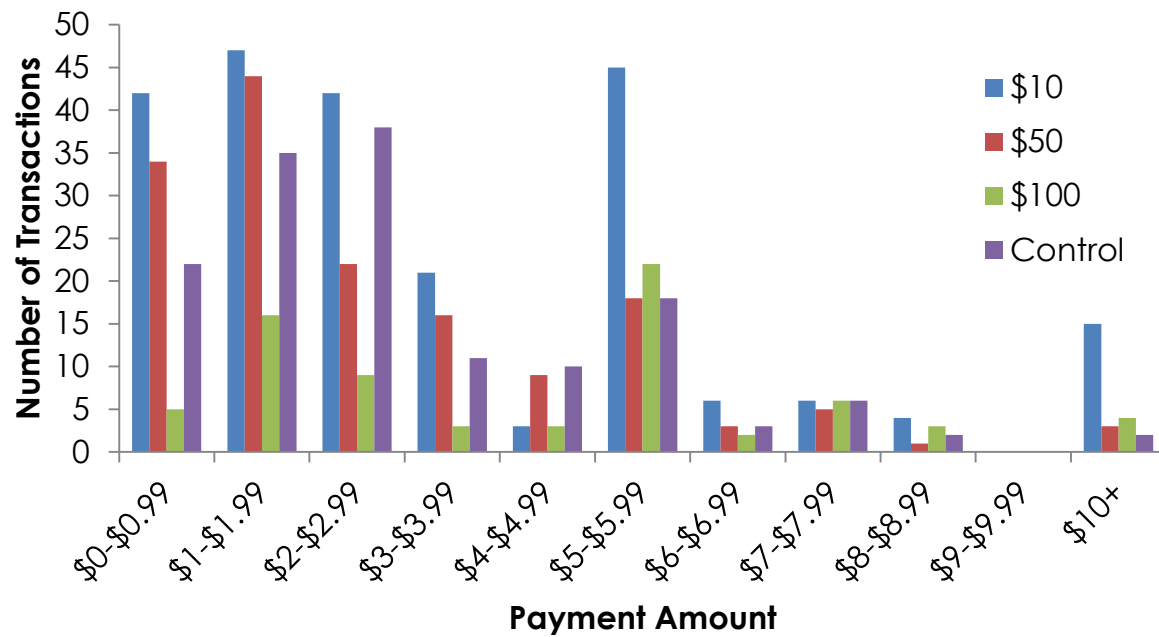
Note: As discussed in the main text, different sets of conditions were run in different months, hence not all conditions have equal sample sizes (e.g., many more overall were in the $10 vs. $100 condition).

2.14.1. Payment Distribution in Study 14a
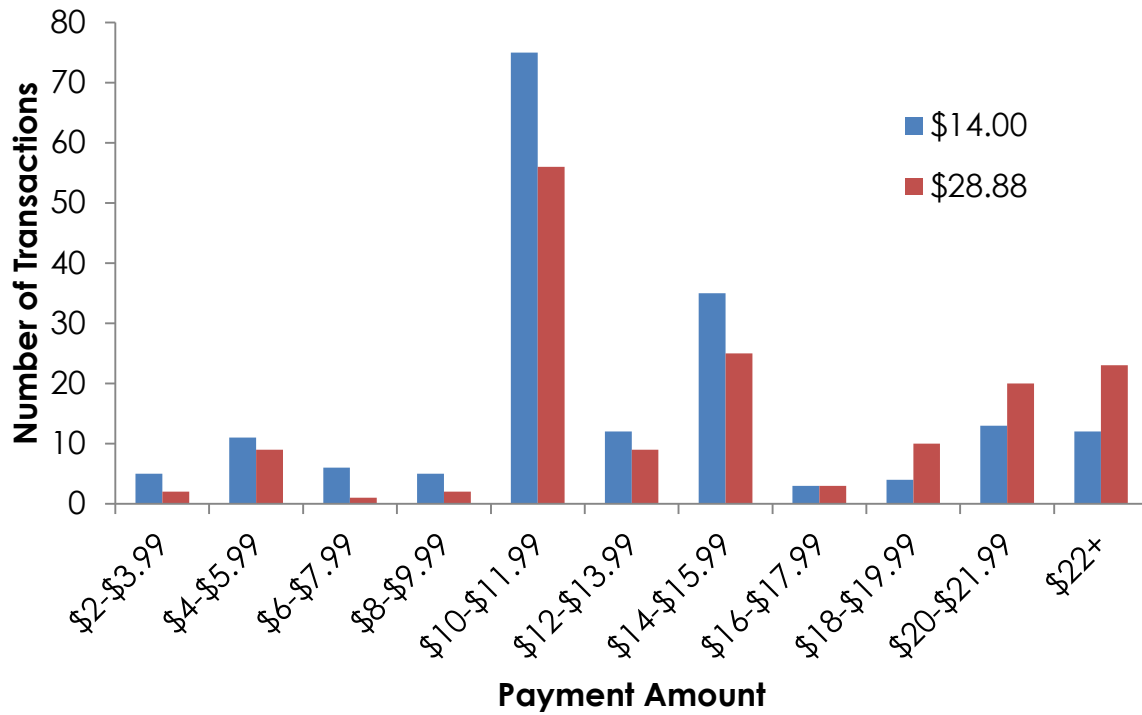
2.14.2. Payment Distribution in Study 14b
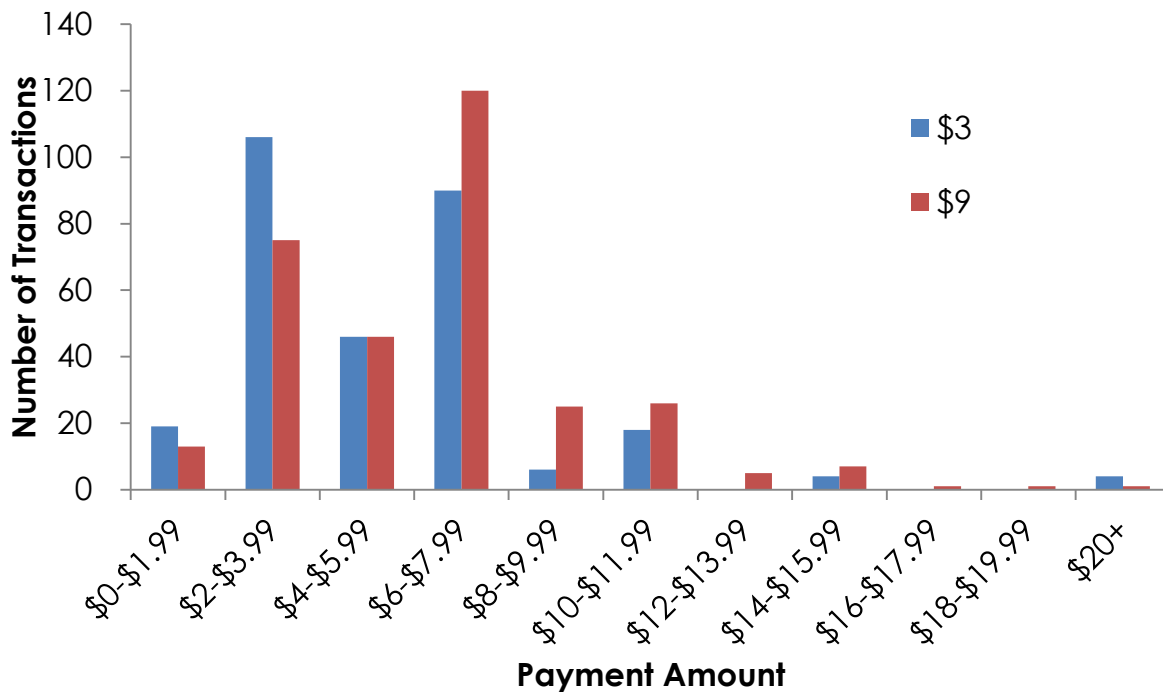
2.14.3. Payment Distribution in Study 14c
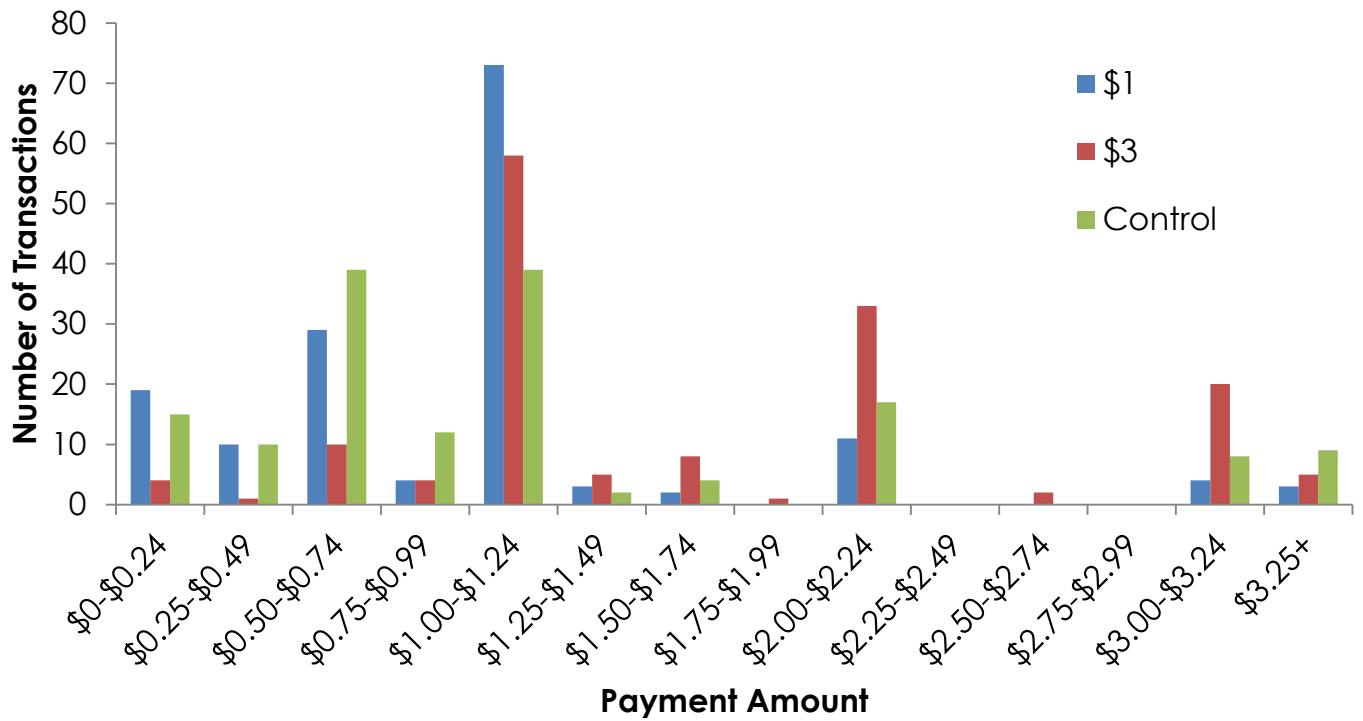
2.14.4. Payment Distribution in Study 14d
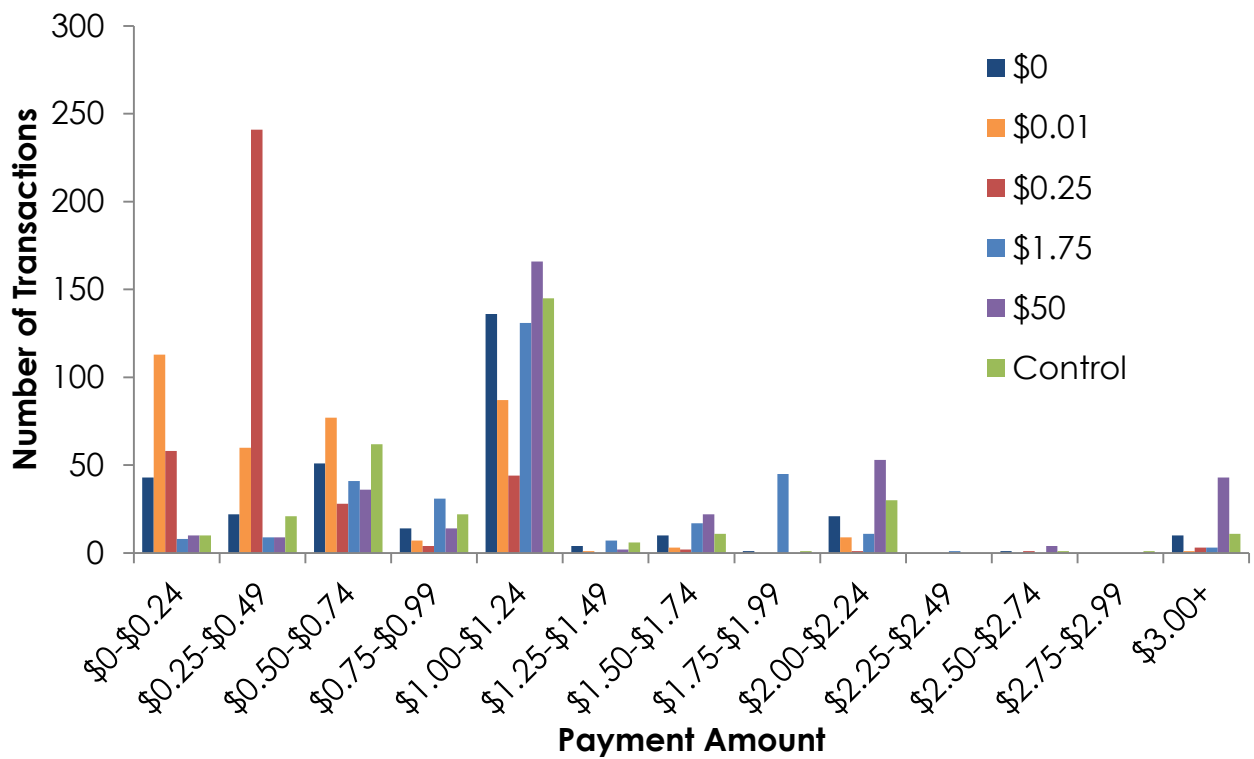
3. ADDITIONAL ANALYSES

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]	$\chi^2=646.66, p<.001$
		$1.75	181	14,548	1.24%[b]	
		No anchor	338	28,073	1.20%[b]	
Study 3b	Media bundle	$8	1,443	19,663	6.8%[a]	$\chi^2=16.47, p<.001$
		$20	1,291	19,739	6.1%[b]	
		$50	1,244	19,737	5.9%[b]	
Study 4	Media bundle	$12	1,602	18,634	7.9%[a]	$\chi^2=.01, p=.937$
		$15	1,612	18,805	7.9%[a]	
Study 6a	Doughnuts	$1	113	9,128	1.24%[a]	$\chi^2=7.67, p=.022$
		$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]	$\chi^2=.50, p=.778$
	bundle	$12	393	4,130	8.7%[a]	
		No anchor	407	4,202	8.8%[a]	
Study 10	Media bundle	$14	790	9,371	7.8%[a]	$\chi^2=3.73, p=.053$
		$28.88	715	9,407	7.1%[a]	
Study 12a	Media bundle	$19.99	140	5,795	2.4%[a]	$\chi^2=.93, p=.334$
		$20.01	158	5,838	2.6%[a]	
Study 12b	Media bundle	$19.99	827	12,998	6.0%[a]	$\chi^2=.00, p=.998$
		$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 ($p < .05$).

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	Doughnuts	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]	$8.71[a]	2.10[a]	1,331	$7.65[a]	1,076	$8.81[a]

22

				($4.88)	($4.45)	(0.70)		($4.47)	($5.97)
5	VODO 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)	

Study	Product	N	Anchor				N		N	
		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)

Study	Item	N	Anchor				N		N	
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.

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	$t(2685)=5.36$, $p<.001$	$t(2685)=4.15$, $p<.001$	$t(2685)=4.43$, $p<.001$	$t(2555)=3.67$, $p<.001$	$t(1090)=5.03$ $p<.001$
		$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.

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	

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	

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	$0.01	$0.25	$1.75	$50	No anchor
$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 Anchors	$0	$0.01	$0.25	$1.75	$50	No 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	$t(556)=.09$, $p=.929$	$t(520)=2.62$, $p=.024$	$t(428)=.67$, $p=.501$	$t(543)=1.45$, $p=.149$	$t(643)=4.49$, $p<.001$	

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.