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

Minah H. Jung[1], Hannah Perfecto[2], Leif D. Nelson[2]

[1]New York University

[2]University of California, Berkeley

Author Note

The first and second authors contributed equally to this research. The first author was supported by the National Science Foundation Graduate Research Fellowship and the Greater Good Dissertation Fellowship. The second author is supported by the University of California, Berkeley Chancellor's Fellowship. The last author was supported by the Barbara and Gerson Bakar Faculty Fellowship. Both of the first and last authors were supported by a grant from the Garwood Center for Corporate Innovation. Correspondence concerning this article should be addressed to Minah H. Jung, the Leonard N. Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012, minah.jung@stern.nyu.edu.

Abstract

Anchoring, the biasing of estimates towards 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: 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 their absolute gap on a number line (Studies 3a-5). Second, low anchors influence payments more than high anchors (Studies 6a-6b). Third, findings from the literature that should enhance anchoring effects—anchor precision, descriptive and injunctive norms, non-suggestions—yield null results in payment (Studies 7-13). Finally, the above patterns do not emerge in hypothetical settings (Studies 14a-14d), where anchoring is as big and reliable as the literature has previously suggested.

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

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

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

A Brief Background on Anchoring

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

The giraffe example demonstrates three critical features of an anchoring paradigm. First, is the anchor, which can be arbitrary or devised by the participant. Second, is the deliberate consideration of the anchor before the final estimate. Though deliberation is widely agreed upon to yield the largest anchoring effects (e.g., Brewer and Chapman 2002; Mochon and Frederick 2013), deliberation-free anchors can seemingly also be influential (e.g., Wilson, Houston, Etling, and Brekke 1996; Critcher and Gilovich 2008). Finally, the value of the target judgment should be uncertain (e.g., anchoring is unlikely to influence estimates of the number of hours in a day[1].) Overall, when evaluating an uncertain numeric entity, higher anchors should produce higher estimates.

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

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

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

Field Anchoring

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

A smaller body of work considers anchoring effects with incentive-compatible designs. Work by Ariely, Loewenstein, and Prelec (2003) as well as Maniadis, Tufano, and List (2014) employed designs with real money and goods at stake. Both of these papers show data consistent with classic anchoring effects[2]. Both retained some contrived characteristics from the lab: the consideration and rejection of an arbitrary price derived from the last three digits of a social security number before eliciting the actual willingness to pay through a Becker, DeGroot, Marschak (BDM) procedure. Nunes and Boatwright (2004) used more naturalistic anchors (featured prices of nearby products at a local concert), but also utilized a BDM procedure. Given that the real world frequently lacks such contrivances and participants frequently misunderstand the BDM procedure's premise (Carson and Plott 2012), these findings offer an incomplete account of how anchoring operates in the field.

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

Charitable donations provide a closer analog, as payment is both elective in amount and guaranteed. A variety of papers examine the influence of both seller-produced (Alpizar, Carlsson, and Johansson-Stenman 2008; Martin and Randal 2008; Desmet and Feinberg 2003; Smith and Berger 1996) and donor-derived reference prices (Croson and Shang 2008; Shang and Croson 2009) on donation behavior. These papers, too, frequently produce significant anchoring effects. Finally, there have even been some investigations of anchoring under PWYW, in which the presence (e.g., Kim, Kaufmann, and Stegemann 2013; Gneezy et al. 2012) of an anchor is varied, also frequently producing significant effects.

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

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

Using this PWYW context, we explore how and why anchoring effects change between the lab and the field. Sixteen studies tested the size and presence of anchoring effects, and four additional studies consider similar manipulations in a hypothetical domain. The persevering reader will see our general inferences: First, although anchoring researchers agree that larger anchor gaps generate larger effects, 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 is 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 field, with both financial and non-financial consequences in Studies 1 and 2. In Study 3, we more closely examine absolute and distributional anchor gaps, an apparently critical distinction which has been generally neglected in the literature. Studies 4-6b 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, as we show anchoring effects are much easier to find in hypothetical judgments. Descriptive statistics for each study are in Table 1, histograms for each study are available in the Web Appendix, along with robustness checks for each study (e.g., log-transforming). 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 paper, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures. In no case did we analyze the data before reaching our pre-determined sample size. All data and materials are posted online.

—Insert Table 1 about here—

EXAMINING DIFFERENT TYPES OF ANCHOR GAPS

Study 1: Field Anchoring with Non-Payments

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

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

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

Results. Customers allocated more of their payment to developers when they saw a higher anchor. A one-way ANOVA revealed differences among the conditions, $F(5, 1322) = 366.98$, $p < .001$. Further analysis showed that this effect was driven primarily by the large differences between the three low ($M = 61.64\%$) and three high anchors ($M = 89.20\%$), $t(1326)$

= 42.82, $p < .001$, $d = 2.35$, in line with our general prediction. The three low anchors were not different from each other, $F(2, 645) = 0.08$, $p = .922$, but the three high anchors showed some differences, $F(2,677) = 4.36$, $p = .013$. The lowest high anchor ($M_{89\%} = 88.23\%$), generated a lower allocation that the medium high allocation ($M_{90\%}=89.27\%$), which in turn generated a lower allocation than the highest anchor ($M_{91\%} = 90.02\%$).

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

Study 2: Field Anchoring with Real Payments

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

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

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

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

We sold glazed doughnuts at an outdoor plaza at the University of California, Berkeley for 27 days in March and April of 2014. People (N = 70,091) passing by the doughnut stand saw our shop's sign, "Dream Fluff Doughnuts!" In the selection conditions, the sign also read "Pay What You Want," "$0.25 or Pay What You Want," or "$1.75 or Pay What You Want" (see 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.

The sign changed 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 the price information. We recorded date/ time of transactions, number of passersby, payment, customer group size, gender, and age. We predetermined to collect data until we had at least 100 observations per condition.

Results. Customers (N = 892 groups; N=1,038 individuals) bought 1,054 doughnuts[6]. We used individual doughnut purchase as unit of analysis with the average payment per doughnut as the dependent measure. We excluded 29 transactions by the experimenters' friends[7], which left us 1,009 purchases for analysis.

—Insert Figure 1 about here—

Consistent with basic self-selection, people were strongly influenced by posted prices on the signs: People were more likely to buy a doughnut when the sign said "$0.25 or PWYW" than when it said "$1.75 or PWYW" (382 out of 14,631 vs. 87 out of 14,548), $\chi^2(1) = 186.89$, $p < .001$. Furthermore, people seeing the former were more likely to purchase than those seeing either the PWYW sign (220 out of 13,639), $\chi^2(N=28{,}270) = 33.73$, $p < .001$ or the more ambiguous "Dream Fluff" sign (320 out of 28,073), $\chi^2(1) = 128.72$, $p < .001$. Even though all participants could pay what they wanted, the invitation to "pay $0.25 or pay what you want" was more motivating than any other sign.

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

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

Study 3a: Considering Different Types of Gaps

Although we replicated a standard anchoring effect with a standard, large anchor gap, Study 2—and most anchoring studies—did not differentiate the types of anchor gaps, e.g., in absolute terms, standardized terms, or in percentile[5] 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 to anchors both absolutely and distributionally far apart on these measures.

Method. We again collaborated with the PWYW media retailer from Study 1, but now focused on payments rather than allocations. This company typically offers a two or three-week promotion in which customers can pay what they want for a collection of thematically organized media goods. There are a few critical features in each promotion: First, there is a minimum price (e.g., \$1 for this bundle of six goods). Second, there is a "bonus" price, above which any elective payment also buys additional 2-4 goods (e.g., $9, for ten in total). Third, prices are identified by either typing in a box or by selecting on a sliding scale, which moves in $1 increments and ranges from the minimum to $100. Customers also may donate 10% of their payment to charity.

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

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

Study 3b: Replicating Study 3a

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

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

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

Discussion

Whereas Studies 1 and 2 conceptually replicated past work on anchoring, Studies 3a and 3b were more complicated. Some anchor gaps produced large and reliable effects (e.g., \$8 vs. \$20) but other gaps did not (e.g., \$20 vs. \$50). As we will elaborate through this paper, the mix of findings suggests that even with a very large 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 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 vary default prices and measure payments.

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

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

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

Study 5: Attempting Wider Gaps

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

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

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

Results. The promotion did better than expected, resulting in average payments between $9 and $12 across the three-week period. Therefore, the three low anchors ($2, $5, and $9) represented similar percentile payments: 30.0, 31.4, and 32.3, respectively, with the highest anchor, $12, generally remaining above average and resulting in a 60.7 percentile rank (using inclusive percentiles). The four means were similar, $F(3, 1599) = 0.21$, $p = .890$. The largest gap of the set, \$2 vs. \$12, actually showed a non-significant trend opposite the anchors ($M_{\$2}$ = \$11.48, $M_{\$12}$ = \$11.29), $t(757) = .34$, $p = .736$. We interpret this null effect as suggesting that the anchor gap needed to be even larger.

Study 6a: Considering High Anchors

Study 6a attempts to widen the 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 \$0.99 payment is at the 66^{th} percentile, but because 25% of customers pay \$1, a \$1 payment might alternatively be interpreted as a 91^{th} percentile payment, a 66^{th} percentile payment, or something in between. In Study 6a, we use the \$1 payment as the lower anchor and compare it to the unambiguously high anchor of \$3 (95^{th} percentile). Three additional conditions considered a true control (PWYW without any anchor) and two fixed price conditions (equivalent to the anchors) allow for assessment of overall demand.

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

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

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

Study 6b: Considering Low Anchors

In order to further probe the potential differential in anchor consideration in Study 6a (i.e., the null effect between a $3 anchor and a PWYW-only control, but the significant difference between the $1 anchor and control), we implemented a somewhat similar design in Study 6b. To better test this and to ensure we would find anchoring effects, Study 6b employs anchor gaps closer to lab-levels of extremity (16^{th} vs. 67^{th} percentiles), while retaining our control condition. This also allows us to test whether participants would reject *all* extreme anchors—both high and low—or whether we would find another asymmetry, in magnitude.

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

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

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

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

Discussion

Study 6 showed that more explicit consideration of the anchor, as in most lab studies, does not facilitate anchoring effects with smaller anchor gaps. In fact, in Studies 4-6, we found overall that the 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 may actually be the minimum requirement.

Study 6a also revealed a previously undiscovered nuance: if the gap is too wide, leaving anchors to be too extreme, customers may reject them from consideration and behave as if they saw an only slightly high anchor. Moreover, this aversion to extremeness appears to be asymmetric in our data: while customers may be put off by an implication that they should make large payments, they embrace the tacit permission to pay a small amount (see Croson and Shang (2008) for a similar effect of low anchors through social information in donations). We investigate this finding further in a hypothetical domain in Study 14d.

TESTING INSIGHTS FROM THE LITERATURE IN THE FIELD

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

Study 7: Anchors That Inform About the Behavior of Others

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

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

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

—Insert Figure 2 about here—

Study 8: Average Payment Information and Price Defaults

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

Method. Groups of visitors (N = 957 groups, 2,761 individual visitors) were randomly assigned to one of eight conditions: 4 (suggested donation amount: No information, $0.50, $1, or $2) x 2 (average donation amount: No information or $1) between-participants study[9]. We selected these numeric values from the distribution of donation amount in the previous month. On average, visitors gave about $0.94 (including 66% of visitors who did not donate).

Each group of visitors was asked to fill out a card that contained our main manipulations. Visitors read, "Today you can donate any amount to support the museum. (The suggested donation amount is [$0.50/$1/$2] per person). (Each visitor to the museum donates $1 on average.) How much would you like to donate?" Visitors also indicated their group size and home zip code on the card. (Examples of the cards can be found in the Web Appendix.) We predetermined to collect data until we had at least 100 groups per condition.

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

Study 9: Anchoring on the Payment of an Identifiable Other

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

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

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

Study 10: Anchoring with Real Retail Prices

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

Method. Study 10 (N=2,190) took place with the same PWYW retailer with the minimum price of $1 and the bonus price of $10. Visitors to the site were randomly assigned to see either a $14 default or a $28.88 default. In addition, half of the visitors saw the message "Full Retail Value: $28.88!" above the payment slider. (Due to a programming error, this 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. On the first day, anchors did not affect payments, $t(683) = .02$, $p = .991$. Although this null persisted for the rest of the promotion, payments did trend marginally in the direction of anchoring, as customers in the $14 condition paid less ($M$ = $10.68) than those in the $28.88 condition ($M$ = $11.26), $t(1503) = 1.91$, $p = .057$. The presence of the $28.88 retail price neither produced a main effect by itself, $F(1, 1501) = .72$, $p = .395$, nor did it interact with the default manipulation, $F(1, 1501) = .01$, $p = .92$. Even employing a high, precise, and justified anchor was not sufficient to elicit significant anchoring effects relative to a lower, imprecise, and unjustified one.

Study 11: Anchoring with Suggested Payments

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

Method. For Study 11 (N = 429 groups, 1,234 individuals), we returned to the children's museum from Study 8 conducted a 2 (regular admission fee ($11): present or absent) x 2 (suggested donation amount ($5): present or absent) between-participants study. For example, participants in the regular-fee or suggested-fee-present condition saw, "Your admission has been sponsored by ScholarShare[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) x 2 (suggested donation amount ($5): present or absent) ANOVA showed that neither variables influenced payments significantly. Visitors donated similar amounts regardless of whether they saw the regular admission fee, ($M_{\$5\&\ \text{regular fee present}}$ = $0.96 vs. $M_{\$5\&\ \text{regular fee absent}}$ = $1.06), $F(1, 425) = 0.56$, $p = .456$. Their payments did not differ whether or not they were provided with a suggested donation amount, ($M_{\$11\ \&\ \text{suggestion present}}$ = $0.96 vs. $M_{\$11\&\ \text{suggestion absent}}$ = $0.77), $F(1, 425) = 1.26$, $p = .263$. Furthermore, the interaction was not significant, $F(1, 425)=.01$, $p= .909$.

Studies 12a and b: Anchoring with Precise Values

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

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

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

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

Study 13: Anchoring with Maximum Possible Payments

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

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

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

Results. Although anchoring effects appeared sporadically over the four month-long periods (see Web Appendix for full discussion of the results), they were conflicting and unreliable: a combined analysis dampened any seeming anchoring effect. We analyzed all of the anchor differences while controlling for month-to-month variation (which was substantial).

There were no differences in payments between the different anchor conditions except the two highest anchor conditions, $50 and $100 (see Table 1). We are disinclined to read too much into the unpredicted pattern between the $50 and $100 anchors as it has not been observed in any other studies (including other studies conducted at the same museum). Our analysis controlled for month-to-month variation, but there is room for an unidentified systematic influence due to the lack of perfect random assignment across four months of data collection. We predict that this pattern would not replicate under better conditions, however, future research could investigate it.

Discussion

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

RETURNING TO THE LAB

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

Study 14a

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

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

Results. Fourteen participants gave an answer less than the minimum price or did not complete the survey and were excluded from analyses, leaving 386 participants. We conducted a 2 (default price: \$14 or \$28.88) x 2 (response format: slider vs. text box) ANOVA on payment. People paid slightly more when using the slider than when using the text box (M_{slider} = \$15.28 vs. $M_{textbox}$ = \$13.35), $F(1, 382) = 3.21$, $p = .074$, but, critically, did not interact with the anchoring effect, $F(1, 382) = .49$, $p = .483$. Most important, participants who saw a \$14 anchor paid significantly less (M=\$13.01) than those who saw a \$28.88 anchor (M = \$15.78), $t(384) = 2.68$, $p = .008$, d=0.28. Although this distribution resembled our field sample (e.g., the minimum, bonus, and default prices were all relatively popular), the upper bound of payments was much higher than what we usually see. Therefore, we winsorized our results at the 95th percentile (\$28.88) and repeated our analysis: we still found no significant interaction, $p = .157$, and the anchoring effect remained highly significant, $F(1,382) = 16.01$, $p < .001$, $d = 0.41$. Importantly, though, response format, was even further from statistically significant, $p = .543$.

Study 14b

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

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

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

Study 14c

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

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

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

Study 14d

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

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

Including $50 and $0.01 allowed us to see how participants used anchors that were beyond or at the boundaries of payments and compare that to how they used anchors that were still extreme, but to a lesser extent. Having a "Pay nothing or PWYW" condition allowed us to examine how participants think of PWYW pricing. One potential concern with using

PWYW pricing to study anchoring is PWYW may essentially generate an anchor of $0, which could mute anchoring effects.

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

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

Although we pre-registered the winsorized specification, it is worth noting that all but one difference was significant when analyzing the untransformed data. That alternative specification is reported in 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 b) also showed us that the peculiarities of one of our study sites likely did not suppress anchoring effects. In addition, in Studies 14a and c, we showed significant effects with the same designs that produced null effects in the field. Finally, Study 14d showed that anchors in a hypothetical context can be fairly close together—both absolutely and distributionally—and still elicit significant differences in payments.

GENERAL DISCUSSION

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

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

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

The Lab Versus the Field

The most obvious of our three speculative explanations is the impact of taking a paradigm out of the lab and into the field and vice versa. Most lab researchers think that their effects will look different (and smaller) in the field. Relative to tidy lab studies, field settings introduce noise in measurement and in manipulation. Consistent with that thinking, when we took some of our null findings out of the field and back into the lab (e.g., Studies 14a and c), we show highly significant results. Nevertheless, we think that the inherent noisiness of the field is an unlikely explanation for the observed effects. In our case, "lab vs. field" stands in for more than just variation in noise and signal. Perhaps it is just that anchoring effects are small when real money is at stake? An unrealistically high anchor was quite influential in hypotheticality but not in reality. When buying hypothetical confections with hypothetical currency, very large anchors are quite influential.

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

Asymmetry of Magnitude Perception

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

Size of Anchor Gap

It may be the case that extremely high anchors are ignored, but it is certainly the case that the high anchor has to be substantially higher than the low anchor. Moreover, that anchor gap needs to be considered in terms of 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 (99^{th} vs. 100^{th}). Consider the dark points (representing field experiments with payments) in Figure 3. For every anchor gap smaller than 50 percentile points, we observed non-significant effects and very small effect sizes. For six of the seven gaps larger than 50 percentile points we observed statistically significant effects (and generally larger effect sizes).

—Insert Figure 3 about here—

Though a focus on something as simple as the distributional anchor gap might seem obvious, it has been missed by previous researchers testing anchoring effects in the field. For example, successful instances of anchoring in the field (e.g., Croson and Shang 2008) derived their anchors from each customer's past donation, placing them somewhat above or below that amount—a strategy likely to maximize effect sizes by ensuring a large distributional gap with anchors still within the range of consideration. Other attempts, such as Smith and Berger (1996) also based anchors off of prior donations, but in the same direction with a much smaller gap (10% vs. 50% of the prior amount), which may have contributed to their null result. Others still (e.g., Alpizar, et al 2008) failed to account for 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, while seemingly arbitrary (i.e., lacking explicit justification), are often perceived as recommendations (McKenzie 2006). Anchors are frequently defined as the exact opposite, with several paradigms even informing participants that they are not useful.

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

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

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), 1047-1060.

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

Armstrong-Soule, Catherine and Robert Madrigal (2014), "Anchors and Norms in Anonymous Pay-What-You-Want Pricing Contexts," *Journal of Behavioral and Experimental Economics,* in press.

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 T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge: Cambridge University Press, 120-128.

Critcher, Clayton R. and Thomas Gilovich (2008), "Incidental Environmental Anchors," *Journal of Behavioral Decision Making,* 21(3), 241-251.

Croson, Rachel, and Jen Yue Shang (2008), "The Impact of Downward Social Information on Contribution Decisions," *Experimental Economics*, 11(3), 221-233.

Danilowitz, Jennifer, Shane W. Frederick, and Daniel Mochon (2014), "Anchoring as Inference," working paper, Yale School of Management.

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), 349-376.

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.

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, Leif D. Nelson, and Amber Brown (2010), "Shared Social Responsibility: A Field Experiment in Pay-What-You-Want Pricing and Charitable Giving," *Science*, 329(5989), 325-327.

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*, 109(19), 7236-7240.

Hardesty, David M., and Tracy A. Suter (2013). "Maximizing willingness to bid within “Buy It Now” auctions." *Journal of Business Research,* 66(4), 554-558.

Jacowitz, Karen E., and Daniel Kahneman, (1995), "Measures of Anchoring in Estimation Tasks," *Personality and Social Psychology Bulletin,* 21, 1161-1166.

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*.

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), 622-628.

Kim, Ju-Young, Martin Natter, and Martin Spann (2009), “Pay What You Want: A New Participative Pricing Mechanism,” *Journal of Marketing*, 73(1), 44-58.

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*, 1-15.

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 Science 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,” *The American Economic Review*, 104(1), 277-290.

Martin, Richard and John Randal (2008), “How is Donation Behaviour Affected by the Donations of Others?,” *Journal of Economic Behavior & Organization*, 67(1), 228-238.

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), 759-763.

Nunes, Joseph C., and Peter Boatwright (2004). "Incidental prices and their effect on willingness to pay." *Journal of Marketing Research*, 41(4), 457-466.

Regner, Tobias and Javier A. Barria (2009). “Do Consumers Pay Voluntarily?: The Case of Online Music.” *Journal of Economic Behavior and Organization* 71, 395-406.

Riener, Gerhard and Christian Traxler (2012), “Norms, Moods, and Free Lunch: Longitudinal Evidence on Payments from a Pay-What-You-Want Restaurant,” *The Journal of Socio-Economics*, 41(4), 476-483.

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,” *The Economic Journal*, 119(540), 1422-1439.

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.

Simonson, Itamar and Aimee Drolet (2004), “Anchoring Effects on Consumers’ Willingness-to-Pay and Willingness-to-Accept,” *Journal of consumer research*, 31(3), 681-690.

Simonsohn, Uri , 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.

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), 219-231.

Smith, Sarah, Frank Windmeijer, and Edmund Wright (2014), “Peer Effects in Charitable Giving: Evidence from the (Running) Field,” *The Economic Journal*.

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), 131-144.

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.

Mussweiler, Thomas, and Fritz Strack (2001), “Considering the Impossible: Explaining the Effects of Implausible Anchors,” *Social Cognition*, 19(2), 145-160.

Tversky, Amos, and Daniel Kahneman (1974), “Judgment under Uncertainty: Heuristics and Biases,” *Science,* 185(4157), 1124-1131.

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.

FOOTNOTES

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

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

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

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

5. 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.

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

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

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.

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

Table 1. Basic Info for Each Study

	Product	N	n	Anchor Type	Anchor	Anchor Percentile	Mean (SD)	Outcome Percentile
Study 1	Allocations	1,328	648	Default	50%	16.9	61.64% (15.36%)[a]	30.9
			680		90%	80.0	89.20% (6.60%)[b]	80.0
Study 2	Doughnuts	1,009	490	Suggestion	$0.25	20.3	$0.44 ($0.38)[a]	90.1
			181		$1.75	92.8	$1.04 ($0.71)[b]	50.4
			338		Control	---	$0.66 ($0.54)[c]	29.9
Study 3a	Media Bundle	303	110	Default	$3	4.6	$6.59 ($2.38)[a]	55.1
			91		$9	69.6	$7.79 ($2.44)[b]	66.3
			102		$20	96.7	$8.29 ($4.57)[b]	69.6
Study 3b	Media bundle	3,978	1,443	Default	$8	21.5	$8.88 ($4.03)[a]	23.5
			1,291		$20	95.3	$9.89 ($4.96)[b]	23.5
			1,244		$50	100.0	$9.88 ($5.63)[b]	23.5
Study 4	Media bundle	3,214	1,602	Default	$12	75.9	$8.80 ($4.64)[a]	23.2
			1,612		$15	87.1	$8.88 ($4.88)[a]	23.2
Study 5	Vodo	1,603	376	Default	$2	29.2	$11.48 ($7.87)[a]	42.7
			425		$5	30.7	$11.24 ($7.65)[a]	42.1
			419		$9	32.1	$11.05 ($7.68)[a]	41.8
			383		$12	45.6	$11.29 ($7.72)[a]	42.3
Study 6a	Doughnuts	395	113	Suggestion	$1	27.9	$0.91 ($0.38)[a]	27.9
			78		$3	94.8	$0.84($0.75)[a,b]	27.9
			106		Control	--	$0.72 ($0.45)[b]	24.9
Study 6b	Museum tickets	431	97	Suggestion	Nothing	0.0	$2.55 ($2.73)[a,c]	57.1
			117		$0.01	16.2	$2.47 ($2.36)[a,c]	52.0
			92		$5	67.3	$3.47 ($2.96)[b]	62.9
			102		Control	---	$3.24 ($2.77)[b,c]	61.3
Study 8	Museum tickets	957	211	Suggestion	$0.50	56.4	$0.93 ($2.20)[a]	64.8
			273		$1	64.9	$0.77 ($1.29)[a]	64.6
			253		$2	83.4	$0.91 ($1.36)[a]	64.8
			220		Control	---	$1.08 ($6.80)[a]	74.8

Study 9	Media bundle	1,175	378	Previous Payment	$8	7.0	$11.10 ($3.97)[a]	68.1
			393		$12	68.2	$11.17 ($3.50)[a]	68.2
Study 10	Media bundle	2,190	1,122	Default	$14	88.1	$10.78 ($5.40)[a]	62.2
			1,068		$28.88	98.5	$11.17 ($6.19)[a]	67.5
Study 11	Museum Tickets	429	121	Suggestion	$5	93.0	$1.06($1.55)[a]	70.2
			88		$11	99.8	$0.77($1.48)[a]	68.1
			126		Control	--	$0.90($1.52)[a]	68.5
Study 12a	Media bundle	714	138	Default	$19.91	84.7	$11.67 ($6.13)[a,b]	25.6
			140		$19.99	85.9	$12.31 ($6.39)[a]	66.1
			135		$20.00	88.1	$11.53 ($6.63)[a,b]	25.6
			158		$20.01	93.8	$10.52 ($7.65)[b]	25.4
			143		$20.09	95.2	$11.60 ($6.89)[a,b]	25.6
Study 12b	Media bundle	4,110	790	Default	$19.91	91.4	$10.56 ($4.81)[a]	75.9
			827		$19.99	92.2	$10.39 ($4.48)[a]	75.4
			861		$20.00	93.5	$10.32 ($4.45)[a]	75.4
			838		$20.01	97.1	$10.60 ($5.67)[a]	75.9
			794		$20.09	97.7	$10.55 ($4.71)[a]	75.9
Study 13	Museum tickets**	590	231	Maximum Price	$10	99.0	$2.72 ($3.25)[a,b]	58.7
			155		$50	100	$2.25 ($2.27)[a]	57.3
			73		$100	100	$2.90 ($2.54)[b]	58.7
			147		Control	--	$2.01($2.94)[a,b]	57.3
Study 14a	Media bundle*	386	205	Default	$14	56.2	$13.01 ($9.82)[a]	56.2
			181		$28.88	94.6	$15.78 ($10.34)[b]	73.1
Study 14b	Media bundle*	581	274	Default	$3	21.7	$5.10 ($3.43)[a]	47.0
			307		$9	85.7	$5.99 ($3.12)[b]	47.0
Study 14c	Doughnuts*	169	61	Suggestion	$1	36.7	$0.82 ($0.82)[a]	36.1
			56		$3	89.3	$1.47 ($0.94)[b]	73.4

Study 14d	Doughnuts*	2,038	313	Suggestion	$0.00	0.0	$0.92 ($0.78)[a]	47.9
			359		$0.01	3.8	$0.50 ($0.52)[b]	29.6
			382		$0.25	11.9	$0.39 ($0.52)[c]	29.0
			304		$1.75	87.6	$1.09 ($0.61)[d]	83.3
			359		$50	100	$1.52 ($1.26)[e]	87.6
			321		Control	--	$1.04 ($0.73)[f]	73.4

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

Figure 1. Study 2 – We replicate the standard anchoring effect using real payments, while failing to find support for any biasing effects of selection.

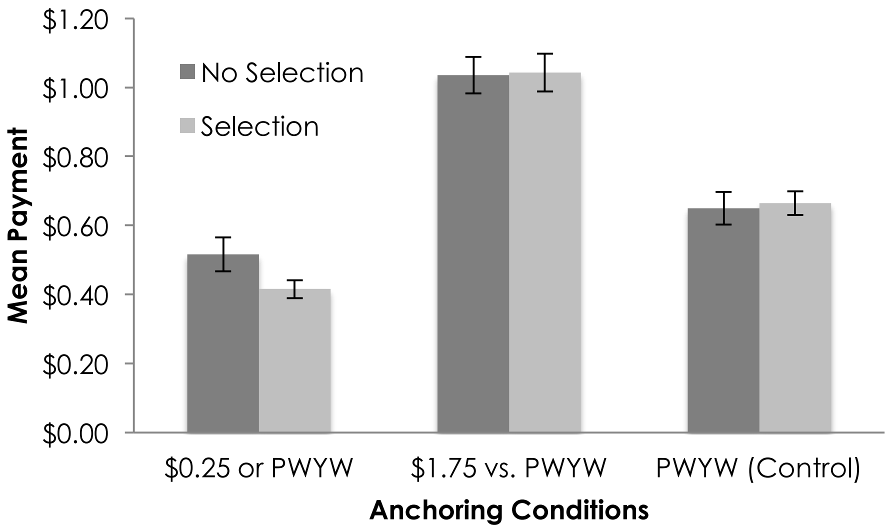

Figure 2. Study 7 – We find payments are not influenced by knowing the average price paid by previous customers, as shown by the regression lines for each group.

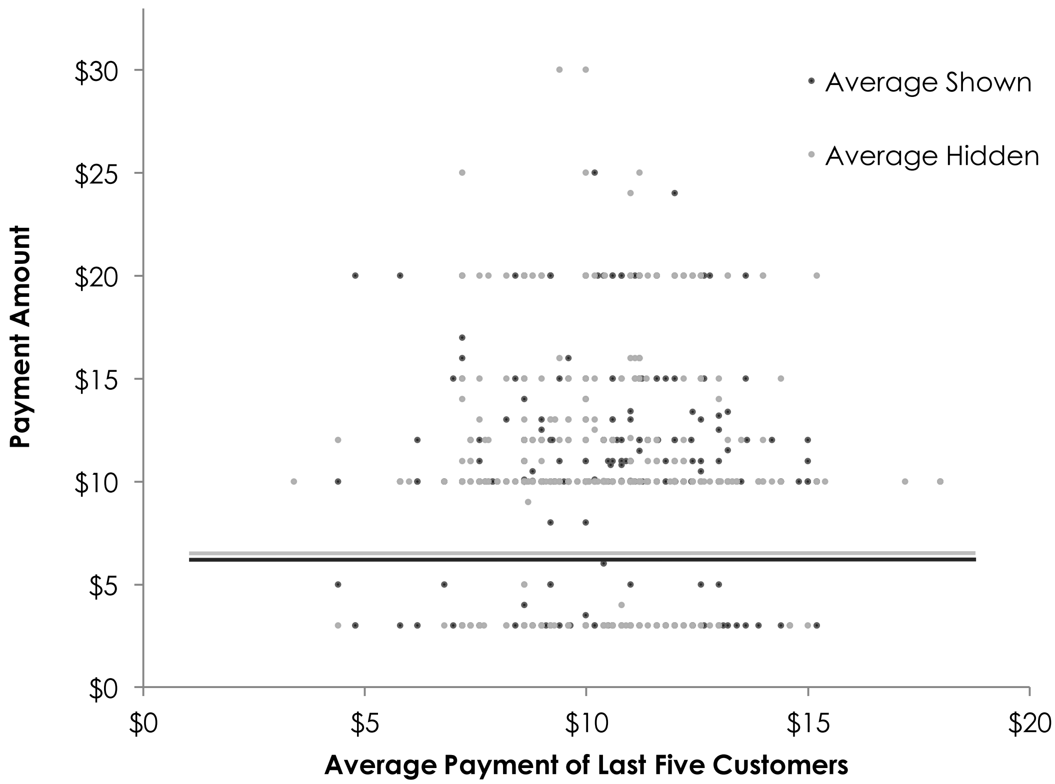

Figure 3. Distributional Anchor Gap versus Cohen's d Across All Studies