Gratuities in a Digital Services Marketplace

Seung Hyun Kim, On Amir and Kenneth C. Wilbur

Abstract

Tip request messages are marketing communications that can help balance platform customer spending with worker compensation and effort. We report the first large-scale field experiments that exogenously manipulate motivations in tip request messages, collectively treating 276,006 transactions by 88,857 buyers. We test how appeals to two common tipping motivations–injunctive norms ("It's customary to leave a tip for the seller's service") and reciprocity–affect tipping and related behaviors in a global freelance services marketplace. On the web, first exposures to the injunctive norms message increased new buyers' tipping rate–defined as the proportion of transactions tipped–by 11.8 percentage points (p.p.), and repeat buyers' tipping rate by 6.8 p.p. In the app, first exposures to the injunctive norms message increased tipping rate by 2.5 p.p. for new buyers and 1.4 p.p. for repeat buyers. Surprisingly, tipping rate increases come without detectable reductions in platform repatronage or spending. The injunctive norms message effect is larger in transactions with higher prices, 5-star ratings and North American buyers. Collectively, the results suggest that tip request messages can motivate customer tipping to better incentivize and compensate gig workers, and that optimization requires careful testing.

Keywords: Digitization, Field Experiments, Gig Economy, Gratuities, Injunctive Norms, Norms, Platforms, Platform Design, Services, Tipping

Kim is a Visiting Assistant Professor of Marketing at Northeastern University. Amir is the Wolfe Family Endowed Chair and Professor of Marketing at the University of California, San Diego. Wilbur is the White Family Endowed Chair and Professor of Marketing and Analytics at the University of California, San Diego.

1. Introduction

Tipping is an important contributor to gig-economy workers' earnings. Surveys show that 16% of U.S adults earned income through online gig platforms, with roughly one-third of those primarily employed as gig workers during the previous year (Anderson et al. 2021). 29% of gig workers surveyed reported earnings below their state's hourly minimum wage (Zipperer et al. 2022). Nearly 80% of rideshare drivers say that tips are essential to their overall earnings (Kilroy 2024). Digital platforms–including point-of-sale payment systems, ridesharing apps, professional service markets, and "gifts" to creators on social media–frequently suggest gratuities of 15%-30%. Customers retain discretion over tipping, but technological and interface changes can affect tipping behaviors (Carr 2013).

Tip requests have therefore become an important design consideration for digital platforms. Platforms have to strike a delicate balance: customers desire quality service at affordable prices, but face considerable uncertainty as to when and how much they should tip; whereas service providers want adequate compensation, work, and reward for service effort. The platform's tip request should preserve goodwill on both sides of the marketplace, so consumers do not feel coerced or alienated, and so service providers are adequately appreciated, motivated and compensated.

Tip requests are important and common, yet we lack causal evidence about how motivational appeals in tip request messages affect tipping. We also need to understand potential counterfactual effects on buyers' purchases: Do tip requests operate like price increases, whereby more tipping leads to fewer sales? Counterfactual downsides of tip requests are difficult to quantify without ecologically valid experimental evidence.

Most public research on digital tipping has focused on how default tip options (e.g., 20% or $5 tip buttons) influence tipping behavior. This is important work to be sure, and well amenable to experimental manipulation and econometric analysis. Yet, so far as we can discern, there is no prior evidence on how *tip request message wording* affects tipping decisions or how tip request messages affect subsequent customer actions. The tip request message is a marketing communication that requires testing in order to help strike the right balance between buyer spending, worker earnings and platform profits.

To address these gaps, we collaborated with a large digital freelance services platform to run two large-scale field experiments to test how tip request messages affect buyer tipping and repurchase. We pose the following research questions:

- How do reciprocity and injunctive norms motivations in tip request messages affect tipping, compared to a Control message? Do injunctive norms tip request messages reduce future business, like a price increase would?
- What contextual factors predict tip request message effects on tipping? Are they more effective at higher or lower prices? How does tipping induced by norm-focused tip request messages affect subsequent buyer and seller actions on the platform?

We first report a Web Field Experiment that jointly manipulated tip request messages and default tips. The message treatments included three motivations for tipping: Impersonal Reciprocity, Personalized Reciprocity, and an Injunctive Norms message ("It is customary to leave a tip for the seller's service," which we refer to as "Norms"). The latter two messages were combined with a default tips menu change that added a third default tip option. The experiment ran in the market for a predetermined four weeks. After the experiment concluded, the platform reverted to its incumbent tip request message ("Would You Like to Leave a Tip To [seller name]?"), which served as the Control during the experiment; but the platform retained the expanded default tips menu. Therefore, the experiment conclusion offers a quasi-experimental discontinuity that enables identification of tip request message effects apart from the default tips manipulation. Our analysis pools the test data with the post-test data and distinguishes causal effects between new buyers and repeat buyers, and between first exposures and all exposures. We subsequently conducted an App Field Experiment, an A/B test that compared the same Control message to the Norms message, without any other change. We first report the main effects of treatments on tipping rate within each field experiment, and then we analyze broader platform implications, including causal effects on buyer orders, spending, tip count and tip spending; treatment effects in price partitions; treatment effects on subsequent buyer and seller platform behaviors; and observable transaction features that correlate with treatment effects.

Our first main result is that tip request messages influence tipping. In the web experiment, the injunctive Norms message increased tipping on first exposure by 11.8 percentage points (p.p.) for new buyers and 6.8 p.p. for repeat buyers, relative to the Control message, holding default tips fixed. Those are economically meaningful increases compared to relatively low baseline tipping rates of 14.6% and 14.9%, respectively. The Personalized Reciprocity message also increased tipping, albeit less, with first-exposure effects of 5.3 p.p. for new buyers and 4.4 p.p. for repeat buyers, while Impersonal Reciprocity effects are near zero. The platform's expansion from 2 to 3 default tip options reduced first-exposure tipping by 4.7 p.p. among new buyers and 4.3 p.p. for repeat buyers. The Norms message effects in the app field experiment are smaller, though still highly significant, and show qualitatively similar patterns across new-buyer and first-exposure subsamples. Second, although Norms-related treatments increased tipping, they did not detectably change buyers' order counts or total spending. Norms treatment effects are larger at higher order prices. Third, we do not detect behavioral changes from Norms-induced tipping in instrumental-variables estimates, though this analysis is not strongly powered. Finally, Norms treatment effects are larger in web transactions with 5-star ratings and North American buyers, but do not differ significantly across gig service categories.

Collectively, the findings suggest that platforms can use tip request messages to increase gig worker income, and remarkably, that the revenue increase can come without decreases in subsequent buyer orders or prices. However, not all tip request messages are equally effective, and the majority of transactions remain untipped across all treatments and all subsamples. Optimization requires careful testing of both tip request messages and default tips.

Section 2 describes how the paper learns from, and adds to, prior literature. Section 3 describes the empirical context and how our research collaboration with Fiverr influenced the experiments and timing. Section 4 describes the Web Field Experiment design and treatment effects on tipping rates, and Section 5 describes the App Field Experiment design and treatment effects on tipping rates. Section 6 analyzes web and app experiments' data to examine how the winning treatments (a) affected platform repatronage, (b) differed across order price partitions, (c) affected subsequent buyer and seller actions on the platform, and (d) interacted with observable transaction features to predict tipping response. Section 7 concludes with implications and other tipping-related topics that merit testing.

2. Contributions and Relationship to Prior Literature

When and why do consumers freely choose to pay extra money? These questions have motivated extensive research on charitable giving, pay-what-you-want (PWYW) business models, and tipping (van Teunenbroek et al. 2023). All three mechanisms involve a discretionary component regarding whether and how much to pay extra, but their underlying motivations and structural characteristics differ (Raghubir and Bluvstein 2024). PWYW enables buyers to set their own price, which is often guided by perceived value and fairness (Kim et al. 2009, Gneezy et al. 2012). Charitable giving is commonly driven by altruistic or prosocial motivations, and may also provide indirect personal benefits such as self-signaling and psychological satisfaction (Small and Loewenstein 2003, Bekkers and Wiepking 2011). Tipping is motivated by social norms, perceived service quality, reciprocity, and service/esteem motives (Azar 2007, 2011, 2020; Lynn 2015, 2016, 2017). Web Appendix A summarizes and compares related empirical papers.

Our focus on tipping occurs as the digitization of services and payments has increased tip request frequency and consumer uncertainty about appropriate tipping norms (Wolfe 2023). Offline tipping norms rely on interpersonal cues, social pressure, and public visibility of payments (Azar 2007, Bluvstein and Raghubir 2021). These cues may differ in computer-mediated transactions as perceived anonymity, reduced identifiability, and the absence or reduction of nonverbal, interpersonal cues may diminish norm-consistent behaviors (Lapidot-Lefler and Barak 2012, Cohn et al. 2022), and greater social distance online may reduce reciprocal behavior (Charness et al. 2007). It remains to be tested how traditional tipping motivations like norm compliance and reciprocity operate in online service transactions.

A growing literature shows that online tipping practices may diverge from offline tipping in several interesting ways. Duhaime and Woessner (2019) show that customers tip gig workers less than traditional employees, as gig workers are perceived as more autonomous. Chen et al. (2023) find that buyers may offer digital ratings as substitutes for monetary tips in online contexts. Yet, recent studies also suggest that platform design choices can partially restore or reshape tipping norms online. Paridar et al. (2025) develop a model of user learning about social norms based on tips received and tips observed, in which learning shapes beliefs about the norm; beliefs drive tipping actions; and collective tipping develops the endogenous norm. They calibrate the model using the entire tipping history of a user-generated content platform. Model-based counterfactual simulations indicate that platform-provided information about overall tipping rates can hasten the convergence of tipping norms.

To our knowledge, no prior research has directly tested motivations in tip request messages. One reason may be because most tips were customarily *given*, rather than *requested*, prior to the advent of modern point-of-sale payment platforms like Square or Stripe. We found one survey study (Karabas et al. 2020) and one field study (Dyussembayeva et al. 2022) that tested explicit tip requests vs. no tip requests, finding that explicitly requesting a tip in a restaurant reduces tipping compared to not requesting a tip. There is also work that has examined related effects of requesting tips and framing defaults in restaurant contexts: field experiments have shown tipping responses to writing “thank you” on customers’ checks (Rind and Bordia 1995), drawing happy faces (Rind and Bordia 1996), writing helpful messages about future specials (Rind and Strohmetz 1999), providing default tips and framing them as custom conveniences (Cabano and Attari 2023), and adding emojis to default tips (Lefebvre et al. 2024).

Similarly, most prior digital tipping research studies default tips in mixed digital/interpersonal services. Examples include digital tipping screens at coffee shops (Bluvstein and Raghubir 2022), in taxi cabs (Haggag and Paci 2014, Donkor 2021), in ride-sharing services (Chandar et al. 2019), and in app-based dry-cleaning services (Alexander et al. 2021). This research is closely related, as default tips indirectly communicate injunctive normative recommendations (Donkor 2021), but these contexts all feature interpersonal interactions interspersed with digital tipping, rather than fully digital settings.

Building on these, our paper contributes to the literature in four ways. First, to the best of our knowledge, we report the first large-scale field experiments that exogenously manipulate motivations in tip request messages–that is, the cues that immediately precede and encourage customer tipping decisions.

Field experiments offer scale, statistical power, and ecological validity, which are all important for studying real-world behaviors that involve actual monetary costs, such as tipping (Harrison and List 2004). Field experiments provide the strongest evidence because they capture agents' actual behavior in natural settings, without limiting assumptions based on stated intentions or environments that are overly controlled, artificial, or unfamiliar. Second, more specifically, this is the first study to test how injunctive norms and reciprocity appeals elicit tips. Third, the transactions we analyze are fully digitally intermediated, without face-to-face interactions, and between consumers and service providers who are distributed worldwide. Fourth, we compare message effects to a default tips change, show how they affect customer repatronage, estimate message-induced tip effects on subsequent buyer and seller platform behaviors, and quantify which transaction features predict treatment effects on tipping. As such, we offer methodological guidance to platforms that want to test tip request message effects.

3. Empirical Context

We describe relevant platform details, our research collaboration, and the outcome measures.

3.1. Platform description and purchase process

Fiverr is a leading global platform for freelance digital services. In 2024, it connected 3.6 million buyers with over 380,000 freelance service providers ("sellers"), generating $303.1 million in marketplace revenue (Fiverr 2024, Saxena 2025). Fiverr helps freelance service buyers transact with sellers, similar to Upwork.com, PeoplePerHour.com, and Freelancer.com, among others. Sellers list "gigs," which are packages of service attributes and prices. Fiverr earns money when sellers complete gigs to buyers' satisfaction, charging buyers $2 per order under $40, or 5% of prices over $40; and retaining 20% of seller revenue, including tips.[1]

A Fiverr transaction involves four steps. First, buyers browse the service directory or search for

[1] Fiverr's transparent fee structure contrasts with revelations of misleading tip policies at other platforms (Hanbury, 2019; Newman, 2019).

services by keyword. Each gig listing displays the service description, seller name, average rating, and price. Second, buyers click through to detailed gig pages featuring service attributes and seller information, including photo, location, historical rating average, and past reviews. An in-platform message system enables buyers and sellers to discuss service specifics in writing. Third, after a buyer agrees to a transaction, the platform takes the buyer's payment, and then the seller completes and delivers the work digitally through the platform. After the buyer confirms gig receipt and completeness, Fiverr releases the seller's payment. Fourth, immediately after the buyer confirms gig completion, Fiverr prompts the buyer to rate the service, and then displays a tip request message and default tips.[2]

All transaction details are determined and completed prior to the display of the tip request message and the buyer's decision to tip. Therefore, buyers' first exposure to a tip request message is fully exogenous to all transaction characteristics. Buyers' subsequent repatronage decisions and subsequent transaction features may be affected by prior exposure to experimental manipulations of tip request messages and default tips. We therefore distinguish between first-exposure and all-exposure samples when reporting findings.

3.2. Research collaboration overview

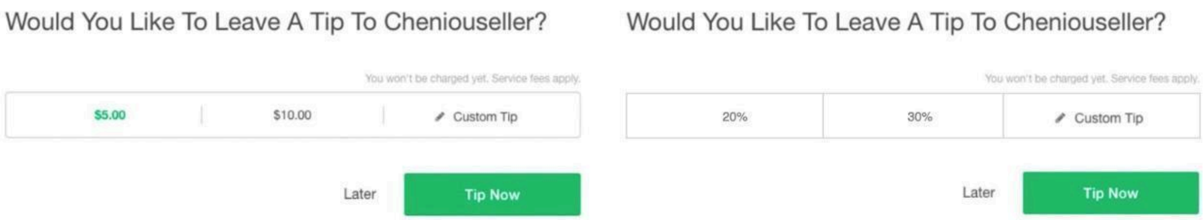
Fiverr sought to increase the average tipping rate on the platform, for three reasons. First, most service requests require some degree of seller customization effort to understand and meet the buyer's specific needs. Tipping offers an extra incentive to elicit and reward seller effort, which may in turn improve buyer satisfaction and platform repatronage. Second, tipping could directly increase seller and platform revenue, independent of buyer satisfaction. Third, tipping can only increase costs to buyers who freely choose to give tips, offering a form of self-selected price discrimination (in which self-selection may depend on buyer willingness to pay, ability to pay, satisfaction, or other factors), and therefore may help to maintain lower prices (Fennell, 2023; Lynn and Withiam 2008; Schmidt, Spann and Zeithammer 2015).

Figure 1 shows the tip request message and default tips as they appeared before the Web Field

[2] Fiverr withholds the tip request when gigs are rated as 1- or 2-stars, which affects less than 1% of transactions.

Figure 1: Original Tip Request Message and Default Tips

A: When price ≤ $25 **B: When price > $25**



Experiment. The status-quo message serves as the control message in both field experiments reported below. Prior to the Web Field Experiment, the status-quo default tips menu depended on whether the gig price exceeded a dollar/percentage threshold of $25. For gig prices below the dollar/percentage threshold, the default tips were $5, $10, or Custom; for prices above the threshold, the default tips were 20%, 30%, or Custom. Approximately 70% of tips used default tip options, but the 30% option was rarely used.

We proposed to randomize the tip request message, as this seemed like the most direct way to potentially influence customer tipping decisions, and as the experimental manipulation had not been tested previously. The company agreed to run the test on its website. Fiverr frequently ran A/B tests on the platform to test how design changes impacted platform customers and workers. It allocated a predetermined four weeks of in-market testing resources to run the test.[3]

Fiverr also decided to simultaneously include a new three-default tipping menu along with two tip request message treatments. Unfortunately, Fiverr's testing apparatus allowed a maximum of four simultaneous treatments. Therefore, two message treatments–Norms and Personalized Reciprocity–were combined with a new default tips menu, and two treatments–Control and Impersonal Reciprocity–retained the two-default tips menu. Section 4.1 describes the three-default tips menu in detail.

After the test concluded, Fiverr management decided to revert to the status-quo tip request message, and also to retain the new three-default tips menu, as it awaited its internal preliminary data analysis. These decisions coincided with the conclusion of the predetermined four-week test period, and

[3] This was the longest in-market test the company had run to date. Our data use agreement required us to obfuscate data if requested by the firm, but did not enable the firm to censor publication.

were not made on the basis of the test results (the data had not been analyzed during the test) or other market-related factors, either predicted or realized. The post-test period therefore enables separation of the message treatment effects and the three-default tip menu effect, as Section 4.4 explains.

The preliminary data analysis was completed several weeks after the test concluded. The Norms/3-Defaults treatment condition maximized tipping rates and revenues during the test, so the Norms tip request message was adopted on all web transactions, and the 3-Defaults menu was maintained. The Norms message adoption corresponded with meaningful increases in average tipping rates among new buyer cohorts who first transacted in weeks 1-3 after the test, and also in weeks 4-6 after the test (Web Appendix Figure B10).

Next, we proposed testing the same injunctive Norms message in Fiverr's mobile app, as it was presenting the legacy tip request message in Figure 1. An App Field Experiment was conducted, this time as a simple A/B test with the same Control message and same injunctive Norms message, both presented with the 3-Defaults tip menu. The main results showed significant tipping increases, though smaller than in the Web Field Experiment, so the Norms message was subsequently adopted in the mobile app as well.

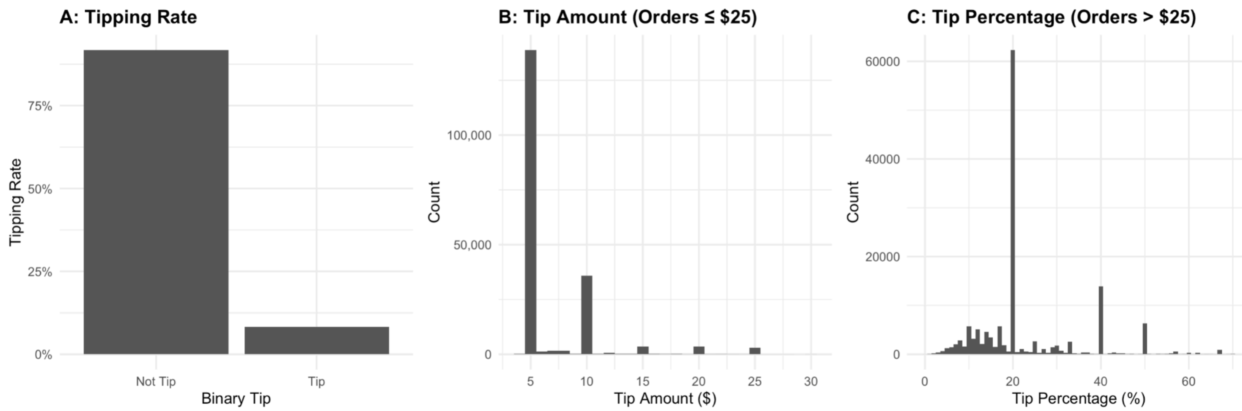
3.3. Measures

Figure 2A shows that 8% of transactions were tipped prior to the Web Field Experiment.[4] The data also suggest that tipping was situational: Among buyers with five or more transactions, collectively accounting for 58% of all transactions, 31.6% of frequent buyers tipped at least once, but only 0.7% always tipped.

70% of the pre-experimental tips used the default tips illustrated in Figure 1. Figure 2B shows that the distribution of tip amounts has large mass points at $5 and $10, as these were the default tip options provided. Figure 2C shows that the distribution of tip percentage (i.e., tip amount divided by price) has a large mass point at 20%, but no corresponding mass point at 30%, so the firm subsequently

[4] The pre-experiment data include 4.1 million transactions between 1.3 million buyers and 171,088 sellers over 5.3 months. The data represent $152.4 million in total spending, including $3.9 million (2.5%) in 341,325 tips.

Figure 2: Tipping Rates and Tip Percentage.



decided to eliminate the 30% default tip option.

The experiments below focus on explaining variation in tipping rates (i.e., the share of transactions that receive a tip), as tipping is the single most important behavior for establishing a tipping norm. Web Appendices F and G show that regressions explaining tip amounts or tip percentages are noisier than regressions explaining tipping rates due to the increased variability in spending decisions, but the point estimates are mostly congruent across outcome variables.

Web Appendix B describes how pre-experimental tipping correlates with observable features of transactions, and relates those patterns to tipping motivations described by classic research. Most notably, it shows that tipping was four times more prevalent in 5-star transactions than in 3-star transactions (14.1% vs. 3.5%). North American buyers, where restaurant tipping norms are prevalent, tipped in 11.5% of orders, compared to 3.9% of buyers in other global regions. We use these features, among others, to predict Norms effects on tipping rates in Section 6.4.

4. Web Field Experiment

Section 4.1 describes the message and default tip treatments. Section 4.2 describes the data, randomization and relevant subsamples. Section 4.3 presents nonparametric estimates of combined message/default-tips treatment effects. Section 4.4 describes how the post-test period enables separate identification of tip request messages and default tips main effects and the model. Section 4.5 reports treatment effects on tipping and robustness checks.

4.1. *Treatments*

The Web Field Experiment tested four treatment conditions that each combined a tip request message with a default tips menu. We first describe the tip request messages, and then the default tips.

The treatments included the status-quo and three new treatment messages:

1. *Control*: "Would You Like to Leave a Tip To (seller name)?"
2. *Impersonal Reciprocity*: "Show your appreciation to your seller by giving a tip."
3. *Personalized Reciprocity*: "Leave (seller name) a tip to show your appreciation for a job well done."
4. *Norms*: "It's customary to leave a tip for the seller's service."[5]

A platform tipping norm is not specific to any individual seller. Therefore, we did not want to include the seller's name in the Norms message condition. Despite that, the removal of the seller's name was controversial within the company, which had historically emphasized usernames in an effort to promote the perception of a friendly, interpersonal marketplace. Therefore, we jointly decided to test two reciprocity messages, one with and one without the seller's name.

The Norms message signals an injunctive social norm, informing buyers of what they ought to do, namely, "leave a tip." Framing tipping as a "customary" injunctive norm conveys moral and social duty, appealing to buyers' shared standards and obligations. It does not report what most buyers currently do; that would be a descriptive norm (e.g., "most buyers leave tips when well satisfied", "31% of frequent buyers leave tips"; Cialdini et al. 1990).

Web Appendix C presents a manipulation check showing that treatment messages map onto theoretical constructs of reciprocity and norms, and that the seller's name inclusion did not change buyer

[5] The injunctive Norms message was motivated by the strong correlation between regional restaurant tipping norms, suggesting duty motives, and tipping (Web Appendix Figure B1). The reciprocity messages were motivated by the strong correlation between buyer satisfaction rating and tipping (Web Appendix Figure B2).

interpretations of the treatment messages. 1,000 online participants were presented with a 4-by-2 set of the four messages, with and without seller names, and asked to map messages to 6 implied reasons to tip, with reasons drawn from the persuasion typology of Cialdini (1993). The manipulation check suggests that participants perceived both the Impersonal Reciprocity and the Personalized Reciprocity messages as "Reciprocity" and the Norms message as "Social Proof (Norms)."[6] The interpretations of the three treatment messages were not affected by the inclusion or exclusion of the seller's name.

As previously mentioned, the platform simultaneously changed the default tips in two experimental treatment conditions. Web Appendix D shows that 90% of tipped transactions for gigs priced under the dollar/percentage threshold used default tips of $5 or $10, with the most common custom (non-default) tip being $15; therefore, the platform added a $15 default tip option in low-price gigs. Meanwhile, about 50% of tips on gigs above the dollar/percentage threshold used the 20% default, but less than 2% of tips used the 30% default. Therefore, the platform removed the 30% default and added default tips of 15% and 25%. We view these as limited changes in default tips and interpret their effects accordingly.[7] We refer to the two sets of default tips as *2-Defaults* and *3-Defaults*; the option to enter a Custom tip remained available in all cases.

The new default tips coincided with the Personalized Reciprocity and Norms tip request messages, but not with the Control or Impersonal Reciprocity tip request messages. Figure 3 illustrates how default tips corresponded to tip request messages within the Web Field Experiment. The Control and Impersonal Reciprocity conditions retained the original default options: $5, $10, and "Custom" for orders under $35, and 20%, 30%, and "Custom" for orders above $35. In contrast, the Personalized Reciprocity and Norms conditions offered options for $5, $10, $15, and "Custom" for orders under $35; and 15%, 20%, 25%, and "Custom" for orders above $35.[8] Figure 4 illustrates how the treatment conditions

[6] The Control message mapped weakly to five of seven answers (Reciprocity, Social Proof, Commitment, Liking, and N/A), with none indicated by more than 27% of respondents.
[7] Imagine that the popular $5 default was replaced by a $500 default; the default tip usage changes would be large because no one would use the new $500 default tip unless they chose it by mistake. Therefore, the potential default effect relates to the size of the default change.
[8] The optimal experiment design might have manipulated the default tips independently of the two tip request messages, but this was not possible due to the testing software constraint of four experimental cells, and firm executives' final authority over the design.

Figure 3: 2-Defaults and 3-Defaults Tips Menus.

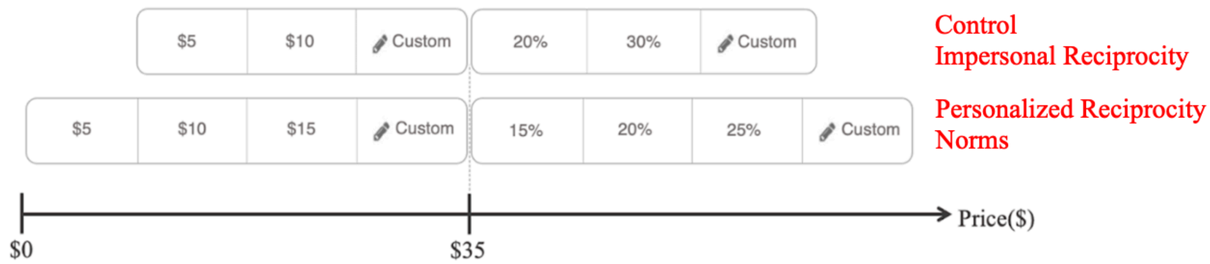


Figure 4: Tip Request Messages.

appeared to buyers, as Figure 1 displays the Control condition.

The platform also increased the dollar/percentage threshold from $25 to $35 during the test for all treatment conditions. The dollar/percentage threshold was set at $25 in 2010 when most Fiverr gigs cost $5. Later, Fiverr moved upmarket into more expensive services, and inflation eroded the value of the original dollar/percentage threshold, motivating the threshold increase to $35. This change affected both the 2-Default tips menu and the 3-Defaults tip menu.

4.2. *Data, Randomization, and Subsamples*

The Web Field Experiment ran on the web for a predetermined four weeks starting June 10, 2019. It treated all buyers who completed transactions through the website, including 7,880 new buyers and 45,497 buyers who had previously transacted.[9] No other tests were run during this period. Buyer accounts were randomly assigned to one of four persistent treatment conditions, so no buyer was exposed to

[9] The Post-Test Period contained a further 39,263 buyers who did not purchase during the Test Period. We include these non-treated buyers in all analyses involving cells T5-T8.

multiple conditions during the test period. Chi-square tests in Web Appendix Tables E1 and E2 confirm successful randomization, with observable buyer characteristics balanced across the four test conditions.

We report aggregate results, and also distinguish results within partitions based on buyer types (all buyers, new buyers only, and repeat buyers only) and based on exposures (all exposures vs. first exposures only). We define **new buyers** as purchasers whose first platform transaction involved a first exposure to a Web Field Experiment treatment condition. New buyers are less familiar with the platform, hence are likely to rely more on design elements such as tip request messages to guide their actions (Limayem, Hirt and Cheung 2007). We therefore expect that new buyers would be more influenced by tipping message treatments. **Repeat buyers** include all other buyers. Repeat buyers' prior experience may guide their actions, hence we expect that they may consider design elements less, and be less influenced by tipping treatments.

Relatedly, within each buyer type, we further distinguish observations according to **first exposures** vs. **all exposures**. The first exposure is defined within each buyer and time period as the first time the buyer encounters a particular combination of tip request message and default tips.[10] First-exposure manipulations were unexpected, novel stimuli occurring after all other transaction arrangements were completed (e.g., buyer's choice of seller, buyer/seller communication, payment, work completion confirmation, buyer rating), and therefore could not have influenced transaction existence or observable transaction features within first exposures. The all-exposures sample contains subsequent transactions whose parameters could conceivably have been influenced by the first exposure to treatment, but it offers a more complete picture of how treatments affect customer behaviors. We expect that treatments affect buyer tipping more in first exposures, since novelty is likely to draw attention and consequent elaboration.

[10] *First exposures* are defined within each combination of individual buyer, test condition, and time period. Hence, each buyer can have at most one *first exposure* per time period, including one corresponding to their first transaction during the test period, and another corresponding to their first transaction during the post-test period. All buyers were exposed to different treatments between the test and post-test periods, so every buyer was eligible to have a first exposure after the test period concluded.

4.3. Nonparametric treatment effects

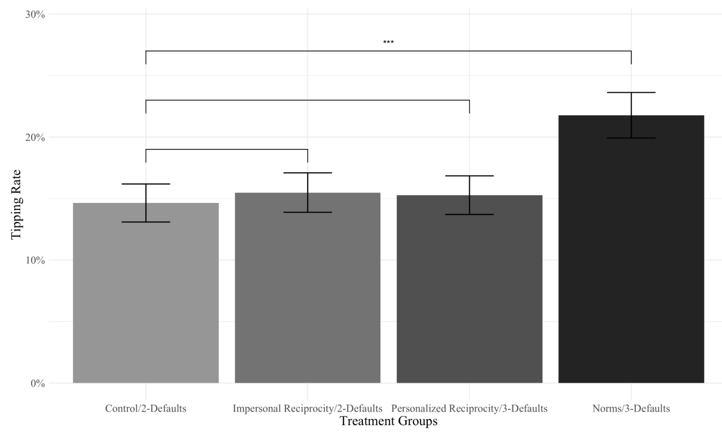
Figures 5a and 5b present the non-parametric effects of combined message/default-tips treatments on tipping in first exposures. We interpret all treatment effect sizes as absolute changes in percentage points, i.e. points on the percentage scale, which we abbreviate as "p.p." to distinguish them from relative changes. The Norms/3-Defaults condition increased new buyer tipping by 7.2 p.p. ($p<0.001$) upon first exposure, from 14.6% in Control to 21.8%. Among repeat buyers, the Norms/3-Defaults treatment increased tipping by 2.4 p.p. ($p<0.001$) upon first exposure, from 14.9% in Control to 17.3%. Tipping in the other two test conditions–Impersonal Reciprocity/2-Defaults and Personalized Reciprocity/3-Defaults–did not differ significantly from Control for either buyer type. Next, we explain how we identify tip request message effects separately from default tips effects.

4.4. Main Effect Identification and Specification

Table 1 summarizes the four treatment condition attributes and post-test changes described previously, including an identifier in each cell. The Web Field Experiment lasted a predetermined four weeks and randomized buyers into four conditions, each of which was characterized by a tip request message and a set of default tips. After the test concluded, all tip request messages reverted to the Control tip request message, and all default tips retained the new 3-Defaults tip menu, during the Post-Test Period.

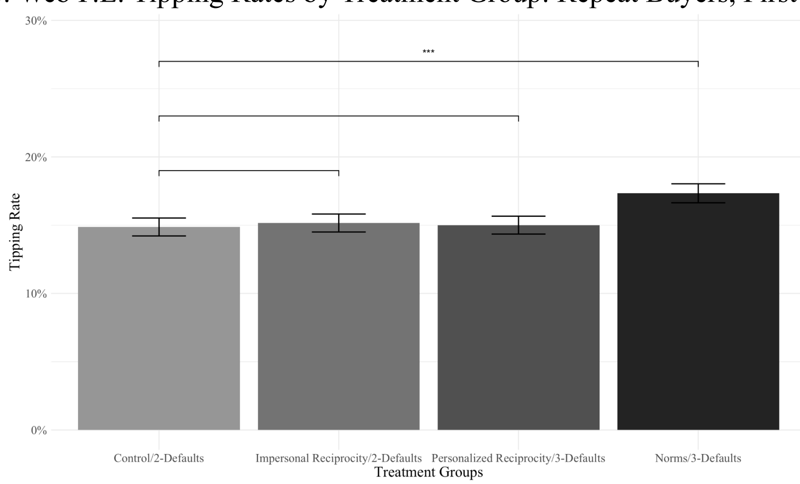
Deconfounding tip-request-message and default-tips effects requires an identification strategy that can separate the two. We use the natural experiment created by the test period conclusion, along with the post-test data, to separately identify message and default tips effects on tipping rates. The *3-Defaults* effect is identified by the comparison of the cells labeled T1 and T5 in Table 1, as both cells share the same tip request message. The *Impersonal Reciprocity* message effect is identified by the comparison of cells T1 and T2, as both cells share the same default tips. The *Personalized Reciprocity* message effect is identified by the comparison of cells T3 and T7, as both cells share the same default tips. The *Norms* message effect is identified by the comparison of cells T4 and T8, as both cells share the same default

Figure 5a. Web F.E. Tipping Rates by Treatment Group: New Buyers, First Exposures



Note. ** $p < 0.01$, *** $p < 0.001$.

Figure 5b. Web F.E. Tipping Rates by Treatment Group: Repeat Buyers, First Exposures



Note. ** $p < 0.01$, *** $p < 0.001$.

tips.[11]

Causal identification requires several assumptions. First, treatment effects on tipping depend on the current treatment only, and do not directly depend on past treatment exposures. Second, the timing of the test period was not chosen to coincide with time-varying unobservables. We are confident in this assumption as our involvement in the test design and implementation clearly indicated that timing decisions depended solely on the firm's internal testing calendar; those timing decisions preceded testing;

[11] We sincerely thank an anonymous reviewer for suggesting, and the editors for encouraging, this identification strategy and explanation of identifying assumptions.

Table 1: Platform Design Change After Web Field Experiment.

	Web Field Experiment Period	**Post-Test Period**
Control + 2-Defaults	Control Message + 2-Defaults & Custom (T1)	Control Message + 3-Defaults & Custom (T5)
Impersonal Reciprocity + 2-Defaults	Impersonal Reciprocity Message + 2-Defaults & Custom (T2)	Control Message + 3-Defaults & Custom (T6)
Personalized Reciprocity + 3-Defaults	Personalized Reciprocity Message + 3-Defaults & Custom (T3)	Control Message + 3-Defaults & Custom (T7)
Norms + 3-Defaults	Norms Message + 3-Defaults & Custom (T4)	Control Message + 3-Defaults & Custom (T8)

and the testing calendar was a queue that did not rely on recent business results or expected future outcomes.[12] Third, there should be no discontinuity in unobservables at the boundary between the test and post-test periods. Fourth, there should be no cross-buyer interference via shared sellers.

We pool all test and post-test data, and use a Linear Probability Model (LPM) to regress a tipping indicator on an intercept, a *3-Defaults* indicator, and three message indicators. The model clusters standard errors within buyers. We interpret results as statistically significant at 99% or 99.9% confidence levels to reduce Type I errors.

4.5. Treatment Effects on Tipping

Table 2 presents the average treatment effects on tipping rates. The intercept estimates show that baseline tipping rates are higher in first-exposure subsamples (14.8%) than in full-data subsamples (10.6%). This is partly because very frequent buyers tipped at lower average rates,[13] and are over-represented in the

[12] The reversion to the Control message and 3-Defaults tip menu was predetermined, as the analytics team did not begin to analyze the experimental data until after the test concluded.
[13] See Web Appendix Figure B4 for details.

Table 2. Treatment Effects on Tipping Rates (Web F.E.)

	(1) All Buyers, All Exposures	(2) All Buyers, First Exposures	(3) New Buyers, All Exposures	(4) New Buyers, First Exposures	(5) Repeat Buyers, All Exposures	(6) Repeat Buyers, First Exposures
Intercept	0.106***	0.148***	0.116***	0.146***	0.105***	0.149***
	(0.003)	(0.003)	(0.006)	(0.008)	(0.003)	(0.003)
Impersonal Reciprocity	0.000	0.004	−0.003	0.008	0.001	0.003
	(0.004)	(0.004)	(0.009)	(0.011)	(0.004)	(0.005)
Personalized Reciprocity	0.015***	0.046***	0.022***	0.053***	0.015***	0.044***
	(0.003)	(0.003)	(0.006)	(0.008)	(0.003)	(0.003)
Norms	0.030***	0.076***	0.066***	0.118***	0.025***	0.068***
	(0.003)	(0.003)	(0.008)	(0.010)	(0.003)	(0.004)
Three Defaults	−0.015***	−0.044***	−0.019**	−0.047***	−0.015***	−0.043***
	(0.003)	(0.003)	(0.006)	(0.008)	(0.003)	(0.003)
Num. Obs.	375,264	116,465	53,494	23,224	321,770	93,241
Num. Buyers	87,966	87,966	20,436	20,436	67,530	67,530
Num. Sellers	67,100	41,500	22,515	13,499	61,169	36,427
R2	0.001	0.007	0.004	0.012	0.001	0.006
Adj R-sq.	0.001	0.007	0.004	0.012	0.001	0.006

Note. **p<0.01; ***p<0.001

all-exposure samples. In-sample fit statistics are small, as should be expected with sparse outcome variables and parsimonious models without control variables.

Table 2 shows that the Impersonal Reciprocity tip request message did not significantly change tipping rates compared to Control, in any of the subsamples. By contrast, the Personalized Reciprocity tip request message significantly increased tipping rates relative to Control, with increases of 1.5 p.p. in the full sample, and 4.6 p.p. in the first-exposure sample with all buyers. These effects do not show significant differences between new buyers and repeat buyers, and cannot be explained by the inclusion of the seller's name alone, since the Control message also included the seller's name. The comparison between Impersonal Reciprocity and Personalized Reciprocity estimates suggests that reciprocity-motivated tip request messages should name the service provider, as this may help remind the buyer of the party that earned their gratitude and would receive their tip.

Table 2 also shows that the injunctive Norms message increased tipping substantially more than the Personalized Reciprocity message, with average tips rising 3.0 p.p. in the all-exposure sample and 7.6 p.p. in the first-exposure subsample. This effect was especially pronounced among new buyers, with effects of 6.6 p.p. in the all-exposure sample and 11.8 p.p. in the first-exposure subsample, both of which are significantly larger than the causal effects in the repeat-buyer subsamples of 2.5 p.p. and 6.8 p.p., respectively. It appears that new buyers are more responsive to information about socially desirable actions than repeat buyers, given their lower contextual familiarity, suggesting the impact of a Norms-driven message might be larger in a developing platform than a mature platform.

The shift from 2-Defaults to 3-Defaults reduced tipping rates by 1.5 p.p. in the full sample, and by 4.4 p.p. in the first-exposure sample. Reductions did not differ significantly between the new-buyer and repeat-buyer subsamples. Web Appendix Tables F1 and F2 show that the 3-Defaults effect also significantly reduced tip amounts in five out of six subsamples, and significantly reduced tip percentages in all six subsamples, with slightly smaller reductions relative to baseline than those in Table 2.

We can compare the causal effects of tip request messages and default tips. We do this with caution, as the default tips manipulation was relatively subtle. Therefore, this relative comparison should be interpreted as local to the range of the default tips manipulation. In the all-exposures sample, the Norms effect was approximately double the absolute 3-Defaults effect, with larger relative effects among new buyers and smaller relative effects among repeat buyers. The Personalized Reciprocity effect was similarly sized to the absolute 3-Defaults tips effect across subsamples. And finally, the Impersonal Reciprocity effect was nearly zero and hence much smaller than the default tips effect. We therefore see that tip request message effects may either be larger, smaller, or similar to default tips effect sizes. This pattern implies that a platform that wishes to increase tipping should consider tip request messages and default tips manipulations in concert and test combinations of each.

For robustness, we checked whether non-pooled regressions reproduce the pooled-regression main effects on tipping rates. Web Appendix Tables F3 and F4 show five two-cell comparisons that isolate individual treatment effects, all of which produce effects that are not significantly different from the main effects in Table 2. For example, when we use data from cells T1 and T5 in Table 1 to estimate a Three-Defaults effect, we find an estimate that is statistically indistinguishable from that in Table 2.[14]

[14] The other four two-cell comparisons identify (i) the Impersonal Reciprocity effect using only data in cells T1 and T2; (ii) the difference between Personalized Reciprocity and Norms effects using only data in cells T3 and T4; (iii)

As another robustness check, we estimated the model using a "donut" sample. Recall that separation of the 3-Defaults and Norms message effects relies on the platform's quasi-experimental change at the end of the test period. The "donut" sample drops the final week of the Web Field Experiment test period, and also drops the first week of the post-test period, and then uses the remaining data to estimate the pooled regressions. Trimming data around the policy change timing might reduce sensitivity to transitory unobservables, which can matter in discontinuity-in-time settings even when broader trends remain smooth (Hausman and Rapson 2018).[15] Web Appendix Table F5 shows that the treatment effect estimates based on "donut" samples are quite similar to those in Table 2.

Next, we describe the App Field Experiment design and main effects.

5. App Field Experiment

This section reports the mobile application A/B test. Section 5.1 describes the test design; Section 5.2 describes the test data, buyer randomization and subsamples; Section 5.3 describes identification and model specification; and Section 5.4 reports the main effects on Tipping Rate. We maintain as much analytical consistency with the Web Field Experiment as possible.

5.1. *Design*

The App Field Experiment was a simple A/B test of the Norms tip request message in Figure 4 vs. the legacy Control message in Figure 1. Both conditions included the 3-Default tips menu. The experiment ran exclusively on mobile app transactions from October 17 to December 10, 2019. The platform refrained from running other tests in the app to avoid potential confounds. The Norms tip request message had previously been adopted on all Web transactions on August 1, 2019, and therefore was familiar to

the Personalized Reciprocity effect using only data in cells T3 and T7; and (iv) the Norms effect using only data in cells T4 and T8.

[15] Donut-style sensitivity analyses are common in discontinuity designs because they reveal whether estimates hinge on observations closest to the cutoff; similar results after trimming strengthens the interpretation that the post-test comparison isolates message effects from the default-tips menu change (Imbens and Lemieux 2008; Lee and Lemieux 2010).

most repeat buyers in the app, as 60% had transacted on the web since Norms message adoption.

5.2. *Data, Randomization and Subsamples*

The test treated 61,705 app transactions by 9,570 new buyers and 25,910 repeat buyers. As before, persistent randomization was carried out at the buyer level, and treatments within the app were held constant across all subsequent transactions during the test period. Randomization checks in Web Appendix Tables E3 and E4 confirm balanced buyer characteristics between groups.

As before, we report results within the full data; within new buyers, defined as purchasers whose first platform transaction involved a first exposure to an App Field Experiment treatment; and within repeat buyers, which include all other buyers. It is important to note that our usage of the term "new buyers" is specific to each field experiment; hence, "new buyers" during the Web Field Experiment are classified as "repeat buyers" in the App Field Experiment. We further distinguish first-exposure samples from all-exposure samples within all three buyer groups.

5.3. *Identification and Estimation of Main Effects*

Causal effect identification relies simply on the comparison of average tipping rates in the randomized Treatment and Control groups. We analyze test-period data exclusively. We use a Linear Probability Model (LPM) to regress tipping rate on an intercept and a Norms message indicator, to maintain consistency with Section 4. The model clusters standard errors within buyers. As before, we interpret results as statistically significant at 99% or 99.9% confidence levels to reduce Type I errors.

5.4. *Treatment Effects on Tipping*

Table 3 reports that the Norms tip request message increased tipping rates by 1.7 p.p. in the all-buyers, first-exposures subsample, and by 1.1 p.p. in the all-buyers, all-exposures sample. As before, the treatment effects on new buyers were larger than on repeat buyers, with first-exposure effects on new buyers of 2.5 p.p. and first-exposure effects on repeat buyers of 1.4 p.p. These effects are all

Table 3. Treatment Effects on Tipping Rates (App F.E.)

	(1) All Buyers, All Exposures	(2) All Buyers, First Exposures	(3) New Buyers, All Exposures	(4) New Buyers, First Exposures	(5) Repeat Buyers, All Exposures	(6) Repeat Buyers, First Exposures
Intercept	0.075***	0.084***	0.088***	0.095***	0.070***	0.080***
	(0.002)	(0.002)	(0.005)	(0.004)	(0.002)	(0.002)
Norms	0.011***	0.017***	0.017**	0.025***	0.010**	0.014***
	(0.003)	(0.003)	(0.006)	(0.006)	(0.003)	(0.004)
Num. Obs.	61,705	35,480	15,194	9,570	46,511	25,910
Num. Buyers	35,480	35,480	9,570	9,570	25,910	25,910
Num. Sellers	26,562	18,327	9,074	6,273	22,637	15,233
R2	0.000	0.001	0.001	0.002	0.000	0.001
Adj R-sq.	0.000	0.001	0.001	0.002	0.000	0.001

Note. **p<0.01; ***p<0.001

proportionally smaller than those in the Web Field Experiment, but the qualitative patterns across subsamples are consistent. Web Appendix Tables G1 and G2 show that the significant increases in tipping rates are partially corroborated, and not contradicted, by analyses that explain tip amounts or tip percentages.

We see two likely reasons why the treatment effect sizes differed between the two field experiments. One is that mobile devices' smaller screens may make it easier for buyers to disregard a tip request message. Another possible reason is that buyers or transactions may differ between channels, for example, web buyers may be more professional or more wealthy on average than app buyers, or buyers may use the web to specify complex services and may use the app to specify simpler services. It is not unusual for buyers to differ systematically in different digital channels (Liu et al. 2019).

The App Field Experiment directly confirms that the Norms tip request message treatment independently and significantly increased tipping rates in all subsamples, with more pronounced effects among new buyers and in first exposures, though the treatment effects were smaller than in the Web Field Experiment. Section 6 analyzes treatment exposure consequences and correlates in both test datasets.

6. Beyond Main Effects: Platform Implications of Norms-related Treatments

In this section we consider what we can learn from the data beyond the main effects on tipping rates. Section 6.1 shows that norms-related treatments significantly increased buyers' tips without significantly reducing buyers' total orders or total spending, suggesting that the Norms-induced revenue does not come at the expense of subsequent business. Section 6.2 partitions the test data by order price, showing that treatment effects on tipping generally increase with order price. Section 6.3 uses treatment assignments as instrumental variables to estimate how Norms-induced tipping changed buyers' and sellers' subsequent actions on the platform, but does not produce reliable evidence due to limited power. Section 6.4 reports interactions between treatments and transaction features to estimate when treatments increased tipping most. All four subsections present results from both field experiments to facilitate comparisons. These findings may help other platforms decide whether to test an injunctive norms message in other contexts.

6.1. *Treatment Effects on Orders and Spending*

An important question for platforms to evaluate before adopting injunctive Norms tip request messages: Do tips discourage future business? Table 4a reports causal effects of Norms / 3-Defaults assignments vs. Control in the Web Experiment on buyers' order count, total spending, total tip spending, and tip count, in all-buyers' and new-buyers' first-exposure subsamples. Table 4b shows similar causal effects of Norms treatment assignment in the App Field Experiment. The qualitative patterns are mostly consistent across the two tables: Norms-related treatments did not significantly change buyers' order counts or total spending compared to controls. It did, however, increase tip count and total spending on tips. We therefore do not detect economic disincentives to adopt injunctive Norms tip request messages.

6.2. *Treatment Effects Within Order Price Partitions*

Default tips were presented as dollar values for gig prices below the $35 dollar/percentage threshold, and as percentage values above the threshold. We wondered if the discontinuity at the dollar/percentage threshold corresponded to treatment effect differences. We further wondered how Norms-related treatment effects covaried with price both above and below the dollar/percentage threshold. Therefore, we partitioned each field experiment's dataset at $10, $35, and $60, and re-estimate the treatment main effects within each partition, with results reported in Tables 5a and 5b.

Table 4a. Web F.E. Norms / 3-Defaults Treatment Effects on Orders, Spending and Tips

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Order Count All Buyers	Order Count New Buyers	Total Spend All Buyers	Total Spend New Buyers	Total Tip Spending All Buyers	Total Tip Spending New Buyers	Tip Count All Buyers	Tip Count New Buyers
Intercept	4.043***	3.282***	140.559***	104.219***	4.915***	3.884***	0.430***	0.380***
	(0.046)	(0.060)	(2.729)	(4.002)	(0.163)	(0.256)	(0.010)	(0.020)
Norms/	−0.023	0.169	1.634	17.876	0.560	2.261***	0.055***	0.181***
3Defaults	(0.063)	(0.092)	(3.643)	(6.987)	(0.237)	(0.465)	(0.014)	(0.033)
Observations	26,644	3,914	26,644	3,914	26,644	3,914	26,644	3,914
R^2	0.000	0.001	0.000	0.002	0.000	0.006	0.001	0.008
Adjust R^2	-0.000	0.001	-0.000	0.001	0.000	0.006	0.001	0.008

Note. **p<0.01; ***p<0.001 New buyers sample (N=3,914) includes only Control (N=2,008) and Norms/3-Defaults (N=1,906) conditions.

Table 4b. App F.E. Norms Treatment Effects on Orders, Spending and Tips

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Order Count All Buyers	Order Count New Buyers	Total Spend All Buyers	Total Spend New Buyers	Total Tip Spending All Buyers	Total Tip Spending New Buyers	Tip Count All Buyers	Tip Count New Buyers
Intercept	1.733***	1.592***	62.827***	61.786***	1.289***	1.362***	0.130***	0.141***
	(0.013)	(0.021)	(0.905)	(1.644)	(0.046)	(0.085)	(0.003)	(0.008)
Norms	0.012	−0.009	1.271	−1.623	0.193**	0.329	0.021***	0.026
	(0.018)	(0.028)	(1.252)	(2.189)	(0.065)	(0.131)	(0.005)	(0.010)
Observations	35,480	9,570	35,480	9,570	35,480	9,570	35,480	9,570
R^2	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001
Adjust R^2	-0.000	-0.000	0.000	-0.000	0.000	0.001	0.000	0.001

Note. **p<0.01; ***p<0.001

There is a general positive relationship between order prices and larger Norms treatment effects, but the patterns differ between the two field experiments. Table 5a shows that cheap Web transactions had much smaller treatment effects below $10, and that increasing price beyond that shows a more muted relationship with treatment effect estimates. Table 5b, by contrast, shows that expensive App transactions had much larger treatment effect estimates, whereas the positive relationship is more muted across lower price partitions.

The 3-Defaults effects on Web Tipping Rates were more negative in the two high-price Web Field Experiment partitions than in the two low-price partitions. This finding coheres with Bluvstein and Raghubir (2022), who found that default tips convey injunctive normative guidance at low prices, but exert less influence on consumer decisions at higher prices when tips are more expensive and therefore receive greater consumer elaboration. It is therefore quite interesting that an injunctive norm tip request

Table 5a. Web F.E. Treatment Effects on Tipping Rates by Price Partition

	P ≤ 10		**10 < P ≤ 35**		**35 < P ≤ 60**		**P > 60**	
	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.107***	0.100***	0.169***	0.160***	0.190***	0.213***	0.187***	0.198***
	(0.004)	(0.010)	(0.006)	(0.014)	(0.010)	(0.026)	(0.009)	(0.025)
Impersonal	0.008	0.016	0.008	0.031	0.003	-0.052	-0.021	-0.019
Reciprocity	(0.006)	(0.015)	(0.008)	(0.021)	(0.014)	(0.035)	(0.012)	(0.034)
Personalized	0.037***	0.038***	0.050***	0.054***	0.068***	0.098***	0.055***	0.069**
Reciprocity	(0.004)	(0.011)	(0.006)	(0.016)	(0.010)	(0.026)	(0.009)	(0.026)
Norms	0.046***	0.067***	0.090***	0.140***	0.102***	0.201***	0.109***	0.157***
	(0.005)	(0.012)	(0.007)	(0.018)	(0.011)	(0.031)	(0.010)	(0.027)
Three Defaults	-0.033***	-0.039***	-0.050***	-0.036	-0.058***	-0.085**	-0.064***	-0.084***
	(0.004)	(0.011)	(0.006)	(0.015)	(0.010)	(0.027)	(0.009)	(0.025)
Num. Obs.	46,274	9,011	37,644	7,538	14,929	3,138	17,618	3,537
Num. Buyers	39,309	8,306	33,796	7,163	14,062	3,060	16,038	3,377
Num. Sellers	22,051	6,204	19,244	5,286	9,543	2,405	10,068	2,572
R2	0.005	0.008	0.008	0.013	0.011	0.026	0.011	0.019
Adj R-sq.	0.005	0.008	0.008	0.013	0.010	0.024	0.011	0.018

Note. **p<0.01; ***p<0.001

Table 5b. App F.E. Treatment Effects on Tipping Rates by Price Partition

	P ≤ 10		**10 < P ≤ 35**		**35 < P ≤ 60**		**P > 60**	
	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures	All Buyers, First Exposures	New Buyers, First Exposures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.055***	0.058***	0.093***	0.106***	0.113***	0.133***	0.105***	0.101***
	(0.003)	(0.006)	(0.004)	(0.007)	(0.006)	(0.013)	(0.006)	(0.011)
Norms	0.013**	0.019	0.013	0.018	0.017	0.013	0.034***	0.066***
	(0.004)	(0.009)	(0.005)	(0.011)	(0.009)	(0.018)	(0.009)	(0.017)
Num. Obs.	12,264	2,932	12,933	3,547	4,982	1,537	5,301	1,554
Num. Buyers	12,264	2,932	12,933	3,547	4,982	1,537	5,301	1,554
Num. Sellers	7,811	2,327	8,112	2,677	3,614	1,213	3,640	1,178
R2	0.001	0.001	0.001	0.001	0.001	0.000	0.003	0.009
Adj R-sq.	0.001	0.001	0.000	0.001	0.000	−0.000	0.002	0.009

Note. **p<0.01; ***p<0.001

message can have larger effects at higher prices, possibly compensating for the declining injunctive effects of default tips at higher prices.

6.3. *Treatment-Induced Tipping Effects on Subsequent Buyer/Seller Platform Behaviors*

We examined whether treatment-induced tipping affected subsequent buyer and seller behaviors. That is, did buyers who tipped because of Norms-related treatments act differently in subsequent transactions? Direct estimation of Norms-driven tipping on subsequent actions would yield biased estimates due to persistent unobservables that drive tipping (e.g., buyer budget, buyer tipping habit, seller effort, seller quality). We therefore use instrumental variables estimated with Two-Stage Least Squares (2SLS). We use buyer assignments to treatment conditions as an instrument for tipping. The instrument satisfies the first-stage relevance condition, as indicated by the main effects in Tables 2 and 3. The instrument satisfies the second-stage exclusion restrictions under the assumption that treatment assignments did not directly affect subsequent actions except via immediate treatment effects on tipping.

We consider the effects of treatment-induced tipping on three subsequent buyer actions during the test period: any buyer repurchase, subsequent total spending, and subsequent count of five-star ratings; all defined within the relevant test period.[16] We also consider effects on four subsequent seller actions, in each case drawing the seller's instrument from the treatment assigned to their first treated buyer during the test period. The four subsequent seller actions are subsequent gigs completed, subsequent total earnings, subsequent count of five-star ratings earned, and count of subsequent seller tip mentions (i.e., gigs with on-platform messages from sellers to buyers containing "tip" or "tips"); again, all defined within the relevant test period. We focus this analysis on the first-exposure data only, as first exposures were exogenous to agents' self-selected actions.

Web Appendix H reports instrumental-variables estimates in all-buyers and new-buyers samples, for both field experiments, for all seven subsequent-behavior variables. First-stage F-statistics indicate strong instruments when they exceed 16.38, a value that prevents first-stage weakness from overly distorting second-stage test sizes in the case of one endogenous variable and one instrument (Dufour 1997, Nguyen 2025, Staiger and Stock 1997), a case that applies to four out of eight agent/channel/buyer-type combinations. However, for all eight cases, we do not find evidence that Norms-related treatment-induced tipping detectably altered subsequent behaviors. Many of these null effects are imprecisely estimated, as the standard errors tend to be sizeable compared to the baselines.

6.4. Which other transaction features predict larger Norm-related effects?

Here we estimate which observable transaction characteristics predicted greater or lesser tipping in

[16] We cannot identify effects on subsequent tips in this analysis because the treatment assignment was persistent, so excluding the instrument would mis-specify the second stage equation.

Norms-related treatments vs. Control, in hopes of pointing toward potential mechanisms. An important caveat is that transaction characteristics were not randomly assigned, and hence may correlate with unobserved drivers of tipping such as buyer budget or seller effort, so these interaction effects cannot rule out alternate explanations (Harrison and List 2004; Bullock, Green and Ha 2010; Imai et al. 2011). However, we do think the results may indicate potential drivers of tipping response to treatment, which could be investigated with further experimentation. We examine the following transaction features.

1. Regional tipping culture: Regional tipping norms are the strongest documented determinant of offline tipping (Azar 2007, 2011, 2020). Web Appendix Figure B1 shows that buyers in North America tipped at much higher rates than buyers elsewhere, even when interacting with sellers from other regions. We operationalize this using a binary indicator for regional tipping culture, i.e. North America vs. other regions.
2. Seller reputation and buyer satisfaction: Customers for offline services tip more when they receive high-quality service (Lynn and McCall 2000, Lynn 2015). On digital platforms, buyers may predict service quality *ex-ante* from a seller's average rating, as reputation systems signal seller quality and influence purchase decisions (Yoganarasimhan 2013); we code high seller reputation with an indicator of above-median rating count and rating average. Buyers also assess service quality *ex post*; we code this with a five-star rating indicator.

Tests do not indicate multicollinearity among these three indicators, as their bivariate correlations are all 0.05 or less, so we include them in a single empirical framework. We estimate models including main norms-related treatment effects, main effects of each transaction feature above, and interactions between transaction features and norms-related treatments, with buyer-clustered standard errors.

Web Appendix Table I1 reports how the Norms treatment effects covary with the three potential moderator variables, with estimates from the first-exposure samples of all buyers and new buyers in each field experiment. The web data show two significant results. Norms-related treatment effects were larger in the all-buyers, first-exposures sample among North American buyers. They also were larger in the new-buyers, first-exposures sample when buyers rated the transaction as 5-stars. Seller reputation did not significantly alter Norms effect sizes on tipping. The app data do not show any significant interactions.

We also checked whether the Norms treatment interacted with gig service categories to predict tipping, as average tipping norms differ across service contexts,[17] and as average tipping rates vary across gig categories in Web Appendix Figure B7. We did not find significant differences in Norms treatment effects across gig categories, due to limited statistical power.

7. Discussion

We offer the first large-scale field experiments to identify how tip request messages motivate tipping. We find that an injunctive norms message generated an economically large, statistically significant increase in tipping rates compared to Control, and that the Personalized Reciprocity tip request message had smaller causal effects on tipping. Remarkably, the treatment effects do not support a business case against injunctive norm tip request messages. Although the tip request message effects are large relative to baseline tipping rates, the most transactions remained untipped, suggesting that a tip request message by itself is insufficient to establish a regular tipping norm. We close with implications for practice, and how the limitations of our research might motivate future research.

7.1. *Implications*

The most significant implications of our research are the findings that online platforms can utilize marketing communications to influence buyers' voluntary payments, even without face-to-face interactions between customers and service providers, and without reliably diminishing customers' future orders or spending. The Norms message influenced tipping more in transactions that have high prices and five-star ratings. High prices are situations when default tips convey less influence, suggesting the tip request message can be an alternate means of influence in cases when default tips are less effective (Bluvstein and Raghubir 2022).

[17] For example, 81% of survey respondents say they always tip in sit-down restaurants; 65% always tip when getting a haircut; 53% always tip when buying a drink at a bar; and 12% always tip when buying a beverage at a coffee shop (Pew Research 2023; see also Zuluaga 2024).

When tips are infrequent, sellers may struggle to infer whether a tip signals buyer satisfaction, the buyer's general tipping tendency, the platform's encouragement, or other factors. Therefore, sellers may struggle to use tips as reliable performance feedback. These results imply that platforms seeking marketplace-wide welfare gains should pair norm-based prompts with complementary seller-focused interventions, such as enabling them to see buyers' past tipping rates prior to entering into transactions. An example might be a statement like "this buyer tipped on 50% of their past transactions" to inform sellers about the potential returns to service quality and customization effort, mirroring how existing platforms typically enable buyers to view sellers' historical ratings. The same information could be publicized to the buyer as well, helping them to understand how they appear to potential sellers and further encouraging normative behavior.

Our findings can help to inform the evolving regulatory agenda around digital tip requests, especially in app-based marketplaces. For example, New York City rules require delivery platforms to request tips at or before checkout and to include a default tip of at least 10% (NYC Consumer and Worker Protection 2025), and Doordash and Uber Eats changed their digital tip requests in response to a 2023 New York City law requiring minimum wages for delivery workers (McCarthy 2023). We find that tip request message and default tips effects on tipping covary with order prices and buyer experience on the platform, showing that careful testing is needed to balance the delicate interplay between tip request messages, default tips, consumer response, and worker compensation. If regulations mandate particular default tips or tip request messages without testing, they may unwittingly decrease workers' tips earned.

7.2. *Limitations and Future Research Directions*

Large-scale field experiments provide rich, ecologically valid insights, but several constraints limit our studies and suggest future research opportunities. First, the results indicate that the Norms message can increase online tipping; however, we tested only one specific injunctive norms message. There is room to explore how other norms messages, among other messaging strategies like reciprocity or fairness, affect short-run and long-run tipping behaviors. We suspect descriptive-norm tip request messages could increase tipping in contexts with higher baseline tipping rates (Paridar et al. 2025), or tip request messages could be personalized to infrequent tippers to help them better understand descriptive tipping norms. Other messaging strategies could include informing buyers about how tips benefit sellers, how tips motivate service quality, or how tips help to maintain low prices on the platform (Lynn and Withiam 2008). It also may be interesting to study rotating tip request messages or quality-contingent messages, to see whether dynamic messages could increase tip response by attracting buyer attention, or whether there is scope to personalize tip request messages based on buyer attributes or behaviors.

Second, we did not intervene on the seller side of the platform. Future research could investigate sellers' awareness of, and reactions to, buyer-side tipping interventions, as well as explore complementary seller-facing interventions. For example, a platform could directly inform sellers about the empirical relationships between buyer satisfaction and tips, thereby illustrating how seller quality investments affect earnings. Experiments could even be designed with two-sided treatments and to identify two-sided feedback effects between buyers and sellers, for example by cluster-randomizing treatments across submarkets on the platform (e.g., gig types, native languages, etc.) rather than limiting tests to one side of the platform (e.g., buyers or sellers alone). Full-time gig workers and occasional gig workers may respond differently to tips, as they may differ in important ways, such as their sensitivity to tips received, or their consistency of quality service provision.

Finally, we observe customer non-tipping and examine repatronage, but we did not directly enable buyers to submit negative feedback about tip request messages. It remains to be seen how buyers' affective reactions to tip request messages predicts their tipping and repatronage decisions, as negative feelings of coercion or guilt could offset desired effects of tip request messages.

Overall, we believe that, so long as buyers remain free to easily decline to tip, tip requests on digital platforms have the potential to benefit all parties by motivating and rewarding seller effort and service quality. Gratuities can even benefit non-tipping buyers by helping to keep service prices low. We have shown that the tip request message itself can influence the buyer's decision to tip or not, and therefore recommend that platforms test this marketing communication carefully to balance all parties' interests. We are confident the industry will continue its momentum toward improving its tip request messages and understanding how to measure their larger consequences for quality provision, buyer satisfaction, market competition and gig worker earnings.

References

Alexander, Damon, Christopher Boone, and Michael Lynn. 2021. "The Effects of Tip Recommendations on Customer Tipping, Satisfaction, Repatronage, and Spending." *Management Science* 67 (1): 146–165.

Anderson, Monica, Colleen McClain, Michelle Faverio, and Risa Gelles-Watnick. 2021. "The State of Gig Work in 2021." Washington, DC: Pew Research Center.

Azar, Ofer H. 2007. "The Social Norm of Tipping: A Review." *Journal of Applied Social Psychology* 37 (2): 380–402.

Azar, Ofer H. 2011. "Business Strategy and the Social Norm of Tipping." *Journal of Economic Psychology* 32 (3): 515–525.

Azar, Ofer H. 2020. "The Economics of Tipping." *Journal of Economic Perspectives* 34 (2): 215–236.

Bekkers, René, and Pamala Wiepking. 2011. "A Literature Review of Empirical Studies of Philanthropy: Eight Mechanisms That Drive Charitable Giving." *Nonprofit and Voluntary Sector Quarterly* 40 (5): 924–973.

Bluvstein Netter, Shirley, and Priya Raghubir. 2021. "Tip to Show Off: Impression Management Motivations Increase Consumers' Generosity." *Journal of the Association for Consumer Research* 6 (1): 120–129.

Bluvstein, Shirley, and Priya Raghubir. 2022. "Nothing Matters: A '0' Tip Option Increases Consumers' Voluntary Payments." Working paper, available at SSRN.

Bullock, John G., Donald P. Green, and Shang E. Ha. 2010. "Yes, but What's the Mechanism? (Don't Expect an Easy Answer)." *Journal of Personality and Social Psychology* 98 (4): 550–558.

Cabano, Frank G., and Amin Attari. 2023. "Don't Tell Me How Much to Tip: The Influence of Gratuity Guidelines on Consumers' Favorability of the Brand." *Journal of Business Research* 159 (April): 2055–2071.

Carr, Austin. 2013. "How Square Register's UI Guilts You into Leaving Tips." *Fast Company*, Dec. 12. https://web.archive.org/web/20141220073029/http://www.fastcodesign.com/3022182/innovation-by-design/how-square-registers-ui-guilts-you-into-leaving-tips

Chandar, Bharat, Uri Gneezy, John A. List, and Ian Muir. 2019. "The Drivers of Social Preferences: Evidence from a Nationwide Tipping Field Experiment." NBER Working Paper 26380. Cambridge, MA: National Bureau of Economic Research.

Charness, Gary, Ernan Haruvy, and Doron Sonsino. 2007. "Social Distance and Reciprocity: An Internet Experiment." *Journal of Economic Behavior and Organization* 63 (1): 88–103.

Chen, Jinjie, Alison Jing Xu, Maria A. Rodas, and Xuefeng Liu. 2023. "Order Matters: Rating Service Professionals First Reduces Tipping Amount." *Journal of Marketing* 87 (1): 81–96.

Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren. 1990. "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places." *Journal of*

Personality and Social Psychology 58 (6): 1015–1026.
Cialdini, Robert B. 1993. Influence: The Psychology of Persuasion. Rev. ed. New York: Morrow.
Cohn, Alain, Tobias Gesche, and Michel André Maréchal. 2022. "Honesty in the Digital Age." *Management Science* 68 (2): 827–845.
Donkor, Kwabena. 2021. "The Economic Value of Norm Conformity and Menu Opt-Out Costs." Working paper, Stanford Graduate School of Business.
Dufour, Jean-Marie. 1997. "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models." *Econometrica* 65 (6): 1365–87.
Duhaime, Erik P., and Zachary W. Woessner. 2019. "Explaining the Decline of Tipping Norms in the Gig Economy." *Journal of Managerial Psychology* 34 (4): 233–245.
Dyussembayeva, Shynar, Giampaolo Viglia, Marta Nieto-García, and Anna S. Mattila. 2022. "Would You like to Add a Gratuity? When Explicit Requests Hamper Tipping." *Journal of Business Research* 139 (February): 908–917.
Fennell, Lee Anne. 2023. "Optional Price Discrimination." *Texas A&M Law Review* 10 (3): 485–548.
Fiverr International Ltd. 2024. "Fiverr Announces Fourth Quarter and Full Year 2023 Results." Press release, February 22. Accessed April 1, 2024.
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): 7236–7240.
Haggag, Kareem, and Giovanni Paci. 2014. "Default Tips." *American Economic Journal: Applied Economics* 6 (3): 1–19.
Hanbury, Mary. 2019. "Instacart Workers Are Fighting Back against a Policy Change They Say Drastically Cut Their Wages." *Business Insider*, January 20. Accessed February 2, 2022.
Harrison, Glenn W., and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–1055.
Hausman, Catherine, and David S. Rapson. 2018. "Regression Discontinuity in Time: Considerations for Empirical Applications." *Annual Review of Resource Economics* 10 (1): 533–552.
Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105 (4): 765–789.
Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615–635.
Karabas, Ismail, Marissa Orlowski, and Sarah Lefebvre. 2020. "What Am I Tipping You For? Customer Response to Tipping Requests at Limited-Service Restaurants." *International Journal of Contemporary Hospitality Management* 32 (5): 2007–2026.
Kilroy, Ashley. 2024. "How Much to Tip Your Uber or Lyft Driver." AAA Club Alliance, September 25. Accessed May 14, 2025.
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.
Lapidot-Lefler, Noam, and Azy Barak. 2012. "Effects of Anonymity, Invisibility, and Lack of Eye Contact on Toxic Online Disinhibition." *Computers in Human Behavior* 28 (2): 434–443.
Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.
Lefebvre, Sarah, Laura Boman, and Marissa Orlowski. 2024. "Look on the Bright Side: Emojis Impact

Tipping Behaviour." *International Journal of Hospitality Management* 117 (February).
Limayem, Moez, Sabine Gabriele Hirt, and Christy M. K. Cheung. 2007. "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance." *MIS Quarterly* 31 (4): 705–737.
Liu, Huan, Lara Lobschat, Peter C. Verhoef, and Hong Zhao. 2019. "App Adoption: The Effect on Purchasing of Customers Who Have Used a Mobile Website Previously." *Journal of Interactive Marketing* 47: 16–34.
Lynn, Michael. 2015. "Service Gratuities and Tipping: A Motivational Framework." *Journal of Economic Psychology* 46: 74–88.
Lynn, Michael. 2016. "Motivations for Tipping: How They Differ across More and Less Frequently Tipped Services." *Journal of Behavioral and Experimental Economics* 65: 38–48.
Lynn, Michael. 2017. "Should US Restaurants Abandon Tipping? A Review of the Issues and Evidence." *Psychosociological Issues in Human Resource Management* 5 (1): 120–159.
Lynn, Michael, and Michael McCall. 2000. "Gratitude and Gratuity: A Meta-Analysis of Research on the Service-Tipping Relationship." *Journal of Socio-Economics* 29 (2): 203–214.
Lynn, Michael, and Glenn Withiam. 2008. "Tipping and Its Alternatives: Business Considerations and Directions for Research." *Journal of Services Marketing* 22 (4): 328–336.
McCarthy, Kelly. 2023. "DoorDash, Delivery Apps Remove Tipping Prompt at Checkout in NYC." Good Morning America. ABC News. December 7, 2023. https://www.goodmorningamerica.com/food/story/doordash-delivery-apps-remove-tipping-prompt-checkout-nyc-105461852 Accessed December 2025.
Newman, Andy. 2019. "DoorDash Changes Tipping Model after Uproar from Customers." *New York Times*, July 24. Accessed July 24, 2019.
Nguyen, Mike. 2025. *Foundations of Data Analysis. Vol. 1*. Cham: Springer.
NYC Consumer and Worker Protection. 2025. "Requirements for Delivery Apps." https://www.nyc.gov/site/dca/businesses/Delivery-Apps-Requirements.page. Accessed December 2025.
Paridar, Mahsa, Mina Ameri, and Elisabeth Honka. 2025. "Online Tipping under an Evolving Social Norm: Implications for Platform Design." Working paper. Available at SSRN 5020156.
Pew Research Center. 2023. "2023 Pew Research Center's American Trends Panel, Wave 133, Final Topline, August 7–27, 2023." Washington, DC: Pew Research Center.
Raghubir, Priya, and Shirley Bluvstein. 2024. "From Bribes to Bequests and Gifts to Gratuities: The Black, White, and Shades of Gray of How and Why Consumers Pay What They Want." *Consumer Psychology Review* 7 (1): 75–92.
Rind, Bruce, and Prashant Bordia. 1995. "Effect of Server's 'Thank You' and Personalization on Restaurant Tipping." *Journal of Applied Social Psychology* 25 (9): 745–751.
Rind, Bruce, and Prashant Bordia. 1996. "Effect on Restaurant Tipping of Male and Female Servers Drawing a Happy, Smiling Face on the Backs of Customers' Checks." *Journal of Applied Social Psychology* 26 (3): 218–225.
Rind, Bruce, and David Strohmetz. 1999. "Effect on Restaurant Tipping of a Helpful Message Written on the Back of Customers' Checks." *Journal of Applied Social Psychology* 29 (1): 139–144.
Saxena, Esha. 2025. "Fiverr in 2025: Key Stats, Trends, and Insights You Need to Know." *GrabOn Blog*, June 24. Accessed May 14, 2025.
Schmidt, Klaus M., Martin Spann, and Robert Zeithammer. 2015. "Pay What You Want as a Marketing

Strategy in Monopolistic and Competitive Markets." *Management Science* 61 (6): 1217–1236.
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.
Staiger, Douglas, and James H Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3): 557–86.
van Teunenbroek, Claire, Carolina Dalla Chiesa, and Laura Hesse. 2023. "The Contribution of Crowdfunding for Philanthropy: A Systematic Review and Framework of Donation and Reward Crowdfunding." *Journal of Philanthropy and Marketing* 28 (3): e1791.
Wolfe, Rachel. 2023. "Tipping at Self-Checkout Has Customers Crying 'Emotional Blackmail'." *Wall Street Journal*, May 9. Accessed April 17, 2025.
Yoganarasimhan, Hema. 2013. "The Value of Reputation in an Online Freelance Marketplace." *Marketing Science* 32 (6): 860–891.
Zipperer, Ben, Celine McNicholas, Margaret Poydock, Daniel Schneider, and Kristen Harknett. 2022. "National Survey of Gig Workers Paints a Picture of Poor Working Conditions, Low Pay." Washington, DC: Economic Policy Institute. Accessed September 23, 2024.
Zuluaga, Tessa. 2024. "How Restaurant Guests Really Feel about Tipping in America." *Toast: On the Line*. Accessed November 12, 2025.

Web Appendix for "Gratuities in a Digital Services Marketplace" by Seung Hyun Kim, On Amir, and Kenneth C. Wilbur (2025)

Web Appendix A. Overview of Research on Voluntary Payments: Contexts, Designs, and Findings

Paper	Payment Type	Research Design (Context)	Effect Tested	Main Findings
Kim et al. (2009)	PWYW	*Field Experiments (Offline)* - Restaurant (N = 253, Survey N = 172) - Cinema (N = 386, Survey N = 247) - Deli (N = 813, Survey N = 271)	Fairness; Altruism; Loyalty; Price consciousness; Income; Satisfaction; Reference Price	In PWYW, consumers paid based on their reference price and the proportion they willingly shared with sellers, driven mainly by fairness, satisfaction, price consciousness, and income, while altruism and loyalty mattered only in specific contexts.
Gneezy et al. (2012)	PWYW	*Field Experiments (Offline)* - Theme park (N = 54,231) - Tour boat (N = 20 cruises × ≥ 50 groups per treatment) - Restaurant (N = 257)	Identity; Self-image consideration	In PWYW, customers prioritize self-image, often avoiding purchases rather than feeling "cheap," and public payments can backfire by diminishing self-signaling value. Yet PWYW remains profitable as people pay fair amounts to maintain a positive identity.
Small and Loewenstein (2003)	Charitable Donation	*Lab Experiment (Offline)* - Dictator-game allocation (N = 76) *Field Experiments (Offline)* - Public Places in Pittsburgh (N = 234)	Identifiability (determined vs. undetermined victims); Social-distance/Empathy	When the beneficiary is already chosen, generosity jumps—lab donors gave about 60 % more and field donors about 25 % more (with higher participation)—showing that even minimal identifiability markedly boosts altruism.
Haggag and Paci (2014)	Tipping	*Quasi-Field Experiment (Offline)* - NYC Taxi (N = 13,820,735 total rides; N = 6,218,196 in primary analysis)	Default tips	Higher defaults increase average tips but reduce the likelihood of tipping
Chandar et al. (2019)	Tipping	*Field Experiments (Online)* - Uber (N = 12,040,801)	Default tips	Higher default tip options raise the average tip amount, but they also increase the share of rides that receive no tip whatsoever.
Duhaime and Woessner (2019)	Tipping	*Lab Experiment (Online)* - Grocery delivery (N = 392) *Archival Data Analysis (Offline)* - NYC Taxi (N = 7,553,909) *Field Experiments (Online)* - Food delivery - Foodler (N = 115) - Grubhub (N = 154)	Effect of worker autonomy on tipping norms	Consumers tip more generously to workers perceived as traditional employees with limited autonomy, but tip less frequently and in smaller amounts when the same work is framed as flexible gig labor.
Donkor (2021)	Tipping	*Archival Data Analysis (Offline)* - NYC Taxi (N = 8,578,501)	Norm conformity cost; Menu-opt-out costs	On-screen tip defaults strongly influence behavior, as most riders use suggested options to avoid norm-nonconformance and effort costs.

Alexander et al. (2021)	Tipping	*Field Experiment (Online)* - Laundry App Service (N = 94,571 orders from 24,637 customers)	Default tips suggestions	Higher default tip recommendations increase total revenue despite reducing tipping frequency, with no negative impact on customer satisfaction, return visits, or spending.
Lu et al. (2021)	Tipping	*Field Experiment (Online)* - Live Streaming Platform (N = 153 broadcasters; 2,222 treatment-period + 2,226 pre-period)	Audience size (social image concern vs. seeking reciprocity)	Larger audiences generate higher PWYW revenues because social image concerns (paying for status) outweigh reciprocity seeking (competition for attention) - but with diminishing returns.
Chen et al. (2023)	Tipping	*Quasi-Field Experiment (Online)* - Uber/Lyft (N = 10 rides) *Archival Data Analysis (Online)* - Uber/Lyft (N = 36) *Lab Experiment (Online)* - Ridesharing (N = 146) - Uber (N = 709) - Instacart (N = 349/ N = 490) *Field Experiment (Offline)* - Restaurant (N = 251)	Order of prompts (Rating → Tip vs. Tip → Rating)	The sequence of evaluations and tips significantly impact tipping behavior. Providing rating before tipping leads customers to reduce their tip amount, while tipping before rating shows no effect on subsequent ratings.
Bluvstein and Raghubir (2022)	Tipping	*Lab Experiment (Online)* - Food delivery app (N = 405) *Lab Experiment (Offline)* - Coffee shop (N = 439/ N = 324) - POS Checkout Screen (N = 623/ N = 401/ N = 250/ N = 463) *Field Experiments (Offline)* - Coffee shop (N = 1,023)	Opt-out Framing: 0% vs. No Tip	Using “0 %” instead of “No Tip” on POS screens nudges more customers to tip and boosts the average amount, because an explicit zero can feel self–image–damaging.
Present study: Kim, Amir & Wilbur (2026)	Tipping	*Field Experiments (Online)* - Fiverr (web: N = 375,264; mobile: N = 61,705) *Archival Data Analysis (Online)* - Fiverr (N = 4,134,928)	Social Norms; Reciprocity	Injunctive norm message increases tipping, especially by new buyers on first exposures, but does not detectably reduce total orders or spending.

Web Appendix B: Digital Tipping – Motives, Patterns, and Seller Response

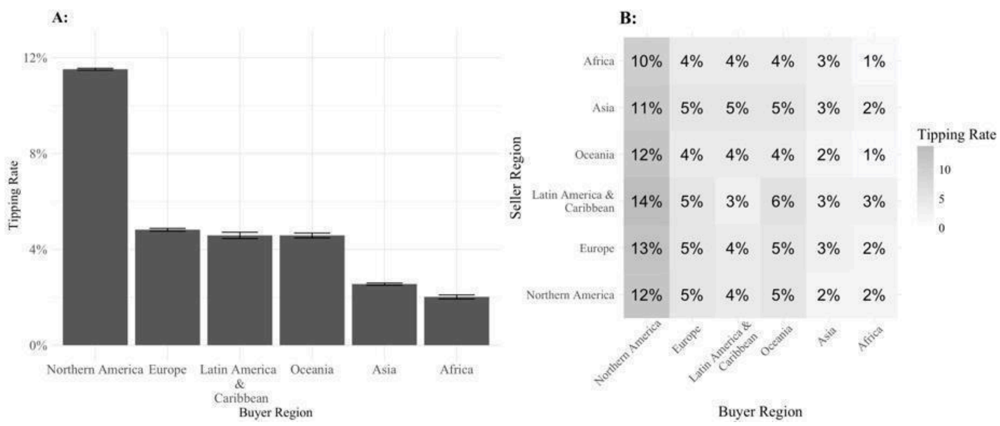
We describe 4.1 million pre-experimental transactions in four steps. First, we investigate indicators that duty, reciprocity, and service/esteem motives (Lynn, 2016) may persist in the digital marketplace. Second, we document how tipping incidence and magnitude vary with buyer, transaction, and seller characteristics. Third, we estimate which buyer, transaction, and seller attributes predict tipping, controlling for buyer and seller fixed effects. Finally, we turn to the supply side and ask whether receiving a tip predicts subsequent seller behavior. This Appendix only reports descriptive, or correlational, findings; no causal interpretations are implied.

B1. Model-Free Tipping Facts and Offline Motive Alignment

We explore duty motives by investigating whether culturally rooted duty norms governing offline tipping persist in fully digital marketplaces. Offline tipping is primarily shaped by social norms that vary across regions (Lynn et al., 1993). However, research suggests that social norms in general are weakened online, as anonymity and reduced social observability attenuate social pressure (Lapidot-Lefler and Barak 2012, Cohn et al. 2022). Thus, it remains unclear whether duty-based motivations survive in digital environments.

Descriptive evidence suggests they do. Figure B1.A shows that buyers from North America, where restaurant tipping is customary and widely expected, tipped in 11.5% of transactions, whereas buyers from other regions tipped 3.9% of transactions on average. The pattern is surprising for two reasons. First, if duty motives depend on social pressure, the anonymity of online interactions should erase regional differences. However, the regional tipping gap instead points to internalized duty norms. Second, even buyers from non-tipping regions occasionally leave tips, suggesting that factors beyond regional customs, such as reciprocity, may also motivate digital tipping. Figure B1.B shows that buyers do not tip more when interacting with sellers from high-tipping regions, suggesting that tipping may be a buyer decision more so than the outcome of a buyer/seller pair.

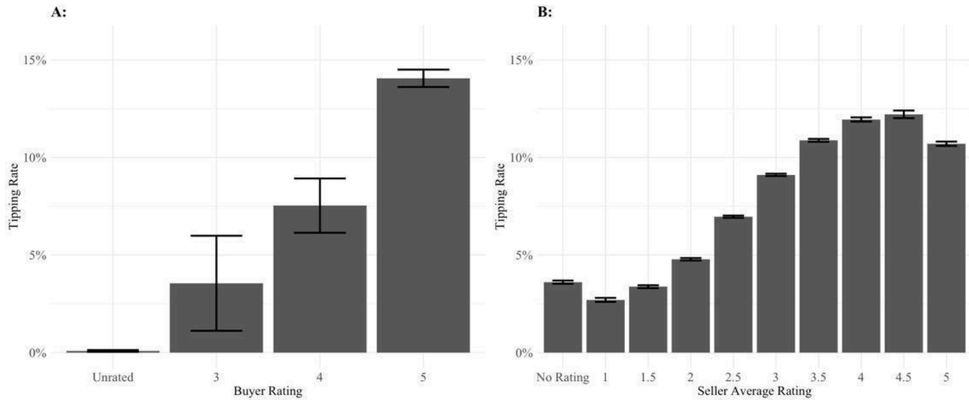
Figure B1: Tipping by Buyer and Seller Regions.



Next, we describe the relationship between buyer satisfaction ratings and tips. Previous literature found that people commonly tip to reward higher-quality service in traditional offline contexts (Azar 2007, Lynn and Sturman 2010), as well as in ride-sharing contexts (Chandar et al. 2019). Buyers rated 61% of pre-experimental transactions, with 93.6% of all ratings being five stars. Figure B2.A shows that tipping rates were four times higher among 5-star transactions than among 3-star transactions.[18] Buyers tipped approximately twice as frequently for 5-star transactions compared to 4-star transactions. Figure B2.B shows that tipping rates increase with the seller's average rating, as calculated using previous transactions. Thus, tipping behavior reflects both immediate buyer satisfaction and the seller's historical performance, consistent with reciprocity motives as a possible driver of tipping.

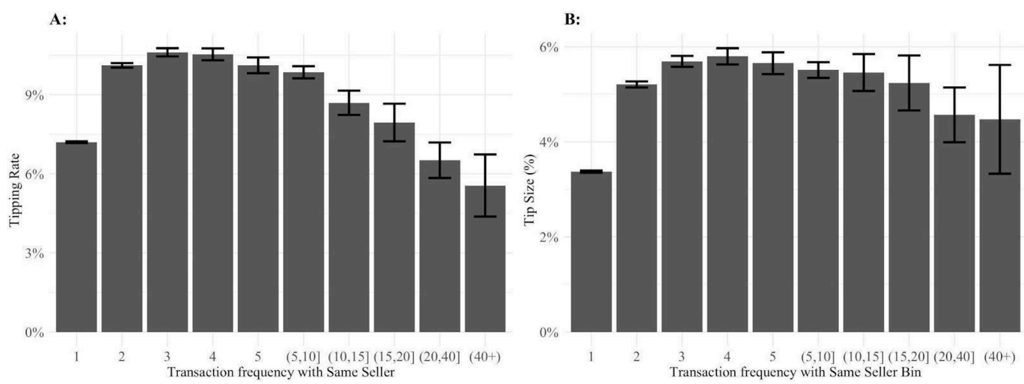
Figure B2: Tipping and Buyer Satisfaction.

[18] Unrated transactions were always untipped, because the platform only displayed a tip request if the buyer rated the service three stars or more.



Next, we describe how tipping patterns vary with transaction frequency within stable buyer-seller pairs. Figure B3 shows that buyers tipped on 7.2% of one-time transactions, and about 10% for transactions involving two to five interactions between the same buyer-seller pairs, then declined after ten interactions. Average Tip Sizes follow a similar pattern.

Figure B3: Tipping by Buyer/Seller Transaction Frequency



B2. How tipping changes with buyer, transaction, and seller characteristics.

Buyer tipping covaries with platform experience. The variable *"Buyer Prior Orders"* counts buyer transactions in 2019. Figure B4 classifies pre-experiment orders into comparably sized bins, showing that buyer prior orders correlate negatively with tipping rates but positively with Average Tip Percentage (ATP). Buyers with more than 100 transactions were substantially less likely to tip than infrequent buyers. Buyer prior orders correlate strongly with lower prices, partly explaining their positive correlation with ATP, as lower-priced gigs were more likely to lead to default tips of $5 or $10.

Figure B4:Tipping by Buyer Prior Orders

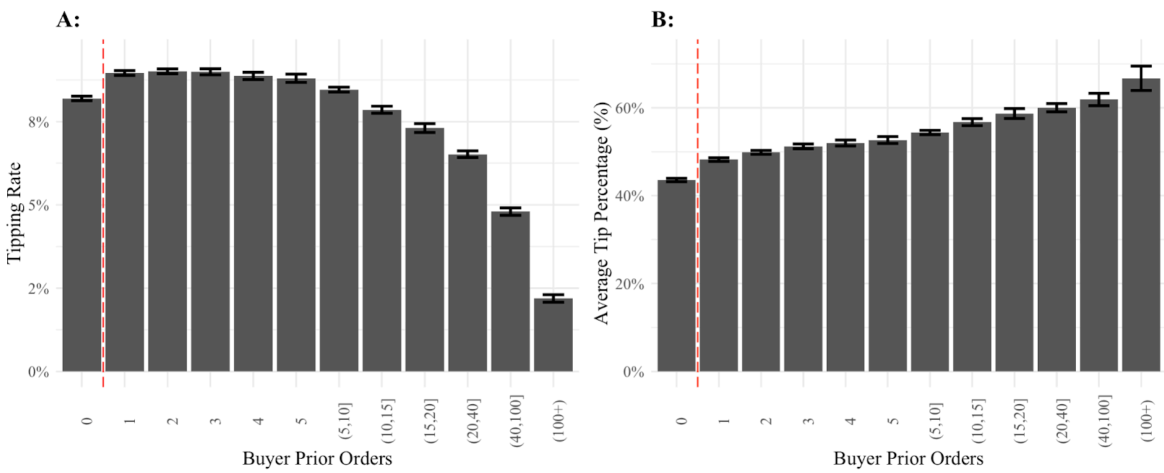


Figure B5 shows that buyers' tipping history also predicts tipping. Buyers with no prior tips tipped on just 3.8% of orders, rising to 27.8% among buyers with one prior tip, up to 84.6% among buyers with 100 prior tips. The figure does not distinguish state dependence from heterogeneity–i.e., learned tipping behavior from individual tipping propensity. However, ATP varies much less with buyer prior tips than tipping rates.

Figure B5: Tipping by Buyer Prior Tips

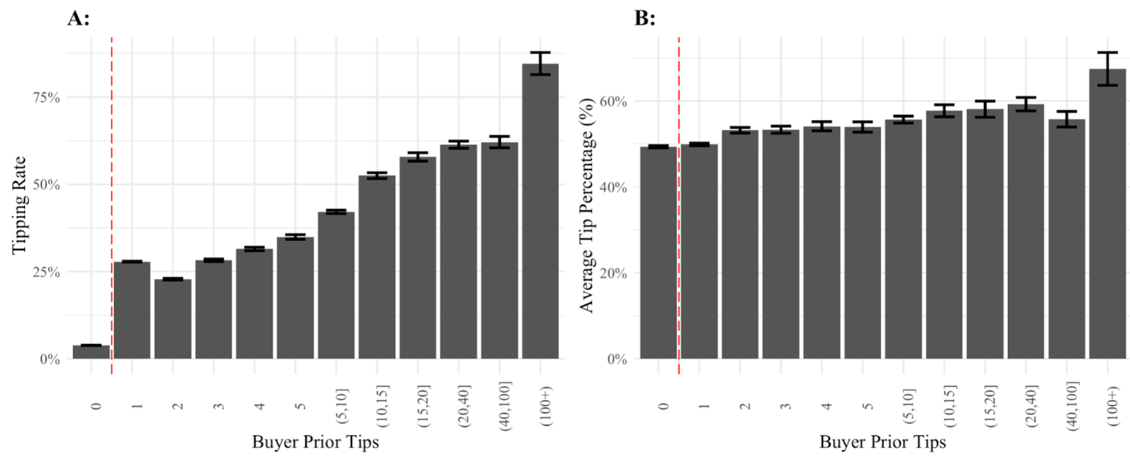
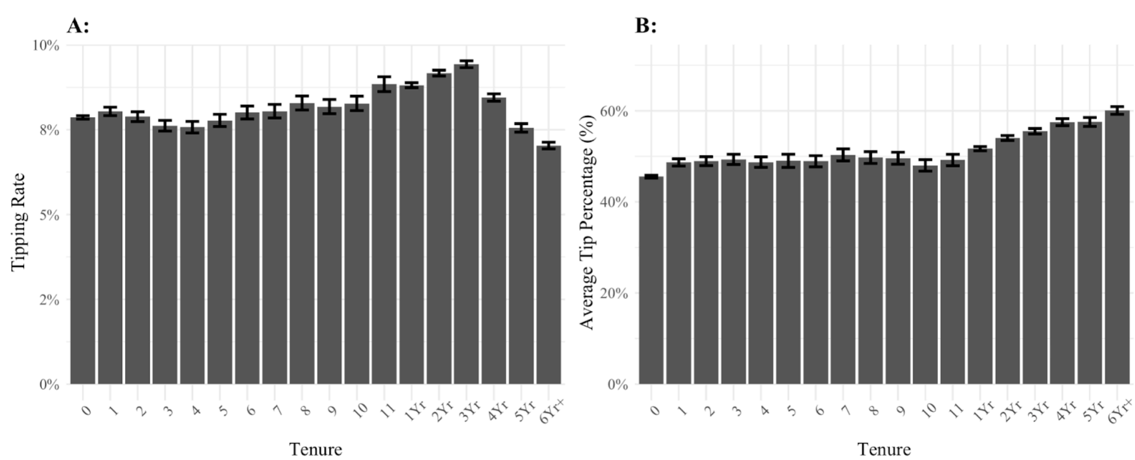


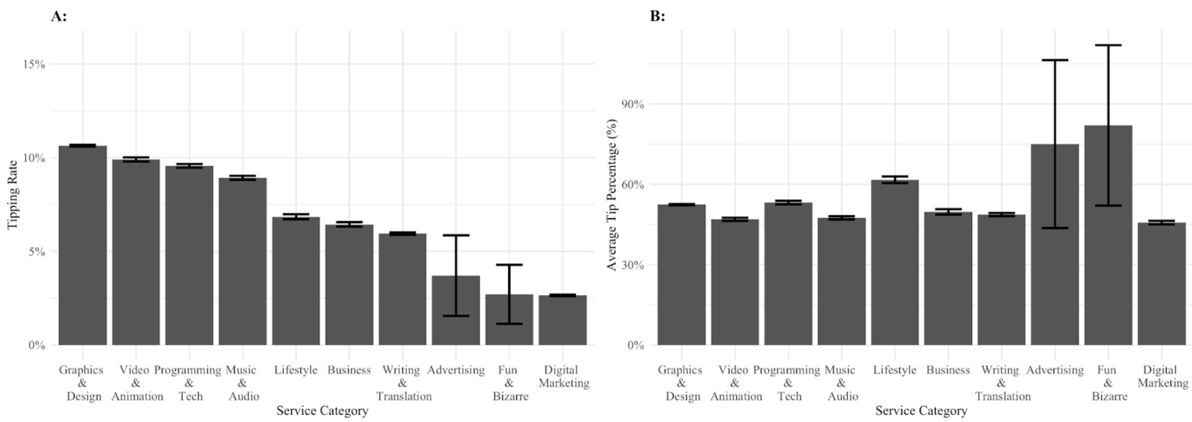
Figure B6 shows how tipping changes with the time elapsed since each buyer's first transaction date. Tipping weakly increases with buyer tenure up to a peak at 3 years, then declines among older buyer cohorts. Based on our understanding of the platform, we believe that is because Fiverr used low prices to attract frugal buyers initially, then later moved upmarket. Average Tip Percentage is nearly flat among newer buyers, but then rises with buyer tenure among older cohorts.

Figure B6: Tipping by Buyer Tenure



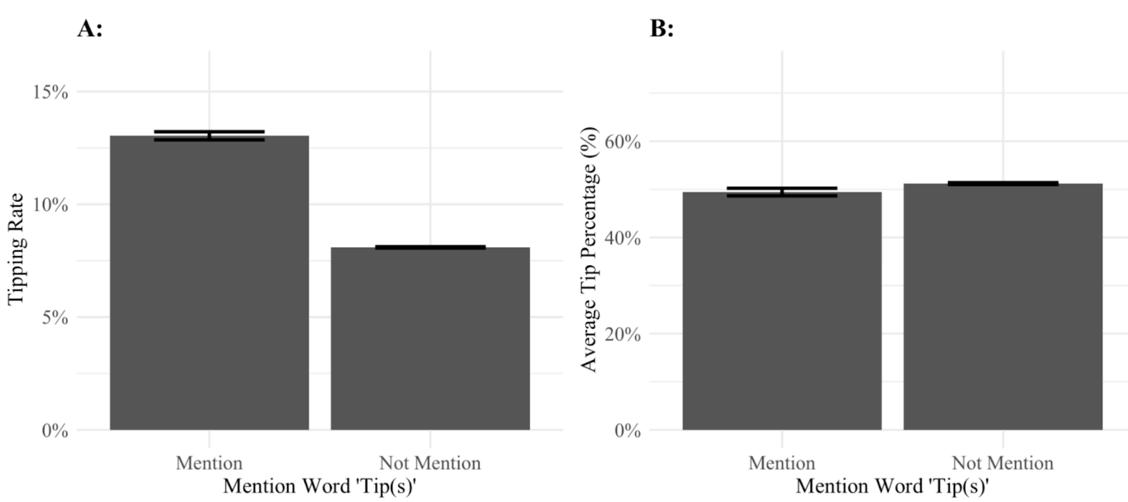
Next we look at how tipping covaries with the type of service provided. Figure B7 shows a nearly four-fold spread in tipping rates, ranging from 2.7% of Digital Marketing gigs up to 10.6% in the Graphics & Design category. ATP shows less variation across categories, relative to its mean, than tipping rates.

Figure B7: Tipping and Service Category



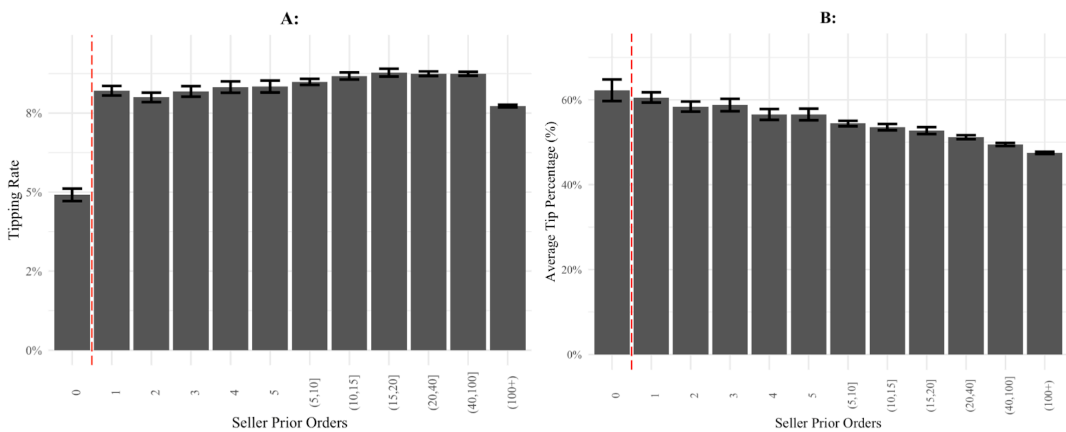
Seller messages to buyers accompanied completed orders in 96.5% of transactions. 3.1% of seller messages mentioned the words "tip" or "tips." Figure B8 shows that tipping rates increased with seller tip mentions from 8.1% to 13%. However, ATP was nearly the same regardless of seller tip mentions.

Figure B8: Tipping by Seller Mention "Tip" or "Tips"



Next we look at how tipping relates to seller experience, as measured by *Seller Prior Orders* delivered since their first transaction in 2019. Figure B9 shows that buyers tipped first-time sellers on just 4.9% of transactions, rising to 8.2% of transactions by one-time sellers. ATP declines slightly with seller prior orders, possibly due to the correlation between both variables and price.

Figure B9: Tipping By Seller Prior Orders



B3. Descriptive Tipping Regression

We estimate descriptive regressions using the pre-experiment data to simultaneously estimate multiple associations between tipping and observable contextual factors. The pre-experiment data contain anonymous account IDs with limited demographic information but extensive behavioral information on purchases, ratings, pricing, tipping, and seller tip mentions. Table B1 summarizes descriptive statistics of observable transaction features, grouping them into 3 sets: Buyer characteristics, Transaction characteristics, and Seller characteristics. We also consider Buyer and Seller fixed effects.

Equation (1) shows the regression framework:

$$y_i = f\left(X_i^T,\ \alpha_{b(i)},\ \gamma_{s(i)}, \epsilon_i\right) \tag{1}$$

where i indexes transactions; y_i is a binary tipping indicator; X_i^T represents transaction characteristics; $\alpha_{b(i)}$ and $\gamma_{s(i)}$ are buyer and seller fixed effects, respectively; and ϵ_i captures other unobserved factors that affect tipping.

Table B2 displays four variations of the general specification to illustrate the incremental contributions of buyer and seller fixed effects. Column (1) includes only transaction characteristics; Column (2) adds buyer fixed effects, but not seller fixed effects; Column (3) adds seller fixed effects but not buyer fixed effects; and Column (4) includes both buyer and seller fixed effects simultaneously. All regressions exclude single-transaction buyers and sellers to maintain consistent effective sample sizes across columns, as buyer and seller fixed effects fully absorb these single-transaction observations. Table B3 reports the same four specifications, including single-transaction participants, with similar findings.

These descriptive regression analyses suggest that buyer, seller, and transaction characteristics each correlate with tipping. However, the magnitudes of the associations change significantly when controlling for buyer and seller fixed effects. For example, the association between tipping and 5-star ratings decreases from 13.2% in the benchmark model (Column 1) to 11.4% when the model includes buyer and seller fixed effects (Column 4). Further, the regressions highlight the relative importance of buyer-specific factors in explaining variations in tipping behavior. Adding buyer fixed effects substantially increases Adjusted R-square from 0.066 to 0.398. However, adding seller fixed effects alone leads to a meaningful, but much smaller, increase from 0.066 to 0.107.

Table B1: Variables and Summary Statistics in Pre-Experiment Transaction Data.

	Mean	**Standard Deviation**
Buyer Characteristics (X^B)		
Buyer Prior Orders	10.51	42.9
Buyer Prior Ratings	10.88	33.97
Buyer Average Rating (Given)	4.96	0.15
Buyer Prior Tips	0.51	2.99
Buyer Tenure (Months)	20.97	25.22
User-defined Female	1.2%	
User-defined Male	3.11%	
Transaction Characteristics (X^T)		
Buyer-Seller Repeat Indicator	29.9%	
Count of Buyer-Seller Repeat Orders	1.78	9.22
Buyer Rating = 3	0.54%	
Buyer Rating = 4	3.34%	
Buyer Rating = 5	56.6%	
Price	$36.86	$82.3
Seller Tip Mention	2.98%	
Seller Characteristics (X^S)		
Seller Prior Orders	227.7	645
Seller Prior Ratings	128.03	346.49
Seller Average Rating (Received)	4.96	0.07

Table B2: Descriptive Regression without Single-Transaction Buyers and Sellers

	(1)	(2)	(3)	(4)
Buyer-Seller Repeat Indicator	0.003***	0.000	0.003***	−0.002***
	(0.0003)	(0.000)	(0.0003)	(0.0003)
Count of Buyer-Seller Repeat Orders	−0.000***	0.000	−0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Buyer Rating = 3	0.030***	0.019***	0.032***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
Buyer Rating = 4	0.066***	0.053***	0.064***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
Buyer Rating = 5	0.132***	0.117***	0.128***	0.114***
	(0.000)	(0.000)	(0.000)	(0.000)
Price	0.0001***	−0.000***	0.0001***	−0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
Price^2	−0.000***	0.000***	−0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Seller Tip Mention	0.047***	0.037***	0.063***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
Buyer FE		Y		Y
Seller FE			Y	Y
Transaction Char.	Y	Y	Y	Y
Observations	3,439,974	3,439,974	3,439,974	3,439,974
Num. Buyers	613,417	613,417	613,417	613,417
R^2	0.066	0.505	0.139	0.533
Adjust R^2	0.066	0.398	0.107	0.407

Note. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table B3: Descriptive Regression with Single-Transaction Buyers and Sellers

	(1)	(2)	(3)	(4)
Buyer-Seller Repeat Indicator	0.004***	0.000	0.003***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Count of Buyer-Seller Repeat Orders	−0.000***	0.000	−0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Buyer Rating = 3	0.030***	0.020***	0.032***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
Buyer Rating = 4	0.067***	0.053***	0.066***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
Buyer Rating = 5	0.134***	0.118***	0.131***	0.114***
	(0.000)	(0.000)	(0.000)	(0.000)
Price	0.000***	−0.000***	0.000***	−0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
Price^2	−0.000***	0.000***	−0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Seller Tip Mention	0.049***	0.037***	0.063***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
Buyer FE		Y		Y
Seller FE			Y	Y
Transaction Char.	Y	Y	Y	Y
Observations	4,134,928	4,134,928	4,134,928	4,134,928
Num. Buyers	1,274,011	1,274,011	1,274,011	1,274,011
R^2	0.067	0.582	0.136	0.610
Adjust R^2	0.067	0.396	0.098	0.404

Note. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

B4. Seller Responses to Tips

How do sellers react after receiving tips? To answer this question, we sample sellers with at least two transactions, so we can describe behavior changes following tips vs. non-tips. We consider three ways in which a seller might respond after a tip. The first one is direct: the seller can mention tips to the buyer on her subsequent transaction. The other two are indirect measures of seller effort on her subsequent transaction: buyer tipping and buyer satisfaction rating. We explain each of the three outcome variables using an indicator for whether a seller received a tip on their previous gig. We also control for time-varying seller characteristics (number of seller prior orders and lagged seller average rating), transaction characteristics (order price and service category), buyer fixed effects, and seller fixed effects.

Table B4 shows that a seller who receives a tip is 0.1% more likely to mention a tip ($p < 0.01$) when they deliver their subsequent job, and 1.3% less likely to receive a tip ($p < 0.001$), relative to no tip received on the previous transaction. Their subsequent public rating also increases by 0.003 stars, but this change is not statistically significant. These results control for persistent seller effects, but they do not separate tipping drivers from other time-varying unobservables, so we again interpret them as descriptive or correlational.

Table B4: Seller Behaviors After Tips

	Seller Tip Mention (1)	Tip Received (2)	Public Rating (3)
Lag(Tip Received)	0.001** (0.0002)	−0.013*** (0.001)	0.003 (0.004)
Seller Prior Order	−0.000*** (0.000)	−0.000*** (0.000)	−0.000*** (0.000)
Lag(Seller Average Rating)	−0.000 (0.000)	−0.002*** (0.000)	−0.073*** (0.002)
Price ≤ $25	0.000 (0.000)	0.000*** (0.000)	0.001*** (0.0002)
Price > $25	−0.000 (0.000)	−0.000*** (0.000)	−0.000*** (0.000)
Buyer FE	Y	Y	Y
Category FE	Y	Y	Y
Seller FE	Y	Y	Y
Observations	3,963,840	3,963,840	3,963,840
Num. Buyers	1,239,353	1,239,353	1,239,353
Num. Sellers	125,040	125,040	125,040
R^2	0.753	0.594	0.600
Adjusted R^2	0.625	0.383	0.391

Note. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

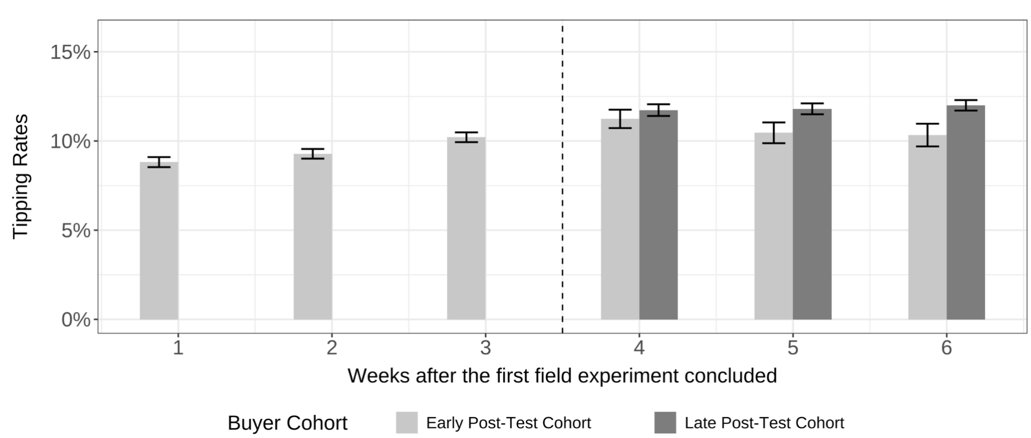
B5. *Platform Rollout of Norms Message After Post-Experiment Reversion to Control Message*

The platform adopted the Norms message as the default for all users 3 weeks after the Web Field Experiment concluded. It continued to offer the 3-Defaults tip menu. Here we describe how tipping behavior changed after the Norms tip request message was adopted for all Web buyers. These results are correlational but suggestive.

Figure B10 summarizes average weekly tipping in each of two cohorts. The "early post-test cohort" includes all new buyers that first transacted in weeks 1, 2 or 3 after the test, which received the Control tip request message in weeks 1-3 after the test. The "late post-test cohort" includes all new buyers that first transacted in weeks 4, 5 or 6 after the test. Both cohorts exclusively received the Norms tip request message in weeks 4-6 after the test conclusion. The 3-defaults tip menu was presented to both cohorts in all six weeks.

We find two notable results. First, the early post-test cohort tipped in 9.4% of transactions while receiving the Control message in Weeks 1-3. After the platform adopted the Norms message, this same cohort's tipping rates increased to 10.8%, a statistically significant increase ($p < 0.001$). Second, the later post-test cohort who experienced only the Norms message tipped at significantly higher rates than the early post-test cohort during their respective Norms periods (11.9% vs. 10.8%, $p < 0.001$).

Figure B10: Tipping Rates during Post-test period.



Web Appendix C: Treatment Manipulation Analysis

To test whether the tip request message treatments indeed tapped the appropriate and anticipated theoretical constructs of reciprocity and implied normative behavior, we designed a test eliciting the interpretation ascribed to the different manipulations by responders from a similar population of online users. In accordance with the firm's decision to use messages that included the seller's name in two of the four treatments, we also tested whether this inclusion (or exclusion) changes the implied construct in meaningful ways, with the hypothesis of no impact.

One thousand participants from Amazon Mechanical Turk were randomly assigned to one of eight conditions: 4 x Message (Control [status-quo], Impersonal Reciprocity, Personalized Reciprocity, and Norms) by 2 x Seller name (present vs. not present). Each participant then responded to a question that asked them to select all the reasons to leave a tip to the seller of a digital service on an online platform, as implied by the message. The six reasons presented followed Cialdini's six "weapons of influence," a widely accepted and widely used typology of techniques to influence customers (Cialdini, 1993). We adapted Cialdini's six influence techniques to online tipping as follows:

- Reciprocity: "This emphasizes that the seller did something good for me, and therefore I need to return the favor."
- Social Proof: "This emphasizes that I should tip because everyone else does it; it is normal to tip."
- Commitment: "This emphasizes that because I chose this seller, I am already committed to this transaction and should tip."
- Authority: "This emphasizes that an authoritative figure says I should tip."
- Liking: "This emphasizes that I should tip because someone I really like is asking me to."
- Scarcity: "This emphasizes that tipping is really valuable because it is rare."

All participants were asked to mark all explanations that fit how they interpreted each tip message. We first checked whether participants' perceptions of the tip messages corresponded to what we expected. Figure C1 shows that the Control message motivates multiple interpretations, and no distinct one emerges. 26% of participants marked that none of the six reasons to tip apply to the Control message, while around 21% of participants interpreted it as "Reciprocity", with another 20% interpreting the message as "Liking." On the contrary, each treatment condition yielded a dominant response as intended ($p < 0.001$). "Reciprocity" was the dominant interpretation for both Impersonal Reciprocity and Personalized Reciprocity tip messages, with "Social Proof" being the dominant interpretation of the Norms message. The manipulation check confirms that the three treatment messages correspond to our intention to test the underlying theoretical constructs motivating people to leave a tip.

We further examined whether the dominant perception for each tip message varied with the inclusion of a seller's name. Figure C2 shows that the inclusion of the seller's name does not significantly change the prevalent perception of the treatment messages. The Control message motivates multiple interpretations, with "none of the six reasons above apply" being the most common. "Reciprocity" was the dominant interpretation for both Impersonal Reciprocity and Personalized Reciprocity tip messages, regardless of the seller's name, with "Social Proof" being the dominant interpretation of the Norms message. The overlapping confidence intervals for the dominant interpretation in four graphs confirm that there is no difference between them with and without the seller's name for all messages.

Figure C1: Perception of Each Tip Message.

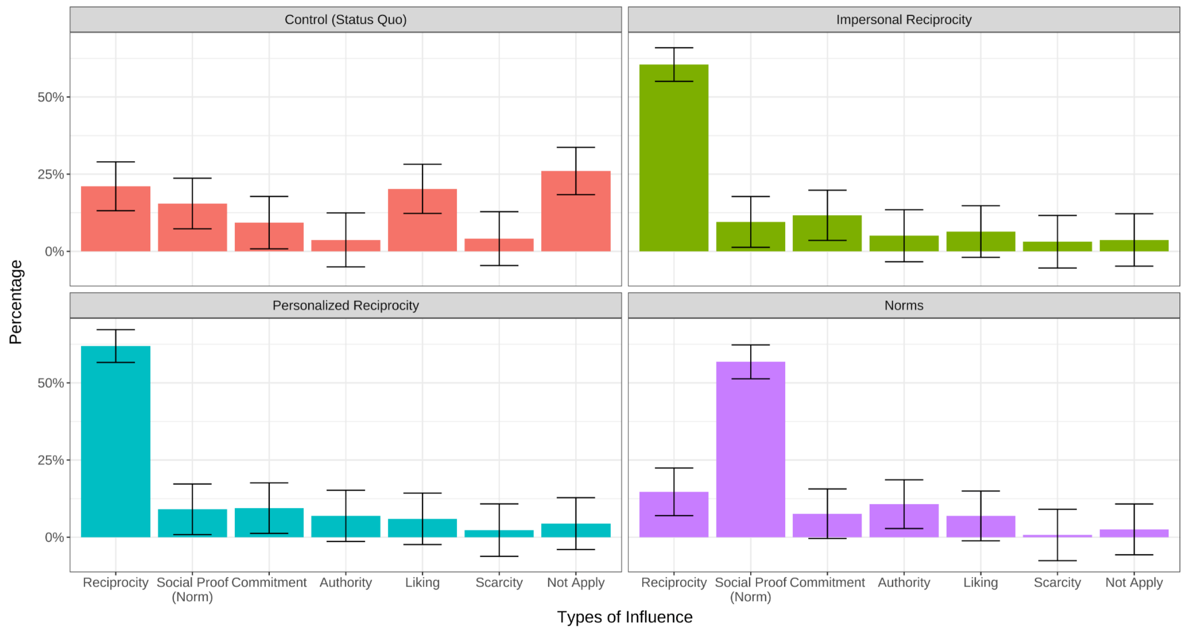
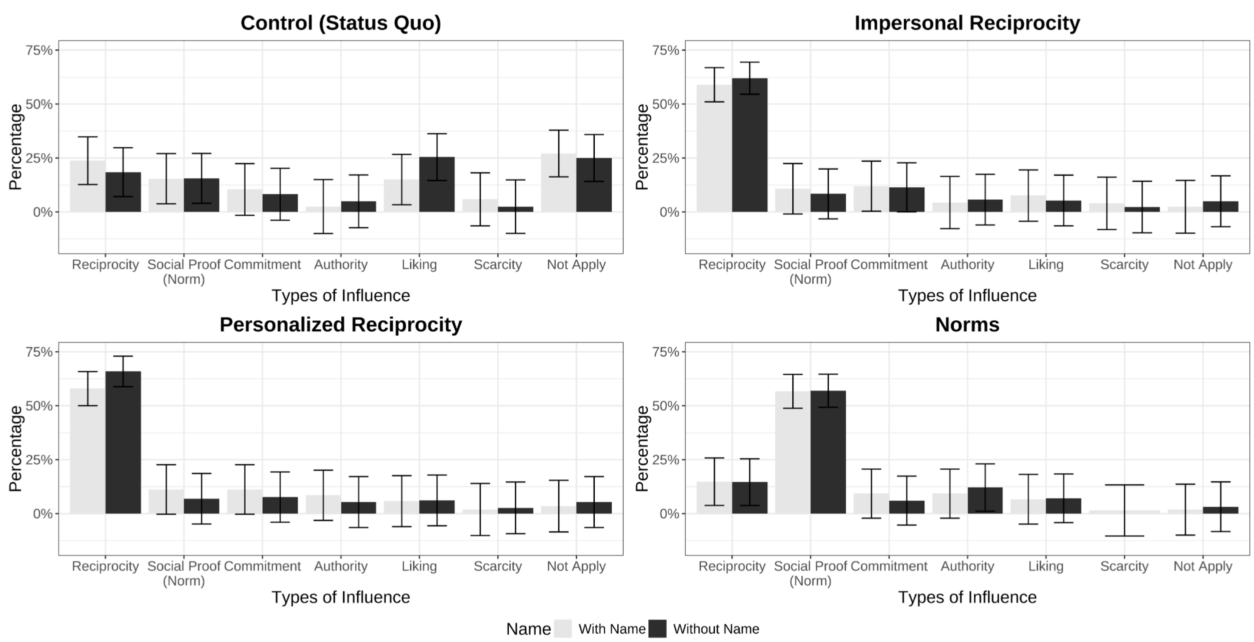


Figure C2: Interaction between Dominant Perception and Seller Name.



Web Appendix D: Pre-Experimental Default Tips Usage

Figure D1 shows how default tip usage changed with order prices under $25 in the pre-experiment data. Over 90% of tips used default tips rather than custom tips for gigs under $25, with approximately 70% choosing the $5 default tip, about 20% choosing the $10 default tip, and the majority of custom tips entered being near $15. Therefore, the company viewed adding a $15 default tip option as a natural way to lower the consumer costs of entering a larger tip on a low-priced order.

Figure D2 shows how default tip usage changed with order prices over $25 in the pre-experiment data. About half of tips used the 20% default tip, with the other half entering a custom dollar amount, and fewer than 2% of tipping consumers chose the 30% default tip. Therefore, the platform decided to eliminate the nearly-unused 30% default tip and it decided to replace it with 15% and 25% default tips.

Figure D1: Pre-Web F.E. Tipping for Order Prices under $25.

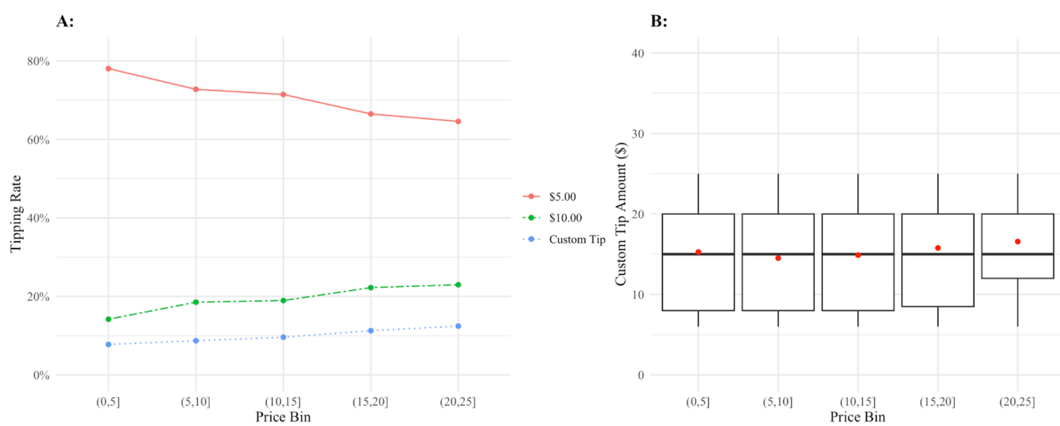
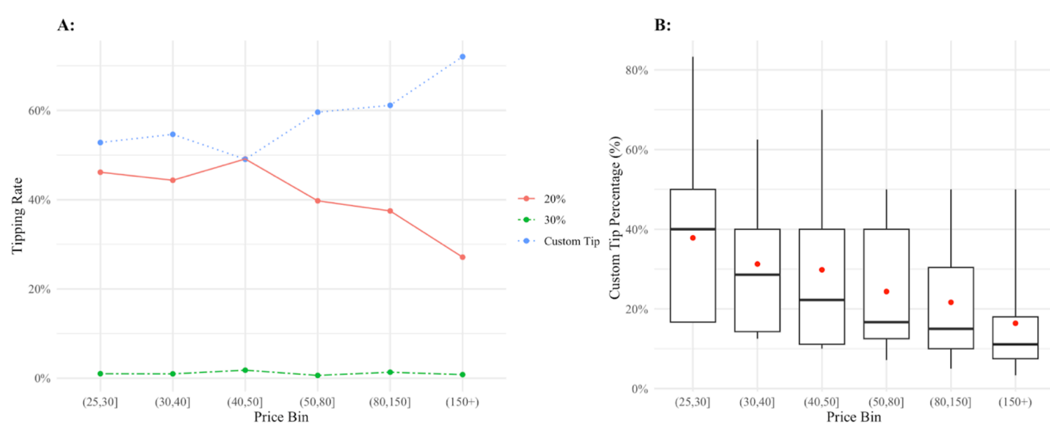


Figure D2: Pre-Web F.E. Tipping for Order Prices over $25.



Web Appendix E: Randomization Checks

Web Field Experiment

The platform randomly assigned 7,880 new buyers and 45,497 repeat buyers to one of four conditions: Control, Impersonal Reciprocity, Personalized Reciprocity, and Norms. We evaluated balance across conditions with Chi-square tests, as reported in Tables E1 and E2. new buyers had no purchase history, so we tested whether the distribution of buyer regions was balanced across the four groups. Repeat buyers were assessed for balance in region and six pre-treatment behavioral variables, confirming that randomization was correctly achieved at the buyer level.

App Field Experiment

The platform randomly assigned 9,570 new buyers and 25,910 repeat buyers to either the Control or the Norms message in the App Field Experiment. 79.5% of repeat buyers completed at least one transaction between January 1, 2019, and October 16, 2019, giving us the pre-experiment data required for balance tests. All randomization checks for repeat buyers, therefore, use this subsample.

Table E3 shows that the regional distribution of new buyers is balanced across conditions. Table E4 shows that repeat buyers are also well balanced overall, confirming that random assignment produced comparable treatment and control groups for both new and repeat buyers in the App Field Experiment.

Table E1: Randomization Check for New Buyers in the Web Field Experiment.

	Control	Impersonal Reciprocity	Personalized Reciprocity	Norms	p-value
N	2,008	1,950	2,016	1,906	
Africa	0.04	0.05	0.04	0.03	0.0607
Asia	0.13	0.12	0.12	0.12	0.6286
Europe	0.26	0.26	0.26	0.27	0.6022
Latin America & Caribbean	0.03	0.02	0.02	0.03	0.6833
North America	0.50	0.50	0.51	0.50	0.9838
Oceania	0.04	0.04	0.04	0.04	0.4962

Table E2: Randomization Check for Repeat Buyers in the Web Field Experiment.

	Control	Impersonal Reciprocity	Personalized Reciprocity	Norms	P-value
N	10,125	10,233	10,226	10,239	
Africa	0.03	0.03	0.03	0.03	0.4697
Asia	0.12	0.12	0.12	0.12	0.4523
Europe	0.24	0.24	0.24	0.24	0.8315
Latin America & Caribbean	0.03	0.03	0.03	0.03	0.7507
North America	0.54	0.54	0.54	0.54	0.7459
Oceania	0.04	0.04	0.04	0.04	0.7776
Buyer Tenure	26.86	26.79	26.87	27.07	0.8987
Pre-treat Orders	14.48	14.60	14.98	14.47	0.5707
Pre-treat Number of Sellers	7.21	7.17	7.29	7.22	0.8601
Pre-treat Expenditure	507.92	519.07	516.09	496.15	0.5595
Pre-treat Tip Frequency	1.21	1.26	1.32	1.23	0.3287
Pre-treat Tip Expenditure	14.35	16.65	15.66	14.83	0.2681

Table E3: Randomization Check for New Buyers in the App Field Experiment.

	Control	Norms	P-value
N	9,712	10,037	
Africa	0.01	0.01	0.185
Asia	0.07	0.07	0.4811
Europe	0.21	0.21	0.6971
Latin America & Caribbean	0.02	0.02	0.3993
North America	0.64	0.65	0.5689
Oceania	0.04	0.04	0.3824

Table E4: Randomization Check for Repeat Buyers in the App Field Experiment.

	Control	Norms	P-value
N	17,755	17,684	
Africa	0.01	0.01	0.4089
Asia	0.07	0.08	0.2715
Europe	0.19	0.19	0.3189
Latin America & Caribbean	0.02	0.02	0.1052
North America	0.67	0.67	0.4943
Oceania	0.04	0.04	0.3733
Buyer Tenure	21.72	21.38	0.1961
Pre-treat Orders	11.92	11.65	0.3046
Pre-treat Number of Sellers	7.14	7.04	0.3532
Pre-treat Expenditure	437.46	437.84	0.9731
Pre-treat Tip Frequency	1.09	1.12	0.5003
Pre-treat Tip Expenditure	12.27	13.16	0.2031

Table E3: Randomization Check for New Buyers in the App Field Experiment.

	Control	Norms	P-value
N	4,707	4,863	
Africa	0.01	0.01	1
Asia	0.08	0.07	0.1113
Europe	0.20	0.19	0.4061

Latin America & Caribbean	0.01	0.02	0.453
North America	0.67	0.68	0.1837
Oceania	0.04	0.04	0.9894

Table E4: Randomization Check for Repeat Buyers in the App Field Experiment.

	Control	**Norms**	**P-value**
N	10,295	10,304	
Africa	0.01	0.01	0.5996
Asia	0.07	0.07	0.2164
Europe	0.17	0.17	0.6772
Latin America & Caribbean	0.02	0.02	0.4088
North America	0.69	0.69	0.9907
Oceania	0.04	0.04	0.6491
Buyer Tenure	19.58	19.23	0.2829
Pre-treat Orders	12.06	11.67	0.2393
Pre-treat Number of Sellers	7.31	7.25	0.7135
Pre-treat Expenditure	432.35	424.89	0.6075
Pre-treat Tip Frequency	1.04	1.05	0.9403
Pre-treat Tip Expenditure	11.36	12.42	0.2292

Web Appendix F: Supplementary Web Field Experiment Results

Tables F1 and F2 supplement the main analysis in section 4.5 by showing treatment effect estimates on tip amounts and tip percentages, using the same estimation procedure described in section 4.4. The point estimates are directionally consistent with Table 2, but noisier due to the increased variability of tip amounts and tip percentages.

Table F1. Web F.E. Treatment Effects on Tip Amounts

	(1)	(2)	(3)	(4)	(5)	(6)
	All Buyers, All Exposures	All Buyers, First Exposures	New Buyers, All Exposures	New Buyers, First Exposures	Repeat Buyers, All Exposures	Repeat Buyers, First Exposures
Intercept	1.216***	1.683***	1.184***	1.464***	1.220***	1.721***
	(0.041)	(0.056)	(0.079)	(0.111)	(0.046)	(0.063)
Impersonal Reciprocity	0.179	0.237	−0.048	−0.007	0.211	0.277
	(0.091)	(0.166)	(0.112)	(0.155)	(0.103)	(0.193)
Personalized Reciprocity	0.097	0.353***	0.135	0.254	0.094	0.354***
	(0.045)	(0.065)	(0.101)	(0.104)	(0.049)	(0.076)
Norms	0.273***	0.750***	0.660***	1.210***	0.221***	0.655***
	(0.044)	(0.070)	(0.117)	(0.167)	(0.048)	(0.078)
Three Defaults	−0.127**	−0.408***	−0.063	−0.306**	−0.138**	−0.410***
	(0.045)	(0.064)	(0.088)	(0.118)	(0.050)	(0.074)
Num. Obs.	375,264	116,465	53,494	23,224	321,770	93,241
Num. Buyers	87,966	87,966	20,436	20,436	67,530	67,530
Num. Sellers	67,100	41,500	22,515	13,499	61,169	36,427
R2	0.000	0.001	0.001	0.004	0.000	0.001
Adj R-sq.	0.000	0.001	0.001	0.004	0.000	0.001

Note. **p<0.01; ***p<0.001

Table F2. Web F.E. Treatment Effects on Tip Percentages

	(1)	(2)	(3)	(4)	(5)	(6)
	All Buyers, All Exposures	All Buyers, First Exposures	New Buyers, All Exposures	New Buyers, First Exposures	Repeat Buyers, All Exposures	Repeat Buyers, First Exposures
Intercept	5.821***	7.397***	5.555***	6.256***	5.859***	7.600***
	(0.193)	(0.227)	(0.367)	(0.457)	(0.214)	(0.254)
Impersonal Reciprocity	−0.026	0.477	0.111	1.249	−0.044	0.337
	(0.256)	(0.331)	(0.575)	(0.754)	(0.280)	(0.366)
Personalized Reciprocity	1.030***	2.385***	0.965	2.192***	1.022***	2.321***
	(0.201)	(0.238)	(0.398)	(0.539)	(0.223)	(0.265)
Norms	1.505***	3.590***	3.149***	5.019***	1.258***	3.246***
	(0.180)	(0.255)	(0.458)	(0.683)	(0.195)	(0.276)
Three Defaults	−1.141***	−2.273***	−1.156**	−1.812***	−1.121***	−2.257***
	(0.191)	(0.237)	(0.381)	(0.482)	(0.212)	(0.266)
Num. Obs.	375,264	116,465	53,494	23,224	321,770	93,241
Num. Buyers	87,966	87,966	20,436	20,436	67,530	67,530
Num. Sellers	67,100	41,500	22,515	13,499	61,169	36,427
R2	0.001	0.003	0.002	0.005	0.001	0.003
Adj R-sq.	0.001	0.003	0.002	0.005	0.001	0.003

Note. **p<0.01; ***p<0.001

Tables F3 and F4 report non-pooled regressions that isolate each of the causal effects in Table 2. In each case, we retain only data from two out of the eight cells displayed in Table 1. In all five cases, the estimates are not significantly different from the analogous estimate reported in Table 2. We conclude that pooling did not adversely affect the treatment effect estimates, compared to non-pooled regressions.

Table F3. Web F.E. Two-Cell Identification of Main Effects on Tipping Rates

	(1) T1&T2 Only All Exposures	(2) T1&T2 Only First Exposures	(3) T3&T4 Only All Exposures	(4) T3&T4 Only First Exposures
Intercept	0.106***	0.148***	0.106***	0.150***
	(0.003)	(0.003)	(0.003)	(0.003)
Impersonal Reciprocity	0.000	0.004		
	(0.004)	(0.004)		
Norms			0.014***	0.029***
			(0.004)	(0.005)
Num. Obs.	106,532	26,647	107,769	26,730
Num. Buyers	26,647	26,647	26,730	26,730
Num. Sellers	34,343	16,904	34,498	16,763
R2	0.000	0.000	0.001	0.002
Adj R-sq.	-0.000	-0.000	0.001	0.002

Note. **p<0.01; ***p<0.001

Table F4. Web F.E. Two-Cell Identification of Main Effects on Tipping Rates

	(1) T1&T5 Only All Exposures	(2) T1&T5 Only First Exposures	(3) T3&T7 Only All Exposures	(4) T3&T7 Only First Exposures	(5) T4&T8 Only All Exposures	(6) T4&T8 Only First Exposures
Intercept	0.106***	0.148***	0.090***	0.099***	0.094***	0.102***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
Three Defaults	−0.015***	−0.049***				
	(0.002)	(0.003)				
Personalized Reciprocity			0.016***	0.052***		
			(0.002)	(0.003)		
Norms					0.027***	0.078***
					(0.003)	(0.003)
Num. Obs.	134,527	53,584	135,017	53,729	134,775	53,737
Num. Buyers	46,431	46,431	46,609	46,609	46,500	46,500
Num. Sellers	40,677	25,821	40,606	25,749	40,642	25,807
R2	0.001	0.005	0.001	0.005	0.002	0.011
Adj R-sq.	0.001	0.005	0.001	0.005	0.002	0.011

Note. **p<0.01; ***p<0.001

Table F5. Web F.E. Estimation Results with "Donut" Samples (14-Day "Holes")

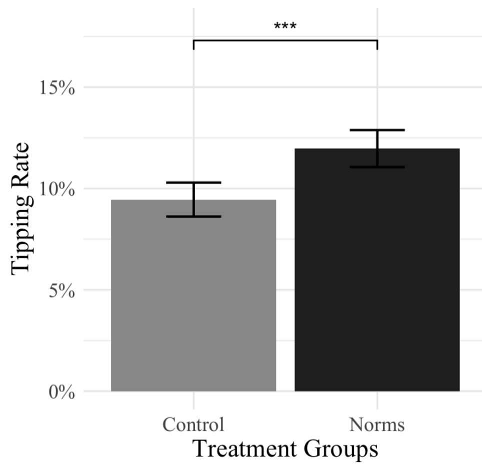
	(1)	(2)	(3)	(4)	(5)	(6)
	All Buyers, All Exposures	All Buyers, First Exposures	New Buyers, All Exposures	New Buyers, First Exposures	Repeat Buyers, All Exposures	Repeat Buyers, First Exposures
Intercept	0.109***	0.148***	0.118***	0.150***	0.107***	0.147***
	(0.003)	(0.003)	(0.007)	(0.008)	(0.003)	(0.003)
Impersonal Reciprocity	0.002	0.005	0.002	0.006	0.002	0.004
	(0.004)	(0.005)	(0.009)	(0.012)	(0.004)	(0.005)
Personalized Reciprocity	0.020***	0.045***	0.027***	0.059***	0.019***	0.041***
	(0.003)	(0.004)	(0.007)	(0.009)	(0.003)	(0.004)
Norms	0.032***	0.073***	0.070***	0.125***	0.027***	0.063***
	(0.003)	(0.004)	(0.009)	(0.011)	(0.003)	(0.004)
Three Defaults	−0.018***	−0.041***	−0.020**	−0.051***	−0.018***	−0.038***
	(0.003)	(0.004)	(0.007)	(0.009)	(0.003)	(0.004)
Num. Obs.	265,588	80,392	36,736	16,056	228,852	64,336
Num. Buyers	76,123	70,578	16,258	15,103	59,865	55,475
Num. Sellers	58,541	34,140	17,920	10,394	53,370	29,773
R2	0.001	0.006	0.005	0.015	0.001	0.004
Adj R-sq.	0.001	0.006	0.005	0.014	0.001	0.004

Note. **p<0.01; ***p<0.001

Web Appendix G: Supplementary App Field Experiment Results

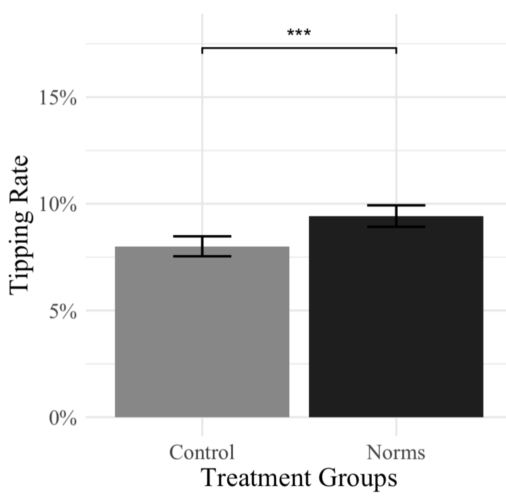
Figure G1 shows that the injunctive Norms treatment increased new buyers' tipping rates by 2.6 p.p. upon first app exposure (12% vs. 9.4%, $p < 0.001$). Figure G2 shows that the Norms treatment effect was 1.4 p.p. in repeat buyers' first app exposure (9.4% vs. 8%, $p < 0.001$).

Figure G1: App F.E. Tipping Rates: New Buyers, First Exposures



Note. $^{**}p < 0.01$, $^{***}p < 0.001$

Figure G2: App F.E. Tipping Rates: Repeat Buyers, First Exposures



Note. $^{**}p < 0.01$, $^{***}p < 0.001$.

Table G1. App F.E. Treatment Effects on Tip Amounts

	(1) All Buyers, All Exposures	(2) All Buyers, First Exposures	(3) New Buyers, All Exposures	(4) New Buyers, First Exposures	(5) Repeat Buyers, All Exposures	(6) Repeat Buyers, First Exposures
Intercept	0.744***	0.853***	0.856***	0.930***	0.708***	0.824***
	(0.026)	(0.036)	(0.052)	(0.057)	(0.030)	(0.044)
Norms	0.105**	0.170***	0.212**	0.303***	0.070	0.119
	(0.037)	(0.049)	(0.081)	(0.088)	(0.042)	(0.058)
Num. Obs.	61,705	35,480	15,194	9,570	46,511	25,910
Num. Buyers	35,480	35,480	9,570	9,570	25,910	25,910
Num. Sellers	26,562	18,327	9,074	6,273	22,637	15,233
R2	0.000	0.000	0.001	0.001	0.000	0.000
Adj R-sq.	0.000	0.000	0.001	0.001	0.000	0.000

Note. **p<0.01; ***p<0.001

Table G2. App F.E. Treatment Effects on Tip Percentages

	(1) All Buyers, All Exposures	(2) All Buyers, First Exposures	(3) New Buyers, All Exposures	(4) New Buyers, First Exposures	(5) Repeat Buyers, All Exposures	(6) Repeat Buyers, First Exposures
Intercept	3.221***	3.414***	3.350***	3.412***	3.179***	3.415***
	(0.112)	(0.124)	(0.220)	(0.211)	(0.130)	(0.151)
Norms	0.437**	0.601***	0.819	0.971**	0.311	0.463
	(0.163)	(0.181)	(0.326)	(0.331)	(0.188)	(0.215)
Num. Obs.	61,705	35,480	15,194	9,570	46,511	25,910
Num. Buyers	35,480	35,480	9,570	9,570	25,910	25,910
Num. Sellers	26,562	18,327	9,074	6,273	22,637	15,233
R2	0.000	0.000	0.001	0.001	0.000	0.000
Adj R-sq.	0.000	0.000	0.001	0.001	0.000	0.000

Note. **p<0.01; ***p<0.001

We checked whether App treatments affected treated buyers' post-treatment tipping in Web transactions. Table G3 summarizes. First-exposure subsamples are not reported, because all first exposures occur in the App, by definition of first exposures.

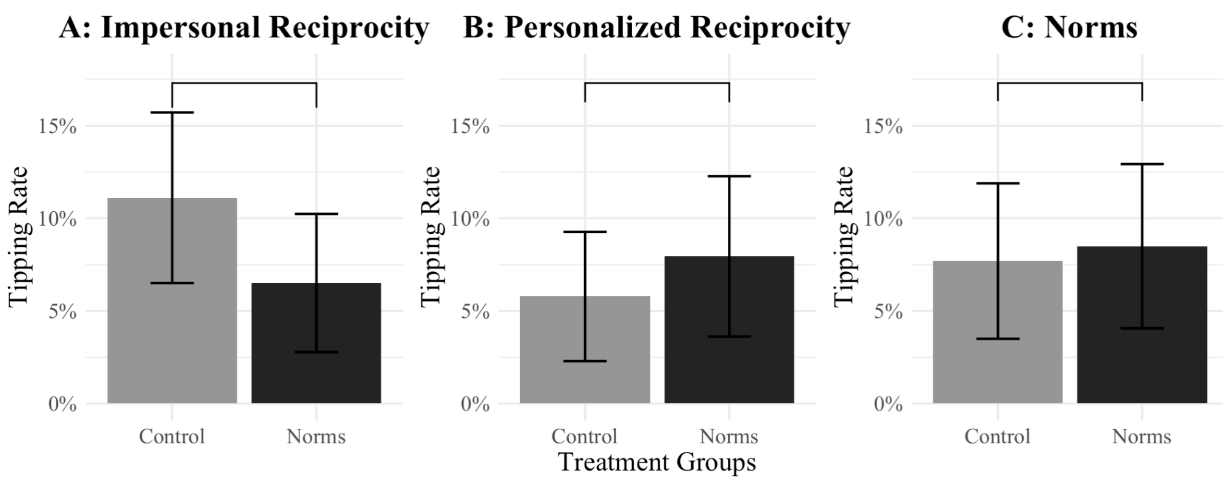
Table G3. App F.E. Treatment Effects on Post-Treatment Web Tipping Rates

	(1) All Buyers, All Exposures	(2) New Buyers, All Exposures	(3) Repeat Buyers, All Exposures
Intercept	0.067***	0.051***	0.069***
	(0.004)	(0.010)	(0.004)
Norms	0.001	0.038	−0.003
	(0.006)	(0.017)	(0.006)
Num. Obs.	12,768	1,180	11,588
Num. Buyers	5,560	641	4,919
Num. Sellers	9,203	1,066	8,519
R2	0.000	0.006	0.000
Adj R-sq.	−0.000	0.005	−0.000

Note. **p<0.01; ***p<0.001

We searched for crossover effects between the two field experiments. That is, did App Field Experiment behaviors differ across Web Field Experiment treatment groups? We focus on those who were initially exposed to Impersonal Reciprocity (N = 349), Personalized Reciprocity (N = 324), or Norms (N = 309) messages in the Web Field Experiment and received the Norms message in the App Field Experiment. These groups did not exhibit significantly different tipping rates on their first app transaction compared to the Control group, indicating that prior exposure to the two Reciprocity messages or the Norms message did not produce detectable spillover effects on later tipping behavior (see Figure G3).

Figure G3: Tipping Rates by App F.E. Treat Group, by Web F.E. Treatment:
Repeat Buyers, First Exposures



Note. $^{**}p < 0.01$, $^{***}p < 0.001$.

Web Appendix H. Norms-Driven Tipping Effects on Subsequent Platform Behaviors

Table H1. Web F.E. First-stage IV Regressions, All Buyers

	Buyer			Seller			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.148***	0.148***	0.148***	0.107***	0.107***	0.107***	0.107***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Norms/3-defaults	0.031***	0.031***	0.031***	0.018***	0.018***	0.018***	0.018***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Num. Obs	26,644	26,644	26,644	25,357	25,357	25,357	25,357
Num. Buyers	26,644	26,644	26,644	14,848	14,848	14,848	14,848
Num. Sellers	16,870	16,870	16,870	25,357	25,357	25,357	25,357
R2	0.002	0.002	0.002	0.001	0.001	0.001	0.001
Adj R-sq	0.002	0.002	0.002	0.001	0.001	0.001	0.001
F	47.84	47.84	47.84	19.89	19.89	19.89	19.89

Note. **p<0.01; ***p<0.001

Table H2. Web F.E. Second-stage IV Regressions, All Buyers

	Buyer			Seller			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.871***	104.221***	2.224***	0.565***	114.337***	2.244***	0.104
	(0.021)	(17.749)	(0.247)	(0.040)	(28.753)	(0.445)	(0.104)
$\widehat{Tip}$	0.051	−5.678	−1.403	0.287	−10.619	0.311	−0.063
	(0.128)	(106.169)	(1.482)	(0.346)	(247.486)	(3.845)	(0.915)
Num. Obs	26,644	26,644	26,644	25,357	25,357	25,357	25,357
Num. Buyers	26,644	26,644	26,644	14,848	14,848	14,848	14,848
Num. Sellers	16,870	16,870	16,870	25,357	25,357	25,357	25,357

Note. **p<0.01; ***p<0.001

Table H3. Web F.E. First-stage IV Regressions, New Buyers

	New Buyer			Seller with New Buyer			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.146***	0.146***	0.146***	0.112***	0.112***	0.112***	0.112***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Norms/3-Defaults	0.071***	0.071***	0.071***	0.041***	0.041***	0.041***	0.041***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Num. Obs	3,914	3,914	3,914	3,160	3,160	3,160	3,160
Num. Buyers	3,914	3,914	3,914	1,995	1,995	1,995	1,995
Num. Sellers	3,293	3,293	3,293	3,160	3,160	3,160	3,160
R2	0.009	0.009	0.009	0.004	0.004	0.004	0.004
Adj R-sq	0.008	0.008	0.008	0.003	0.003	0.003	0.003
F	33.80	33.80	33.80	11.79	11.79	11.79	11.79

Note. **p<0.01; ***p<0.001

Table H4. Web F.E. Second-stage IV Regressions, New Buyers

	New Buyer			Seller with New Buyer			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.913***	37.670	1.300***	0.579***	62.835	1.691***	−0.056
	(0.021)	(16.550)	(0.192)	(0.058)	(24.415)	(0.476)	(0.239)
$\widehat{Tip}$	0.125	213.990	1.418	−0.205	82.909	−0.424	1.184
	(0.112)	(95.936)	(1.065)	(0.437)	(187.454)	(3.648)	(2.085)
Num. Obs	3,914	3,914	3,914	3,160	3,160	3,160	3,160
Num. Buyers	3,914	3,914	3,914	1,995	1,995	1,995	1,995
Num. Sellers	3,293	3,293	3,293	3,160	3,160	3,160	3,160

Note. **p<0.01; ***p<0.001

Table H5. App F.E. First-stage IV Regressions, All Buyers

	Buyer			Seller			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.084***	0.084***	0.084***	0.081***	0.081***	0.081***	0.081***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Norms	0.017***	0.017***	0.017***	0.010**	0.010**	0.010**	0.010**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Num. Obs	35,480	35,480	35,480	26,562	26,562	26,562	26,562
Num. Buyers	35,480	35,480	35,480	18,893	18,893	18,893	18,893
Num. Sellers	18,327	18,327	18,327	26,562	26,562	26,562	26,562
R2	0.001	0.001	0.001	0.000	0.000	0.000	0.000
Adj R-sq	0.001	0.001	0.001	0.000	0.000	0.000	0.000
F	31.48	31.48	31.48	8.86	8.86	8.86	8.86

Note. **p<0.01; ***p<0.001

Table H6. App F.E. Second-stage IV Regressions, All Buyers

	Buyer			Seller			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.336***	20.341***	0.377***	0.421***	39.842	0.749***	0.113
	(0.027)	(5.324)	(0.065)	(0.051)	(15.781)	(0.227)	(0.067)
$\widehat{Tip}$	0.097	40.837	0.486	−0.189	57.531	0.481	−0.859
	(0.294)	(56.967)	(0.705)	(0.593)	(183.813)	(2.652)	(0.770)
Num. Obs	35,480	35,480	35,480	26,562	26,562	26,562	26,562
Num. Buyers	35,480	35,480	35,480	18,893	18,893	18,893	18,893
Num. Sellers	18,327	18,327	18,327	26,562	26,562	26,562	26,562

Note. **p<0.01; ***p<0.001

Table H7. App F.E. First-stage IV Regressions, New Buyers

	New Buyers			Sellers with New Buyers			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.095***	0.095***	0.095***	0.095***	0.095***	0.095***	0.095***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Norms	0.025***	0.025***	0.025***	0.020	0.020	0.020	0.020
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)

Num. Obs	9,570	9,570	9,570	5,765	5,765	5,765	5,765
Num. Buyers	9,570	9,570	9,570	4,465	4,465	4,465	4,465
Num. Sellers	6,273	6,273	6,273	5,765	5,765	5,765	5,765
R2	0.002	0.002	0.002	0.001	0.001	0.001	0.001
Adj R-sq	0.002	0.002	0.002	0.001	0.001	0.001	0.001
F	15.80	15.80	15.80	6.16	6.16	6.16	6.16

Note. **p<0.01; ***p<0.001

Table H8. App F.E. Second-stage IV Regressions, New Buyers

	New Buyers			Sellers with New Buyers			
	Any Repurchase (1)	Total Spending (2)	Count 5stars (3)	Gigs Completed (4)	Total Earnings (5)	Count 5stars (6)	Count Tip Mention (7)
Intercept	0.273***	25.978***	0.387***	0.461***	45.294	1.009***	0.035
	(0.041)	(7.489)	(0.088)	(0.071)	(24.071)	(0.301)	(0.082)
$\widehat{Tip}$	0.243	−59.229	−0.455	−0.443	50.102	−1.829	0.058
	(0.380)	(68.114)	(0.802)	(0.674)	(227.132)	(2.846)	(0.790)
Num. Obs	9,570	9,570	9,570	5,765	5,765	5,765	5,765
Num. Buyers	9,570	9,570	9,570	4,465	4,465	4,465	4,465
Num. Sellers	6,273	6,273	6,273	5,765	5,765	5,765	5,765

Note. **p<0.01; ***p<0.001

Web Appendix I. Transaction Features that Predict Norms Effects

Table I1 reports the parameter estimates that are discussed in Section 6.4.

Table I1. Web and App F.E. Potential Moderators of Norms Effects on Tipping

	Web F.E		App F.E.	
	All Buyers, First Exposures (1)	New Buyers, First Exposures (2)	All Buyers, First Exposures (3)	New Buyers, First Exposures (4)
Intercept	0.004	0.014	−0.032***	−0.047***
	(0.008)	(0.019)	(0.003)	(0.005)
Norms	−0.012	−0.016	0.006	0.031***
	(0.011)	(0.028)	(0.004)	(0.008)
NorthAmerica	0.135***	0.139***	0.061***	0.088***
	(0.006)	(0.015)	(0.004)	(0.007)
5stars	0.079***	0.058**	0.123***	0.130***
	(0.008)	(0.020)	(0.003)	(0.007)
HighReputation	0.002	0.031	−0.010	−0.010
	(0.006)	(0.016)	(0.026)	(0.041)
Norms × North America	0.024**	0.036	0.009	−0.011
	(0.009)	(0.024)	(0.006)	(0.011)
Norms × 5stars	0.029	0.089**	0.010	0.002
	(0.012)	(0.029)	(0.005)	(0.010)
Norms × HighReputation	0.007	−0.024	0.007	0.045
	(0.009)	(0.025)	(0.038)	(0.070)
Observations	26,644	3,914	35,480	9,570
Num. Buyers	26,644	3,914	35,480	9,570
Num. Sellers	16,870	3,293	18,327	6,273
R^2	0.047	0.062	0.060	0.060
Adjust R^2	0.047	0.060	0.059	0.059

Note. ** $p < 0.01$, *** $p < 0.001$.

References

Azar, Ofer H. 2007. "The Social Norm of Tipping: A Review." *Journal of Applied Social Psychology* 37, no. 2: 380–402.

Chandar, Bharat, Uri Gneezy, John A. List, and Ian Muir. 2019. "The Drivers of Social Preferences: Evidence from a Nationwide Tipping Field Experiment." NBER Working Paper 26380. Cambridge, MA: National Bureau of Economic Research.

Chen, Jinjie, Alison Jing Xu, Maria A. Rodas, and Xuefeng Liu. 2023. "Order Matters: Rating Service Professionals First Reduces Tipping Amount." *Journal of Marketing* 87, no. 1: 81–96.

Cialdini, Robert B. 1993. *Influence: The Psychology of Persuasion*. Revised ed. New York: William Morrow.

Cohn, Alain, Tobias Gesche, and Michel André Maréchal. 2022. "Honesty in the Digital Age." *Management Science* 68, no. 2: 827–845.

Lapidot-Lefler, Noam, and Azy Barak. 2012. "Effects of Anonymity, Invisibility, and Lack of Eye-Contact on Toxic Online Disinhibition." *Computers in Human Behavior* 28, no. 2: 434–443.

Lu, Shijie, Dai Yao, Xingyu Chen, and Rajdeep Grewal. 2021. "Do Larger Audiences Generate Greater Revenues under Pay What You Want? Evidence from a Live Streaming Platform." *Marketing Science* 40 (5): 964–984.

Lynn, Michael. 2016. "Motivations for Tipping: How They Differ across More and Less Frequently Tipped Services." *Journal of Behavioral and Experimental Economics* 65: 38–48.

Lynn, Michael, and Michael J. Sturman. 2010. "Tipping and Service Quality: A Within-Subjects Analysis." *Journal of Hospitality & Tourism Research* 34, no. 2: 269–275.

Lynn, Michael, George M. Zinkhan, and Judy Harris. 1993. "Consumer Tipping: A Cross-Country Study." *Journal of Consumer Research* 20, no. 3: 478–488.