Tipping in a Digital Services Marketplace

Seung Hyun Kim* On Amir† Kenneth C. Wilbur ‡

Abstract

Digital tipping behavior is opaque so norms are unclear. We describe tipping behavior in a digital services freelance marketplace, finding that tipping depends mostly on buyer factors including geography and satisfaction. Next, we report a field experiment that tested an injunctive norm message about tipping against two reciprocity-related messages and a status quo message. The Norms treatment increased tipping rate by 40% for new buyers and 10% for repeat buyers, relative to the status quo message. The two reciprocity treatments did not significantly change tipping behavior. The Norms-driven tipping increase did not significantly change buyers' subsequent platform usage, seller pricing or seller effort. Two post-experiment platform design changes show that reverting to the status-quo message causally reduced tipping by Norms-treated buyers, and then platform-wide adoption of the Norms message corresponded with meaningfully higher tipping rates. Collectively, the results indicate the importance of digital platform design in establishing market norms, the greater impressionability of new platform users, and also the inherent limitations and measurement challenges of platform design on economic behavior.

Keywords: Field Experiment, Market Design, Platforms, Pricing, Services, Tipping

*Doctoral Student, Rady School of Management,Seung.Kim@rady.ucsd.edu
†Professor of Marketing, Rady School of Management, oamir@ucsd.edu
‡Professor of Marketing and Analytics, Rady School of Management, kennethcwilbur@gmail.com

1. Introduction

Suppose you purchased a digital service–like copy editing, graphic design, or professional translation–in an online gig marketplace. You receive the work, finalize the payment, rate the provider, and then the platform asks you to leave a tip. How would you respond–would you voluntarily increase your payment? What factors would affect your decision? Would you know how often other buyers leave tips? What would the service provider expect you to do? How much would the tip affect the provider's income? How would leaving a tip affect future choices–by you, the provider or the platform?

Digital tipping is opaque: most platforms that request tips do not share tipping data with buyers. Buyers can gather some information from default tips provided by a platform, but they know the platform's economic incentives may influence those defaults. Further, buyers and sellers may have diverse expectations, as a single digital platform may serve diverse participants and purposes in global markets. Yet tipping may have substantial positive effects, as research in traditional contexts shows that voluntary tipping can motivate service effort, enable price discrimination and help retain talented workers. Uncoordinated tipping norms affect many people: in the U.S. alone, 16% of those surveyed said they had earned money from online gig platforms, 31% of whom said it was their main job within the past year (Anderson et al. 2021).

What motivates buyer tipping and how does tipping affect downstream actions? This paper investigates the possibility that platform messages could help shape participants' economic behaviors. Collectively, the results indicate that (i) some platform messages can meaningfully increase tipping; (ii) digital platform design influences economic behavior with larger effects on new users than on returning users; and (iii) platform design effects can be limited in size and difficult to estimate precisely even with full population data.

We first describe tips observed in 4.1 million transactions on Fiverr and document key facts about online tipping. Tipping correlates strongly with buyer region but not with seller region. Tipping Rate increases with buyer satisfaction, rising from 3.5% in 3-star transactions up to 14.1% in 5-star transactions. Buyer fixed effects explain nearly half of buyer tipping behavior even after controlling for other transaction characteristics, whereas seller fixed effects explain just 6% of buyer tipping behavior.

Next, we report a field experiment we designed to help the platform encourage tipping. The experiment manipulated the default tips and tip requests in three conditions labeled Reciprocity, Implicit Reciprocity, and Norms. A manipulation check showed that the three treatment messages relate to the two most common tipping motivations, namely, gratitude and compliance with social norms, but it was not clear ex-ante which treatments would affect digital tipping behavior the most.

We find that the Norms treatment increased new buyer tipping propensity by 7.1% more than the status quo message, and repeat buyer tip propensity by 0.5% upon first exposure. New buyers who tipped after the Norm treatment were more likely to tip on their next transaction. Tipping rate increases came without any detectable reductions in repeat purchase behaviors. Tipping by Implicit Reciprocity- and Reciprocity-treated customers were statistically indistinguishable from Control.

We then investigate two discontinuities in post-test platform design. First, the platform reverted all tip messages to the status quo after the test period ended, reducing Norms-treated new buyers' tipping rate by 3.9%. Then, three weeks later, the platform adopted the Norms message for all buyers, corresponding to meaningfully higher tipping rates among both new and returning buyer cohorts.

We explore some mechanism evidence in the field experiment data by interacting treatments with moderators related to tipping motivations. The Norms treatment effect on new buyers increased with buyer satisfaction rating but did not change with buyer region.

Finally, we examine the indirect effects of treatment on sellers. Norms treatment increased seller revenue, but we find no significant effects of treatment on sellers' repeat transactions. We also find no significant treatment effects on sellers' subsequent prices, transaction ratings, tip mentions, or tips received.

1.1. Literature Review

Numerous papers in economics, marketing, psychology, and other fields study tipping in traditional offline contexts, especially restaurants, transportation, tourism, and leisure services. Many papers investigate consumer motivations for tipping in traditional contexts (Azar 2007, 2011, 2020, Lynn 2015, 2016, Donkor 2021, Bluvstein Netter and Raghubir 2021). Much of this literature has identified two major factors - Norms and Reciprocity - that motivate consumers to pay tips. Less frequent motivations included monitoring future service quality, feeling guilty, repeated interactions, and impression management strategy.

Despite a shared focus on tipping behavior, digitally intermediated transactions differ greatly from traditional contexts, so we are unsure how much of the previous literature will carry over to newer settings. For example, the Fiverr platform we study consists of mostly business-to-business transactions; global market participants; purely digital service delivery; rarely repeated interactions between buyers and sellers; opaque tipping practices, so expected norms are unclear; and asynchronous online buyer-seller communication without face-to-face interactions. All of those factors differ from traditional in-person tipping contexts like restaurants or transportation. Consider as an example Donkor (2021), who estimates a structural model to size consumers' costs of opting out of suggested tips in order to estimate consumers' ideal tipping points. He interprets the opt-out costs as "guilt or shame of not tipping the norm." In our setting, by contrast, we report an experiment run by the platform to understand tipping motivations. We therefore try to avoid imposing a particular behavioral theory on the data, and instead try to use statistical models to learn what the data can teach us about how the platform can motivate tipping behavior online.

Our study is closely related to papers on tipping in digitally mediated transactions using field data. Chandar et al. (2019) use Uber data to characterize how tipping covaries with demand, supply, and transaction features. A field experiment treated default tip options and increased average tip percentage by 2.5%. Duhaime and Woessner (2019) found that buyers tip employees more than they tip autonomous gig workers, suggesting that seller autonomy in the gig economy can partially explain lower tipping rates. Alexander et al. (2021) reports a field experiment in which a laundry service manipulated the default tips in a smartphone application used for customer communication and payment. Larger default tip options decreased tipping rate but increased tip revenues overall. Lu et al. (2021) report a field experiment in which a platform manipulated audience size during livestreaming sessions, finding that viewer tipping exhibited an inverted-U shape in audience size. We contribute to this literature with the first evidence of how and why tip request messages can increase consumers' voluntary, post-transaction payments to sellers.

We also contribute to a small but growing literature on digital platform design changes, as digital tips often occur within platform contexts. Chen et al. (2022) propose that platforms should request tips before requesting ratings, documenting that buyers view ratings as a reward for service, which reduces their obligation to tip. Blake et al. (2021) experimentally manipulated fee disclosure timing on StubHub, finding that late fee disclosure increases both the quantity and quality of ticket purchases compared to early fee disclosure. We add to this digital marketplace design literature by documenting heterogeneity in the tipping message effect by buyer cohort and experience; by showing how buyers' repeat purchases change with tipping conditions; and by documenting seller reactions to induced changes in buyer tipping.

2. Empirical Context

We study data from Fiverr, a marketplace for freelance digital services that matches buyers with sellers. Fiverr describes its mission as "to change how the world works together. The Fiverr platform connects businesses of all sizes with skilled freelancers offering digital services in more than 300 categories across 8 verticals including graphic design, digital marketing, programming, video, and animation."

Sellers list "gigs," which are packages of service attributes and prices. Seller listings and buyer search are free, but Fiverr charges transparent fees on both sides of the market, adding 5% on to buyers' order price and taking 20% from total seller revenue.[1]

[1] Fiverr's transparent fee structure contrasts with revelations of misleading tip policies used by Doordash and Instacart (Newman 2019, Hanbury 2019).

Fiverr connected 2.4 million buyers with 305,000 sellers in 2019, generating \$408 million in marketplace transactions and \$107 million in platform income, up 42% from 2018. Fiverr competed with Upwork.com, PeoplePerHour.com, and Freelancer.com, among others. A typical transaction has four steps.

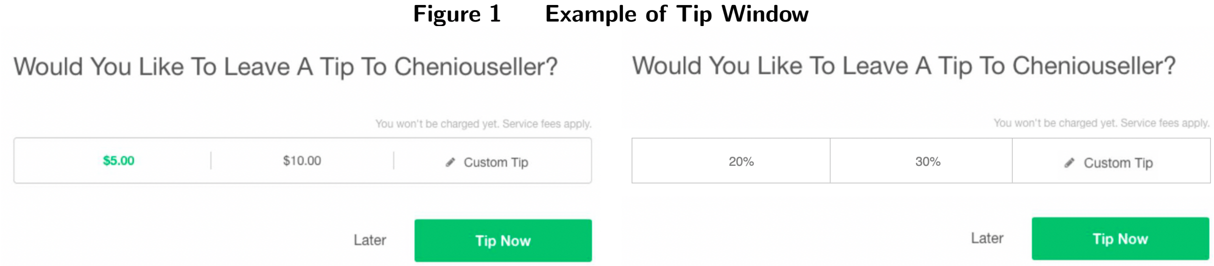
I. Search Buyer explores the directory of service listings or searches specific keywords and views results. Each gig listing includes the seller's name, service description, average rating, and price.

II. Order Buyer clicks one or more gig listings, leading to a dedicated page that describes service details, seller information including image and country, past ratings and reviews, and other details. Buyer can contact the seller digitally via in-platform messages regarding gig attributes, price, delivery date, and requirements. Buyer pays Fiverr upon ordering.

III. Delivery The seller completes the service, and Fiverr digitally delivers the work to the buyer. The buyer confirms receipt, and Fiverr pays the seller.

IV. Rating and Tip Fiverr asks the buyer to rate the service. Fiverr then asks the buyer to tip the seller with default options shown in Figure 1.[2] Prices up to \$25 led to default options of \$5, \$10, and Custom; prices above \$25 led to default options of 20%, 30%, and Custom. Fiverr also emailed rating requests 24 hours after service delivery.

Figure 1 Example of Tip Window



(a) When price $\leq$ \$25 (b) When price $>$ \$25

3. Tipping Facts

This section empirically describes buyer tipping behavior and establishes some key facts. We analyze all marketplace transactions from January 1, 2019, until the start of the field experiment in June 2019. The pre-experiment data include about 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.

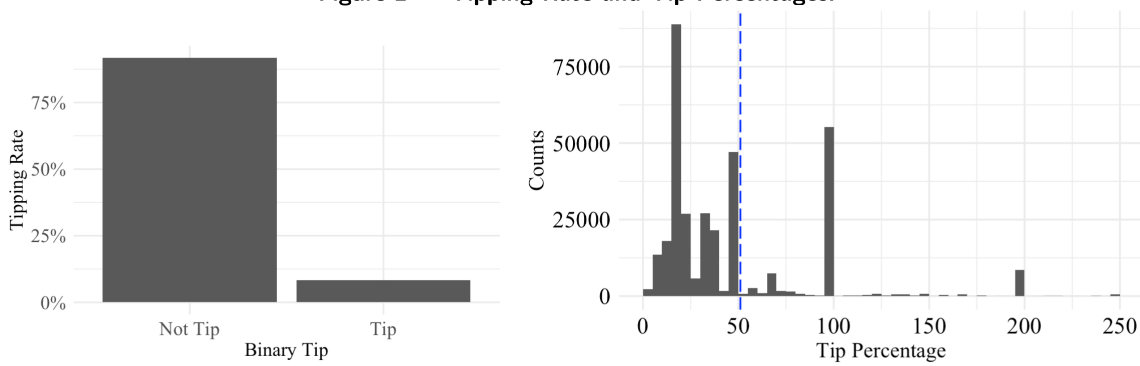
[2] Fiverr did not request a tip after 1- or 2-star ratings, representing 1.1% of transactions.

3.1. Tipping Measures and Distributions

Buyers tipped sellers in 8.3% of transactions overall on the platform. Among tipped transactions, the average tip observed is 51.1% of the order price.

At the buyer level, 85.4% of buyers never tipped, while 5.3% of buyers always tipped. The remaining 9.3% of buyers tipped in 23.9% of their transactions. However, buyer-level analysis confounds tipping with the number of orders observed for each buyer. If we limit attention to the 15.7% of buyers with 5 or more transactions, collectively accounting for 58% of all transactions, then we find that 68.4% of frequent buyers never tipped, and only 0.7% of frequent buyers always tipped.

Figure 2 Tipping Rate and Tip Percentages.



(a) Tipping Rate (b) Tip Percentages

Note. Vertical line in Panel (b) represents the mean. Tip Percentage axis is truncated from its maximum of 2,800%.

Next, we look at how tipping covaries with price. Figure 3(a) classifies orders into comparably-sized bins and shows that 5% of \$5 dollar transactions are tipped, rising to 11% of \$30-\$40 transactions, and then falling back to 9% of transactions priced over \$150. Figure 3(b) displays how box plots of Tip Percentage distributions change with price, with red dots representing the average tip percentage. The average tip percentage decreases with price up to \$25 but changes little with price after that point. Similarly, the median tip percentage is nearly constant at the 20% default for prices above \$25. We also found that buyers chose default tips in 68% of all tipped transactions even though they could easily enter custom amounts. In what follows, we mainly focus on one measure of tipping behavior: *Tipping Rate*, the proportion of transactions that were tipped.

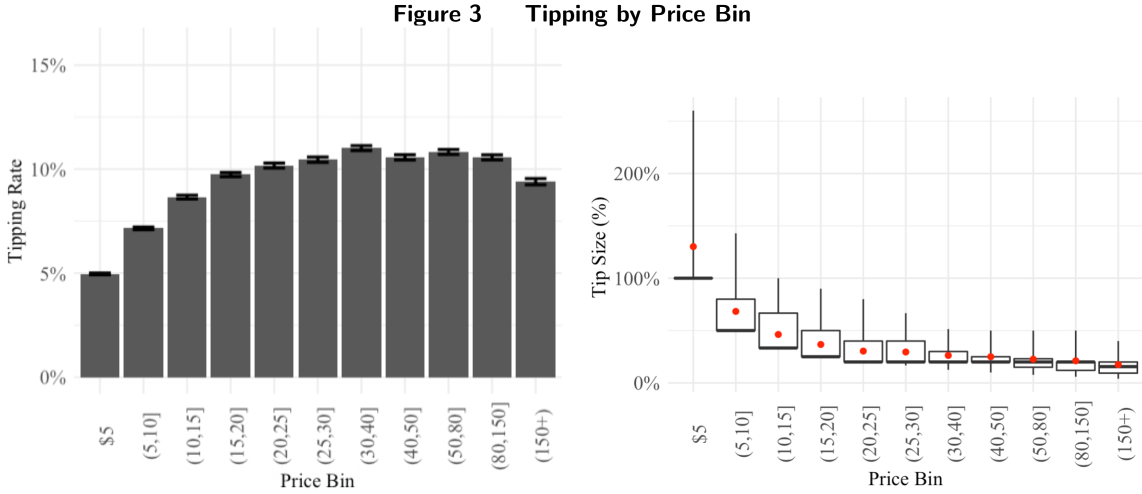
3.2. Key Observations about Online Tipping

Here we provide model-free evidence to motivate key features of the experiment and the analysis.

Key Fact I. Buyer region predicts tipping more strongly than seller region.

Do offline tipping norms affect online tips? Figure 4(a) shows that North Americans tipped 11.5% of all transactions, whereas buyers in other regions tipped in 3.9% of transactions on average. One

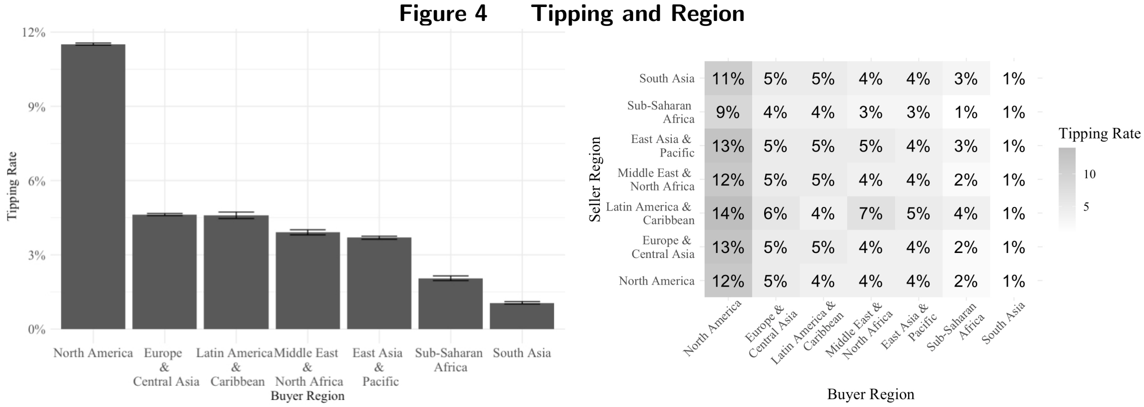
Figure 3 Tipping by Price Bin



(a) (b)

might hypothesize a home-region bias in which buyers tip more to culturally similar sellers or a "tourist" effect in which buyer tipping matches local expectations in the seller's region. Figure 4(b) shows that buyer region affects Tipping Rate much more than seller region, and there is little evidence of a home-region bias or tourist effect. For example, North American buyers tipped North American sellers at about the same rate as most other regions. In fact, no region's buyers tipped sellers from their home region much more frequently than other regions, and global buyers tipped North American sellers at rates similar to tips to non-North American sellers. This key fact indicates that buyer regional social norms could predict online tipping.

Figure 4 Tipping and Region



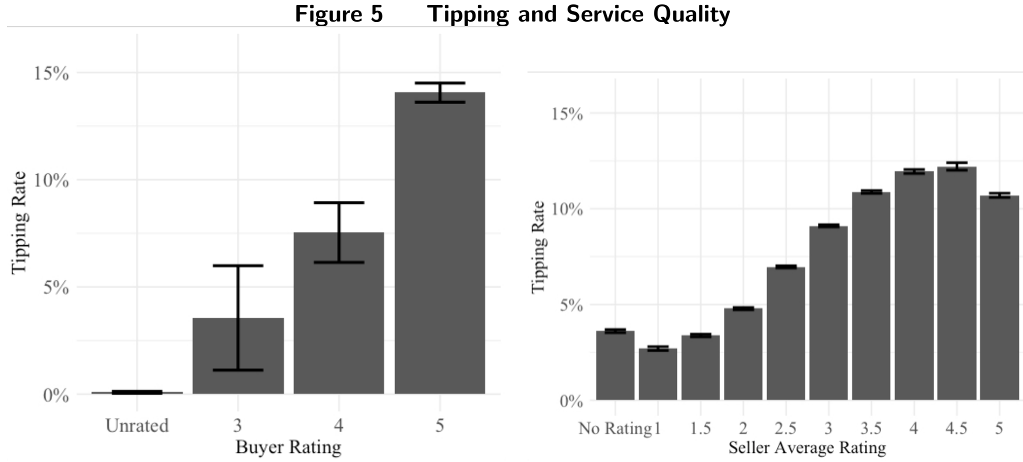
(a) (b)

Key Fact II. Buyer tipping increases with satisfaction rating.

Next, we look at how tipping covaries with buyer satisfaction rating. Buyers rated 61% of transactions, with 93.6% of all ratings being five stars. Figure 5(a) shows that Tipping Rate was four times higher among 5-star transactions than 3-star transactions.[3]

We also consider how tipping covaries with sellers' previous average ratings, as seller quality is probably autocorrelated across gigs. Figure 5(b) shows that Tipping Rate increases rapidly with sellers' average rating, suggesting that buyer satisfaction rating is partly driven by the seller's identity. Thus, past seller ratings could predict future tips in addition to buyer satisfaction ratings. Key fact II suggests that buyers may reciprocate high-quality service with tips.

Figure 5 Tipping and Service Quality



(a) (b)

Key Fact III. Buyer characteristics, seller characteristics, and transaction characteristics all help to predict tipping, but buyer effects matter most.

We run descriptive regression with the pre-experiment data to simultaneously estimate multiple associations between tipping and buyers, sellers, and transaction characteristics. Buyer characteristics include Buyer Region, Buyer Prior Orders (a count of transactions since the first day of use in 2019), Buyer Prior Tips (a count of tips since the first day of use in 2019), and Buyer Tenure (time elapsed since the buyer's first transaction date). Seller characteristics include Seller Average Rating and Seller Prior Orders (a count of prior orders delivered since the first day of use in 2019). Lastly, transaction characteristics include transaction price, service category, and Seller Tip Mention, which is a binary variable including whether the seller's gig delivery message to the buyer includes "tip" or "tips." We also report specifications that replace buyer and seller characteristics with individual buyer and seller fixed effects.

Equation (1) specifies the descriptive regression:

[3] 99.9% of unrated transactions were also untipped, likely because the tip request immediately followed the rating request.

$$y_i = f(X_i^B, X_i^T, X_i^S, \epsilon_i) \tag{1}$$

where i indexes transactions; y_i is a binary tipping indicator; X_i^B, X_i^T and X_i^S represent buyer, transaction, and seller characteristics, respectively; and ϵ_i captures all unobserved factors that affect tipping.

Table 1 displays parameters estimated using linear regression. Column (1) reports a model with buyer and seller characteristics but no buyer or seller fixed effects. Column (2) replaces buyer characteristics with buyer fixed effects, whereas Column (3) replaces seller characteristics with seller fixed effects. Column (4) includes both buyer and seller fixed effects.

Tipping correlates positively with Buyer Rating, Seller Tip Mention, and Buyer Prior Tips. However, the magnitudes of the relationships change when we control for buyer and seller fixed effects. For example, tipping changes less with buyer rating when we control for buyer and seller fixed effects. Similarly, price does not significantly predict tipping when both buyer and seller characteristics are included in the model.

The model's explanatory power increases substantially when it includes buyer and seller fixed effects. Most of the model fit improvement comes from buyer fixed effects, with the Adjusted R-square statistic rising from .118 in column 1 to .396 in column 2. Seller fixed effects add a much smaller incremental of .021 in column (3) and .008 in column (4).

Therefore, we conclude that controlling for unobserved buyer characteristics is essential to size the associations. Next, we describe the field experiment design, data, and results.

4. Field Experiment: Design and Data

We designed a field experiment to understand and motivate tipping on the online platform. The experiment ran on the web and treated 7,880 new buyers and 40,823 repeat buyers. Treatment started in June 2019 and lasted four weeks. The platform refrained from running other tests in-market during the sample period to avoid potential confounds, even though it had often run concurrent experiments beforehand.

The control and three treatment messages were:

Control/Status Quo: "Would You Like to Leave a Tip To (Seller name)?"
Implicit Reciprocity: "Show your appreciation to your seller by giving a tip."
Reciprocity: "Leave (Seller name) a tip to show your appreciation for a job well done."
Norms: "It's customary to leave a tip for the seller's service."

Each treatment message emphasized a different theoretical motivation for tipping. Implicit Reciprocity says that tips convey the buyer's appreciation for the service, but does not name the seller

Table 1 **Descriptive Regression**

	(1)	(2)	(3)	(4)
Buyer Rating = 3	0.031***	0.020***	0.033***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
= 4	0.069***	0.052***	0.068***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
= 5	0.126***	0.117***	0.124***	0.114***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Price	0.0001***	−0.00002***	0.0001***	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Price^2	−0.000***	0.000***	−0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Seller Tip Mention	0.047***	0.038***	0.061***	0.054***
	(0.001)	(0.001)	(0.001)	(0.001)
Buyer Prior Order	−0.0003***		−0.0003***	
	(0.00000)		(0.00000)	
Buyer Prior Tips	0.017***		0.017***	
	(0.00004)		(0.00005)	
Buyer Tenure	−0.0001***		−0.0001***	
	(0.00001)		(0.00001)	
Seller Prior Order	−0.00000***	−0.00001***		
	(0.00000)	(0.00000)		
Seller Average Rating	0.001***	0.002***		
	(0.0001)	(0.0001)		
Category FE	Y	Y	Y	Y
Buyer Region FE	Y		Y	
Seller Region FE	Y	Y		
Buyer FE		Y		Y
Seller FE			Y	Y
Observations	4,134,928	4,134,928	4,134,928	4,134,928
R^2	0.118	0.582	0.175	0.610
Adjusted R^2	0.118	0.396	0.139	0.404

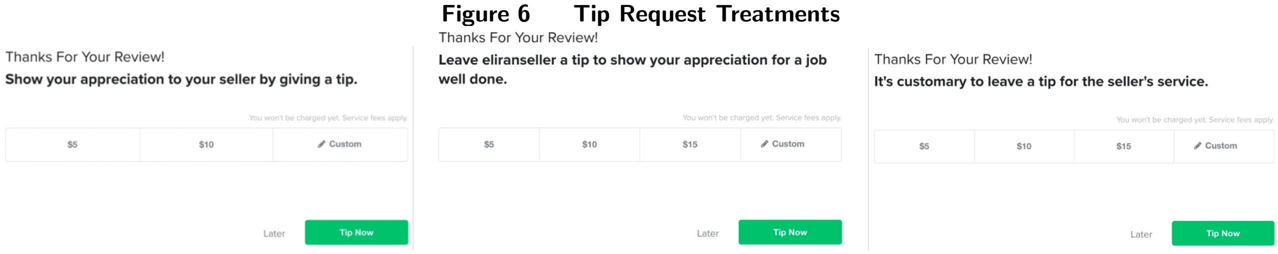
Note: **p<0.05; ***p<0.01

nor does it call out the explicit quid-pro-quo of the transaction.; this treatment also captures some sense of goodwill as an antecedent for tipping. Reciprocity explicitly recommends a tip in exchange for a "job well done," and retains the seller's name. The Norms message implies the buyer should conform to an injunctive social norm about tipping because they received the seller's service. The injunctive norm was a subjective view about what buyers should do as opposed to a descriptive norm about what buyers usually do. Figure 6 shows how the treatment messages looked to buyers, as Figure 1 showed how the control message appeared to buyers.

Appendix A presents a manipulation check of whether treatment messages map into the theoretical constructs of reciprocity and norm and how buyer interpretations of treatment messages change with seller name inclusion. 1,000 online participants were presented with a 4x2 set of messages, with and without seller names, and asked to map messages to 6 implied reasons to tip, with reasons drawn from Cialdini (2009). The manipulation check suggests that participants perceived both Implicit Reciprocity and Reciprocity messages as "Reciprocity" and the Norms message as "Social Proof (Norm)." The interpretation of all three treatment messages is similar regardless of the inclusion of the seller's name.

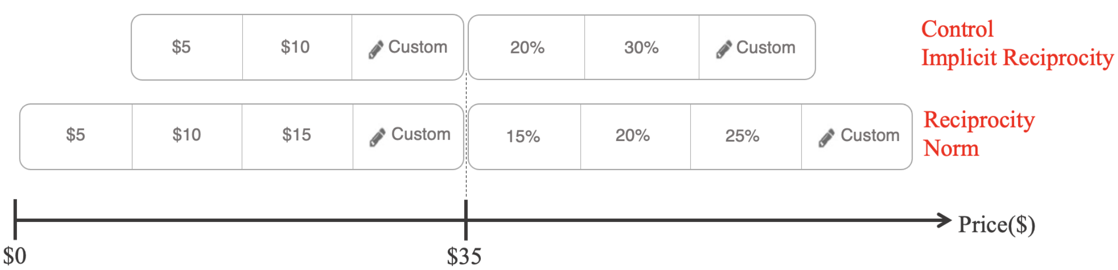
The platform simultaneously changed the default tips and threshold in the Reciprocity and Norms conditions, as shown in Figure 7. The new default tips contained three options rather than two, with custom tips. Further, the price threshold between the dollar and percentage default tips increased from \$25 to \$35[4]. We refer to the treatment groups by their treatment messages (e.g., Norms, Reciprocity, etc.). Later, we use two post-experiment platform design changes to disentangle the causal effects of messages from default tips.

Figure 6 Tip Request Treatments



(a) Implicit Reciprocity (b) Reciprocity (c) Norm

Figure 7 Default Tip Manipulation



4.1. Outcomes

We focus on Tipping Rate as the main outcome variable. We distinguish new buyers from repeat buyers because establishing a new tipping behavior may differ from changing existing behavior. We also examine buyers' repeat purchase behavior for three reasons. First, experimental treatments could directly affect buyers' transaction utility and repeat purchases. Second, tipping is costly. Thus, if buyers respond to the tipping treatments with more tips, that could operate as a price

[4] Firm executives had final control over the design. An optimal design might have manipulated the default tips and threshold independently from the messages, but that did not occur.

increase which may reduce subsequent demand. Third, platform revenues depend more on order volume than on tipping.

We analyze results in multiple ways. We estimate treatment effects on tipping in pooled transaction data. We also partition the sample by buyers' order number after the first treatment, pooling buyers' first rated transactions in the treatment period; buyers' second transactions; and buyers' third transactions.

Treatments were displayed on the tip request screen after the buyer confirmed receipt and rated the transaction. Therefore, all characteristics of each buyer's first treated transaction are fully exogenous to the treatment. However, subsequent transactions could have been affected by repeated exposure to treatment or differential platform usage after the first treatment.

4.2. Randomization Checks

We checked whether the firm successfully randomized buyers to treatments. We did this separately for new buyers and for repeat buyers because repeat buyers have more covariates available. Table 2 reports Chi-square tests indicating that the new buyer regions were well balanced across treatment groups. Table 3 reports Chi-square tests indicating that repeat buyers' characteristics were well balanced across treatment groups.

Table 2 Randomization Check - New Buyers

Variable	Control	Implicit Reciprocity	Reciprocity	Norms	P-value
Buyer Region					
East Asia & Pacific	0.08	0.09	0.09	0.08	0.3272
Europe & Central Asia	0.26	0.27	0.26	0.28	0.6664
Latin America & Caribbean	0.03	0.02	0.02	0.03	0.6833
Middle East & North Africa	0.04	0.04	0.04	0.03	0.4007
North America	0.50	0.50	0.51	0.50	0.9838
South Asia	0.06	0.05	0.05	0.05	0.178
Sub-Saharan Africa	0.03	0.04	0.03	0.02	0.1234

Note: The table shows new buyer regions by treatment group and Chi-square tests of the null hypothesis of random assignment.

5. Treatment Effects on New Buyers

5.1. Nonparametric Treatment Effects

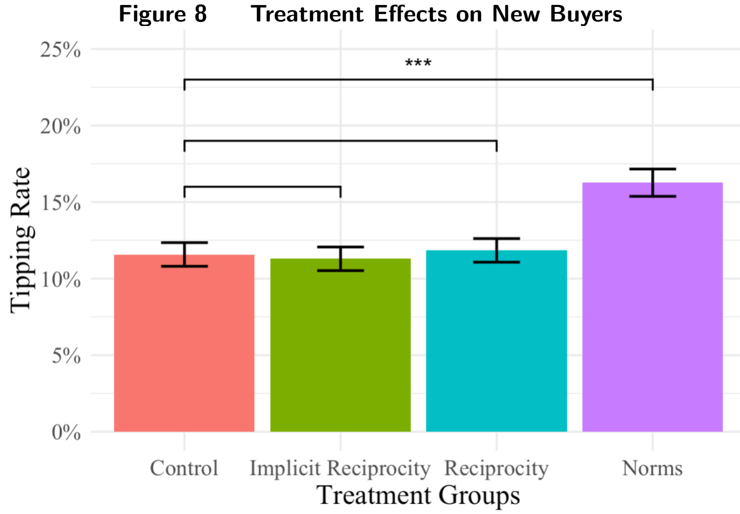
Figure 8 displays mean differences in new buyer Tipping Rate by treatment groups. New buyers who received the Norm treatment tipped about 40% more often than Control (16.3% vs. 11.6%). There were no significant differences in tipping among new buyers who received the Control, Implicit Reciprocity, or Reciprocity treatments.

Table 3 Randomization Check - Repeat Buyers

Variable	Control	Implicit Reciprocity	Reciprocity	Norms	P-value
Buyer Region					
East Asia & Pacific	0.09	0.09	0.09	0.09	0.3814
Europe & Central Asia	0.24	0.24	0.24	0.25	0.4157
Latin America & Caribbean	0.03	0.03	0.03	0.03	0.9626
Middle East & North Africa	0.04	0.04	0.03	0.03	0.6649
North America	0.55	0.54	0.55	0.54	0.437
South Asia	0.04	0.04	0.04	0.04	0.1231
Sub-Saharan Africa	0.02	0.02	0.02	0.02	0.2728
Buyer Characteristics					
Buyer Tenure	26.58	26.41	26.50	26.64	0.9264
Pre-Treatment Orders	14.07	14.15	14.51	14.05	0.6091
Pre-Treatment Expenditure	499.15	507.07	506.90	489.39	0.66
Pre-Treatment Tip Frequency	1.20	1.21	1.28	1.20	0.4138
Pre-Treatment Tip Expenditure	14.14	15.93	15.24	14.44	0.3877

Note: The table shows repeat buyer characteristics by treatment group and Chi-square tests of the null hypothesis of random assignment.

Figure 8 Treatment Effects on New Buyers



5.2. Fixed Effect Regression

We use regression to size treatment effects while controlling for covariates. Table 4 reports linear regressions of tipping on treatment effects in four transaction samples, controlling for buyer, seller, and transaction characteristics.[5] Across all transactions, new buyers exposed to the Norm condition had a 3.5% ($p<.01$) increased probability of tipping compared to those exposed to Control, all other things equal. The Norms effect was especially pronounced when new buyers were first exposed to the treatment, leading to an absolute increase of 7.1% ($p<.01$) compared to the Control.

[5] Appendix Table B1 shows that the qualitative results are robust to a Logit model specification

Second and third exposures to the Norms treatment are directionally positive but smaller and not significant. The Norms findings contrast with the Reciprocity and Implicit Reciprocity conditions, which resulted in smaller, non-significant effects across all transaction subsamples.

There are some interesting covariates in Table 4. Buyer satisfaction ratings and seller tip mentions become weaker predictors of tipping as new buyers become more experienced in using the platform. Tipping Rate rises with price. New buyers who tip are more likely to tip again, with prior tips increasing second-transaction tipping rates by 34% and third-transaction tipping rates by 28%. The number of past orders the seller received and the seller overall rating do not reliably predict tipping.

We also estimate treatment effects on new buyers' repeat purchases during the sample period, using the same regression specification. Table 5 shows the results. The outcome variable in column (1) equals one if a new buyer completed a second transaction during the sample period and zero otherwise. The outcome variable in column (2) analogously equals one if a third transaction occurred during the test period.

The results show that the Norms and Implicit Reciprocity treatments led to slightly more repeat purchases on average, whereas the Reciprocity treatment led to fewer repeat purchases, but none of the effects were statistically significant. The table also shows that a higher buyer satisfaction rating predicts that buyers will revisit the platform. Therefore, the platform viewed the Norms condition as increasing tipping without reducing new buyers' return visits to the platform.

6. Treatment Effects on Repeat Buyers

6.1. Nonparametric Treatment Effects

Figure 9 shows the mean differences in repeat buyer Tipping Rate by treatment groups. Repeat buyers exposed to the Norm treatment tipped about 10% more often than Control (11.5% vs. 10.5%). The effect is smaller but directionally similar to the treatment effect on new buyers. Like before, the other two treatments were very similar to Control, with 10.6% average tipping in the Implicit Reciprocity and Reciprocity treatment groups.

6.2. Difference-in-differences Analysis

We use a differences-in-differences regression model with buyer fixed effects, seller fixed effects, and transaction characteristics to control for confounding variables. We pool pre-test transactions with test period transactions, then specify Equation (2) to identify causal effects of treatments on Tipping Rate of repeat buyers.

$$y_i = \beta_1 \mathrm{Post}_i + \beta_2^{M_i}(\mathrm{Treat}_{M_i} \cdot \mathrm{Post}_i) + \beta_3 X_i^T + \mu_{b_i} + \lambda_{s_i} + \epsilon_i \tag{2}$$

Table 4 New Buyers Tipping Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity	0.001	0.011	−0.013	0.003
	(0.006)	(0.011)	(0.009)	(0.012)
Reciprocity	0.005	0.006	−0.003	0.006
	(0.006)	(0.011)	(0.010)	(0.012)
Norms	0.035***	0.071***	0.015	0.008
	(0.007)	(0.012)	(0.010)	(0.013)
Buyer Satisfaction = 3	0.009	0.042**	0.044**	0.002
	(0.012)	(0.017)	(0.019)	(0.022)
= 4	0.074***	0.110***	0.079***	0.073***
	(0.008)	(0.014)	(0.015)	(0.018)
= 5	0.141***	0.186***	0.148***	0.133***
	(0.005)	(0.010)	(0.006)	(0.009)
Price	0.0003***	0.0005***	0.0002**	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Price^2	−0.00000***	−0.00000***	−0.00000	−0.00000**
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Seller Tip Mention	0.038***	0.051	0.034	0.047
	(0.014)	(0.027)	(0.026)	(0.030)
Buyer Prior Order	−0.010***			
	(0.001)			
Buyer Prior Tips	0.135***		0.345***	0.286***
	(0.009)		(0.014)	(0.012)
Seller Prior Order	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00001)
Seller Average Rating	0.002	0.004	0.001	0.003
	(0.002)	(0.003)	(0.003)	(0.004)
Category FE	Y	Y	Y	Y
Buyer Region FE	Y	Y	Y	Y
Seller Region FE	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545
Number of Buyers	7,880	7,880	7,258	3,545
R^2	0.161	0.077	0.242	0.335
Adjusted R^2	0.160	0.073	0.238	0.329
Note:				**p<0.05; ***p<0.01

i indexes transactions; b_i and s_i indicate the unique buyer and seller in transaction i; M_i is the treatment randomly assigned to buyer b_i; y_i is a binary tipping indicator; Post_i is a treatment period dummy; Treat_{M_i} is a treatment dummy; X_i^T contains transaction characteristics–buyer satisfaction rating, order price, seller tip mention, and category dummies; μ_{b_i} is a buyer fixed effect; λ_{s_i} is a seller fixed effect; and ϵ_i is an error that captures all unobserved features that affect tipping, such as unobserved service quality. $\beta_2^{M_i}$ is the key coefficient that captures the causal effects of treatment by comparing the change in Tipping Rate across time within treatment groups.

Table 5 **New Buyer Repeat Transaction Regression**

	Second Order Exists	Third Order Exists
	(1)	(2)
Implicit_Reciprocity	0.005	0.011
	(0.008)	(0.016)
Reciprocity	−0.006	−0.017
	(0.009)	(0.016)
Norms	0.008	0.012
	(0.008)	(0.017)
Buyer Satisfaction = 3	0.342***	0.059
	(0.048)	(0.057)
= 4	0.329***	0.055**
	(0.046)	(0.027)
= 5	0.305***	0.058***
	(0.045)	(0.013)
Price	−0.00003	−0.001***
	(0.0001)	(0.0002)
Price^2	−0.00000	0.00000
	(0.00000)	(0.00000)
Seller Tip Mention	−0.019	0.058
	(0.019)	(0.037)
Buyer Prior Tips		−0.027
		(0.016)
Seller Prior Order	0.00000	−0.00002***
	(0.00000)	(0.00001)
Seller Average Rating	−0.0002	−0.024***
	(0.003)	(0.005)
Category FE	Y	Y
Buyer Region FE	Y	Y
Seller Region FE	Y	Y
Observations	7,880	7,258
Number of Buyers	7,880	7,258
R^2	0.024	0.028
Adjusted R^2	0.020	0.024
Note:		**p<0.05; ***p<0.01

Table 6 presents parameter estimates from Equation (2). The Norm treatment effect on repeat buyers was about 0.5% ($p < .01$) when all treated transactions were pooled, significantly less than the treatment effect on new buyers. Nonetheless, the Norms effect was also pronounced when repeat buyers were first exposed to treatment, leading to an increase of about 2% ($p < .01$) compared to the Control. Treatment effects on buyers' second and third treated transactions did not differ significantly from zero, so the data do not show that repeated exposure to the treatment reliably increases tipping among repeat buyers. We think these comparisons speak to the relative difficulty of changing established behaviors.

Figure 9 Treatment Effects on Repeat Buyers

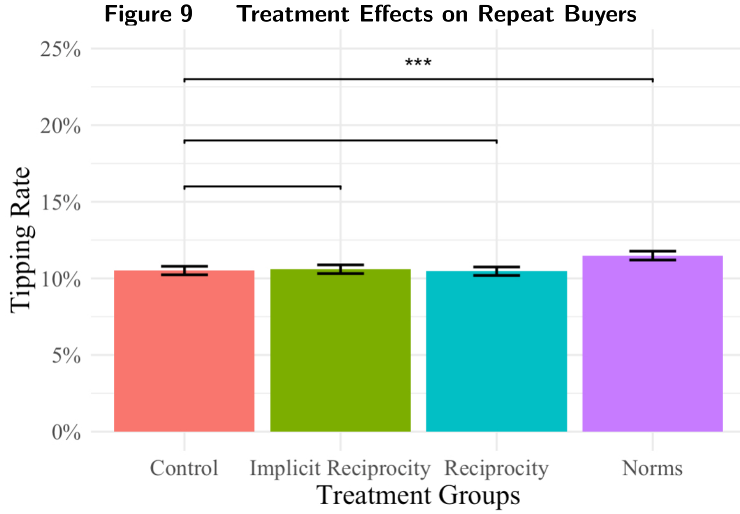


Table 6 Repeat Buyer Tipping Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity:Post	−0.001	0.003	−0.006	−0.006
	(0.002)	(0.004)	(0.004)	(0.005)
Reciprocity:Post	0.002	0.006	0.002	−0.004
	(0.002)	(0.004)	(0.004)	(0.005)
Norms:Post	0.005**	0.020***	0.005	−0.005
	(0.002)	(0.004)	(0.004)	(0.005)
Post	−0.0004	0.002	0.002	0.0003
	(0.002)	(0.003)	(0.003)	(0.004)
Buyer Satisfaction = 3	0.010**	0.008	0.013**	−0.004
	(0.004)	(0.005)	(0.006)	(0.007)
= 4	0.042***	0.040***	0.040***	0.034***
	(0.002)	(0.003)	(0.003)	(0.003)
= 5	0.095***	0.093***	0.092***	0.084***
	(0.001)	(0.002)	(0.002)	(0.002)
Price	0.00000	0.00001	0.00001	0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Price^2	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Seller Tip Mention	0.053***	0.055***	0.051***	0.049***
	(0.004)	(0.004)	(0.004)	(0.004)
Category FE	Y	Y	Y	Y
Buyer FE	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y
Observations	760,135	638,196	581,651	464,295
Number of Buyers	40,823	40,823	34,871	22,226
R^2	0.544	0.562	0.562	0.569
Adjusted R^2	0.453	0.462	0.461	0.469

Note: **p<0.05; ***p<0.01

Table 7 uses the same model specification to estimate the causal effects of treatment on whether we observe repeat buyers completing a second or third transaction during the sample period. There were no significant differences across treatment groups in repurchase rates. The effect of the Norms treatment on repurchase was directionally positive but not significantly different from Status Quo, indicating no evidence that the Norms treatment message reduced subsequent demand. Repeat transactions also increased with satisfaction ratings but did not change significantly with order price or seller tip mentions.

Overall, we have learned that the Norms treatment increases tipping by new buyers and repeat buyers upon first exposure and does not reduce their subsequent platform demand. Next, we exploit post-experiment platform design changes to separately identify causal effects of treatment messages and default tips treatment.

Table 7 Repeat Buyer Repeat Transaction Regression

	Second Order Exists	Third Order Exists
	(1)	(2)
Implicit_Reciprocity:Post	−0.001	−0.010
	(0.005)	(0.007)
Reciprocity:Post	−0.001	−0.008
	(0.005)	(0.007)
Norms:Post	0.005	0.007
	(0.005)	(0.007)
Post	-0.142^{***}	-0.334^{***}
	(0.003)	(0.005)
Buyer Satisfaction = 3	0.011^{***}	−0.0003
	(0.003)	(0.004)
= 4	0.010^{***}	0.005^{***}
	(0.001)	(0.002)
= 5	0.006^{***}	0.003^{***}
	(0.0003)	(0.0004)
Price	0.00000	−0.00000
	(0.00000)	(0.00000)
Price^2	−0.000	0.000
	(0.000)	(0.000)
Seller Tip Mention	−0.001	−0.001
	(0.001)	(0.001)
Category FE	Y	Y
Buyer FE	Y	Y
Seller FE	Y	Y
Observations	638,196	581,651
Number of Buyers	40,823	34,871
R^2	0.455	0.594
Adjusted R^2	0.330	0.500
Note:		$^{**}p<0.05$; $^{***}p<0.01$

7. Separating Message and Scale Effects Using Post-test Data

The main limitation of the experimental design was that the message treatments were confounded with the default tip treatments because the default tip manipulation coincided with the Reciprocity and Norm message treatments. Here we investigate post-test data to separately identify a main effect of the Norms message treatment and indicate what happened after the platform adopted the Norms message treatment for all buyers.

7.1. First Platform Design Intervention after Test

Immediately after the field experiment concluded, the platform reverted to showing Status Quo messages to all users with the new default tips. Table 8 shows the tip message and the default tips seen by each of the four treatment groups during the test period and the 3 weeks immediately following the test period. Two discontinuities at the end of the test period contribute to separating the default tips effect from the treatment message effects. First, the control group had no variation in tip message between the test and post-test periods, but it was exposed to different default tips. Second, the Reciprocity and Norm treatment groups experienced a change in tip message but no change in default tips. We can use the same identification strategy as before for the new and repeat buyers who transacted during the test period but not for new buyers who arrived during the post-test period for whom we have no control group.

Table 8 Post-experiment Platform Design Change

	Control	Implicit Reciprocity	Reciprocity	Norms
Test Period	Status Quo Message & Two Defaults + Custom Tip	Implicit Reciprocity Message & Two Defaults + Custom Tip	Reciprocity Message & Three Defaults + Custom Tip	Norms Message & Three Defaults + Custom Tip
Weeks 1-3 After Test Period I	Status Quo Message & Three Defaults + Custom Tip	Status Quo Message & Three Defaults + Custom Tip	Status Quo Message & Three Defaults + Custom Tip	Status Quo Message & Three Defaults + Custom Tip

We combine the test data and post-test data, then specify Equation (3) to separate the default tip effect from the Norm and Reciprocity message effects.[6]

$$y_i = \beta_1 \text{Post\_Test}_i + \beta_2^{M_i}(\text{Treat}_{M_i} \cdot \text{Post\_Test}_i) + \alpha_1 X_i^T + \mu_{b_i} + \lambda_{s_i} + \epsilon_i \tag{3}$$

Post_Test_i is a binary indicator for weeks 1-3 immediately after the test period. The control group had no change in tip message and is excluded from the interaction for identification, so β_1 is the causal effect of the default tips on tipping. The Norm and Reciprocity treatment groups had no change in default tips. However, they did revert to status quo messages, so β_2^{Norm} and $\beta_2^{\text{Reciprocity}}$ indicate causal effects of Norm and Reciprocity messages on tipping, respectively. The Implicit

[6] We also report Figure B1 and Figure B2 that summarise the mean difference in Tipping Rate between the test and post-test periods across treatment messages in the appendix.

Reciprocity group experienced both message and default tip changes, so $\beta_2^{\text{Implicit Reciprocity}}$ is the joint effect of those two changes on tipping. Equation (3) also controls for buyer fixed effects, seller fixed effects, and transaction fixed effects as in previous analyses.

Table 9 shows that removing the Norm treatment message without changing the default tips reduced new buyers' tipping during the post-test period by 3.9%. The results also show that removing the Reciprocity treatment message without changing the default tips reduced new buyers' tipping during the post-test period by 5.1%.Both causal effects were highly significant. Therefore, we have clear evidence that treatment messages can affect tipping independent of default tips. However, Reciprocity and Norm message removal only decreased repeat buyers' tipping slightly, with non-significant point estimates of 0.4% and 0.3%, respectively. The muted effects are consistent with earlier analysis showing that platform design changes affect new buyers' behavior more than repeat buyers.

The results also show that changing the default tips for Control-group buyers, without changing the tip request messages, increased new buyers' tipping rate from 12.4% to 13.1%. The effect is statistically significant, but the result suggests that the default tips' effect is smaller than the effect of the Norm treatment message. Once again, repeat buyers showed a more muted response to default tips, with a non-significant decrease of 0.2% in Tipping Rate.[7]

7.2. Platform Adoption of the Norms Message

The platform adopted the Norm message as the default for all users about 3 weeks after the test period concluded. The default tips interface did not change at that time.

Figure 10 shows weekly tipping rates for new buyers who first transacted after the test period and before the Norms message adoption. It also shows weekly tipping rates for new buyers who first transacted after Norms message adoption.

We find two interesting results. First, new buyers who arrived three weeks after the test period and hence received the status quo message tipped less than 10% of their transactions on average for 3 weeks after the test period. However, after the platform adopted the Norm message for all buyers, they tipped in about 10.8% of their transactions, a statistically meaningful increase. Second, new buyers who first transacted after the platform adopted the Norm message for all buyers tipped at meaningfully higher rates than other recent cohorts of new buyers (11.9% vs. 10.8%).

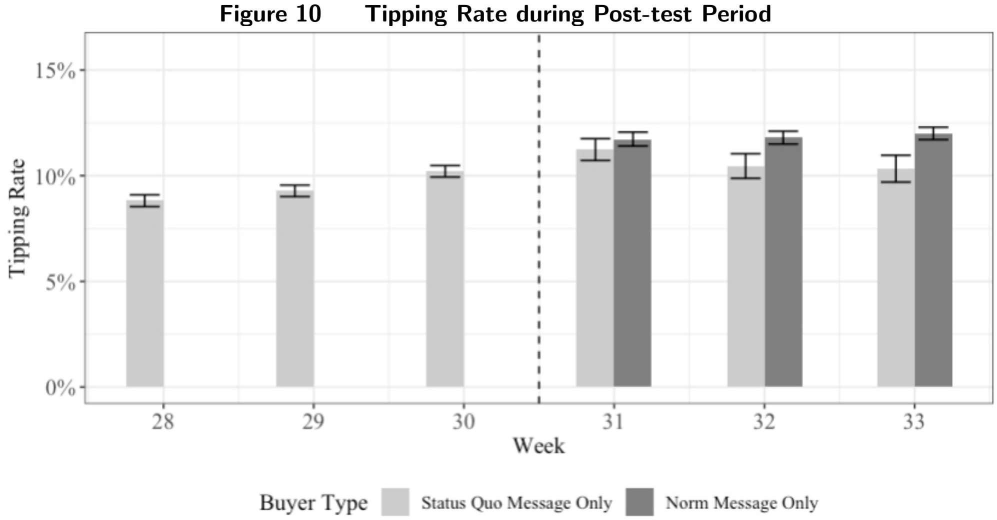
The platform Norms adoption timing does not coincide with an exogenous control group, but both data descriptives are directionally consistent with a positive causal effect of the Norms message on new buyer tipping rate. Next, we consider what more experimental data can teach us about when the treatment effect occurs and how it works.

[7] We tried to test whether the sum of the Norms treatment effect and the default tips effect is consistent with their joint causal effect reported in sections 5 and 6. The estimates are not very different from each other, but the comparison is not statistically powerful due to greater estimation error in the post-test analysis.

Table 9 Post-test Regression

	New Buyers (1)	Repeat Buyers (2)
Implicit_Reciprocity:Post_Test	−0.012	−0.001
	(0.016)	(0.003)
Reciprocity:Post_Test	-0.051^{***}	−0.004
	(0.017)	(0.003)
Norms:Post_Test	-0.039^{**}	−0.003
	(0.017)	(0.003)
Post_Test	0.024^{**}	−0.002
	(0.012)	(0.002)
Buyer Satisfaction = 3	−0.010	0.010
	(0.032)	(0.008)
= 4	0.077^{***}	0.039^{***}
	(0.018)	(0.005)
= 5	0.119^{***}	0.090^{***}
	(0.009)	(0.002)
Price	0.0001	−0.00001
	(0.0001)	(0.00001)
Price^2	−0.00000	0.000
	(0.00000)	(0.000)
Seller Tip Mention	0.087	0.044^{***}
	(0.061)	(0.006)
Category FE	Y	Y
Buyer FE	Y	Y
Seller FE	Y	Y
Observations	19,676	221,352
Number of Buyers	2,946	26,850
R^2	0.791	0.667
Adjusted R^2	0.447	0.501
Note:		$^{**}p<0.05$; $^{***}p<0.01$

Figure 10 Tipping Rate during Post-test Period



8. Field Experiment Mechanism Evidence

This section examines what we can learn from the data about why the Norms message significantly increased tipping, but the two Reciprocity messages did not. We interact treatment effects with potential moderators related to tipping motivations. The findings below shed some light on why the Norms manipulation affected tipping behavior.

A few caveats are in order. Interactions often require more statistical power to estimate precisely than main effects, so a fixed sample size may not yield a conclusive analysis. Further, these are not subsequent experiments on the platform, so we are constrained by the observable features of the current data.

We focus on two potential moderators of the treatment message effect on tipping:

1. *Regional tipping culture: North America vs. other regions*: Buyer region tipping cultural Norms affect tipping much more than seller region, and the Norm treatment message increased tipping. Therefore we could expect that North American buyers are more likely to pay tips when the platform gives the Norm message.

2. *5-star Satisfaction rating*: Buyer satisfaction strongly predicts tipping, so we consider how message treatments interacted with 5-star buyer satisfaction ratings.

We define a binary indicator for each potential moderator. In distinct regressions, we extend the new buyers analysis to include interactions between treatment effects and each indicator. We also extend the repeat buyers analysis to include three-way interactions between treatment period, treatment groups, and each indicator (again, in distinct regressions), in addition to all main effects and two-way interactions.

Table B2 shows how treatment effects vary with the two moderators for new buyers. The interactions between North America and message treatments are not statistically significant. However, the Norms main effect remains positive and significant, suggesting that local tipping culture is not a first-order determinant of how buyers respond to treatment messages. The interaction between 5-star rating and the Norms treatment is positive and significant, suggesting that buyer satisfaction does play a role in predicting compliance with the injunctive tipping norm. Additionally, first exposures to Implicit Reciprocity affected satisfied new buyers on tipping, but the pooled effect across transactions was not statistically significant. These effects suggest that new buyer satisfaction predicted tipping response to treatment messages.

Table B3 presents analogous moderator results estimated using repeat buyer data. The three-way interactions between the Norms treatment and the tipping moderators are not significant.

9. Do Tips Affect Seller Behavior?

This section investigates the seller side of the market. First, we investigate the pre-treatment data to estimate associations between tips and subsequent seller behavior. We consider three ways in which a seller might behave after a tip. The first one is direct: the seller can mention tips to the buyer on her subsequent transaction after a tip. The other two are indirect measures of seller effort on her subsequent transaction: buyer tipping and buyer satisfaction rating. We capture 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, lagged seller average rating), transaction characteristics (order price, category fixed effects), 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 when they deliver their subsequent job, 1.3% less likely to receive a tip, and likely to get rated 0.02 more stars, relative to no tip received on the previous transaction. Sellers, who had received tips at least once, tended to request tips directly by mentioning the word "tip(s)" on the final delivery message. These results control for persistent seller effects, but they do not separate tipping drivers from other time-varying unobservables, so we interpret them as descriptive.

Next, we look for evidence of whether experimental treatments had causal effects on seller behavior during the experiment. We define each seller's "indirect treatment" as the treatment group that was randomly assigned to the buyer in the seller's first transaction during the treatment period. Our first question is whether seller indirect treatments predict repeated seller usage of the marketplace. Figure B3 shows no significant differences in repeat platform usage between seller indirect treatment groups. In other words, sellers who first transacted with Norms-treated buyers were not significantly more likely to use the platform again during the test period than sellers who first transacted with other buyer treatment groups.[8]

We also consider causal effects of indirect seller treatments on seller behaviors in subsequent transactions. We restrict attention to sellers with at least two completed transactions during the sample period, then examine how sellers' first indirect treatment exposure predicts outcomes of the seller's second treated transaction. Figure B5 shows that, overall, indirect seller treatments do not reliably predict changes in subsequent tips received, buyer satisfaction rating, seller tip mention, or price. Finally, we restricted attention to sellers who first interacted with new buyers during the treatment period, because we previously found that the Norms treatment effect was larger for new buyers. However, Figure B6 shows no evidence that indirect treatments influenced sellers' subsequent transaction characteristics or tips received.[9]

[8] Figure B4 shows no significant differences if we restrict attention to sellers who first interacted with new buyers.

[9] We also looked at new sellers' interactions with all buyers; and new sellers who first interacted with treated new buyers only. Both groups showed no significant differences across indirect treatment groups.

10. Discussion

We study voluntary, post-transaction tipping in a large marketplace for freelance digital services. First, we document key facts of online tipping behavior, showing that tipping is strongly associated with buyer region and satisfaction rating, and that buyer characteristics explain tipping much better than seller factors. Next, we reported a field experiment that shows that an injunctive Norm message increases tipping on the first treatment exposure, with effects that are 3.5 times larger among new buyers than among repeated buyers. Treatment effects on new buyers increase with buyer satisfaction. The Norms treatment increased buyers' spending without any apparent reduction in platform usage and had no detectable indirect impact on seller behavior. After the test, the removal of the Norms message reduced treated buyer tipping; and then the subsequent platform-wide adoption of the Norms message corresponded with sharply higher tipping rates.

Taken as a whole, the results show that platform messages about economic norms can influence economic behavior. However, the effect sizes are limited; larger for new users than returning users; and difficult to estimate precisely, even with full population data.

10.1. Implications

The most important implication of our research is the finding that platforms can use messages to influence buyers' voluntary economic payments. Digital marketplace designers, economists, and analysts need to carefully predict how their designs affect market behaviors before implementation and test those predictions afterward.

At the same time, there is ample evidence that platforms have limited ability to nudge buyers' payment habits. For example, two treatment messages did not have reliable effects on tipping during the experiment, and effects were smaller and appeared less durable for repeat buyers compared to new buyers. The result in section 7.2 highlights the importance of first impressions and calls for careful original designs.

Further, some of the results here offer salient reminders that digital marketplaces can be noisy environments in which average treatment effects are difficult to detect. Therefore, statistical power limitations may conceal full information about marketplace design consequences. For example, our analyses showed that unobserved buyer and seller characteristics are strong predictors of tipping and that pre-treatment controls can help to pin down treatment effect estimates in a difference-in-differences regression design. Other experimental designs are possible, such as sampling clusters within networks or making random assignments within pairs of observationally similar agents.

10.2. Future Research and Limitations

Platforms enable interactions between distinct groups of heterogeneous agents, so it is important to consider how marketplace design affects indirect network externalities. For example, how does increased buyer tipping affect sellers' future service efforts? Should the platform screen seller delivery messages and warn sellers in situations where tip mentions may have adverse consequences? Should the platform tailor tip requests to order categories or buyer types? Should the platform leave tip requests to seller's discretion? The full set of interactions is incredibly rich and offers nuanced possibilities to increase market efficiency and seller effort.

Numerous tipping interventions remain available. One interesting direction could be to relax the opaque nature of digital tipping. Platforms could test social tipping messages, for instance announcing recent tips (e.g., "20 minutes ago, a buyer tipped \$12 on a \$28 order"), reporting aggregate tipping statistics (e.g., "18% of global buyers tipped yesterday"), or seller-specific tips (e.g., "Seller 123 has earned tips from 3 out of 4 satisfied buyers this year").

Heterogeneous treatments offer another interesting direction. Does the optimal tipping message vary with buyer cohort? Would it be better to maintain a constant tipping message or rotate across a series of related messages? How would buyers react if the platform reminds them of their own tipping behavior (e.g., "your last tip was 12 transactions ago")? Adaptive policies could be designed to efficiently explore, learn and exploit information about what messages maximize both tipping, service quality and repeat patronage. Of course, such algorithms should not neglect endogenous effects on agents' platform satisfaction, usage, quality, and pricing, in addition to more obvious effects on tipping.

Finally, an important limitation of this research is its inability to detect off-platform seller characteristics. Design changes may affect full-time gig workers more than occasional gig workers, and their implications may be relatively greater for low-income workers than high-income workers. An important unexplored direction is how platform design rewards or penalizes agents of various types in addition to overall revenue and competition goals.

10.3. Conclusion

Digital marketplaces offer unprecedented opportunities to study and nudge participants' decisions and resulting market implications. We have learned that an injunctive message about tipping norms increased buyers' voluntary post-transaction payments to sellers, with maximal detectable effects on new buyers at first treatment, with no apparent changes in subsequent platform usage or seller behavior. We hope this research will help point the way toward future platform design improvements, efforts to quantify how those changes affect agent behavior, and ultimately improved rewards for seller effort.

References

Alexander D, Boone C, Lynn M (2021) The Effects of Tip Recommendations on Customer Tipping, Satisfaction, Repatronage, and Spending. *Management Science* 67(1):146–165, ISSN 0025-1909, 1526-5501, URL http://dx.doi.org/10.1287/mnsc.2019.3541.

Anderson M, Mcclain C, Faverio M, Gelles-Watnick R (2021) The State of Gig Work in 2021. URL https://www.pewresearch.org/internet/2021/12/08/the-state-of-gig-work-in-2021/.

Azar OH (2007) The Social Norm of Tipping: A Review. *Journal of Applied Social Psychology* 37(2):380–402, ISSN 0021-9029, 1559-1816, URL http://dx.doi.org/10.1111/j.0021-9029.2007.00165.x.

Azar OH (2011) Business strategy and the social norm of tipping. *Journal of Economic Psychology* 32(3):515–525, ISSN 01674870, URL http://dx.doi.org/10.1016/j.joep.2011.03.018.

Azar OH (2020) The Economics of Tipping. *Journal of Economic Perspectives* 34(2):215–236, ISSN 0895-3309, URL http://dx.doi.org/10.1257/jep.34.2.215.

Blake T, Moshary S, Sweeney K, Tadelis S (2021) Price Salience and Product Choice. *Marketing Science* 40(4):619–636, ISSN 0732-2399, 1526-548X, URL http://dx.doi.org/10.1287/mksc.2020.1261.

Bluvstein Netter S, Raghubir P (2021) Tip to Show Off: Impression Management Motivations Increase Consumers' Generosity. *Journal of the Association for Consumer Research* 6(1):120–129, ISSN 2378-1815, URL http://dx.doi.org/10.1086/710239.

Chandar B, Gneezy U, List JA, Muir I (2019) The Drivers of Social Preferences: Evidence from a Nationwide Tipping Field Experiment URL http://dx.doi.org/10.3386/w26380.

Chen J, Xu AJ, Rodas MA, Liu X (2022) Order Matters: Rating Service Professionals First Reduces Tipping Amount. *Journal of Marketing* 002224292210986, ISSN 0022-2429, 1547-7185, URL http://dx.doi.org/10.1177/00222429221098698.

Cialdini RB (2009) *Influence: science and practice* (Boston: Pearson Education), 5th ed edition, ISBN 9780205609994, oCLC: ocn227205869.

Donkor K (2021) The Economic Value of Norm Conformity and Menu Opt-Out Costs. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.3955553.

Duhaime EP, Woessner ZW (2019) Explaining the decline of tipping norms in the gig economy. *Journal of Managerial Psychology* 34(4):233–245, ISSN 0268-3946, URL http://dx.doi.org/10.1108/JMP-06-2018-0270.

Hanbury M (2019) Instacart workers are fighting back against a policy change they say drastically cut their wages. *Business Insider* URL https://www.businessinsider.com/instacart-workers-request-22-cent-tips-protest-2019-1.

Lu S, Yao D, Chen X, Grewal R (2021) Do Larger Audiences Generate Greater Revenues Under Pay What You Want? Evidence from a Live Streaming Platform. *Marketing Science* 40(5):964–984, ISSN 0732-2399, 1526-548X, URL http://dx.doi.org/10.1287/mksc.2021.1292.

Lynn M (2015) Service gratuities and tipping: A motivational framework. *Journal of Economic Psychology* 46:74–88, ISSN 01674870, URL `http://dx.doi.org/10.1016/j.joep.2014.12.002`.

Lynn M (2016) Motivations for tipping: How they differ across more and less frequently tipped services. *Journal of Behavioral and Experimental Economics* 65:38–48, ISSN 22148043, URL `http://dx.doi.org/10.1016/j.socec.2016.09.001`.

Newman A (2019) DoorDash Changes Tipping Model After Uproar From Customers. *The New York Times* ISSN 0362-4331, URL `https://www.nytimes.com/2019/07/24/nyregion/doordash-tip-policy.html`.

Appendix A: Treatment Manipulation Analysis

To test whether the manipulations agreed upon with the company indeed tapped the appropriate and anticipated theoretical constructs of reciprocity and of implied normative behavior, we designed a pre-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], implicit reciprocity, explicit 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 2009). 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 the norm 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."

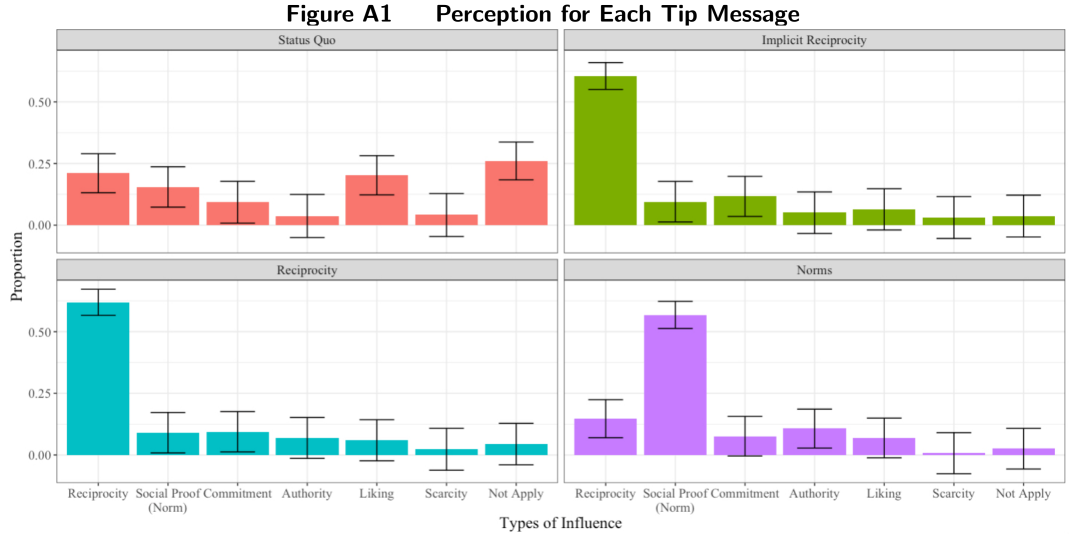
Liking: "This emphasizes that I should tip because someone I really like is asking me to."

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 A1 shows that the control (Status Quo) message motivates multiple interpretations, and no distinct one emerges. 26% of participants marked that none of six reasons to tip apply to Status Quo 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 Implicit Reciprocity and 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.

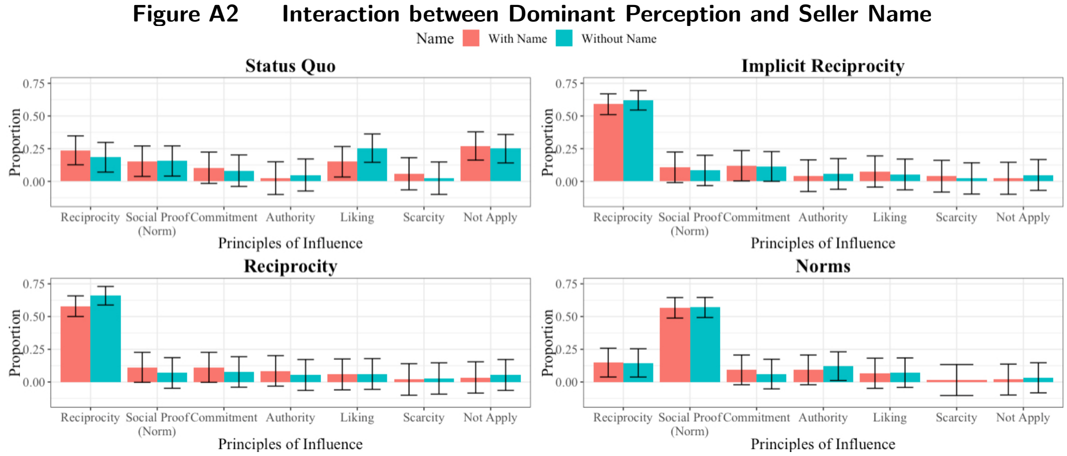
We further examined whether the dominant perception for each tip message varied with the inclusion of a seller's name. Figure A2 shows that the inclusion of the seller's name does not significantly change the prevalent perception of the treatment messages. The Status Quo message motivates multiple interpretations, with "none of the six reasons above apply" marked the most. "Reciprocity" was the dominant interpretation

Figure A1 Perception for Each Tip Message



for both Implicit Reciprocity and 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 A2 Interaction between Dominant Perception and Seller Name



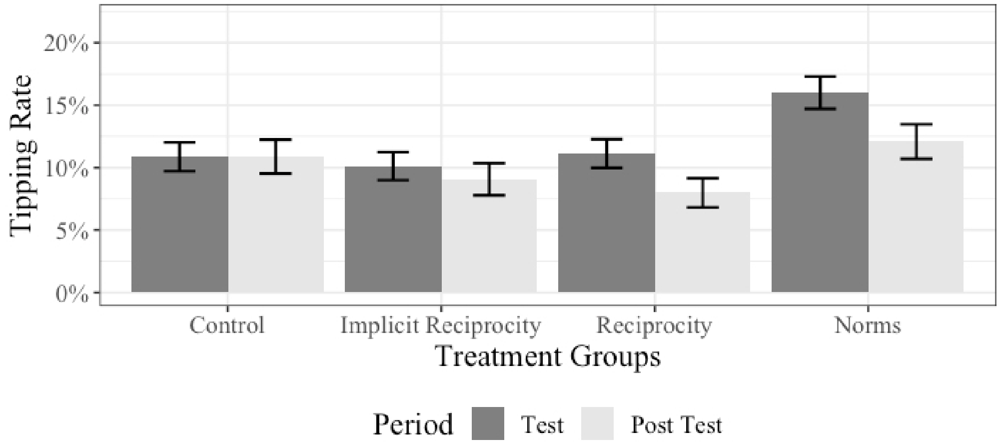
Appendix B: Additional Tables and Figures

Table B1 New Buyers Tipping Logit Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity	0.017	0.087	−0.129	0.113
	(0.062)	(0.092)	(0.123)	(0.194)
Reciprocity	0.053	0.056	−0.027	0.092
	(0.061)	(0.092)	(0.120)	(0.194)
Norms	0.384***	0.527***	0.214	0.126
	(0.059)	(0.088)	(0.116)	(0.193)
Buyer Satisfaction = 3	0.648	12.987	0.674	0.583
	(0.412)	(212.328)	(0.771)	(1.156)
= 4	2.192***	14.364	1.873***	2.139***
	(0.157)	(212.328)	(0.244)	(0.406)
= 5	2.911***	15.059	2.640***	3.065***
	(0.130)	(212.328)	(0.168)	(0.280)
Price	0.005***	0.006***	0.002	0.010***
	(0.001)	(0.001)	(0.001)	(0.003)
Price^2	−0.00001***	−0.00001***	0.00000	−0.00003**
	(0.00000)	(0.00000)	(0.00000)	(0.00001)
Seller Tip Mention	0.362***	0.337**	0.407	0.565
	(0.114)	(0.165)	(0.244)	(0.349)
Buyer Prior Order	−0.309***			
	(0.015)			
Buyer Prior Tips	1.341***		2.384***	2.103***
	(0.037)		(0.088)	(0.100)
Seller Prior Order	−0.00003	−0.00002	−0.00005	−0.0001
	(0.00002)	(0.00003)	(0.00005)	(0.0001)
Seller Average Rating	0.027	0.033	0.029	0.035
	(0.019)	(0.029)	(0.040)	(0.064)
Category FE	Y	Y	Y	Y
Buyer Region FE	Y	Y	Y	Y
Seller Region FE	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545
Number of Buyers	7,880	7,880	7,258	3,545
Log Likelihood	−7,327.836	−3,209.431	−1,915.481	−767.855
Akaike Inf. Crit.	14,731.670	6,484.861	3,902.963	1,605.709

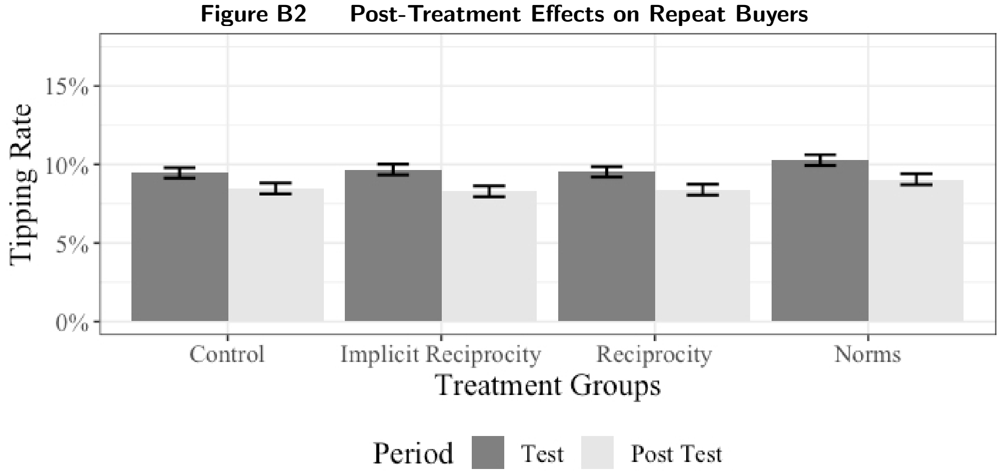
Note: **p<0.05; ***p<0.01

Figure B1 Post-Treatment Effects on New Buyers



Note. This figure shows the mean difference in new buyer Tipping rate between the test and post-test periods across treatment messages. Removing the Norm treatment message without changing the default tips reduced new buyers' tipping during the post-test period from 16% to 12%. Removing the Reciprocity treatment message without changing the default tips reduced new buyers' tipping during the post-test period from 11% down to 7%.

Figure B2 Post-Treatment Effects on Repeat Buyers



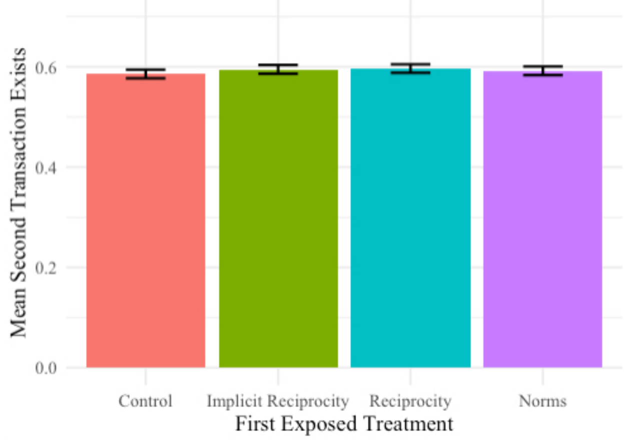
Note. This figure shows the mean difference in repeat buyer Tipping rate between the test and post-test periods across treatment messages. Removing the Norm treatment message without changing the default tips reduced repeat buyers' tipping during the post-test period from 10% to 9%.

Table B2 New Buyer Tipping Mechanism

	Regional Tipping Culture				5-star Satisfaction rating			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Implicit_Reciprocity	0.003	0.009	−0.009	0.001	−0.003	−0.039	−0.013	−0.003
	(0.007)	(0.012)	(0.011)	(0.015)	(0.009)	(0.023)	(0.011)	(0.014)
Reciprocity	0.002	−0.010	−0.005	0.017	−0.006	−0.016	−0.021**	0.001
	(0.007)	(0.012)	(0.010)	(0.015)	(0.008)	(0.025)	(0.010)	(0.014)
Norms	0.024***	0.051***	−0.001	0.009	−0.006	−0.009	−0.008	−0.013
	(0.007)	(0.014)	(0.011)	(0.015)	(0.009)	(0.025)	(0.012)	(0.015)
Implicit_Reciprocity:I_North_America	−0.003	0.004	−0.007	0.005				
	(0.013)	(0.022)	(0.019)	(0.025)				
Reciprocity:I_North_America	0.006	0.031	0.006	−0.021				
	(0.013)	(0.022)	(0.019)	(0.025)				
Norms:I_North_America	0.023	0.039	0.032	−0.002				
	(0.014)	(0.024)	(0.020)	(0.025)				
Implicit_Reciprocity:I_5stars					0.006	0.058**	−0.0002	0.009
					(0.013)	(0.026)	(0.017)	(0.022)
Reciprocity:I_5stars					0.014	0.025	0.028	0.008
					(0.012)	(0.027)	(0.017)	(0.023)
Norms:I_5stars					0.058***	0.091***	0.037**	0.033
					(0.014)	(0.028)	(0.018)	(0.023)
Buyer Satisfaction = 3	0.009	0.041**	0.045**	0.002	0.008	0.040**	0.043**	0.004
	(0.012)	(0.017)	(0.019)	(0.021)	(0.011)	(0.017)	(0.019)	(0.022)
= 4	0.074***	0.110***	0.079***	0.074***	0.073***	0.108***	0.078***	0.072***
	(0.008)	(0.014)	(0.015)	(0.018)	(0.008)	(0.014)	(0.015)	(0.018)
= 5	0.141***	0.185***	0.148***	0.134***	0.121***	0.141***	0.132***	0.120***
	(0.005)	(0.010)	(0.006)	(0.009)	(0.009)	(0.019)	(0.012)	(0.016)
Price	0.0003***	0.0005***	0.0002	0.0001	0.0003***	0.0005***	0.0002**	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Price^2	−0.00000***	−0.00000***	−0.00000	−0.00000**	−0.00000***	−0.00000***	−0.00000	−0.00000**
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Seller Tip Mention	0.038***	0.052	0.034	0.047	0.038***	0.050	0.034	0.047
	(0.014)	(0.027)	(0.026)	(0.030)	(0.014)	(0.027)	(0.026)	(0.030)
Buyer Prior Order	−0.010***				−0.010***			
	(0.001)				(0.001)			
Buyer Prior Tips	0.135***		0.345***	0.286***	0.135***		0.345***	0.285***
	(0.009)		(0.014)	(0.012)	(0.009)		(0.014)	(0.012)
Seller Prior Order	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00001)
Seller Average Rating	0.002	0.004	0.001	0.003	0.002	0.004	0.001	0.003
	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)
Category FE	Y	Y	Y	Y	Y	Y	Y	Y
Buyer Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Seller Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545	25,215	7,880	7,258	3,545
R^2	0.161	0.077	0.242	0.336	0.162	0.077	0.242	0.336
Adjusted R^2	0.160	0.073	0.238	0.329	0.161	0.073	0.238	0.329

Note: **p<0.05; ***p<0.01

Figure B3 Seller Repeat Platform Usage by Indirect Treatment Groups



Note. This figure shows the mean likelihood of second transaction existence by seller indirect treatment groups, that is, the treatment group assigned to each seller's first matched buyer during the test period (N = 50,562).

Table B3 Repeat Buyer Tipping Mechanism

	Regional Tipping Culture				5-star Satisfaction rating			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	-0.005^{***}	-0.015^{***}	-0.001	-0.006	-0.001	-0.007	-0.007	-0.018^{***}
	(0.002)	(0.004)	(0.004)	(0.004)	(0.002)	(0.010)	(0.004)	(0.005)
Implicit_Reciprocity:Post	-0.001	-0.00001	-0.010	-0.002	-0.003	-0.001	-0.007	0.004
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.014)	(0.006)	(0.007)
Reciprocity:Post	0.005	0.005	0.006	0.002	-0.002	0.0005	-0.007	-0.004
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.014)	(0.006)	(0.007)
Norms:Post	0.005	0.017^{***}	0.007	0.002	-0.001	0.002	-0.006	0.001
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.013)	(0.006)	(0.007)
Implicit_Reciprocity:Post:North_America	0.001	0.004	0.006	-0.008				
	(0.004)	(0.008)	(0.008)	(0.010)				
Reciprocity:Post:North_America	-0.005	0.002	-0.007	-0.013				
	(0.005)	(0.008)	(0.008)	(0.010)				
Norms:Post:North_America	-0.0002	0.006	-0.004	-0.012				
	(0.004)	(0.008)	(0.008)	(0.010)				
Implicit_Reciprocity:North_America	-0.008	-0.008	-0.008	-0.006				
	(0.007)	(0.008)	(0.008)	(0.008)				
Reciprocity:North_America	-0.004	-0.001	-0.004	-0.005				
	(0.007)	(0.008)	(0.008)	(0.007)				
Norms:North_America	0.00002	-0.005	-0.003	0.005				
	(0.007)	(0.008)	(0.008)	(0.007)				
Post:North_America	0.009^{***}	0.030^{***}	0.005	0.012				
	(0.003)	(0.006)	(0.006)	(0.007)				
Implicit_Reciprocity:Post:I_5stars					0.003	0.004	0.002	-0.015
					(0.004)	(0.014)	(0.008)	(0.010)
Reciprocity:Post:I_5stars					0.006	0.006	0.013	-0.0004
					(0.004)	(0.015)	(0.008)	(0.010)
Norms:Post:I_5stars					0.008	0.020	0.016	-0.008
					(0.004)	(0.014)	(0.008)	(0.010)
Implicit_Reciprocity:I_5stars					-0.001	0.0002	-0.0002	0.004
					(0.004)	(0.004)	(0.005)	(0.005)
Reciprocity:I_5stars					0.0003	0.001	-0.001	0.002
					(0.004)	(0.004)	(0.005)	(0.005)
Norms:I_5stars					-0.002	-0.002	-0.003	0.0005
					(0.004)	(0.004)	(0.004)	(0.005)
Post:I_5stars					0.001	0.010	0.013^{**}	0.027^{***}
					(0.003)	(0.010)	(0.006)	(0.007)
North_America	0.004	0.004	0.006	0.003				
	(0.004)	(0.005)	(0.005)	(0.004)				
Buyer Satisfaction = 3	0.010^{**}	0.008	0.013^{**}	-0.004	0.010^{**}	0.010	0.014^{**}	-0.003
	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)	(0.006)	(0.007)
= 4	0.042^{***}	0.040^{***}	0.040^{***}	0.034^{***}	0.042^{***}	0.041^{***}	0.040^{***}	0.034^{***}
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
= 5	0.095^{***}	0.093^{***}	0.092^{***}	0.084^{***}	0.095^{***}	0.093^{***}	0.092^{***}	0.081^{***}
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Price	0.00000	0.00001	0.00001	0.00001	0.00000	0.00001	0.00001	0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Price^2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Seller Tip Mention	0.053^{***}	0.055^{***}	0.051^{***}	0.049^{***}	0.053^{***}	0.055^{***}	0.051^{***}	0.049^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Category FE	Y	Y	Y	Y	Y	Y	Y	Y
Buyer FE	Y	Y	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	760,135	638,196	581,651	464,295	760,135	638,196	581,651	464,295
R^2	0.544	0.562	0.562	0.569	0.544	0.562	0.562	0.569
Adjusted R^2	0.453	0.462	0.461	0.469	0.453	0.462	0.461	0.469

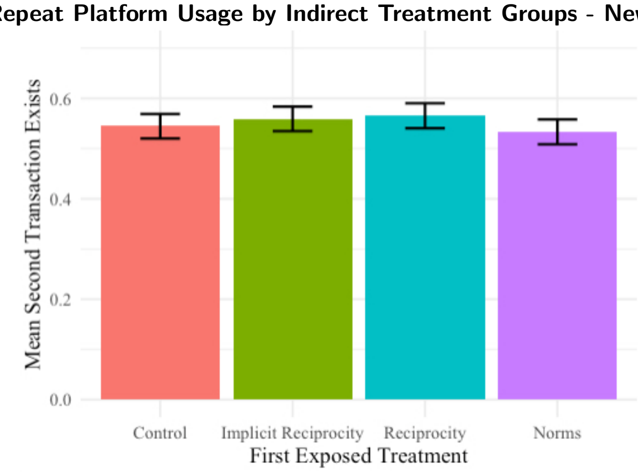
Note: **p<0.05; ***p<0.01

Table B4 How do sellers respond to tips?

	Seller Tip Mention	Tip Received	Buyer Rating
	(1)	(2)	(3)
Lag(Tip Received)	0.001***	−0.013***	0.021***
	(0.0002)	(0.001)	(0.004)
Seller Prior Order	−0.00000***	−0.00000***	−0.0001***
	(0.00000)	(0.00000)	(0.00000)
Lag(Seller Average Rating)	−0.0004***	−0.005***	−0.204***
	(0.0001)	(0.0003)	(0.003)
Price	−0.00000	−0.00001***	−0.001***
	(0.00000)	(0.00000)	(0.00002)
Price^2	−0.000	0.000	0.00000***
	(0.000)	(0.000)	(0.000)
Category FE	Yes	Yes	Yes
Buyer FE	Yes	Yes	Yes
Seller FE	Yes	Yes	Yes
Observations	3,963,840	3,963,840	3,963,840
Number of Sellers	125,040	125,040	125,040
R^2	0.753	0.594	0.600
Adjusted R^2	0.625	0.383	0.393

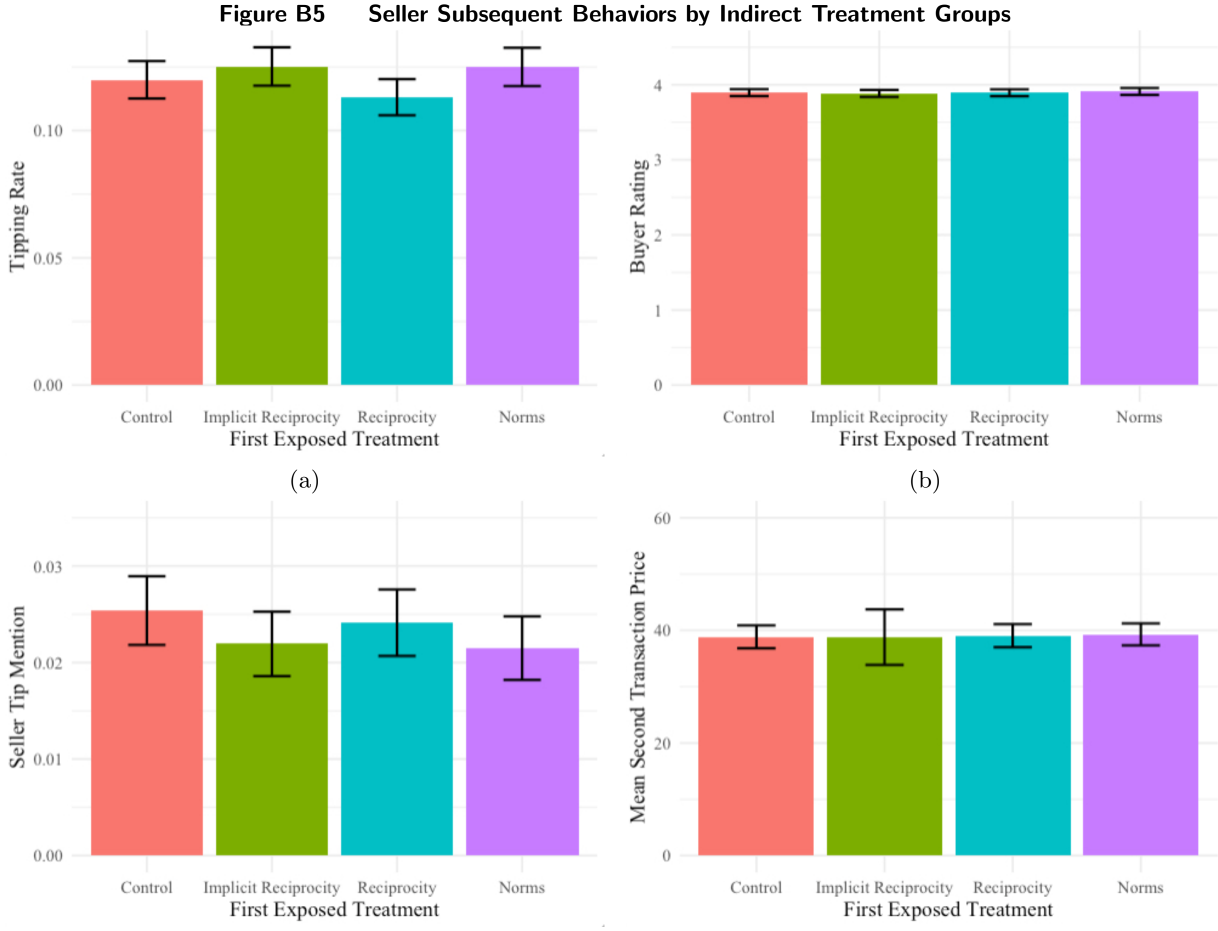
Note: **p<0.05; ***p<0.01

Figure B4 Seller Repeat Platform Usage by Indirect Treatment Groups - New Treated Buyers Only



Note. This figure shows the mean likelihood of second transaction existence by seller indirect treatment groups for sellers who first transacted with a new buyer during the test period (N = 6,268).

Figure B5 Seller Subsequent Behaviors by Indirect Treatment Groups

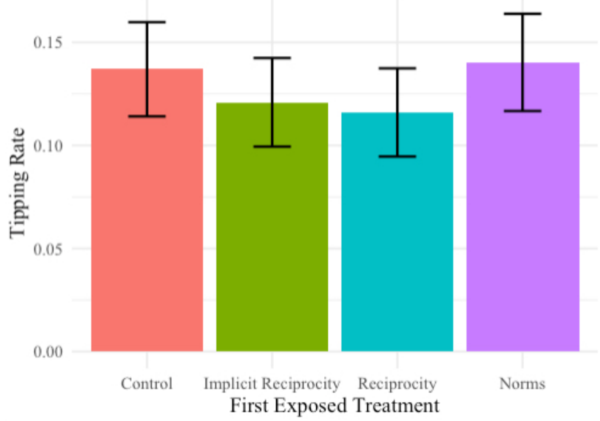


(a) (b)

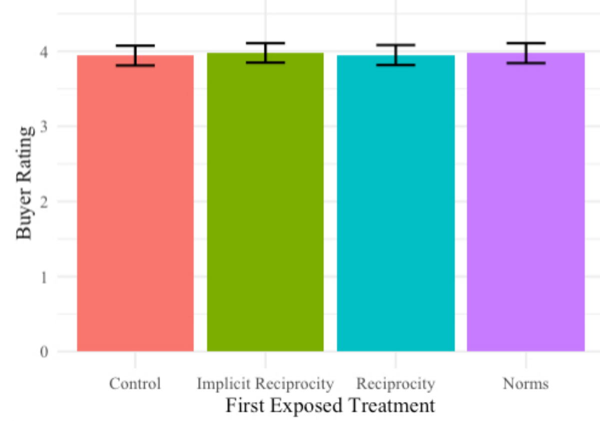
(c) (d)

Note. Figures show average outcomes of (a) tipping, (b) rating, (c) seller tip mention, and (d) price in all sellers' second transactions during the sample period, differentiated by random treatments assigned to the first buyer each seller interacted with (N = 29,944).

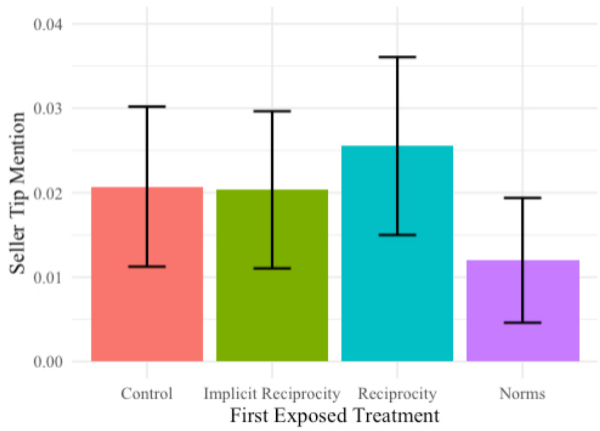
Figure B6 Seller Subsequent Behaviors by Indirect Treatment Groups - New Treated Buyers Only



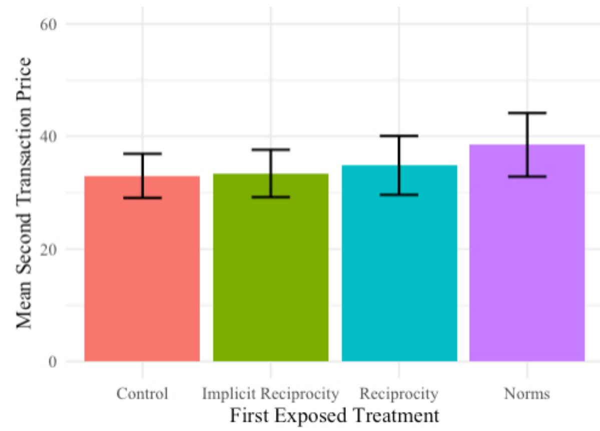
(a)



(b)



(c)



(d)

Note. Figures show average outcomes of (a) tipping, (b) rating, (c) seller tip mention, and (d) price in all sellers' second transactions during the sample period, differentiated by random treatments assigned to the first buyer each seller interacted with; filtered only to sellers whose first interaction was with a new buyer (N = 3,450).