UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Essays on Digital Consumer Behavior: Gen-AI Subscriptions and Tipping Motivations

Permalink

https://escholarship.org/uc/item/6g63h8n6

Author

Kim, Seung Hyun

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Digital Consumer Behavior: Gen-AI Subscriptions and Tipping Motivations

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Management

by

Seung Hyun Kim

Committee in charge:

Professor Kenneth C. Wilbur, Chair
Professor On Amir
Professor Julian McAuley
Professor Kanishka Misra
Professor Robert Evan Sanders
Professor Hyoduk Shin

2024

Copyright

Seung Hyun Kim, 2024

All rights reserved.

The Dissertation of Seung Hyun Kim is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

I dedicate this dissertation to my beloved and supportive parents.

TABLE OF CONTENTS

Dissertation Approval Page iii

Dedication iv

Table of Contents v

List of Figures vii

List of Tables x

Acknowledgements xii

Vita xiv

Abstract of the Dissertation xv

Chapter 1 How did ChatGPT Subscriptions Change Early Adopters' Other Digital Expenditures? 1
1.1 Introduction 1
1.2 Relevant Literature 2
1.3 Data 4
1.3.1 Merchant Sample 5
1.3.2 Consumer Sample 5
1.4 Descriptive Trends 6
1.5 Experimental Setting and Challenges 7
1.6 Methodology 9
1.6.1 Overview 9
1.6.2 Matching 10
1.6.3 The Empirical Model 10
1.7 Main Result 13
1.8 Evaluating the Direction of Adoption 17
1.9 Conclusion and Discussion 18

Acknowledgements 21

Chapter 2 Gratuities in a Digital Services Marketplace 22
2.1 Introduction 22
2.1.1 Conceptual Framework 25
2.2 Setting and Data 30
2.2.1 Empirical Context 30
2.2.2 Tipping Measures 31
2.2.3 Tipping Facts 33
2.2.4 Descriptive Regressions 35

2.3 First Pilot Field Experiment: Design and Data . . . 37
2.3.1 Tip Request Conditions . . . 37
2.3.2 Sample . . . 40
2.3.3 Outcomes . . . 40
2.3.4 Randomization Checks . . . 41
2.4 Experimental Treatment Effects on Tips . . . 41
2.4.1 Effects on New Buyer Tipping . . . 41
2.4.2 Effects on Repeat Buyer Tipping . . . 44
2.4.3 Post-Experiment Changes in Tip Requests and Default Tips . . . 47
2.5 A Second Experiment: Design and Data . . . 52
2.5.1 Tip-Request Conditions . . . 52
2.5.2 Sample and Outcome . . . 53
2.5.3 Randomization Check . . . 53
2.5.4 Treatment Effects on New Buyers . . . 54
2.5.5 Treatment Effects on Repeat Buyers . . . 56
2.6 Laboratory Experiment . . . 59
2.6.1 Message manipulation . . . 60
2.6.2 Method . . . 61
2.6.3 Results . . . 63
2.7 Discussion . . . 65
2.7.1 Implications . . . 66
2.7.2 Limitations and Future Research . . . 68
2.7.3 Conclusion . . . 69

Acknowledgements . . . 70

Appendix A for Chapter 1 . . . 71
A.1 Details of Merchant Selection . . . 71
A.2 Descriptive Statistics . . . 73
A.3 Matching Process . . . 75

Appendix B for Chapter 2 . . . 78
B.1 Additional Data Descriptives . . . 78
B.2 Default Tips Conditions . . . 78
B.3 Field Experiment Mechanism Evidence . . . 87
B.4 Do Tips Affect Seller Behavior? . . . 94
B.5 Treatment Manipulation Analysis . . . 96
B.6 Additional Tables and Figures . . . 99
B.7 Online Experiment Design . . . 113

Bibliography . . . 124

LIST OF FIGURES

Figure 1.1. ChatGPT New Customers and Retention by Adoption-Month Cohort 7

Figure 1.2. Monthly Customer Comparison: OpenAI and Other Digital Services 8

Figure 1.3. Pre-Treatment Trend in Matched Data with Feb Cohort 11

Figure 1.4. Treatment Cohort Group Design 12

Figure 1.5. Subcategory Level Effect (Y_{it}^{j} =Usage) 14

Figure 1.6. Subcategory Level Effect (Y_{it}^{j} =Payment) 15

Figure 1.7. Merchant Level Effect (Y_{it}^{j} =Usage) 16

Figure 1.8. Adoption Timing Comparison: OpenAI vs. Other AI Tools 17

Figure 1.9. OpenAI Subscription Timing vs. Grammarly Renewal Rates 19

Figure 2.1. Example of Tip Window 31

Figure 2.2. Tipping Rate and Tip Percentages 32

Figure 2.3. Tipping by Price Bin 33

Figure 2.4. Tipping and Region 34

Figure 2.5. Tipping and Service Quality 35

Figure 2.6. Tip Request Treatments 38

Figure 2.7. Treatment Effects on New Buyers' First Treated Transactions 42

Figure 2.8. Treatment Effects on Repeat Buyers 44

Figure 2.9. Tipping Rate during Post-test Period 51

Figure 2.10. Treatment Effects on New Buyers' First Treated Transactions 54

Figure 2.11. Treatment Effects on Repeat Buyers' First Treated Transactions 57

Figure 2.12. Treatment Effects on First Transactions of Three Categories of Repeat Buyers 58

Figure A.1. Monthly Customers in Top 10 States 73

Figure A.2. Total Card Spending by Adoption Cohort 74
Figure A.3. Distribution of Months with Positive Spending 75
Figure A.4. Pre-Treatment Trend in Matched Data with March Cohort 76
Figure A.5. Pre-Treatment Trend in Matched Data with April Cohort 76
Figure A.6. Pre-Treatment Trend in Matched Data with May Cohort 77
Figure A.7. Pre-Treatment Trend in Matched Data with June Cohort 77
Figure B.1. Tipping by Buyer Prior Orders 78
Figure B.2. Tipping and Service Category 79
Figure B.3. Tipping by Buyer Prior Tips 79
Figure B.4. Tipping by Buyer Tenure 80
Figure B.5. Tipping by Seller Mention “Tip” or “Tips” 80
Figure B.6. Tipping By Seller Prior Orders 84
Figure B.7. Tipping Behavior Analysis by Price Bin I 85
Figure B.8. Tipping Behavior Analysis by Price Bin II 85
Figure B.9. Default Tip Manipulation 86
Figure B.10. Perception for Each Tip Message 98
Figure B.11. Interaction between Dominant Perception and Seller Name 98
Figure B.12. Post-Treatment Effects on New Buyers 102
Figure B.13. Post-Treatment Effects on Repeat Buyers 103
Figure B.14. Treatment Effects on 42,153 Repeat Buyers’ First Treated Transactions . . 105

Figure B.15. Seller Repeat Platform Usage by Indirect Treatment Groups 108

Figure B.16. Seller Repeat Platform Usage by Indirect Treatment Groups 108

Figure B.17. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups 109

Figure B.18. Seller Subsequent Behaviors during Second Experiment by Indirect Treatment Groups 110

Figure B.19. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups - New Treated Buyers Only 111

Figure B.20. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups - New Treated Buyers Only 112

Figure B.21. Actual Tip Window Example 113

Figure B.22. Example of Fiverr Stimuli 115

Figure B.23. Example of Ride-sharing Stimuli 116

Figure B.24. Example of Food Delivery Stimuli 117

Figure B.25. Frequency Graph (Part 1) 118

Figure B.25. Frequency Graph (continued) 120

Figure B.26. Perception about Control Messages 121

Figure B.27. Perception about Descriptive Messages 122

Figure B.28. Perception about Injunctive Messages 123

LIST OF TABLES

Table 1.1. Summary of Balance for both Not Matched and Matched Data 11

Table 2.1. New Buyers Tipping Regression 43

Table 2.2. New Buyer Repeat Transaction Regression 43

Table 2.3. Repeat Buyers Tipping Regression 46

Table 2.4. Repeat Buyer Repeat Transaction Regression 47

Table 2.5. Post-experiment Platform Design Change 48

Table 2.6. Post-test Regression 50

Table 2.7. New Buyers Tipping and Repeat Transaction Regression 55

Table 2.8. Repeat Buyers Tipping and Repeat Transaction Regression 59

Table 2.9. Structure of Norm Messages 62

Table 2.10. Laboratory Experiment Tipping Regression 64

Table A.1. Brands and market share information 72

Table B.1. Variables and Summary Statistics in Pre-Treatment Transaction Data 81

Table B.2. Descriptive Regression without Single Transaction Buyers and Sellers 82

Table B.3. Descriptive Regression with Single Transaction Buyers and Sellers 83

Table B.4. New Buyer Tipping Mechanism with First Field Experiment data 90

Table B.5. New Buyer Tip Amount Mechanism with First Field Experiment data 91

Table B.6. New Buyer Tip Percentage Mechanism with First Field Experiment data . . 92

Table B.7. New Buyer Tipping Mechanism with Second Field Experiment data 92

Table B.8. Repeat Buyer Tipping Mechanism with First Field Experiment data 93

Table B.9. Repeat Buyer Tipping Mechanism with Second Field Experiment 94

Table B.10. First Pilot Experiment Randomization Check - New Buyers 99

Table B.11. First Pilot Experiment Randomization Check - Repeat Buyers 99

Table B.12. Second Experiment Randomization Check - New Buyers 100
Table B.13. Second Experiment Randomization Check - Repeat Buyers 100
Table B.14. New Buyers Tip Amount Regression 101
Table B.15. New Buyers Tip Percentage Regression 101
Table B.16. Repeat Buyers Tip Amount Regression 102
Table B.17. Repeat Buyers Tip Percentage Regression 103
Table B.18. New Buyers Tip Amount Regression 104
Table B.19. New Buyers Tip Percentage Regression 104
Table B.20. All Repeat Buyers Tipping and Repeat Transaction Regression 105
Table B.21. Frequent Repeat Buyers Tipping and Repeat Transaction Regression 106
Table B.22. Repeat Buyers Tip Amount and Tip Percentage Regression 106
Table B.23. How do sellers respond to tips? 107
Table B.24. Message Manipulation 114
Table B.25. Hypothetical Scenario Example 119

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Kenneth C. Wilbur, who has profoundly impacted my academic journey. He has been an exceptional mentor, fostering in me the values of curiosity, passion, and enthusiasm for quality research. Thanks to Ken's guidance, I have been inspired to conduct research that positively impacts society and educates many. His expertise to rigorous analysis have been invaluable in shaping my growth as a researcher. I am deeply grateful for his support, patience, and encouragement, especially during the challenging times of COVID and the job market. I would not be where I am now without his support and guidance.

I would also like to thank On Amir, who kindly decided to collaborate closely with me on my first project, even though I am not a consumer behavior marketing student. On has been a fantastic coauthor and mentor, I was touched by his detailed guidance despite not being his advisee. His positivity and resilience have shown me that even when research is tough, there is always a bright side.

I am also thankful to Daniel Minh McCarthy, who gave me the opportunity to work on exciting research related to Gen AI. This opportunity allowed me to continue conducting research during a challenging time despite the tough job market. I look forward to continuing our work together.

Kanishka Misra is one of the most insightful professors I know. Whenever I sought advice on my research ideas, he provided excellent feedback, and I learned a great deal from him. His PhD course significantly contributed to my development as a researcher.

I am also thankful to Robert Sanders. Although I did not have many opportunities to discuss research with him, I learned a lot about research and teaching by being a TA for his class. His passion for both was evident, and his feedback on my presentations greatly improved my skills.

I also extend my sincere thanks to my other committee members, Hyoduk Shin and Julian McAuley, for generously giving their valuable time to advise me on my research.

I am grateful to the current and former Rady faculty who supported and encouraged me throughout grad school: Vincent Nijs, Karsten Hansen, Ayelet Gneezy, Uri Gneezy, Wendy Liu, Uma Karmarkar, Rachel Gershon, Ania Jaroszewicz, and Zhe Zhang.

Thank you to the current and former PhD students who have been supportive colleagues and encouraging friends throughout my journey. I want to give a special shout-out to all the marketing Ph.D. students: Michelle Yoosun Kim, Meenakshi Balakrishna, Brianna Chew, PJ Didwania, Ariel Fridman, Katie Hillegass, Kohei Hayashida, Jessica Kim, Giulia Maimone, Carolina Raffaelli, Gal Smitizsky, and Wujing Zhang. It's been amazing to grow together intellectually and also have fun. Thank you to Xiaofeng Liu, Youngjae Choi, Anindo Sarkar, Jinhyeon Han, Weilin Chen, Hannah Bae, and Sookyung Ko for their support throughout the years. Special thanks to Seungjun Kim, Na Jung Lim, Jinhee Huh, and Yoon Woo (Billy) Byun for organizing and joining the Pomodoro sessions during my PhD. Without the Pomodoro group, I could not have finished this journey. Thank you for making this Ph.D. journey enjoyable. It has been a pleasure working with you all.

Last but not least, I want to give an extra special shoutout to my family and friends, who have been my biggest supporters. My PhD journey would not have been possible without your support, and I am very grateful for that.

Both Chapter 1 and Chapter 2 contain unpublished material. Chapter 1 was produced in a joint manner between myself, Daniel Minh McCarthy, and Kenneth C. Wilbur. Chapter 2 was produced in collaboration with On Amir and Kenneth C. Wilbur.

VITA

2013	Bachelor of Arts, Economics, Hankuk University of Foreign Studies, Seoul, South Korea
2016	Master of Arts, Economics, Duke University, Durham
2017–2024	Teaching Assistant, University of California San Diego
2024	Doctor of Philosophy, Management, University of California San Diego

PUBLICATIONS

Joo, Mingyu, **Kim, Seung Hyun**, Ghose, Anindya and Wilbur, Kenneth C., (2023), "Designing Distributed Ledger technologies, like Blockchain, for advertising markets," *International Journal of Research in Marketing*, 40, issue 1, p. 12-21.

FIELDS OF STUDY

Major Field: Quantitative Marketing

ABSTRACT OF THE DISSERTATION

Essays on Digital Consumer Behavior: Gen-AI Subscriptions and Tipping Motivations

by

Seung Hyun Kim

Doctor of Philosophy in Management

University of California San Diego, 2024

Professor Kenneth C. Wilbur, Chair

This dissertation comprises two papers that shed light on how digital platforms are shaping consumer preferences and behavior.

Chapter 1 investigates a large card spending panel to understand how ChatGPT adopter changed their spending on other digital services. We employ Coarsened Exact Matching to balance early ChatGPT subscribers—those who began their subscriptions in distinct monthly cohorts from February to June—with those of later ChatGPT subscribers who joined between September and December. Then we use a triple-difference identification strategy to predict counterfactual digital service payments. We find an increase in consumers' spending on other AI products following the adoption of ChatGPT across five monthly adopter cohorts. The estimates

rule out market share changes of 2% for the majority of brands, with a few exceptions.

Chapter 2 examines online tipping in the gig economy, focusing on how digital platforms can boost tipping. We analyze 4.1 million transactions and find that a buyer's nationality and satisfaction levels are key tipping factors. We also run two field experiments to test if psychologically-based messages about norms and reciprocity could lead to more frequent tipping. The results show that informing buyers about social norms effectively increases tipping. Additionally, in a MTurk survey, we test various normative messages and find that descriptive norms, like highlighting high user tipping rates, effectively raise tipping intentions. This paper highlights the importance of understanding and shaping online tipping practices, not just for helping consumers navigate tipping decisions but also for improving gig workers' financial stability.

Chapter 1

How did ChatGPT Subscriptions Change Early Adopters' Other Digital Expenditures?

1.1 Introduction

Generative artificial intelligence (AI) technologies are becoming a big part of our daily lives, influencing both industries and consumer behavior. The advanced capabilities of gen AI in understanding and generating human-like text, images, and other content are transforming various sectors and reshaping the digital economy. Major tech companies, such as GitHub, Adobe, and Microsoft, are at the forefront of this revolution, upgrading their offerings by integrating AI-powered features to enhance user experiences and streamline workflows.

The rapid adoption of generative AI is evident. OpenAI's ChatGPT, a state-of-the-art language model, gained 100 million consumers within just two months of its launch in November 2022 (Hu, 2023). McClain (2024) shows that one in four U.S. adults reports that they have used ChatGPT at least once. Employed Americans report using ChatGPT for productivity, learning new things, and entertainment. The use of ChatGPT at work has also seen a significant increase, rising from 8% in 2023 to 20% by February 2024. This substantial growth indicates an ongoing trend, suggesting that the adoption of generative AI will continue to expand.

The growing trend highlights the versatility and wide range of applications of generative

AI, with consumers finding these services more effective than some existing paid services. For example, Chegg, an educational company providing study resources, acknowledged the impact of ChatGPT on their new customer growth due to increased student interest. Despite exceeding revenue expectations, Chegg's shares fell by half on the day of the earnings call and have remained flat since then, highlighting the potential of generative AI to disrupt certain industries.

While much discussion has focused on the possibility of generative AI taking human jobs and its impacts on user-created content within online communities, less attention has been given to its potential to replace competing products or services. This paper thus aims to address two main questions: How does the adoption of ChatGPT influence consumers' other digital payments? Which subcategories and merchants experience the most significant changes?

To answer these questions, we use individual consumer credit and debit card panel data from 2022 to 2023, covering one year before and after the release of ChatGPT. We leverage the unique shock of ChatGPT adoption by individual consumers to identify its impact on their purchasing behavior.

Our study adds to the growing literature on how Generative AI technology affects the digital economy, particularly focusing on the demand side dynamics. The findings will inform forecasters, investors, policymakers, and competitors about the broader effects of generative AI on the digital economy, offering valuable insights into the ongoing transformation driven by this new technology.

1.2 Relevant Literature

Our work is related to three streams of prior literature on technology adoption and its impact on consumers and workers, focusing on the implications of ChatGPT4.

First, our study is related to the literature on digital innovations' impact on consumer behavior and well-being. Allcott et al. (2022) explore digital addiction using a randomized experiment. They discover that offering temporary incentives to reduce social media usage leads

to lasting behavioral changes, suggesting that social media usage is habit-forming. Moreover, their findings indicate that individuals' difficulties with self-control account for 31% of social media usage. We contribute to this literature by exploring how the adoption of ChatGPT may change consumers' payment patterns and preferences for various digital services.

Second, our paper is connected to the literature examining the role of generative AI in modifying workplace dynamics, productivity, and firm value. Brynjolfsson et al. (2023) investigate the effect of generative AI tools in the customer service industry. They find that AI assistance adoption significantly increases agent's productivity, particularly among low-skill agents. Additionally, the deployment of AI assistance improves consumer sentiment and employee retention. Similarly, Dell'Acqua et al. (2023) ran a field experiment where consultants are either endowed with ChatGPT, ChatGPT plus training, or no technological aid. Their findings suggest that ChatGPT, particularly when coupled with training, significantly enhances task performance and speed. These indicate that generative AI can be a potent tool for improving work quality and efficiency. Eisfeldt et al. (2023) also find that the release of ChatGPT has a significant positive effect on firms with higher exposure to generative AI, demonstrating that productivity improvements from generative AI not only enhance individual task performance but also translate into substantial financial gains for firms, particularly those with a higher proportion of non-routine cognitive tasks. While prior research focuses on the impact of generative AI on firm-side dynamics, our paper extends this literature by focusing on demand dynamics. We aim to gain insights into how consumers adapt their digital fulfillment spending in response to the adoption of ChatGPT, which is expected to increase their productivity.

Lastly, our research contributes to a growing body of literature studying the role of generative AI in transforming business processes. Li et al. (2024) show that large language models (LLMs) can effectively substitute for human participants in market research exercises, particularly in drawing perceptual maps. Their study demonstrates that LLMs like ChatGPT can generate perceptual analysis outputs closely matching those from human surveys, suggesting that automated market research using LLMs can significantly enhance efficiency and reduce

costs. del Rio-Chanona et al. (2023) and Burtch et al. (2023) find that ChatGPT's release leads to a decline in user-generated content on platforms like Stack Overflow, which may reduce future data availability for training new language models. Additionally, Hui et al. (2023) highlights that adopting generative AI reduces the demand for certain types of freelancers, such as content creation and basic programming, which are increasingly automated by generative AI tools. This shift underscores the broader economic impact of generative AI on the labor market, particularly for gig economy workers. We contribute to this literature by estimating the extent to which the adoption of ChatGPT influences consumer digital purchases, specifically whether they opt for products and services that complement ChatGPT or those that substitute it.

1.3 Data

We use anonymous credit and debit card transactions from an analytics firm with access to individual-level transaction data for approximately 2.3 million individuals, who would like to remain anonymous. This data provides individual-level purchase information that would be expected on standard consumer card statements, including account identifiers, merchant names, categories and subcategories defined by the company, transaction dates, transaction amounts, and limited location details specifying the city and state where the transaction occurred.

Our analysis draws from two distinct consumer panels in the United States, each with different time spans and geographic distributions. The first, the 'shorter panel,' comprises approximately 8 million U.S adults and provides transaction data up to September 2023. The second, termed the 'longer panel,' includes around 2 million U.S. adults with transaction information available until December 2023. These panels cover different periods and feature diverse geographic samples, enabling us to analyze a broad spectrum of consumer transactions. For the main analysis, we combined these two panels into one, referred to as the 'combined panel.'

Additionally, the data includes payments made to about 5,000 consumer-facing brands.

Brands are treated as distinct entities in our analysis, distinguishing between multiple brands under a single parent company. For instance, Amazon, Zappos, Twitch, and Shopbop are treated as separate merchants despite being subsidiaries of Amazon.

1.3.1 Merchant Sample

We focus on 'digital fulfillment' merchants that operate entirely online. Our selection is guided by two primary criteria. First, we focus on merchants that consistently accept digital payments, which ensures consistent product quality and service availability that does not vary geographically. Conversely, we exclude merchants providing hybrid service models like DoorDash, which show significant geographical variation in service delivery and consumer value, to maintain consistency in service quality across our research. Second, we consider digital fulfillment merchants because the dataset presents challenges. For example, sales made through intermediaries like Apple Pay or Google Play lack detailed disclosure about purchased products or services. This issue underlines the importance of focusing on merchants who manage their own payment processes, as it allows for clearer insights into consumer spending behaviors and patterns. Online Appendix A.1 put further details of our merchant selection.

1.3.2 Consumer Sample

We focus our analysis exclusively on consumers who made payments for ChatGPT from February to December 2023 to assess the impact of ChatGPT usage on other digital spending. The data include over 70,000 OpenAI subscribers and logs approximately 354,000 transactions, summing up to around $7 million. A significant majority, 98.4% of these transactions, are regular subscription payments with an average amount of about $20 each. The data contains insufficient non-standard API payments, limiting the scope for further analysis in this area.

1.4 Descriptive Trends

Figure 1.1 examines the subscription patterns of ChatGPT using data from both shorter (top) and longer (bottom) panels. The left figures display the number of new ChatGPT subscriptions each month in 2023, with each line representing a cohort of consumers who started to subscribe in a given month. The right figures show the retention rates of these monthly cohorts over time, indicating the percentage of subscribers from each cohort who continue to maintain their ChatGPT subscriptions in the following months.

Regardless of the panel type, the trends for new subscriptions and retention rates are similar. In the left figures, we observe a notable spike in new subscriptions in March 2023, coinciding with significant promotional activities for ChatGPT. Following this initial surge, the number of new subscriptions experiences a moderate decline until June, after which steady growth is observed in the subsequent months.

The right figures illustrate that the trend of retention rates is consistent across different monthly cohorts. Each monthly cohort experiences a steep decline, dropping to around 75% by the second month and continuing to decrease in the following months. This pattern suggests that a significant portion of consumers who initially subscribe to ChatGPT do not maintain their subscriptions over an extended period. However, despite the declining retention rates, the number of new subscribers continues to grow, indicating ongoing interest in ChatGPT. Further details on the characteristics of the ChatGPT adoption cohorts are detailed in Online Appendix **??**.

To further investigate the impact of ChatGPT adoption, we present model-free evidence by comparing the monthly subscription trends of ChatGPT in 2023 to those of several competitors and examining the number of multi-homing customers (i.e., subscribing to multiple services simultaneously). Figure 1.2 shows that the number of customers using both ChatGPT and another service is relatively small, suggesting that most customers prefer to subscribe to a single service exclusively.

The figures also reveal some notable trends for competitor platform consumers. The

number of Grammarly-only subscribers fluctuates over the months without a clear trend. The education platforms (i.e., Chegg, Coursera, Course Hero) show a moderate decreasing trend in the number of customers who exclusively use this service, indicating that fewer people are subscribing to or maintaining their payment over the months. In contrast, ChatGPT-only subscribers consistently increase throughout the year, suggesting a growing adoption of ChatGPT.

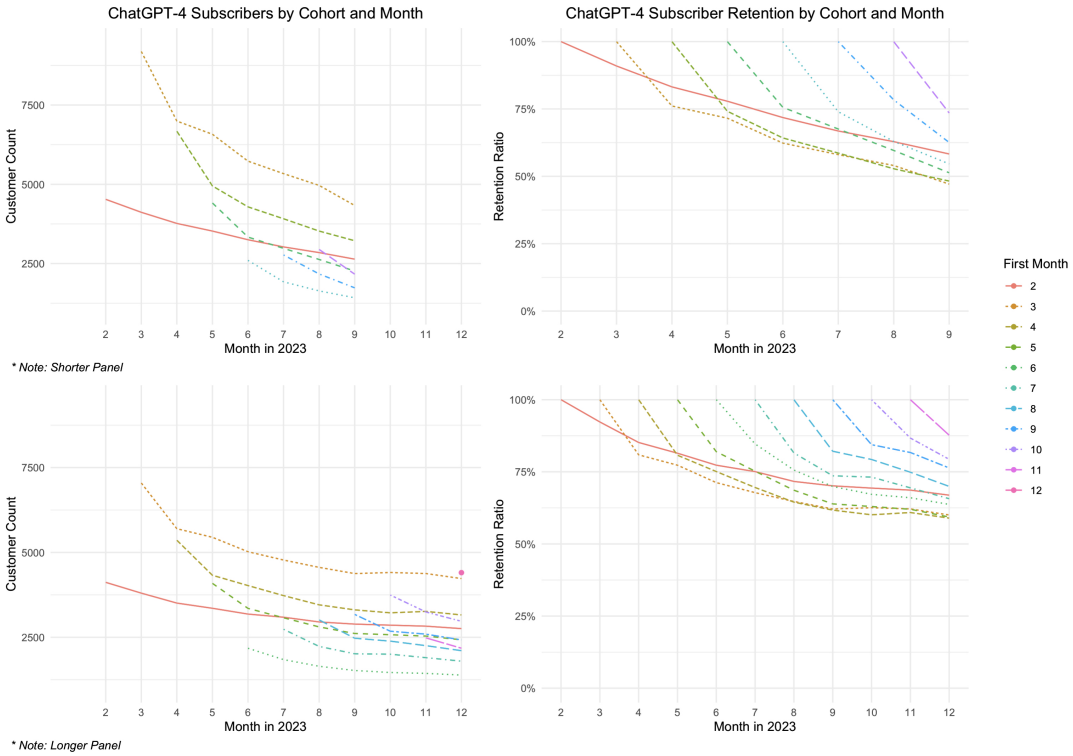


Figure 1.1. ChatGPT New Customers and Retention by Adoption-Month Cohort

1.5 Experimental Setting and Challenges

Descriptive evidence suggests that time-series data alone may not sufficiently reveal the dynamics of customer relationships between different services, and it becomes necessary to design an ideal experiment that can isolate the causal effect of ChatGPT adoption on using other digital services. In the ideal setting, we could randomly provide some consumers with ChatGPT access and collect their comprehensive spending data. This setup allows us to estimate the intent-to-treat (ITT) effects by comparing these consumers with others who did not use ChatGPT.

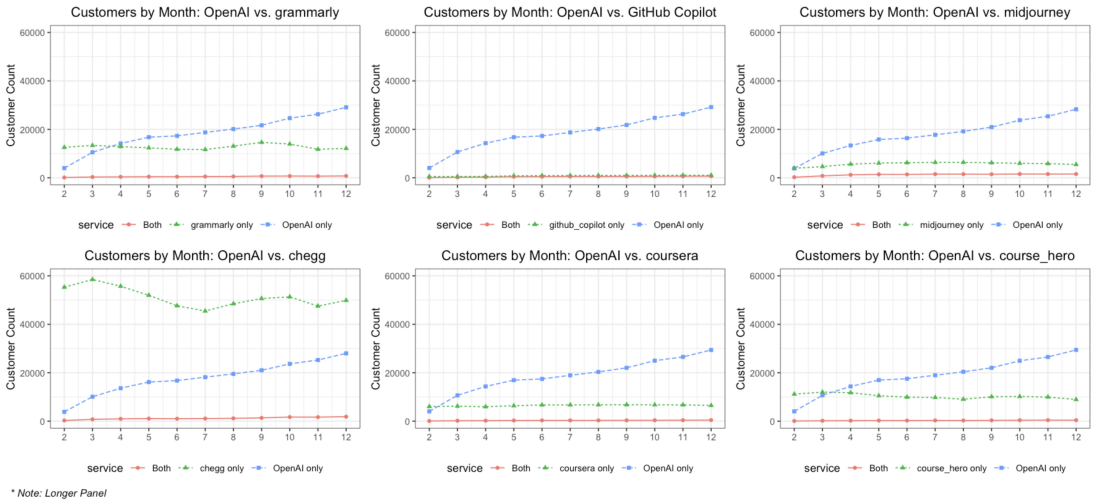


Figure 1.2. Monthly Customer Comparison: OpenAI and Other Digital Services

However, conducting such an experiment is nearly impossible, and providing someone with ChatGPT access does not guarantee they would fully adopt and utilize the technology. Therefore, we aim to use existing data to measure the average treatment effect on the treated (ATT).

For instance, ESPN conducted a six-month study to observe changes in television usage after providing 100 households with free DVRs. This study occurred during a period of concern that DVRs could disrupt viewer attention to TV advertisements. However, the study faced significant challenges: over time, 90% of households dropped out and were replaced by others, and by the end, only 68 households remained in the study. Additionally, many households never set up the DVR, rendering any potential change in their viewing habits unobservable (Kerschbaumer, 2005).

This example highlights the limitations of traditional field experimental approaches in estimating the ITT, especially when new technology is involved. Due to high potential non-adoption rates similar to those in the DVR study, a quasi-experimental framework might be more suitable. This approach would allow us to adjust for factors like technology setup and engagement, providing a more reliable estimate of how changes in technology adoption influence consumer actions across various brands.

1.6 Methodology

1.6.1 Overview

Given the limitations of traditional experimental approaches, we employ a quasi experimental framework to isolate the causal effects of ChatGPT adoption by leveraging the timing of consumers' initial adoption. The timing of consumer's first ChatGPT adoption varies widely throughout 2023, which can be explained by consumers' diverse responses to new technologies based on their specific needs. To account for these varied responses, we divide customers into five groups based on their month of first adoption, ranging from February to June, and considered them as the treatment group. We treat September to December adopters as a control group. This group serves as a useful baseline for comparison, reflecting the behavior of these consumers who adopted ChatGPT later in the year, which may represent more typical consumers compared to early adopters.

However, one potential concern in our setting is the non-randomized nature of ChatGPT adoption, which may lead to self-selection bias since the decision to adopt is inherently endogenous. We address this concern by using Coarsen Exact Matching (CEM) (Blackwell et al., 2009). CEM balances the treatment and control groups based on observed characteristics, making them more comparable and reducing the impact of self-selection bias, thus enhancing the robustness of our model estimates.

Under the ignorability assumption, we assume that the matching variables used in CEM capture all significant differences between the treatment and control groups. If this assumption holds, the observed differences in outcomes between the matched groups can be interpreted as causal effects, allowing us to attribute changes in spending behavior to ChatGPT adoption with more confidence. However, we acknowledge that unobserved confounders affecting both ChatGPT adoption and outcomes of interest could potentially bias the estimated effects, even after applying CEM. Also, we remain cautious about the potential influence of unobserved variables and interpret our results with this limitation in mind.

After matching, the analysis focuses on the local average treatment effect on the treated within the matched consumer groups. We use a Triple Differences (DDD) approach to estimate the causal effect of ChatGPT adoption on other digital payments. First, we use a Difference-in-Differences (DID) to estimate counterfactual predictions of what ChatGPT adopters would have spent if they had not adopted the technology. This helps identify the impact of ChatGPT adoption on purchasing behavior. Then, to account for the influence of seasonal variations in consumer purchasing behavior, we subtract the analogous DID estimate from the previous year Bandiera et al. (2005). This step ensures that our results are not confounded by seasonal trends specific to each brand.

1.6.2 Matching

The matching process is based on whether consumers had any spending in each of the 16 subcategories during 2022. Figure A.3 details of matching variables. Table 1.1 presents both unmatched and matched data, comparing the treated and control groups regarding the matching variables. The matching process has successfully balanced the spending across different subcategories, reducing mean differences and resulting in a more uniform distribution of spending behaviors between the treated and control groups. Figure 1.3 also shows that the pre-treatment trends for the average market share between the matched treated and control groups are similar, which is crucial for validating the parallel trends assumption required for Difference-in-Differences (DID) analysis [1].

The CEM matching effectively produces matched samples of treated groups and control groups with comparable purchasing patterns. This balance gives credibility to the causal inference drawn from our subsequent analysis.

1.6.3 The Empirical Model

We conduct the following empirical analysis to quantify the impact of ChatGPT adoption:

[1] In appendix, Figure A.4 - A.7 display average market share trends between the rest of treated cohort and their corresponding matched controls

Table 1.1. Summary of Balance for both Not Matched and Matched Data

	Mean Difference (Treated-Control)									
	Unmatched Data					Matched Data				
	February	**March**	**April**	**May**	**June**	**February**	**March**	**April**	**May**	**June**
Administrative Tools	0.14	0.03	0.01	0.02	0.07	-0	0	0	0	0
AI Tools	0.39	0.14	0.03	0.08	0.03	0	0	0	0	0
Broker-Dealers	0.31	0.19	0.03	0.05	0.02	0	0	0	0	0
Business & Software	0.89	0.26	-0.07	-0.06	-0.31	0	0	0	0	0
Community Platforms	0.76	0.37	0.21	0.18	0.08	-0	0	0	0	0
E-Books	0.38	-0.1	-0.24	-0.11	-0.24	0	0	0	0	0
Education Resources	0.08	0.02	0.0	-0.08	-0.28	-0	0	0	0	-0
Gaming	0.41	0.3	0.17	-0.06	-0.4	0	0	-0	0	0
Internet Security	0.1	-0.04	-0.01	0.01	-0.04	0	0	0	0	0
Music Streaming & Audio	-0.27	-0.25	-0.35	-0.23	-0.38	-0	0	0	0	-0
News & Print Media	0.19	0.10	0.02	0.1	-0.0	0	0	0	0	0
Online Gambling	-0.01	0.03	0.03	0.02	0.02	0	0	0	0	0
Online Travel Agency	0.0	-0.04	-0.04	-0.01	-0.02	0	0	0	0	0
Others	0.03	0.02	0.01	0.01	-0.0	0	0	0	0	0
Social Media	0.58	0.37	0.16	0.17	0.04	0	0	0	0	0
Video Streaming	0.05	-0.55	-0.54	-0.36	-0.42	-0	0	0	0	0
Panel	-0.3	-0.35	-0.34	-0.3	-0.33	0	0	0	0	0
Sample Sizes (Control)	15,600	15,600	15,600	15,600	15,600	15,600	15,600	15,600	15,600	15,600
Matched (Control)	15,600	15,600	15,600	15,600	15,600	11,509	12,571	12,165	11,497	10,354
Unmatched (Control)	0	0	0	0	0	4,091	3,029	3,435	4,103	5,246
Sample Sizes (Treated)	7,527	13,985	10,417	7,357	4,068	7,527	13,985	10,417	7,357	4,068
Matched (Treated)	7,527	13,985	10,417	7,357	4,068	5,362	10,500	8,055	5,545	3,209
Unmatched (Treated)	0	0	0	0	0	2,165	3,485	2,362	1,812	859

Notes: Combined Panels

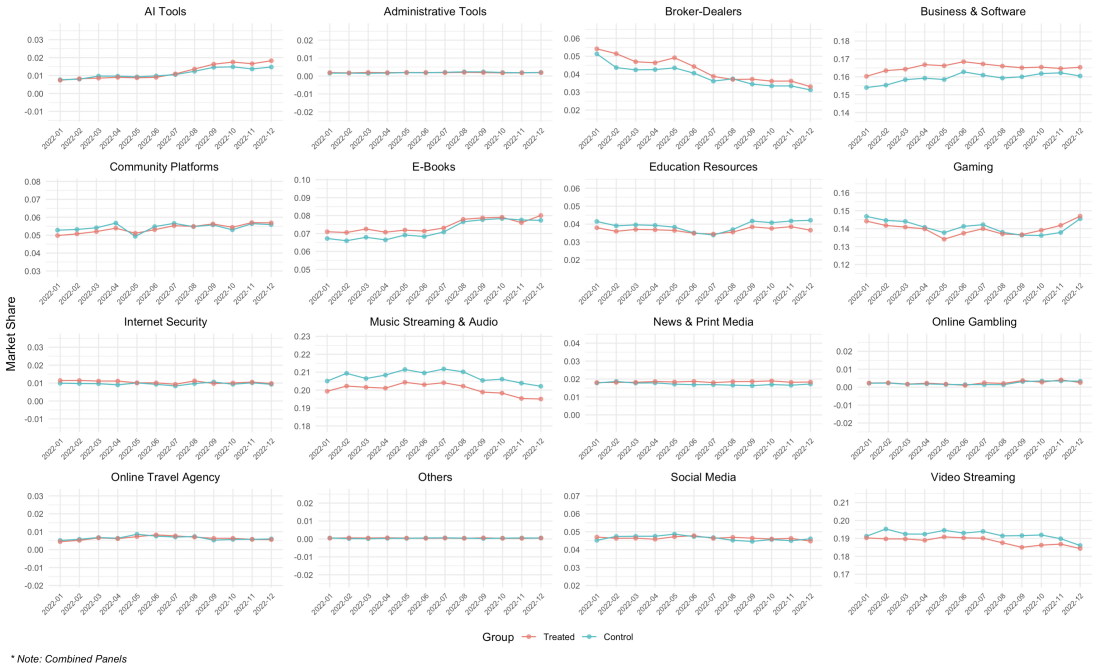


Figure 1.3. Pre-Treatment Trend in Matched Data with Feb Cohort

$$
\begin{aligned}
Y_{it}^{j} = & \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{Period}_t + \beta_3 \text{Year}_t + \\
& \beta_4 (\text{Group}_i \times \text{Period}_t) + \beta_5 (\text{Group}_i \times \text{Year}_t) + \beta_6 (\text{Period}_t \times \text{Year}_t) + \\
& \beta_7 (\text{Group}_i \times \text{Period}_t \times \text{Year}_t) + \varepsilon_{it}
\end{aligned} \tag{1.1}
$$

Where Y_{it}^{j} denotes either a binary variable indicating whether user i made a payment to subcategory/merchant j at time t or the amount paid by user i to subcategory/merchant j at time t. Group_i is a dummy that equals one if consumer i adopts ChatGPT in treated months from February to June and zero if consumer i adopts ChatGPT anytime between September and December. Period_t is a binary variable that is one after the adoption month of each group and zero otherwise. Year_t is a binary variable that equals one if the transaction occurs in the adoption year and zero if it occurs before the adoption year. Lastly, ε_{it} captures any idiosyncratic random errors.

The analysis uses a three-month time window before and after the adoption month for each of the five treatment cohorts, as shown in Figure 1.4. This period provides enough time to observe behavior changes following ChatGPT adoption.

The coefficient β_7 is the primary interest, capturing the change in average digital payment after the ChatGPT adoption compared to the period before adoption while accounting for seasonal variation that happened in the previous year.

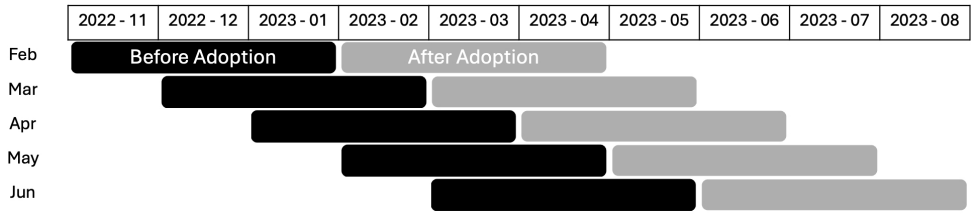


Figure 1.4. Treatment Cohort Group Design

1.7 Main Result

Figure 1.5 shows the results comparing ATT estimates for all five treatment cohorts across 16 subcategories. The most consistent finding is that adopting ChatGPT is associated with a positive and significant increase in the usage of other AI tools. The February, March, and April treatment cohorts increased their usage of other AI tools by around 7% compared with matched control consumers. This effect decreases to 3.6% and 2.7% for May and June cohorts, respectively, but remains significant. These results indicate that subscribing to ChatGPT can positively impact the usage of other AI-related services, suggesting that exposure to ChatGPT may encourage users to explore and adopt other AI tools, even if there might be some overlapping functionality. The experience with ChatGPT seems to stimulate curiosity and interest in AI technology, leading to increased engagement with a broader range of AI-powered services.

Apart from the consistent positive effects on AI tool usage, the results for other subcategories are mixed. For most estimates, we can rule out a market share decrease of 2% or more, as the confidence intervals do not include such a decrease. However, we observe significant negative effects on online gambling in the May cohort and video streaming in the June cohort.

The presence of seven significant effects out of 80 comparisons (5 cohorts × 16 subcategories) raises concerns about false positives due to multiple comparisons. Controlling for the false positive rate is crucial to ensure the reliability of the findings. Focusing on the significant positive effects in the 'Business & Software' while being cautious about potential Type I errors in other non-AI subcategories may be a careful approach, given the risk of false positives.

Figure 1.6 shows the impact of ChatGPT adoption on payment amounts for services across 16 subcategories. The service price or subscription fee for digital fulfillment services generally remains stable throughout the year, leading to low variability in the amounts paid per service by consumers, which complicates the detection of significant effects. Nevertheless, the result shows the significant and positive impact of ChatGPT adoption on payments for services in AI Tools, and the effects are consistent across most cohorts, except for the June cohort. The

February cohort significantly increased spending on AI Tools by $2.55. The March cohort shows a similar increase of $2.24, and the April cohort showed an even more pronounced increase of $3.93. The May cohort also experienced a spending increase of $1.87. Overall, these results suggest that the positive impact of adopting ChatGPT on usage and payments for services in AI Tools subcategory is robust across most cohorts.

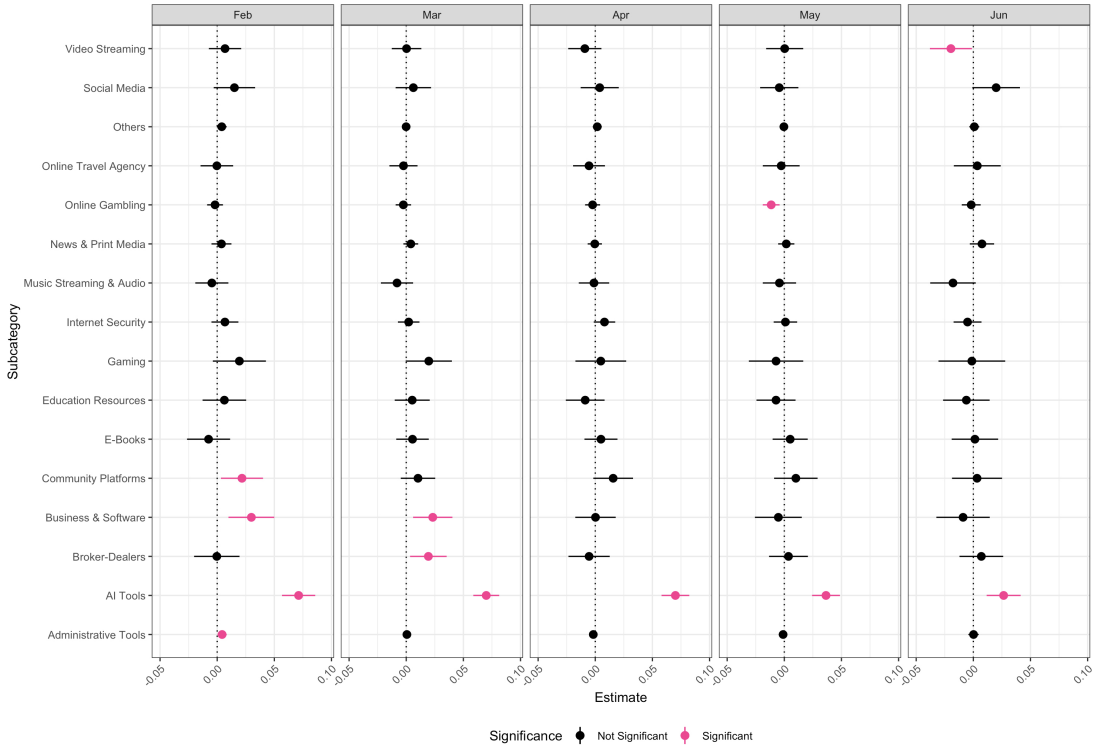


Figure 1.5. Subcategory Level Effect (Y_{it}^{j} =Usage)

Next, we break down the analysis from the subcategory to the merchant level by only considering the top 100 merchants with the highest market share. Figure 1.7 shows that the positive impact of ChatGPT adoption on usage in the AI Tools subcategory is consistent across three major merchants: GitHub Copilot, Grammarly, and Midjourney[2].

On average, after adopting ChatGPT, users significantly increased their usage of GitHub Copilot by 1.73% to 1.66% across the February, March, and April cohorts. Although the increase

[2]GitHub Copilot provides AI-based coding suggestions, Grammarly offers AI-based writing suggestions, and Midjourney provides an AI-based text-to-image service.

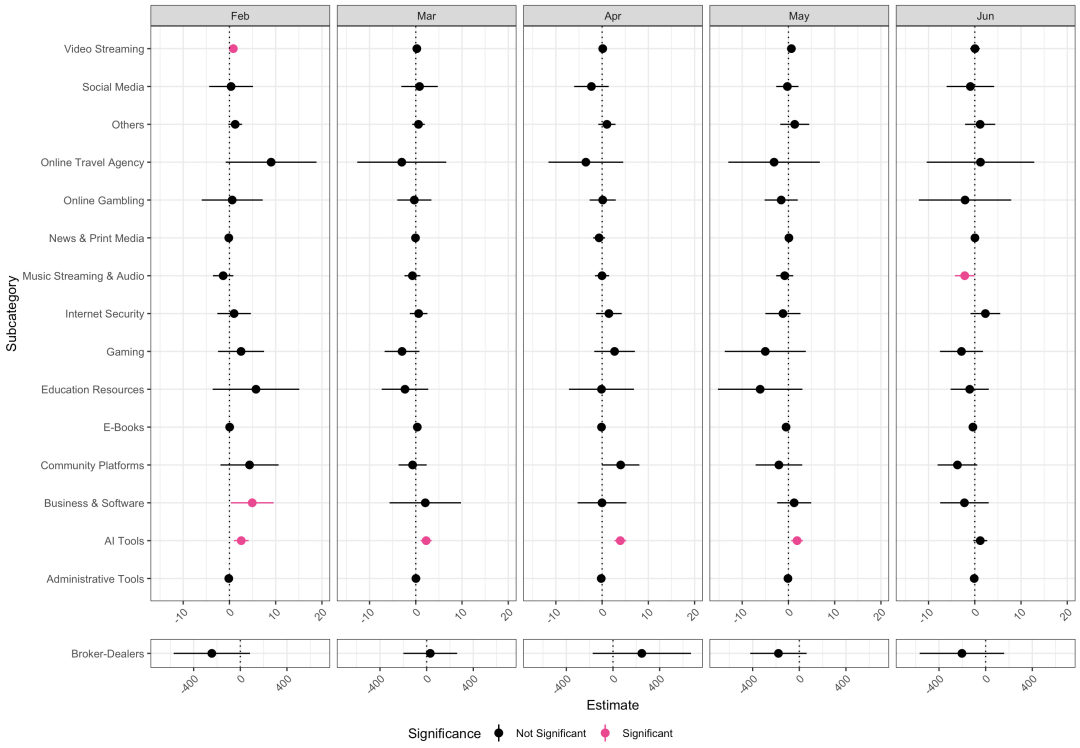


Figure 1.6. Subcategory Level Effect (Y_{it}^{j} =Payment)

slightly diminished for the May and June cohorts, with usage rising by 0.74% and 0.49%, respectively, these effects remain significant.

Similarly, the adoption of ChatGPT led to a substantial rise in Grammarly usage, with users in the February cohort increasing their use by 1.47%. However, results for subsequent cohorts showed mixed outcomes, with significant increases of 0.94% and 1.19% in the April and June cohorts, respectively, but a non-significant decline of 0.19% in the March cohort and a negligible increase of 0.05% in the May cohort.

In the case of Midjourney, since it was initially released on July 12, 2022, there are no estimation results for the February, March, and April cohorts. Additionally, the estimates for the May cohort are based on a Difference-in-Differences estimate rather than a triple-difference estimate. Despite this, Midjourney shows a significant uptick in usage after ChatGPT adoption, with users in the May cohort increasing their use by 2.74% and those in the June cohort raising their usage by 1.27%.

We also observe significant effects for several other merchants in different subcategories. For instance, the June cohort shows a significant increase in Robinhood usage by 2.11% after adopting ChatGPT. In the Business & Software subcategory, several significant positive effects are observed for non-AI-based GitHub services. Similarly, in the Education Resources subcategory, the February cohort exhibits increased usage of Coursera services following the adoption of ChatGPT.



Figure 1.7. Merchant Level Effect (Y_{it}^{j} =Usage)

1.8 Evaluating the Direction of Adoption

The main results suggest that ChatGPT adoption tends to increase the usage of other AI tools. However, one potential concern is the possibility of reverse causation. It is unclear whether ChatGPT drives interest in other AI tools, whether other AI tools increase interest in ChatGPT, or if an unobserved third factor simultaneously increases interest in multiple AI services, including ChatGPT.

We address this concern by examining the direction of adoption within pairs of AI tools, including ChatGPT and the top three merchants that show a significant increase in usage at the merchant level analysis. The analysis focuses on users who made their first payment to one of these AI tools between February 1, 2023, and June 30, 2023. For each pair of AI services, we identify the first AI service each user adopted based on their initial payment date and count the number of users who subscribed to one service before the other. Lastly, we run a chi-square test to check if the number of users who subscribed to each service first is significantly different.

Figure 1.8 shows that for Midjourney users, there is a highly significant trend ($\chi^2 = 39$, $p < 0.001$) of adopting OpenAI before Midjourney. The likelihood of adopting OpenAI first is directionally positive for Grammarly and GitHub Copilot services, but the differences are not significant.

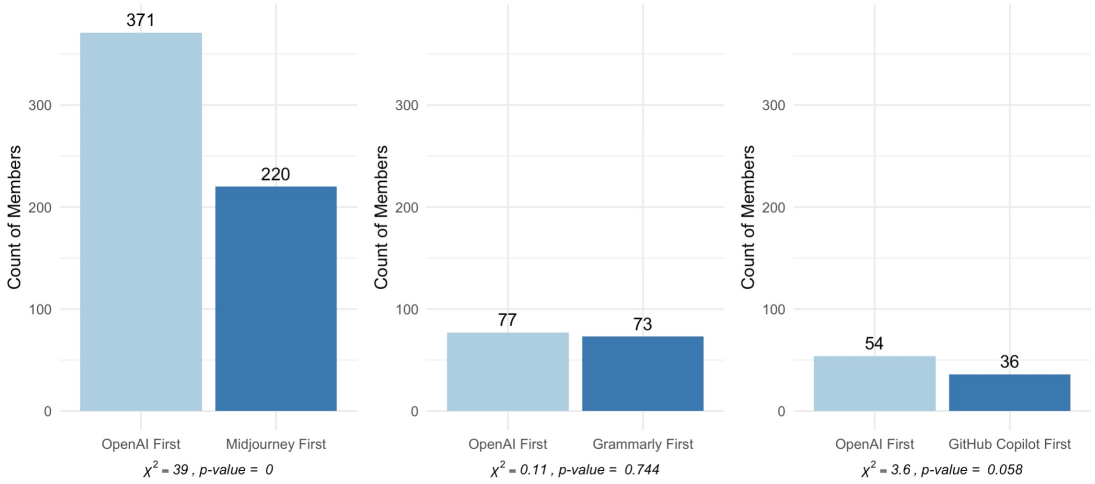


Figure 1.8. Adoption Timing Comparison: OpenAI vs. Other AI Tools

Additionally, we look at how the annual renewal rate for Grammarly co-varied with ChatGPT adoption. Many digital platforms offer both monthly and annual subscription plans, with annual plans often providing a discounted price. In fact, 42% of consumers in our dataset opt for the Grammarly annual plan.

However, because our main analysis uses a three-month time window, the results may not fully capture the behavior change for annual subscribers. To address this, we employ a locally estimated scatterplot smoothing (LOESS)[3] method to visualize the mean renewal rate against the time difference between the Grammarly annual renewal date and the initial ChatGPT subscription date, as shown in Figure 1.9.

The LOESS curve shows that the renewal rate remains relatively steady at around 75% before ChatGPT adoption. However, after ChatGPT adoption, there is a decrease in the renewal rate, dropping to approximately 65% as the time difference increases. This trend suggests that adopting ChatGPT may have a substitution effect on Grammarly's annual subscriptions, as users who adopted ChatGPT were less likely to renew their Grammarly subscription afterward. These findings indicate potential competition between ChatGPT and Grammarly, emphasizing the need for further investigation into how these tools interact and influence user behavior.

1.9 Conclusion and Discussion

Our research explores the implications of ChatGPT adoption on consumers' other digital payments. Utilizing anonymized credit and debit card transactions, we demonstrate that ChatGPT adoption significantly impacts consumers' spending behaviors, particularly by increasing the use of other AI tools such as Grammarly, GitHub Copilot, and Midjourney. This result suggests that initial exposure to generative AI can stimulate further engagement with new technology and related goods. This finding aligns with the broader literature on digital addiction and habit formation, highlighting how new technologies can reshape consumer preferences and behaviors.

[3]LOESS is a non-parametric regression method that fits a smooth curve to the data points, allowing us to observe trends without assuming a specific functional form.

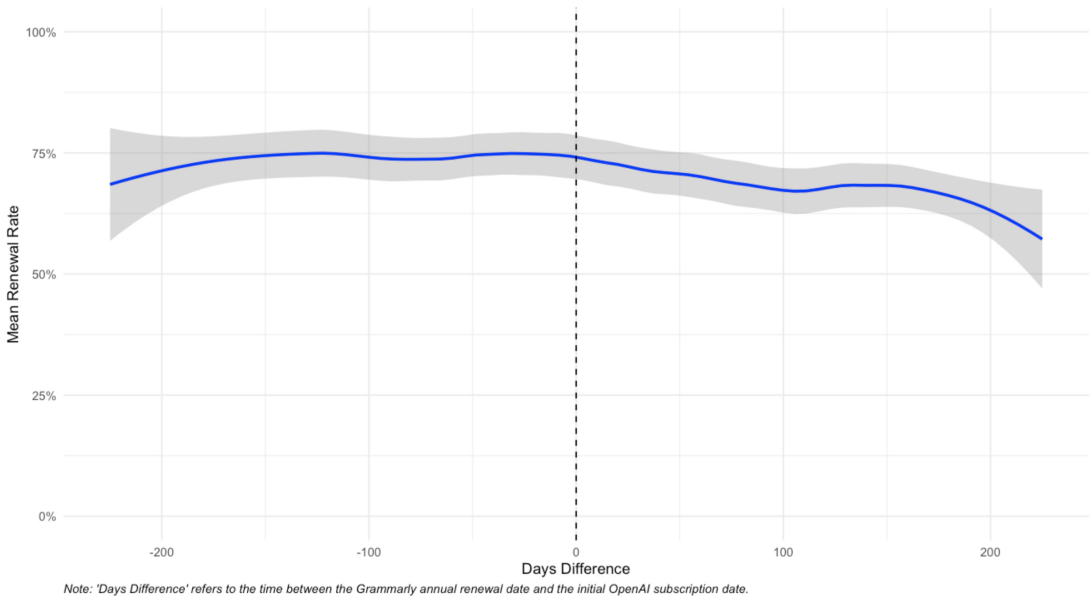


Figure 1.9. OpenAI Subscription Timing vs. Grammarly Renewal Rates

Additionally, we examine the direction of adoption within pairs of AI tools to address potential reverse causality issues. We find a highly significant trend of adopting OpenAI before other AI tools, particularly Midjourney. For Grammarly and GitHub Copilot, there is a directional tendency for users to adopt ChatGPT first before subscribing to other services. Furthermore, we analyze how the annual Grammarly renewal rates co-varied with ChatGPT adoption, suggesting that ChatGPT subscribers tend to cancel their Grammarly annual plan as the time difference between the Grammarly annual renewal date and the initial OpenAI subscription date increases. This trend indicates that adopting ChatGPT may have a substitution effect on Grammarly's annual subscriptions, as users who adopted ChatGPT were less likely to renew their Grammarly subscription afterward.

Although our research shows the possible spillover effect of ChatGPT subscription on other digital purchases, it has certain limitations. First, our specification cannot control the adoption timing or other possible unobservable factors that can increase interest in generative AI in general, including ChatGPT. Thus, our interpretation may not be fully causal. Second, OpenAI recently released a newer version of ChatGPT, ChatGPT-4o. This research cannot explain the

long-term effects of this upgraded version with only one year of transaction data. Future research may investigate how this upgraded version will change consumer preferences.

The potential impact of this new technology is still unclear, but this paper shows a possible directional effect. Despite concerns that generative AI could have negative effects, such as privacy issues or adverse impacts on other industries like education, our research suggests that while new technology increases consumers' purchases related to new technology, it does not significantly alter other behaviors.

ACKNOWLEDGEMENTS

We thank the analytics firm for giving access to the individual-level transaction data. We also appreciate the insightful feedback from other Rady faculty members on this work.

Chapter 1 contains unpublished material coauthored with Daniel Minh McCarthy and Kenneth C. Wilbur. The dissertation author conducted the primary investigation, which involved aggregating the data, performing the descriptive analysis, and conducting the main estimation. Additionally, the dissertation author wrote this paper.

Chapter 2

Gratuities in a Digital Services Marketplace

2.1 Introduction

Digitization is changing traditional tipping practices and norms. Historically, tipping has been a voluntary post-transaction payment given in appreciation of service providers in specific settings like restaurants or personal care establishments, and it is especially prevalent in North America. However, tipping online has brought new complications. Digital platforms and payment providers now frequently request tips. Gig economy workers face uncertain opportunities to work and variable income, and hence often rely on tips to supplement their main income. Customers face more and more tip requests. The evolution in tipping practices requires digital platform designs to adapt to changing consumer behaviors and social norms around tipping.

Online gig work is increasingly common but is not lucrative. In the U.S., 16% of people surveyed have earned income through online gig platforms, with 31% treating it as their main occupation in the previous year (Anderson et al., 2021). However, about 14% of gig workers earned less than the hourly federal minimum wage, and 29% earned less than their home state's minimum wage (Zipperer et al., 2022). Many platforms request consumers to tip gig workers after transactions are complete as a means of supplementing worker income and incentivizing quality work.

Hence, digital platforms bring tip requests into online gig markets without clear tipping

norms. At the same time, consumers have been subjected to changing tipping norms in offline environments (e.g., coffee shops, dry cleaners, etc.) as digital payment processors like Square and Stripe became prevalent. The increasing prevalence of tip requests can cause consumer discomfort: 36% of consumers surveyed said they felt pressured to tip, and 26% were annoyed by the growing trend of tip requests (Miranda and Main, 2023). Further, digital platforms increasingly connect people from diverse tipping cultures, both within countries and across national borders, increasing uncertainty around tipping norms.

This paper aims to answer the following key questions:

- What drives online tipping behavior?
- Can platforms appeal to tipping motivations to increase tipping? If yes, would this reduce subsequent demand?
- Which normative messages are most effective in increasing tipping intention?

We partner with a leading freelance services platform, Fiverr. To begin, we analyze 4.1 million transactions to identify the drivers of online tipping. The results reveal a strong correlation between tipping and buyer nationality, suggesting social norms play a significant role in influencing tipping behavior. Tipping rates increase with buyer satisfaction, rising from 3.5% in 3-star transactions to 14.1% in 5-star transactions, suggesting buyers offer tips to reciprocate for high-quality service. Among various buyer, seller, and transaction characteristics, buyer characteristics explain–by far–the most variation in tipping behavior.

We conducted two field experiments to examine whether digital platforms can leverage specific tipping motivations to encourage tipping. Our focus was primarily on two psychological drivers: norms and reciprocity, as we believed these to be first-order issues in online settings.

A pilot field experiment ran on the web and manipulated default tips and tip requests, reflecting three consumer tipping motivations: norms, reciprocity, and implicit reciprocity. The results indicated that the norms approach significantly boosted tipping rates, by about half among

new platform users on first exposure. However, the reciprocity approaches did not significantly affect tipping behavior.

A subsequent field experiment ran on the mobile app and focused exclusively on the norms manipulation with a more careful test. The findings showed an 18% increase in tipping propensity among new buyers and a 14.1% increase among repeat buyers when first exposed to the norms message compared to the control message. The norms message had a larger treatment effect on highly satisfied new buyers but did not significantly change with buyer nationality or repeated buyer/seller interactions.

The two field experiments demonstrated how a norms message could increase tipping, but their broader application beyond the online freelance sector remains uncertain. We conducted an online experiment using an online subject pool to test various norm messages, including injunctive norms (indicating socially approved or disapproved behaviors) and descriptive norms (indicating typical behavior), against a baseline control message commonly used across different platforms. The results showed that injunctive norm messages generally increased the likelihood of tipping more than the control message. In comparison, descriptive norms messages that emphasized a high tipping rate (e.g., stating millions of riders or about 95% of users tip) had a more significant impact on tipping intentions. These findings suggest that platforms can improve tipping likelihood by customizing normative messages and describing their existing tipping culture. For instance, platforms with a high user tipping rate should use descriptive norms messages with specific numbers.

Finally, we examine the indirect effects of the treatment on sellers using data from the two field experiments. We observe that although the norms treatment increased seller revenue, it had no significant effect on sellers' repeat transactions. This finding suggests that increased tipping benefits sellers without negatively impacting transaction volume on the platform.

2.1.1 Conceptual Framework

Consumers voluntarily leave tips as additional payments after receiving services. Tips therefore raise the interesting and difficult questions of why consumers willingly choose to pay extra (Azar, 2007; Lynn, 2006). Numerous papers study tipping, especially in frequently tipped contexts such as restaurants, transportation, tourism, leisure, and household services (see, e.g., (Azar, 2007, 2011, 2020; Lynn, 2015, 2016; Donkor, 2021; Bluvstein Netter and Raghubir, 2021)). We describe tipping theory and evidence, recent digital tipping research, digital-platform-related issues in tipping, and the current paper's specific contributions.

Lynn (2015) integrated previous literature and proposed the "Tipping Motivation Framework" (TMF), a theory that tipping motivations encompasses helping servers, rewarding service, buying future service, buying social esteem, and sense of duty to tip. This theory can explain the emergence and continuing evolution of tipping norms. Initially, small numbers of individual customers use tipping to help servers, reward service, gain esteem and secure preferential service. Subsequently, as servers begin to rely on and expect tips, their service to and expectations of non-tipping customers begin to change. Tipping norms develop when servers expect tips in exchange for previously customary services. Consumers can strongly internalize tipping norms, even to the point that they continue to tip a customary amount even when they receive poor service.

Lynn (2016) used data to refine and validate the TMF, surveying 378 people about their tipping frequency and motivations across 21 consumer service contexts. Each participant indicated the effect of 14 potential motives for tipping, such as "to reward good service," "to get the server's respect," and "to fulfill a social obligation." A principal components analysis showed four main consumer motivations to tip:

1. Service/esteem motives, that is, gaining the server's liking and respect, getting better service next time, and getting the server to remember me;

2. Altruistic motives, that is, helping the server and increasing their income;

3. Duty motives, that is, complying with tipping norms and fulfilling felt obligation to tip; and

4. Reciprocity motives, that is, rewarding service.

Descriptive regressions further showed consumer motivations differed significantly across contexts, with service/esteem motivations best predicting tipping in rarely tipped occupations, and duty motives best predicting tipping in frequently tipped occupations. The data showed consumer tipping motivations differed significantly between frequently tipped and infrequently tipped consumer service contexts.

Empirical support for service/esteem, duty, and reciprocity motivations illustrates the large differences between tipping and other giving behaviors such as gift giving and charitable giving. Prior work has also found tips can incentivize service effort, helping to explain why tips are more often given after high-quality service (Conlin et al., 2003; Kerr and Domazlicky, 2009; Lynn et al., 2012), reflecting service/esteem and reciprocity motivations. Tipping can also affect non-tipping outcomes such as service-worker retention(Lynn, 2022) and optimal restaurant pricing (Lynn, 2017). To summarize, consumers tip for various reasons, and survey evidence indicates consumer tipping motivations are both heterogeneous and situational (Lynn, 2016; Lynn and Sturman, 2010). Tipping is a complex, dynamic, and fascinating phenomenon.

Service providers increasingly accept digital payments, so consumers increasingly tip using digital systems (Hadero, 2023). Digital systems often request tips, requiring the consumer to actively choose whether to tip or not before completing payment. This development is important because it can change the tipping default; many older payment systems allowed consumers to ignore or even forget the option to give a tip. A burgeoning literature examines how digital tip requests affect tipping. The primary focus of this literature has been how default tip recommendations ("default tips," for short) affect tipping incidence and amounts. We review a few key papers here:

- Haggag and Paci (2014) analyzed New York City taxicab tipping data, a setting in which

98% of rides are tipped. They exploited quasi-experimental discontinuities and found shifting from default tips of ($2, $3, $4, Custom) to (20%, 25%, 30%, Custom) at a threshold of $15 increased average tips by about $0.27-$0.30 (or +2%, a 10% increase relative to baseline) at the threshold. An analysis of quasi-random assignment to default tip menus between (15%, 20%, 25%) and (20%, 25%, 30%) increased average tips but also increased the share of zero tips.

- Donkor (2021) analyzed a similar, but larger, New York City taxi dataset in which 97% of rides were tipped. He proposed a theoretical model of tipping norm compliance and tip calculation, and uses the tipping data to calibrate parameters of the model. He concluded consumers face large costs of norm noncompliance and tip calcuation, and used the model to calculate welfare-maximizing tip menus and to show raising default tips too high would reduce tipping.

- Chandar et al. (2019) analyzed tipping on Uber. Despite similarities between Uber and taxis, only 16% of rides on Uber were tipped in the app, possibly suggesting app design could influence tipping-norms development. They reported numerous Uber tipping descriptives and a field experiment in which default-tips manipulation had a "surprisingly modest" effect on tipping. For example, comparing default tips of ($1, $3, $5) to ($2, $4, $6) increased average tips from $0.467 to $0.479. The authors concluded that "defaults affect tip levels, but are much less powerful than comparable estimates from the literature exploring tips in the taxi cab industry." They speculated that because tips were given after rides were complete, when the rider had left the driver's presence, tipping became more private and less interpersonal. In the context of Lynn (2016), we interpret this finding as suggesting service/esteem and altruistic motives may be stronger when the consumer and service provider are co-located when the consumer tips.

- Alexander et al. (2021) report 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 but increased tip revenues overall. For example, going from a default tip menu of (10%, 15%, 20%) to default tips of (15%, 20%, 25%) increased the average tip from $4.70 to $5.63, but lowered the tipping rate from 58% to 50%. Default tips did not significantly change customer satisfaction, repatronage, or spending.

Digital tip requests clearly affect tipping, but traditional services such as transportation and laundry retain face-to-face interactions between service providers and consumers. What happens when services are fully digitally intermediated? Digital platforms increasingly either match consumers and service providers or enable them to find each other in digital marketplaces. Some of the most recent literature has considered platform-design-related issues in default tips.

- Duhaime and Woessner (2019) found across several contexts that buyers tip employees more than they tip autonomous gig workers, suggesting seller autonomy in the gig economy can partially explain lower tipping rates. However, they did not consider tip-request messages or indirect effects of tipping on other buyer or seller behaviors.

- Chen et al. (2023) document that buyers view ratings as a reward for service, showing convincingly across a range of contexts that tips are greater when tip requests precede rating requests. However, the authors did not consider variation in tip-request messages or broader consequences for platform usage, such as effects on repeat usage or seller behaviors.

- Lu et al. (2021) report a field experiment in which a platform manipulated audience size during livestreaming sessions, and found viewer tipping exhibited an inverted-U shape in reported audience size.

Research to date leaves open the important question of how digital tip requests can appeal to tipping motivations, and how tipping behaviors may respond. Default tips are important, but communicating reasons to tip may also be important, especially when services are fully digitally

intermediated. Tip requests can indicate why tips would be socially desirable, and consumers have fewer opportunities to use social interactions to learn market norms or observe service provider responses to tips.

The current article contributes to the literature in several ways. First, it is two field experiments that manipulate tip-request messages–appealing to two common tipping motivations– as well as default tips *in situ* on a large platform in a fully digitally mediated service context. The findings indicate adding two messages intended to activate consumer reciprocity motivations did not reliably change tipping rates relative to a status quo message, whereas an injunctive norm message about tipping did meaningfully increase tipping. Therefore, this research provides the first field evidence that altering a tip-request message can motivate consumer tipping behavior, suggesting this approach as a strategy that digital platforms can investigate to improve service provider incomes. The null results on the reciprocity messages are precisely estimated, which we view as surprising in light of previous literature.

Second, our paper examines the effects of two different types of tipping norms – injunctive norms (information about what others approve of) and descriptive norms (information about what others do) – on the likelihood of online tipping. Building on the research by Lynn (2021), who found consumers generally conform to injunctive norms about who should be tipped and how much, and that occupational differences in tipping norms can moderate the effects of individual differences in tipping motives, our findings show injunctive norms are generally effective in increasing tipping. However, when a majority of users on digital platforms tend to leave a tip, using descriptive-norms messages with a specific number proved to be more effective.

Third, although we do not claim our findings necessarily apply to other platforms, we do offer an empirical framework that other platforms can adopt when testing their own tip requests. We show how to use standard platform data to detect a range of indirect effects on repeat buyer and seller usage, ratings, satisfaction, and price. We think the current diversity and frequent absence of tipping messages on many platforms suggest this area could likely be developed further, with the potential to increase service provider incomes and service quality.

2.2 Setting and Data

This section describes the platform and tipping measures and then reports key tipping facts and descriptive regressions that inform our subsequent analyses.

2.2.1 Empirical Context

We study data from Fiverr, a marketplace for freelance digital services that enables buyers to search and transact with sellers. Fiverr says its mission is to "change how the world works together." Fiverr connected 2.4 million buyers with 305,000 sellers in 10 digital freelance categories 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.

Sellers list "gigs," which are packages of service attributes and prices. Seller listings and buyer search are free, but Fiverr adds 5% on to buyers' order price and takes 20% from seller revenue.[1] A typical transaction has four steps.

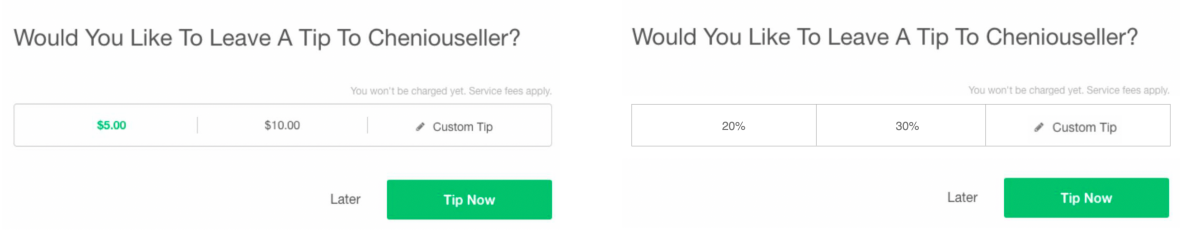
I. Search. Buyer explores the directory of service listings or searches specific keywords and views results. Each gig listing includes the seller's name, service description, seller average rating, and price.

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 releases payment to the seller.

[1] Fiverr's transparent fee structure contrasts with revelations of misleading tip policies that have been used by some other platforms (Newman, 2019; Hanbury, 2019).

IV. Rating and Tip. Fiverr asks the buyer to rate the service. Ratings of three stars or more led to the tip request message and default tips shown in Figure 2.1. 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 emailed rating reminders to buyers who did not rate sellers immediately.



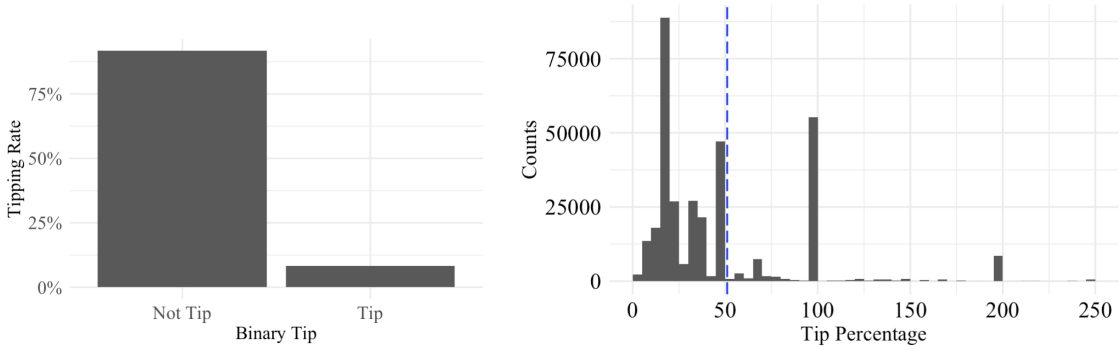
(a) When price $\leq$ $25 **(b)** When price $>$ $25

Figure 2.1. Example of Tip Window

2.2.2 Tipping Measures

This section describes buyer tipping behavior. 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.

Buyers tipped sellers in 8.3% of transactions overall on the platform. The low tipping rate likely reflects the business-to-business nature of most freelance transactions and lack of established tipping norms. Among tipped transactions, the average tip observed is 51.1% of the order price. Figure 2.2(b) shows that the distribution of tip percentage across transactions has large mass points at 20-25%, 50%, and 100%. This is most buyers tipped using default tips, leading to numerous $5 tips on prices of $5, $10 and $20, and frequent tips of 20% on prices above $25.



(a) Tipping Rate **(b)** Tip Percentages

Figure 2.2. Tipping Rate and Tip Percentages

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

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.

Next, we look at how tipping covaries with price. Figure 2.3(a) classifies orders into comparably-sized bins. Like Chandar et al. (2019), tipping rate rises with price, from 5% among \$5 dollar transactions to 11% among \$30-\$40 transactions. However, unlike Chandar et al. (2019), tipping rate then falls with price at particularly high prices, from above 10% to 9% of transactions priced over \$150.

Figure 2.3(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.

Buyers chose default tips in 68% of all tipped transactions even though they could easily enter custom amounts. In what follows, we focus primarily on *Tipping Rate*, the proportion

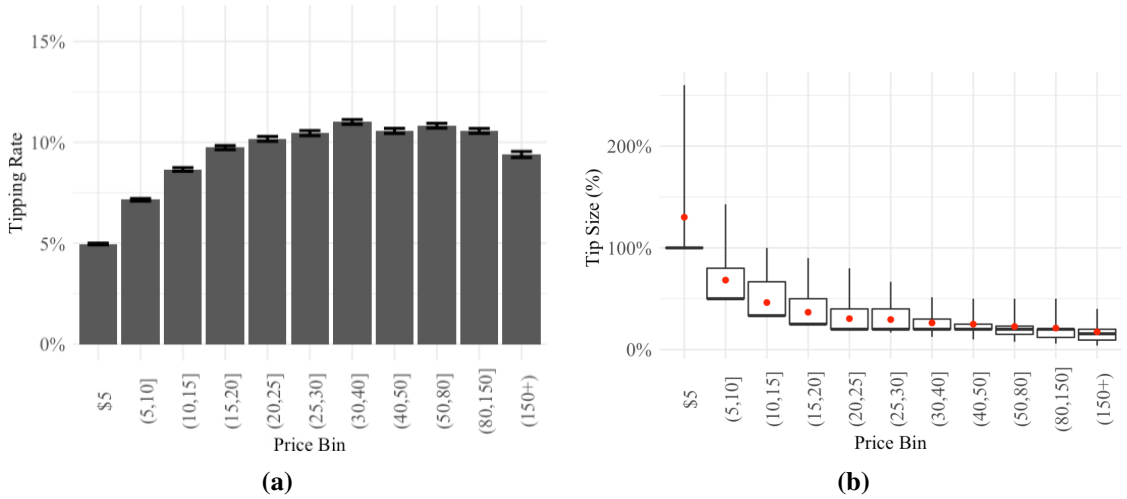


Figure 2.3. Tipping by Price Bin

of transactions that were tipped, simply because this is the single most important behavior for establishing a tipping norm. We also report results for tip amounts and tip percentages.

2.2.3 Tipping Facts

We provide model-free evidence to motivate key features of the analysis.

Do offline tipping norms affect online tipping? Do seller region or buyer/seller region interactions affect tipping? We are unaware of any previous evidence on this topic.

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

Figure 2.4(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 might hypothesize a home-region bias in which buyers tip more to culturally similar sellers or instead a "tourist" effect in which buyer tipping matches local expectations in the seller's region. Figure 2.4(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 influence online tipping.

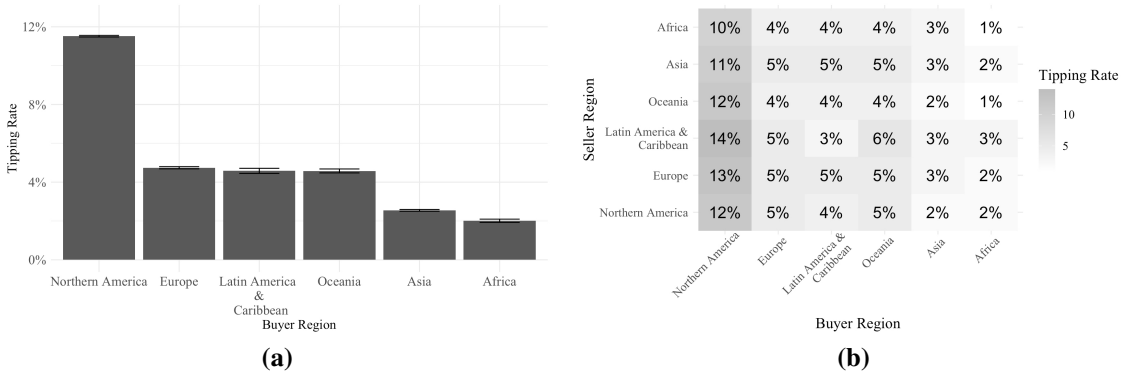


Figure 2.4. Tipping and Region

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 2.5(a) shows that Tipping Rate was four times higher among 5-star transactions than 3-star transactions.[2] We also consider how tipping covaries with sellers' previous average ratings, as seller quality is probably autocorrelated across gigs. Figure 2.5(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.

Online Appendix B.1 provides a series of additional data descriptive showing that Tipping Rates: (i)vary across service categories, (ii) exhibit an inverted "U" shape in buyer prior orders and time since buyer account inception, (iii) increase with buyer prior tips and buyer tenure, (iv) increase with seller tip mentions, and (v) are typically lower for first-time sellers compared to existing sellers. In all of those cases, tipping rate shows more variation relative to its mean than average tipping percentages. As we show later, this is likely the case because of the infrequent usage of custom tips.

[2]99.9% of unrated transactions were also untipped, likely because the tip request immediately followed the rating request.

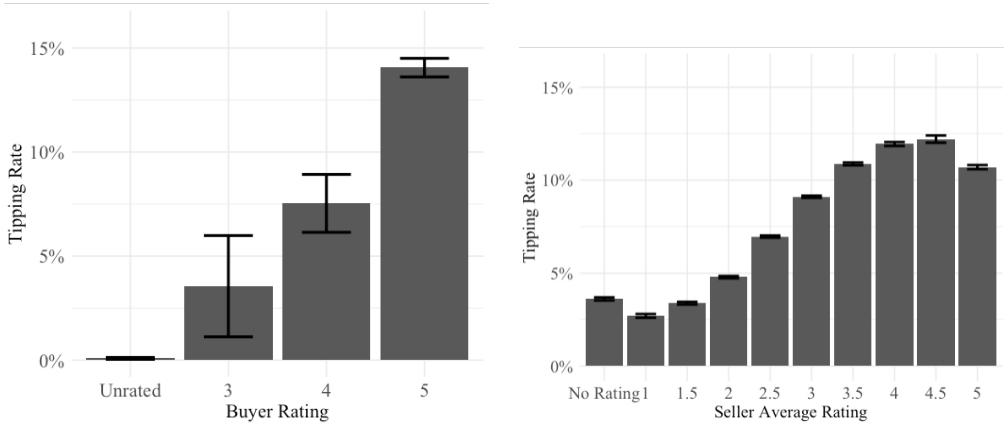


Figure 2.5. Tipping and Service Quality

2.2.4 Descriptive Regressions

We run descriptive regression with the pre-experiment data to simultaneously estimate multiple associations between tipping and observable contextual factors.

First we describe the available predictors. The pre-experiment data contain anonymous account IDs with limited demographic variables. The buyer characteristics include buyer account tenure and user-defined gender, but gender is sparsely populated and non-verifiable.[3] We also observe behavioral data related to purchase, rating, pricing, tipping, and seller tip mentions. Seller tip mention indicates whether the seller's gig delivery message to the buyer included the words "tip" or "tips." Online Appendix Table B.1 provides descriptive statistics of all observable features of the transaction data, in 3 sets: Buyer Characteristics, labeled X^B; Transaction Characteristics, labeled X^T; and Seller Characteristics, labeled X^S.[4] Many of our analyses replace buyer characteristics and seller characteristics with fixed effects; all analyses include transaction characteristics.

Equation (2.1) specifies the descriptive regression:

[3]We do not observe account tenure or user-defined gender for service providers.

[4]Transaction data over-represent frequent buyers, so Table B.1 stats are not the same as user population means.

$$y_i = f(X_i^B, X_i^T, X_i^S, \varepsilon_i) \tag{2.1}$$

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.

Online Appendix Table B.2 displays parameters estimated using linear regression. Column (1) reports a specification with all observable features except buyer region, and column (2) adds buyer region to that model. Column (3) replaces buyer characteristics with buyer fixed effects, whereas Column (4) replaces seller characteristics with seller fixed effects. Column (5) includes both buyer and seller fixed effects. The regressions exclude single-transaction buyers and sellers to hold effective sample size constant across columns; otherwise, buyer and seller fixed effects absorb all information from single-transaction participant observations. Online Appendix Table B.3 reports the same 5 specifications including single-transaction participants, with similar findings.

The descriptive regressions offer three main results:

1. Buyer, transaction, and seller characteristics all help to predict tipping.

2. Buyer factors predict tipping far better than seller factors. Replacing buyer characteristics with buyer fixed effects increases the model's Adjusted R^2 from .127 to .398. Replacing seller characteristics with seller fixed effects increases the model's Adjusted R^2 from .127 to .155.

3. Inclusion of buyer and seller fixed effects can change the magnitudes and estimation error of other associations in important ways. For example, price does not significantly predict tipping when both buyer and seller characteristics are included in the model.

From these findings, we form Key Fact III:

Key Fact III. Buyer factors are stronger predictors of tipping than seller factors.

Key facts I-III inform the analyses below. Next, we describe the field experiment design, data, and results.

2.3 First Pilot Field Experiment: Design and Data

We collaborated with Fiverr on field experiments to understand and motivate tipping. Field experiments have the great advantages of scale, statistical power, and ecological validity. Ecological validity is particularly important in domains, like tipping, where action is costly, because stated intentions and actual behavior may diverge (Melnyk et al., 2022). Of course, field experiments are also often influenced by uncontrolled variables and implementation constraints. Contextual information is required for proper interpretation. The experiment we report was imperfect but, as the first evidence on how tipping request messages motivate tipping, it still offers important and valid findings.

2.3.1 Tip Request Conditions

We have previously introduced the status-quo tip request message, which offered the consumer a tipping opportunity and reminded them of their agency, but did not appeal to any specific tipping motivation. Fiverr interpreted our proposed treatments relative to that status quo. Fiverr's testing apparatus was able to test a maximum of four simultaneous treatments during the pilot, so we were able to add three treatment messages.

We first reviewed the scholarly literature on tipping and thought about how tipping motivations might affect tipping on Fiverr. This is a partially speculative exercise, by necessity, as there is limited research on digital freelance contexts. We thought about the Chandar et al. (2019) result that tips were infrequent on Uber compared to their near ubiquity in taxis, and the possibility that the lack of shared space may severely degrade service/esteem and altruistic tipping motivations. We also thought about the Chen et al. (2023) argument that online buyers could help sellers by giving ratings instead of tips, again offering a free tipping substitute for consumers who have service/esteem or altruistic tipping motivations.

Therefore, we decided to focus on reciprocity and duty motives. In particular, we thought that since the tip request came immediately after the service was delivered and rated, and well after payment was originally made, the timing could make reciprocity tipping motivations particularly salient. Further, most of the services offered on Fiverr are customized to the buyer's needs, offering the possibility of encouraging the buyer to reward the seller's personalized service. We also had a goal to establish a platform norm of tipping for good service, so appealing to the duty motivation directly seemed promising.

Firm executives agreed to the following control and three treatment messages:

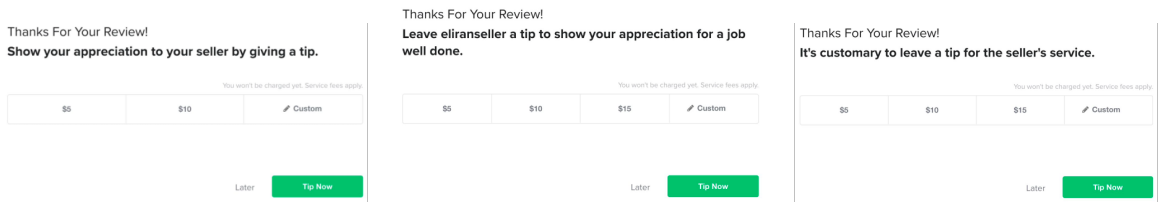
Control/Status Quo: "Would You Like to Leave a Tip To (Seller name)?"

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."

Figure 2.6 shows how the treatment messages looked to buyers, as Figure 2.1 showed how the control message appeared to buyers.



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

Figure 2.6. Tip Request Treatments

A platform tipping norm is not specific to any individual seller. Therefore, we did not want to include the seller's name in Norms message. Despite that, removal of the seller's name was controversial within the company, which had historically emphasized user names in an effort to promote a friendly interpersonal environment. Further, if we removed seller name from the

status quo message, we could not compare control users between time periods, limiting the usefulness of the control group. Therefore, we reached a compromise to include two reciprocity messages, one with and one without the seller's name.

Norms message implies that 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.

Online Appendix B.5 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 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 (2009). The manipulation check suggests that participants perceived both Implicit Reciprocity and Reciprocity messages as "Reciprocity" and Norms message as "Social Proof (Norm)." The interpretations of the three treatment messages did not change with inclusion or exclusion of the seller's name.

The platform simultaneously changed the default tips and threshold in the Reciprocity and Norms conditions as shown in Appendix Figure B.9. While the optimal design might have entailed manipulating the default tips and thresholds separately from the messages, this did not occur in the pilot study because firm executives had final control over the design. Fortunately, we could separate the effects of default tips from message effects by utilizing two post-test platform design changes. In addition, a second field experiment was conducted by Fiverr on mobile apps, eliminating the confounding of tip messages and default tips. This allowed for a more transparent analysis. For a comprehensive understanding of default tip options, Appendix B section B.2 discusses pre-experiment default tip usage based on order price and the adjustments made during the pilot experiment in detail.

2.3.2 Sample

The experiment ran on the web. It began in June 2019 and lasted a predetermined four weeks. During that time it treated 7,880 new buyers and 40,823 repeat buyers. The platform refrained from running other tests in-market during the sample period to avoid potential confounds, despite the platform's normal practice of running shorter, concurrent tests.

2.3.3 Outcomes

We focus on Tipping Rate as the main outcome variable. This is due to the platform's motivation to establish a tipping norm despite low tipping frequency and the greater relative variation in tipping than in average tip percentage. We also report results for models that explain tip percentages and tip amounts.

We distinguish new buyers from repeat buyers because establishing a new tipping behavior may differ from changing existing behavior. We also examine buyer and seller repeat behaviors for three reasons. First, experimental treatments could directly affect participants' transaction utility or satisfaction, and therefore future behaviors like orders, expenditures or ratings. Second, tipping is costly. Thus, if buyers respond to the tipping treatments with more tips, that could operate as a price increase which may reduce subsequent demand (Lynn, 2017). Third, platform and seller incomes depend more on overall expenditures than on tipping. Any intervention that increases tips but decreases overall spending would be contrary to platform and seller interests. 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.

2.3.4 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. Online Appendix Table B.10 reports Chi-square tests indicating that the new buyer regions were well balanced across treatment groups. Online Appendix Table B.11 reports Chi-square tests indicating that repeat buyers' characteristics were well balanced across treatment groups.

2.4 Experimental Treatment Effects on Tips

We first analyze treatment effects on new buyer tipping, then repeat buyer tipping, and then post-test data that distinguish between tip message effects and default tip effects.

2.4.1 Effects on New Buyer Tipping

Figure 2.7 displays mean differences in new buyer Tipping Rate by treatment groups, in the first treated transaction. New buyers who received Norms treatment tipped about 49% more often than Control (21.8% vs. 14.6%, $p<.01$). New buyer tipping rates did not significantly differ between Control, Implicit Reciprocity, or Reciprocity treatment groups.

We use regression to estimate treatment effects more precisely by controlling for covariates. Table 2.1 reports linear regressions of tipping on treatment effects in four transaction samples, controlling for all of the buyer, seller, and transaction characteristics listed in Table B.1.[5] Standard errors in this and all subsequent regressions are clustered at the buyer level. The qualitative findings are robust to a Logit model specification.

Table 2.1 shows that, across all transactions, new buyers exposed to Norms condition had a 3.4% ($p<.01$) higher likelihood of tipping compared to those exposed to Control. Norms effect

[5]We considered including seller fixed effects, but these 7,880 transactions involved over 6,000 distinct sellers, so we instead used seller characteristics, gig category and seller region dummies.

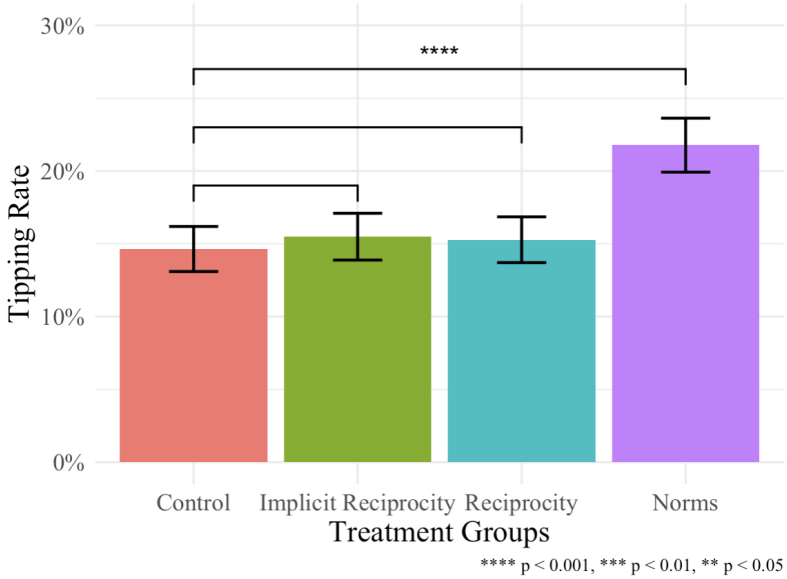


Figure 2.7. Treatment Effects on New Buyers' First Treated Transactions

was especially pronounced when new buyers were first exposed to the treatment, leading to an absolute increase of 7.0% ($p<.01$) compared to the Control. The second and third exposures did not significantly differ from the Control group. Norms findings contrast with the Reciprocity and Implicit Reciprocity conditions, which resulted in smaller, non-significant effects across all transaction subsamples.

Online Appendix Table B.14 shows that Norm treatment increased new buyers' absolute tip amount by \$0.84 on the first exposure, and by \$0.42 in the pooled transaction data. Online Appendix Table B.15 shows that Norm treatment increased new buyers' tip percentage by 3.2% on the first exposure, and by 1.5% in the pooled transaction data. The Implicit Reciprocity and Reciprocity conditions did not significantly differ from Control.

We also estimate treatment effects on new buyers' repeat purchases during the sample period, using the same regression specification. Table 2.2 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 Norms and Implicit Reciprocity treatments led to slightly more repeat purchases on average, whereas the Reciprocity

Table 2.1. New Buyers Tipping Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity	0.001	0.010	−0.012	0.004
	(0.006)	(0.011)	(0.009)	(0.012)
Reciprocity	0.003	0.005	−0.002	0.007
	(0.006)	(0.011)	(0.010)	(0.012)
Norms	0.034***	0.070***	0.016	0.008
	(0.007)	(0.012)	(0.010)	(0.013)
Buyer Char.	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category FE	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545
R^2	0.163	0.076	0.241	0.339
Adjusted R^2	0.162	0.072	0.237	0.332
Note:				**p<0.05; ***p<0.01

Table 2.2. New Buyer Repeat Transaction Regression

	Second Order Exists	Third Order Exists
	(1)	(2)
Implicit_Reciprocity	0.005	0.012
	(0.008)	(0.016)
Reciprocity	−0.005	−0.015
	(0.009)	(0.016)
Norms	0.009	0.007
	(0.008)	(0.016)
Buyer Char.	Y	Y
Transaction Char.	Y	Y
Category FE	Y	Y
Seller Char.	Y	Y
Observations	7,880	7,258
Number of Buyers	7,880	7,258
R^2	0.026	0.090
Adjusted R^2	0.022	0.086
Note:		**p<0.05; ***p<0.01

treatment led to fewer repeat purchases, but none of the effects were statistically significant. Therefore, the platform viewed Norms condition as increasing tipping without reducing new buyers' return visits to the platform.

2.4.2 Effects on Repeat Buyer Tipping

Figure 2.8 shows the mean differences in repeat buyer Tipping Rate by treatment groups. Repeat buyers exposed to Norms treatment tipped about 10% more often than Control (16.7% vs. 14.4%). 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 14.7% transactions tipped in the Implicit Reciprocity and Reciprocity treatment groups.

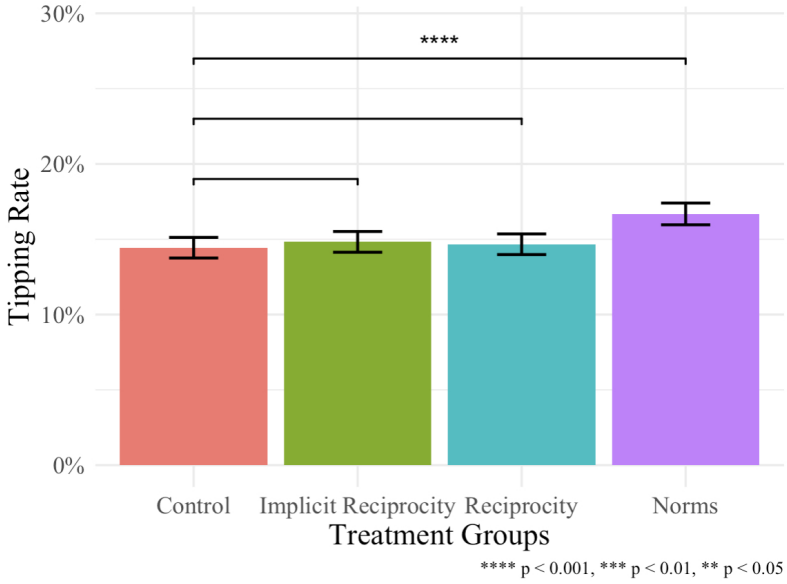


Figure 2.8. Treatment Effects on Repeat Buyers

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 (2.2) to identify causal effects of treatments on Tipping Rate of repeat buyers.

$$y_i = \beta_1 \text{Post}_i + \beta_2^{M_i}(\text{Treat}_{M_i} \cdot \text{Post}_i) + \beta_3 X_i^T + \mu_{b_i} + \lambda_{s_i} + \varepsilon_i \tag{2.2}$$

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 the transaction characteristics listed in Table B.1; 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 2.3 shows that, across all transactions, repeat buyers exposed to Norms condition had a 0.5% ($p<.01$) higher likelihood of tipping compared to those exposed to Control. Norms effect was again larger when repeat buyers were first exposed to the treatment, leading to an absolute increase of 2.0% ($p<.01$) compared to Control. Second and third exposures did not significantly differ from the Control group. Norms findings contrast with the Reciprocity and Implicit Reciprocity conditions, which resulted in non-significant effects across all transaction subsamples.

Online Appendix Table B.16 shows that Norm treatment did not significantly change repeat buyers' tip amounts, but repeat buyers in the Reciprocity condition who transacted three times in the sample period tipped \$0.30 less ($p<.05$) on the third transaction than Control. Online Appendix Table B.16 shows that Norm treatment did not significantly change repeat buyers' tip percentages, but repeat buyers in the Reciprocity condition who transacted three times in the sample period tipped 1.6% less ($p<.01$) on the third transaction than Control, and repeat buyers in the Implicit Reciprocity condition who transacted three times in the sample period tipped 1.2% less ($p<.01$) on the third transaction than Control.

Table 2.4 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 Norms treatment on repurchase was directionally positive but not significantly different from Status Quo, indicating no evidence that Norms treatment message reduced subsequent

Table 2.3. Repeat Buyers Tipping Regression

	Pool (1)	First (2)	Second (3)	Third (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	−0.00001	0.002	0.002	0.001
	(0.002)	(0.003)	(0.003)	(0.004)
Buyer FE	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category 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

demand.

Overall, we have learned that Norms treatment increases tipping by new buyers and repeat buyers upon first exposure and does not reduce their subsequent platform demand.

Next, we explore post-experiment platform design changes to separately identify causal effects of treatment messages and default tips.

Table 2.4. 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)
Buyer FE	Y	Y
Transaction Char.	Y	Y
Category FE	Y	Y
Seller FE	Y	Y
Observations	638,195	581,650
Number of Buyers	40,823	34,871
R^2	0.455	0.594
Adjusted R^2	0.330	0.501
Note:		**p<0.05; ***p<0.01

2.4.3 Post-Experiment Changes in Tip Requests and Default Tips

We explain two post-experiment platform design changes on Fiverr web and how they enable us to isolate tip request effects from default tip effects.

Post-Experiment Reversion to Status-Quo Tip Request and Adoption of New Default Tips

Immediately after the field experiment concluded, the platform reverted to showing Status Quo messages to all users *with the new default tips*. Table 2.5 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. A nuance in interpretation here is that here that we estimate the *removal* of the Reciprocity and Norms messages, whereas before we estimated the combined effects of introducing the Reciprocity/Norms messages and new default tips.

Table 2.5. 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	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

Online Appendix Figures B.12 and Figure B.13 show mean differences in Tipping Rate between the test and post-test periods across treatment conditions. New buyers in the Reciprocity and Norm treatment conditions tipped statistically significantly less in the post-test period than in the test period. Repeat buyers in all treatment conditions tipped statistically significantly less in the post-test period than in the test period, but repeat buyers showed more muted differences.

We can use the same identification strategy as before to identify quasi-experimental effects of post-test platform design changes on tipping. We combine the test data and post-test data, then specify Equation (3) to separate the default tip effect from Norms and Reciprocity message effects.

$$y_i = \beta_1 \text{Post\_Test}_i + \beta_2^{M_i}(\text{Treat}_{M_i} \cdot \text{Post\_Test}_i) + \alpha_1 X_i^T + \mu_{b_i} + \lambda_{s_i} + \varepsilon_i \tag{2.3}$$

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 new default tips on tipping. Norms 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 removing Norm and Reciprocity messages on tipping, respectively. The Implicit 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 (2.3) also controls for buyer fixed effects, seller fixed effects, and transaction characteristics as in previous analyses.

Table 2.6 shows that removing Norms treatment message without changing the default tips reduced new buyers' tipping during the post-test period by a 3.5%, or about 28% (12.5% vs. 16%). 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 5.3%. 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 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 by 3.5%. The effect is statistically significant, but the result suggests that the default tip change had a smaller effect than Norms treatment message. Repeat buyers also showed a more muted response to the default tips change.[6]

[6]We checked whether the sum of Norms treatment effect and the default tips effect is consistent with their joint causal effect reported in sections 4.1 and 4.2. The estimates are not very different from each other, but the comparison is not statistically powerful due to greater estimation errors in the post-test analysis.

Table 2.6. Post-test Regression

	New Buyers	Repeat Buyers
	(1)	(2)
Implicit_Reciprocity:Post_Test	−0.012	−0.002
	(0.017)	(0.003)
Reciprocity:Post_Test	−0.053***	−0.004
	(0.018)	(0.003)
Norms:Post_Test	−0.035**	−0.003
	(0.017)	(0.003)
Post_Test	0.035***	0.00005
	(0.013)	(0.002)
Category FE	Y	Y
Buyer Region FE	Y	Y
Seller Region FE	Y	Y
Observations	19,255	210,199
Number of Buyers	2,920	25,315
R^2	0.796	0.668
Adjusted R^2	0.442	0.495
Note:		**p<0.05; ***p<0.01

Platform Adoption of Norm Message

The platform adopted Norms message as the default for all users 3 weeks after the test period concluded. The default tips interface did not change at that time, keeping the new default tips of ($5, $10, $15) for orders under $35, and (15%, 20%, 25%) for orders over $35.

Figure 2.9 shows weekly tipping rates for new buyers who first transacted after the test period and before the platform-wide Norms message adoption. It also shows weekly tipping rates for new buyers who first transacted after the platform-wide Norms message adoption.

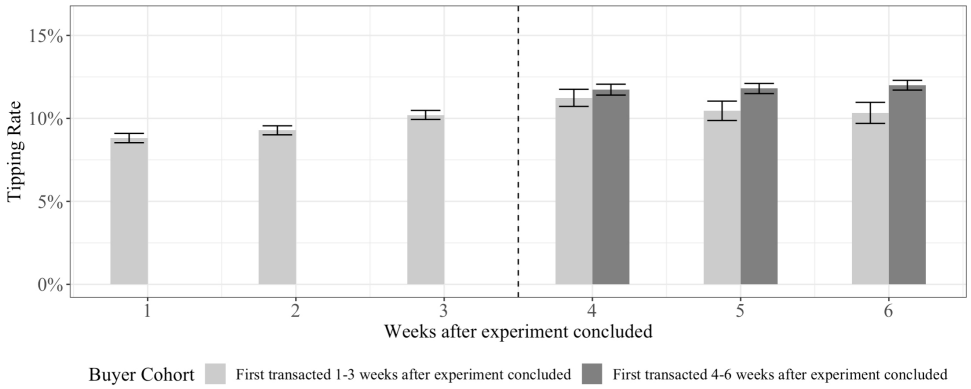


Figure 2.9. Tipping Rate during Post-test Period

We find two interesting results. First, new buyers who arrived in the first three weeks after the test period–and hence first received the status quo message–tipped in 9.4% of their transactions during that three-week period. However, after the platform adopted Norms message for all buyers, that cohort of buyers' tipping rate rose to about 10.8% of transactions, a statistically meaningful increase. Second, new buyers who first transacted after the platform adopted Norms message for all buyers tipped at meaningfully higher rates than the immediately previous cohort of new buyers (11.9% vs. 9.4%).

We consider these results correlational, as new cohorts of buyers who arrived after platform-wide Norm adoption may have differed from new cohorts of buyers who arrived before platform-wide Norm adoption. Still, we know of no other explanation for the difference, and the correlations are directionally consistent with a positive causal effect of Norms message on new

buyer tipping rate. Building upon initial findings from the first pilot experiment, the platform conducted a second field experiment focused on testing the effects of Norms. In the following section, we analyze data from this second field experiment and explore whether the observed Norm effects persist across different channels.

2.5 A Second Experiment: Design and Data

2.5.1 Tip-Request Conditions

The initial pilot experiment and post-test analysis demonstrate the impact of the norms treatment on tipping. Building on these findings, we conducted the second field experiment that focused solely on the norms treatment. The tested control and norms messages were as follows:

Control/Status Quo: "Would You Like to Leave a Tip To (Seller name)?"

Norms: "It's customary to leave a tip for the seller's service."

The second experiment focuses on determining the robustness of norms messages' effectiveness, especially once we remove potential confounding factors. The design of the initial pilot study was constrained by the platform's system, allowing only four messages to be tested, and additionally led to alterations in default tip options for two of these messages, due to the firm's interest in evaluating different tip options. This confounding design introduced complexities in distinguishing the impact of tip messages from the effects of default tip options on tipping behavior. To address this, the second field experiment maintained consistent default tip options irrespective of the message type. For instance, services priced below $35 used default tips of ($5, $10, $15), whereas services priced above $35 had percentage-based defaults (15%, 20%, 25%). This refined design allows us to directly assess the influence of messages on tipping behavior, independently of the influences of default tip options.

2.5.2 Sample and Outcome

The second experiment was conducted exclusively on mobile app transactions,[7] spanning from October 17 to December 10, 2019. It treated 19,749 new buyers and 42,153 repeat buyers. Similar to the first pilot study, the platform refrained from running other in-market tests during the period to avoid potential confounds.

In the analysis of the second field experiment, we use the same specifications detailed in section 2.4. We continue to focus on the tipping rate as the primary outcome variable, due to its direct relevance to our main objectives of increasing tipping frequency and establishing a tipping norm. We differentiate between new buyers and repeat buyers, acknowledging the unique dynamics of introducing new tipping behaviors versus adjusting established tipping habits. We also analyze results in multiple ways, estimating the treatment effects on tipping using pooled transaction data. Furthermore, we extend our analysis to observe how the treatment effect changes with repeated exposure, by segmenting the samples based on each buyer's order number. This approach provides a comprehensive understanding of tipping behaviors and the effects of intervention across different buyer segments, and the adaptive response to recurring exposures.

2.5.3 Randomization Check

We check whether the firm successfully randomized buyers who placed the order via the mobile app to treatments. We do these tests separately for new buyers and for repeat buyers. Chi-square tests reported in Online Appendix Table B.12 show a well-balanced distribution of new buyer regions between the control and norm groups. Similarly, Online Appendix Table B.13 presents Chi-square tests indicating a balanced distribution of repeat buyers' characteristics between the control and norm groups.

[7]The firm carried out randomization when buyers placed orders using the mobile app. However, buyers could have switched to other channels, such as the web or mobile web, when they finalized the service and tip.

2.5.4 Treatment Effects on New Buyers

Figure 2.10 displays the mean difference in the tipping rate among new buyers between control and norms groups for the first treated transactions on the mobile app. New buyers in the norms treatment tipped approximately 18% more often than those in the control group (9.8%vs. 11.6%, $p<.01$).

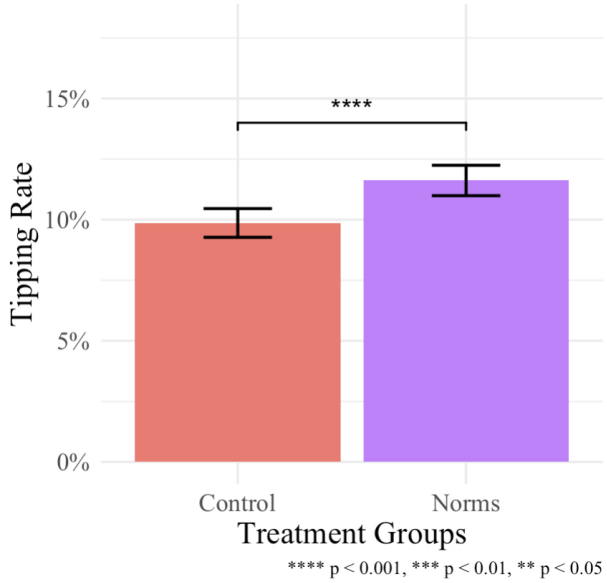


Figure 2.10. Treatment Effects on New Buyers' First Treated Transactions

We further estimate the norms treatment effect more precisely by controlling for covariates, using the same specification discussed in section 2.4.1. Column (1) in Table 2.7 shows that, on average, new buyers exposed to the norms message condition were 2.0% ($p<.01$) more likely to tip than the control group. When considering only the first transactions, as shown in column (2), new buyers who first encountered the norms message via the mobile app displayed a 2.1% ($p<.01$) increased tipping rate compared with the control group. Columns (3) and (4) show subsequent exposures to the norms message (second and third encounters) did not significantly alter tipping behavior relative to the control group. Thus, the second field experiment confirms normative messages effectively enhance tipping behavior among new buyers, with the effect being free of confounding factors.

Online Appendix Table B.18 shows the norms treatment led to an increase in new buyers' absolute tip amount by $0.12 on the first exposure and $0.16 in the pooled transaction data. Furthermore, Table B.19 shows the norms treatment boosted new buyers' tip percentage by 0.76% on the first exposure and 0.79% in the pooled transaction data.

We also estimated the treatment effects on new buyers' repeat purchases during the second sample period, using the same regression specification as before, but with two different outcome variables. The outcome variable in column (5) is 1 if a new buyer completed a second transaction during the sample period, and 0 otherwise. Similarly, the outcome variable in column (6) is 1 if a third transaction took place during the test period. The results suggest the norm treatment has a slight, albeit statistically insignificant, decrease in repeat purchases via the mobile app following the first transaction for new buyers. Yet, when considering instances where a third order was made, the norms treatment seems to marginally increase the frequency of repeat purchases, although the effect is not statistically significant. These findings suggest the norms treatment does not discourage new buyers from returning to the platform, irrespective of whether they access the platform via the mobile app or the web.

Table 2.7. New Buyers Tipping and Repeat Transaction Regression

	Pool	First	Second	Third	Second Order Exists	Third Order Exists
	(1)	(2)	(3)	(4)	(5)	(6)
Norms	0.020***	0.021**	0.042	0.018	−0.00004	0.005
	(0.006)	(0.008)	(0.022)	(0.048)	(0.007)	(0.012)
Buyer Char.	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y	Y	Y
Observations	36,073	19,749	7,216	3,402	19,749	7,216
Number of Buyers	19,749	19,749	7,216	3,402	19,749	7,216
R^2	0.529	0.582	0.803	0.888	0.033	0.032
Adjusted R^2	0.116	0.096	0.151	0.172	0.031	0.028

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

2.5.5 Treatment Effects on Repeat Buyers

In the second field experiment, the platform randomized 42,153 repeat buyers who placed orders through the Fiverr mobile app. Among these, 2,760 repeat buyers had previously been a part of the first pilot experiment, accessing Fiverr through the web. When these buyers placed an order via the mobile app during the second treatment period, they were included in the new randomization process, regardless of their participation in the first experiment. To prevent potential confounding effects due to their prior exposure to multiple conditions and to provide a more accurate estimate of the independent norms treatment effect, we excluded these users from our main analysis.

The final sample for our analysis consists of 39,393 repeat buyers, with the data further divided into three distinct buyer groups based on the primary Fiverr access channel of each buyer. This categorization allows for separate analysis, because tipping behaviors may be influenced by different channel usage patterns. The first group (N_1 = 6,146) includes repeat buyers who had no transaction history prior to the second field experiment in 2019, but began placing orders via mobile during the second treatment period. The second group (N_2 = 9,368) consists of repeat buyers who consistently accessed Fiverr exclusively via the mobile channel prior to the second experiment. Lastly, the third group (N_3 = 23,879) includes repeat buyers who accessed the platform through multiple channels, including both web and mobile.

Figure 2.11 shows the mean difference in the tipping rate of the 39,393 repeat buyers by treatment groups during the first treated mobile app transactions. Repeat buyers in the norms treatment tipped approximately 14.1% more often than the control group (10.5% vs. 9.2%, $p < .01$)[8].

We then explore the potential variation in the tipping rate of repeat buyers, based on their channel usage patterns before the treatment. Figure 2.12(a) shows the tipping rate for

[8]We also summarize the mean difference in the tipping rate with transactions for all 42,153 repeat buyers, which includes the 2,760 repeat buyers who were treated in both the first pilot and second experiments. As shown in Figure B.14, a significant effect of the norms treatment on tipping is still observed, even when including repeat buyers who were treated twice.

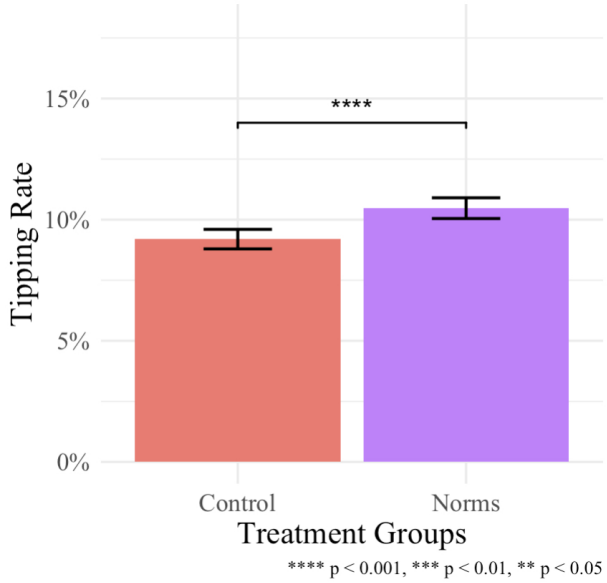


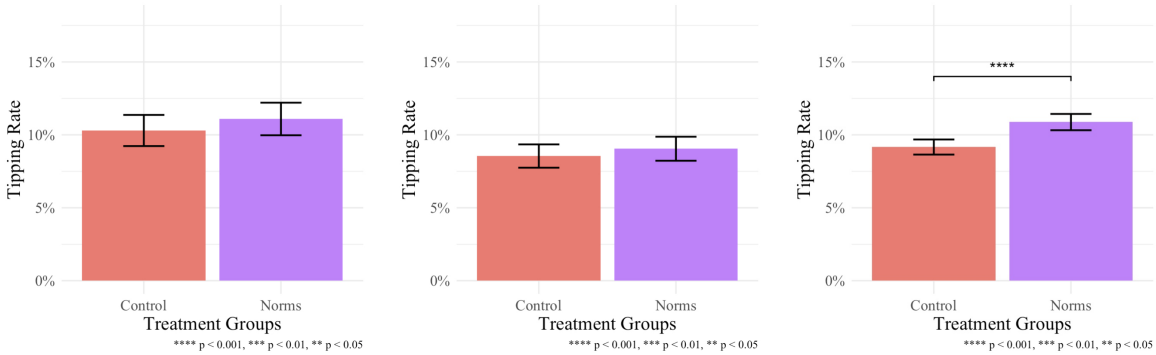
Figure 2.11. Treatment Effects on Repeat Buyers' First Treated Transactions

repeat buyers who had not accessed the platform prior to the second experiment does not show a significant difference in tipping between the norms and control treatments. When we consider repeat buyers who exclusively accessed the platform via the mobile channels, as presented in 2.12(b), we again observe no significant difference in the tipping rate. On the other hand, Figure 2.12(c) shows a significant increase in the tipping rate under the norms treatment compared with the control groups among repeat buyers who accessed the platform via multiple channels. The result suggests more actively engaged repeat buyers, particularly those using multiple channels, are more likely to conform to the suggested tipping norms than less frequent buyers.

We further estimate the norms treatment effect more precisely, employing the same identification strategy specified in equation (2.2)[9]. We merge transactions from both the pre-test period, before October 17, and the second treatment period. Utilizing a difference-in-differences regression model with buyer fixed effects, seller fixed effects, and transaction characteristics allows us to control for the context in which each tip occurs and to control for repeat buyers' tipping habits.

In Table 2.8, column (1) indicates that across all transactions, repeat buyers exposed to

[9] Treat_{M_i} is used here to represent a norms treatment dummy.



(a) First group **(b)** Second group **(c)** Third group

Figure 2.12. Treatment Effects on First Transactions of Three Categories of Repeat Buyers

Figure (a) displays the mean difference in the tipping rate for the first group, representing repeat buyers who had no prior transaction history but began ordering via mobile during the second treatment period. Figure (b) depicts the mean difference in the tipping rate for the second group, corresponding to repeat buyers who exclusively accessed Fiverr via mobile before the second experiment. Figure (c) illustrates the mean difference in the tipping rate for the third group, representing repeat buyers who interacted with the platform via various channels.

the norms message were 0.8% ($p<0.01$) more likely to tip than those exposed to the control. The tipping rate increased by 0.9% upon their first exposure to the norms message. [10]. Online Appendix Tables B.22 indicate that although the norms treatment boosted the tipping rate, it did not significantly change the amount of each tip among repeat buyers.

In a further analysis, we estimated the causal effects of the treatment on the likelihood of observing repeat buyers completing a second or third transaction during the sample period, as presented in columns (5) and (6). We found no significant differences in repurchase rates between the norms and control groups. These results suggest the norms treatment does not discourage repeat buyers from continuing to make subsequent purchases on the platform.

Summarizing the results of the second field experiment, the norms message clearly directly encouraged tipping behavior among both new and repeat buyers without confounding

[10]We run two additional difference-in-differences regressions First, we include transactions from all repeat buyers, even those who were exposed to the treatment twice. The results, outlined in Online Appendix Table B.20, show the significance of the estimates remains unchanged, even when including repeat buyers treated twice. Second, we considered only transactions from frequent buyers who regularly interact with the platform through multiple channels. We observed a stronger effect of the norms message upon the first exposure for these frequent repeat buyers. As displayed in Online Appendix Table B.21 column (2), the tipping rate increased by 1.1% upon their first exposure to the norms message via mobile.

Table 2.8. Repeat Buyers Tipping and Repeat Transaction Regression

	Pool (1)	First (2)	Second (3)	Third (4)	Second Order Exists (5)	Third Order Exists (6)
Norms:Post	0.008***	0.009***	0.009	0.003	−0.002	0.007
	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.008)
Post	−0.023***	−0.020***	−0.027***	−0.024***	−0.479***	−0.396***
	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.006)
Buyer FE	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y
Observations	412,115	364,821	343,422	335,013	364,821	343,422
Number of Buyers	39,393	39,393	35,298	34,141	35,298	34,141
R^2	0.552	0.570	0.572	0.573	0.793	0.778
Adjusted R^2	0.372	0.385	0.389	0.391	0.705	0.683

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

factors. Moreover, the norms treatment had a more significant effect when both buyers were exposed to the norms treatment initially.

We further explore two additional aspects to deepen our understanding of the effects of the norms message treatment. First, we examine the potential moderating effect of the norms message using two sets of field experiment data, aiming to understand why this specific message significantly increased tipping while the two reciprocity messages did not. Second, we examine the seller's side of the market, investigating the possibility of the treatment having an indirect effect on sellers' behavior on the platform. Detailed discussions regarding these explorations are provided in Appendix B.3 and Appendix B.4.

2.6 Laboratory Experiment

Previous field experiments have shown norm messages, particularly injunctive ones, can effectively increase tipping in online freelance marketplaces. These experiments generally focus on testing specific norm messages in that particular context. However, the use of social-norm messages is not limited to this context; they can be varied and broadly applied in other areas, such as ride-sharing or food-delivery services. Considering this wide potential for application, understanding which message can most effectively increase tipping across different platforms

becomes crucial.

Existing literature emphasizes the distinction between descriptive and injunctive social norms, as they tap into different aspects of human motivation (Cialdini et al., 1990; Cialdini and Trost, 1998; Deutsch and Gerard, 1955). Descriptive norms refer to the perception of what people commonly do, thus informing a social standard for behavior. By contrast, injunctive norms refer to the perception of what people should do, by giving a sense of obligation. The influence of these norms on individuals' perceptions and responses has been extensively studied, yet the results can vary depending on the context in which they are applied. For example, descriptive-norm messages have been observed to significantly increase hotel towel reuse rates (Goldstein et al., 2008) and vegetable consumption (Robinson et al., 2014). Conversely, injunctive norms have been found to be more effective in raising both the likelihood and the amount of tipping than descriptive norms in offline environments, such as coffee shops and convenience stores. (Lynn, 2021).

Building upon both field experiments and the existing literature, this section is dedicated to finding optimal messages that digital platforms can use to increase tipping. We present a laboratory experiment that allows us to directly apply theoretical insights from the literature to practical scenarios, deepening our understanding of how to use social norms best to stimulate online tipping.

2.6.1 Message manipulation

We generate a total of 10 messages, including eight distinct norm-treatment messages and two control messages, designed for each of the three service contexts: Fiverr, ride-sharing, and food delivery, resulting in 30 message conditions.

For norms-treatment messages, we manipulate the type of norm and the sentence attributes. We first distinguish between descriptive and injunctive norms through verb choice, according to each norm's definition. For instance, for descriptive norms, we use the verb "leave" in conjunction with "tips." For the injunctive norm, we add "should," resulting in "should leave

tips."

Descriptive Norm Condition: ... leave tips ...

Injunctive Norm Condition: ... should leave tips ...

We then modify the details of the treatment sentences by altering the specific attributes, considering the numerous ways to implement norm messages. For descriptive norms, we contrast quantifiable and non-quantifiable elements, and for injunctive norms, we vary the message tone, comparing active with passive. We also consider all manipulated messages from both buyers' and sellers' perspectives. This approach allows us to see if the participants' reactions differ depending on whether the message emphasizes the payer's or the service provider's viewpoint. A detailed structure of our message manipulation can be found in Table 2.9.

In addition to treatment conditions, we incorporate two control messages. The first control, referred to as the "no-norm message," reflects real-world prompts used by platforms such as Uber and Lyft, stating, "Add a tip." The second control, labeled as the "unclear-norm message," employs a norm message tested in previous field experiments: "It is customary to pay a tip for the seller's service."[11] Further details on message manipulations and stimuli examples across contexts can be found in our online appendix, Table B.24 and Figures B.22, B.23, and B.24.

2.6.2 Method

This study was preregistered(https://aspredicted.org/P8L_22F). We recruited 3,000 MTurk workers and randomly assigned them to one of 30 conditions (10 messages across 3 contexts) in a between-subjects design. Thirty-two participants were excluded from the sample: three participants did not provide consent, and 29 participants had incomplete or inconsistent

[11]We termed the second control as the "unclear-norm message." Our initial aim was to test an injunctive norm from prior field experiments. However, the term "customary" led to varied perceptions among participants, with some interpreting the tested message as either injunctive or descriptive. Consequently, we labeled it as an "unclear-norm message" to account for this ambiguity.

Table 2.9. Structure of Norm Messages

Type	Quantifiable	Tone	Subject
Descriptive Norm	Yes		Buyer
Descriptive Norm	Yes		Seller
Descriptive Norm	No		Buyer
Descriptive Norm	No		Seller
Injunctive Norm		Active	Buyer
Injunctive Norm		Passive	Buyer
Injunctive Norm		Active	Seller
Injunctive Norm		Passive	Seller
No Norm			
Unclear Norm			

responses due to pausing and resuming the survey. Thus, the resulting sample consisted of 2,968 participants.

In the laboratory experiment, each participant was presented with a hypothetical scenario involving the use of an online service: either Fiverr, a ride-sharing platform, or a food delivery service (see Online Appendix Table B.25 for detailed descriptions of the hypothetical scenarios). After participants read each scenario, they were directed to a prompt offering the option to leave a tip with messages and default tip options. Thereafter, each participant responded to the question, "How likely are you to leave a tip for this online service" using a seven-point scale (1 = *Very Unlikely*, 7 = *Very Likely*)[12]. Finally, to ensure the manipulated messages were interpreted as intended, we asked all participants to respond to the question, "How do you perceive the sentence '(insert manipulated messages)" as?' The available responses included "Describing what people commonly do," "Prescribing what people should do," "Both," "Neither," and "Not sure."

[12]Due to the hypothetical nature of the scenario, we opted for a 7-point scale rather than a binary indicator to capture more detailed information on the likelihood of tipping.

2.6.3 Results

We begin our analysis by constructing frequency graphs for each of the 30 different messages to get insights into variations in tipping behavior as they display how participants distributed their responses.

Figures B.25 (a), (b), and (c) separately present frequency distributions of the likelihood of tipping on a 7-point scale in response to various messages across three different digital service contexts: Fiverr, ride-sharing, and food delivery. Regardless of the contexts, response distributions are varied within each message type, implying the different messages may influence tipping behavior in diverse ways when people use digital services.

In the Fiverr context, tipping likelihood spreads evenly across the scale, indicating no clear preference for low or high tipping. Conversely, ride-sharing and food-delivery contexts skew toward higher tipping rates. The frequency of low tips (scores of 1 or 2) is notably lower across all message types in these two contexts. This trend suggests messages within ride-sharing and food-delivery services may be more effective in promoting generous tipping than those in the Fiverr context.

We further use regression analysis to identify the effect of sentence attributes on the likelihood of tipping.

Equation (2.4) specifies the regression

$$y_i = \alpha_0 + \alpha_1 \text{Type}_i + \alpha_2 \text{Quantifiable}_i + \alpha_3 \text{Active}_i + \alpha_4 \text{Buyer}_i + \varepsilon_i, \tag{2.4}$$

where i indexes each participant; y_i denotes the likelihood of tipping on the 7-point scale; Type$_i$is a categorical variable indicating the type of norms, namely, no norms, unclear norms, descriptive norms, or injunctive norms,with the no-norm message serving as the baseline category. Quantifiable$_i$is a dummy variable denoting the presence of a quantifiable factor within the descriptive norm, whereas Active$_i$ is a dummy variable that indicates an active tone within the injunctive norm. ;I_Buyer$_i$is a dummy variable indicating the buyer's perspective, and Context$i$

represents the three contexts: Fiverr, ride-sharing, and food delivery. Lastly, εi captures all unobserved factors that might influence the likelihood of tipping.

Table 2.10. Laboratory Experiment Tipping Regression

	Likelihood of Tipping
Unclear Norm	0.088
	(0.141)
Descriptive	−0.157
	(0.127)
Injunctive	0.008
	(0.127)
Quantifiable (Descriptive Only)	0.271***
	(0.100)
Active (Injunctive Only)	−0.081
	(0.100)
Buyer Factor	0.092
	(0.070)
Context_Food_Delivery	1.711***
	(0.077)
Context_Ride_Sharing	2.132***
	(0.077)
Constant	3.949***
	(0.109)
Observations	2,968
R^2	0.226
Adjusted R^2	0.224
Note:	**p<0.05; ***p<0.01

Table 2.10 shows the results of the regression analysis. The findings indicate that compared with the baseline no-norm message, descriptive-norm messages tend to decrease the likelihood of tipping. By contrast, both injunctive-norm messages and unclear-norm messages appear to have a positive, albeit statistically insignificant, impact on tipping. Notably, the inclusion of quantifiable elements in descriptive-norm messages significantly boosts the likelihood of tipping, underlining the importance of explicit numerical information in shaping tipping behavior. Regarding the tone of injunctive-norm messages, those with a passive tone tend to be more

effective than those with an active tone. The analysis also suggests messages that adopt the buyer's perspective are generally more persuasive than those employing the seller's perspective, even though this effect is not statistically significant.

Based on these results, platforms should customize the tip messages to align with their existing tipping culture. On platforms with a strong culture of tipping, quantifiable descriptive-norm messages that underscore the prevailing tipping behavior (e.g., "About 95% of users leave tips," "Millions of riders tip") may be more effective. Conversely, on platforms where tipping is less prevalent or non-existent, injunctive-norm messages could be utilized to boost online tipping motivation.

We further evaluate how participants perceive their assigned manipulated message, which enables us to verify if the participants' understanding of the message corresponds with our intentions and design. Figure B.26 shows participants' perceptions of two control messages across three contexts. A significant number of participants interpreted the "Add a tip" message as conveying both descriptive and injunctive norms. Notably, participants interacting within the Fiverr context were most likely to perceive this message as an injunctive norm. Figure B.27 illustrates participants' perceptions of descriptive messages across three contexts. Regardless of the context, most participants perceived the descriptive messages as presenting either descriptive norms or a combination of both descriptive and injunctive norms. Lastly, Figure B.28 depicts how participants perceive each injunctive message across three contexts. Injunctive-norm messages were largely perceived as injunctive or as conveying both descriptive and injunctive norms. In conclusion, participants' interpretation of the messages aligns well with our intended design and effectively communicates norms across different platforms.

2.7 Discussion

We study voluntary, post-transaction gratuities in a large marketplace for freelance digital services. First, we document key facts of online tipping behavior, showing tipping is strongly

associated with buyer region and the service category. Our findings suggest buyer characteristics explain tipping behavior more effectively than seller factors, as shown by our regression analysis.

We report results from our initial pilot field experiment demonstrating an injunctive-norm message increases tipping. However, two reciprocity messages do not seem to change tipping behavior. Furthermore, the norms treatment increases buyers' spending without any apparent reduction in platform usage. After the first pilot experiment, removing the norms message reduced the tipping behavior of treated buyers, whereas the subsequent platform-wide adoption of the norms message corresponded with a significant increase in tipping rates.

A second field experiment run on the app confirms the positive effect of an injunctive-norm message on the tipping rate. We also explore the heterogeneity in the effects of norms, discovering that the impact of norm messages on new buyers correlates positively with the buyer satisfaction rating but does not vary with buyers' nationality. Interestingly, we did not detect any indirect impact on seller behavior, suggesting sellers behave consistently, irrespective of tipping rates.

The survey experiment further reveals injunctive-norm messages and descriptive-norm messages with specific tipping numbers significantly influence tipping likelihood.

Taken together, the results suggest platform messages about tipping norms can affect voluntary payments across multiple channels, such as web and app. However, the effect sizes are limited. The effect size is larger for new users than for returning users, and estimating precisely is difficult, even with comprehensive data from the platform. To encourage tipping and further shape norms, platforms should tailor their messaging strategies to align with their established tipping culture, potentially utilizing injunctive-norm messages to boost tipping where tipping is less common.

2.7.1 Implications

The most important implication of our research is the potential for online platforms to strategically use diverse messaging types to augment buyers' voluntary payments, based on the

platform's existing tipping culture. Digital marketplace designers, economists, and analysts should carefully predict potential market behavior effects before implementing designs and validate these predictions afterward. Our study illustrates how normative messages can motivate online tipping and shows strategic tip-message placement can enhance tipping frequency and potentially foster long-term tipping habits.

However, our findings also underscore 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 than new buyers. The result in section 2.4.3 highlights the importance of creating first impressions and calls for careful, well-vetted original designs. And the result in section **??** highlights the importance of using different messages according to tipping situations.

Although we found seller behavior is not indirectly affected by treatment or by tipping, the results do suggest high-quality service delivery by sellers, coupled with well-designed tip messages, encourages more frequent tipping by buyers.

Furthermore, 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 unobserved buyer and seller characteristics are strong predictors of tipping and that pre-treatment controls can help 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.

Finally, as digital platforms become pervasive, understanding the dynamic of online transactions, including tipping, becomes more critical. Platforms could get insights from our research to create guidelines related to financial transactions and user engagement.

2.7.2 Limitations and Future Research

Although our online experiment is designed to assess the generalizability of our findings, how directly applicable these results might be to real-world situations remains unclear. To better validate our findings, field experiments should be conducted based on our laboratory experiment results. Furthermore, the context of the field experiment introduces several limiting factors that might affect the broader applicability of our results:

- The business-to-business nature of most transactions on Fiverr might feature different tipping motivations than business-to-consumer contexts.
- The long-term effects of the tip-request messages are difficult to discern in the available data. In particular, sellers might need more time to observe, understand, and adjust to the new tipping equilibrium.
- The platform always asks for ratings before tips. The results might be different if the tip request precedes the rating request (Chen et al., 2023).

The dynamic and heterogeneous nature of platforms that facilitate interactions between diverse groups necessitates a thorough exploration of how marketplace design impacts indirect network externalities. Questions such as how increased buyer tipping influences seller retention and service efforts, or whether the platform should tailor tip requests to specific order categories or buyer types, need addressing. In addition, the decision to allow sellers to customize their tip message can have significant implications on market efficiency and seller effort.

Exploring various tipping interventions is a promising direction. For instance, alleviating the opaque nature of digital tipping could have potential benefits. Interventions designed to appeal to service/esteem and altruistic tipping motivations, such as providing social proof through aggregate tipping statistics or seller-specific tips, could be effective.

Studying heterogeneous treatments might offer insightful results. Could the optimal tipping message change based on buyer cohort or transaction frequency? How might buyers

react to reminders about their own tipping behavior? Designing adaptive policies that efficiently explore, learn, and exploit information about effective messages could maximize tipping, service quality, and repeat patronage, considering potential effects on user satisfaction, usage, service quality, and pricing.

A notable limitation of this research is its inability to discern off-platform seller characteristics. Changes in design might affect full-time gig workers differently than occasional gig workers, and their implications might be more pronounced for low-income workers than high-income workers. A critical yet unexplored direction is understanding how platform design influences agents of different types concerning overall revenue and competition goals.

2.7.3 Conclusion

Digital marketplaces present unparalleled opportunities to study and influence participant decisions. Our research demonstrates that an injunctive message regarding tipping norms could enhance buyers' voluntary post-transaction payments to sellers, with the most prominent effects observed among new buyers at the initial treatment. This intervention did not appear to significantly alter subsequent platform usage or seller behavior. Additionally, our analysis shows descriptive-norm messages with quantifiable factors can significantly enhance tipping likelihood, suggesting platform messages need to be tailored to fit their established tipping culture. We hope this research will provide valuable insights for future platform design improvement, quantifying the effects of design changes on agent behavior and ultimately fostering more efficient rewards for seller efforts.

ACKNOWLEDGEMENTS

We thank the online platform company for giving access to the individual-level transaction and tipping data. We would like to thank you other Rady faculty members for providing feedback on this work.

Chapter 2 contains unpublished material coauthored with On Amir and Kenneth C. Wilbur. The dissertation author served as the lead researcher, performing the data collection, data analysis, online experiment design, and estimation. Additionally, the dissertation author wrote the majority of the manuscript, including drafting the initial versions and incorporating feedback from coauthors.

Appendix A

for Chapter 1

A.1 Details of Merchant Selection

We focus on 'digital fulfillment' merchants that operate entirely online, selecting a subset of categories such as AI tools, Business & Software, community platforms, and internet security. Table A.1 shows the number of merchants, the number of OpenAI adopters who used any service at each subcategory, and the market shares of the top 3 companies based on user interaction for each subcategory.

Initially, we considered a broader range of categories but found it challenging to link changes in consumption patterns post-GPT adoption for some categories. Consequently, we exclude purely offline merchants and those that provide service via both offline and online, which might have geographical variations in service quality. As a result, we focus on 16 subcategories for our main analysis, enabling a clearer understanding of consumer purchase behavior changes after adopting the new technology.

Table A.1. Brands and market share information across 16 selected subcategories

Subcategory	Number of Brands	Unique consumers	Market Share	Largest Brands (2022 Market Share)
AI Tools	26	10,415	15.0%	grammarly (9.10%), midjourney (2.71%), jasper (1.13%)
Administrative Tools	8	2,590	3.72%	calendly (1.75%), evernote (0.81%), efax (0.71%)
Broker-Dealers	23	23,430	33.6%	coinbase (15.12%), robinhood (11.39%), fidelity_investments (3.91%)
Business & Software	77	50,626	72.7%	microsoft (42.29%), adobe (23.38%), google_one (12.94%)
Community Platforms	25	25,066	36.0%	patreon (16.91%), canva (9.40%), fiverr (7.15%)
E-Books	6	26,801	38.5%	audible (28.15%), kindle (20.59%), epic (0.28%)
Education Resources	49	23,659	34.0%	chegg (12.49%), udemy (7.49%), course$_{h}ero$(5.05%)
Gaming	74	51,127	73.4%	google_play (33.82%), steam_games (30.19%), nintendo (25.91%)
Internet Security	16	11,376	16.3%	norton (4.38%), expressvpn (3.87%), mcafee (3.25%)
Music Streaming & Audio	13	52,907	76.0%	itunes (62.20%), spotify (35.46%), siriusxm (11.67%)
News & Print Media	17	9,031	13.0%	ny_times (7.91%), wsj (2.72%), washington_post (1.63%)
Online Gambling	21	4,485	6.44%	fanduel (3.20%), draftkings (1.41%), betmgm (1.10%)
Online Travel Agency	20	13,702	19.7%	expedia (9.19%), priceline (4.63%), viator (2.32%)
Social Media	28	23,881	34.3%	onlyfans (11.20%), facebook_other (9.93%), discord (8.64%)
Video Streaming	45	49,811	71.5%	netflix (39.31%), hulu (27.27%), disney_plus (19.11%)
Others	10	2,036	2.92%	headspace (0.85%), upwork_gig_pay (0.59%), calm (0.58%)

Notes: The above numbers are derived from combined panel data of OpenAI adopters.

A.2 Descriptive Statistics

Figure A.1 shows the monthly customer counts from the top 10 states with the highest adoption of ChatGPT, highlighting regional variations in adoption rates. The geographic representation varies slightly between the panels. For example, in the shorter panel, Pennsylvania, California, and Maryland are the top three states with the highest number of OpenAI adopters. In the longer panel, the leading states are Texas, California, and Florida. Despite these differences, the increasing adoption trend is consistent across all states and both panels.

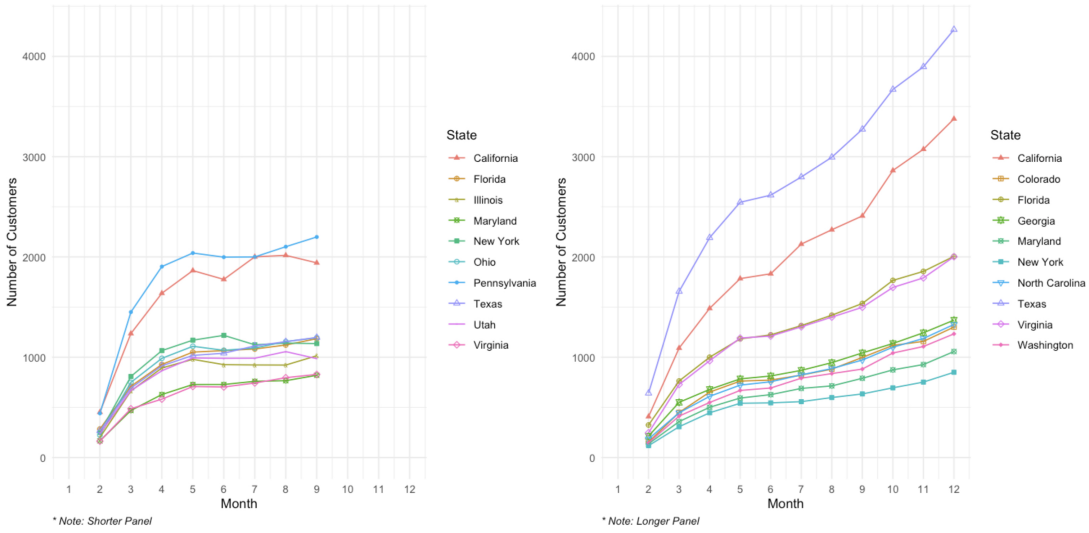


Figure A.1. Monthly Customers in Top 10 States

Next, we examine the distribution of total card spending across adoption cohorts. Figure A.2 shows that the distributions are right-skewed, indicating while most adoption cohorts spend smaller amounts, a few high-spenders significantly contribute to the total. The total spending patterns are consistent across both panels, suggesting a general similarity in spending patterns among OpenAI consumers, regardless of their month of adoption.

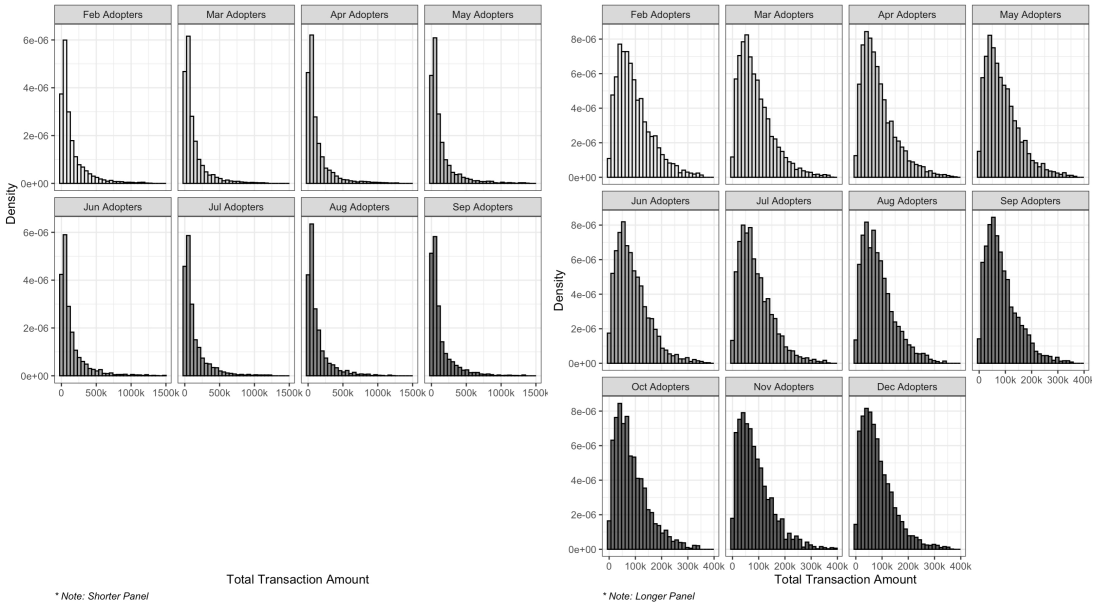


Figure A.2. Total Card Spending by Adoption Cohort

A.3 Matching Process

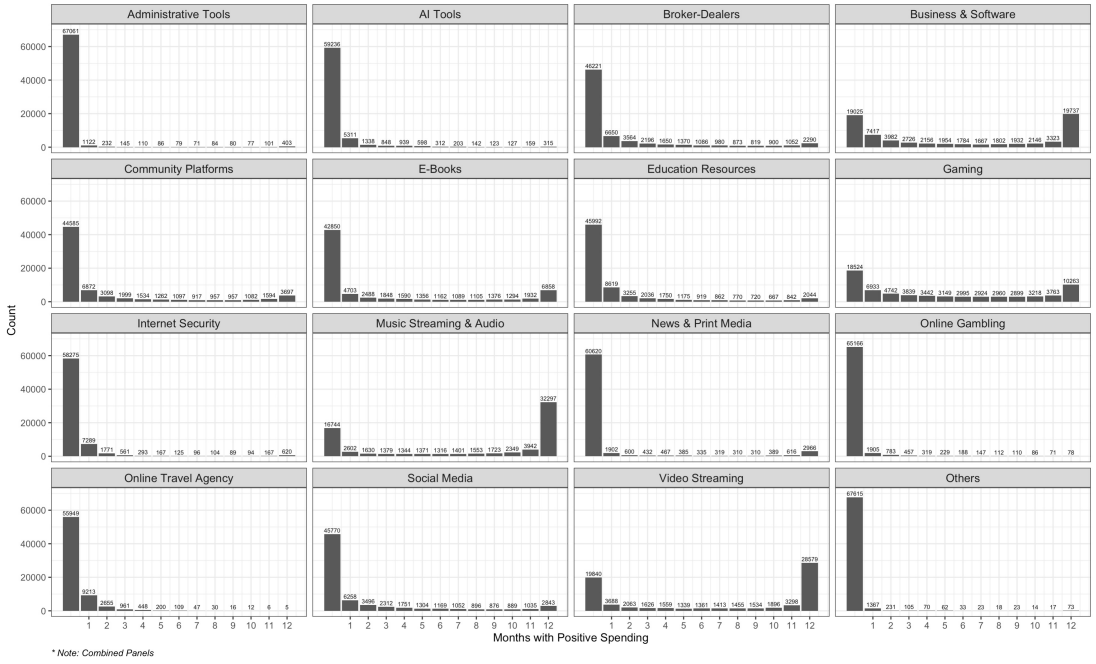
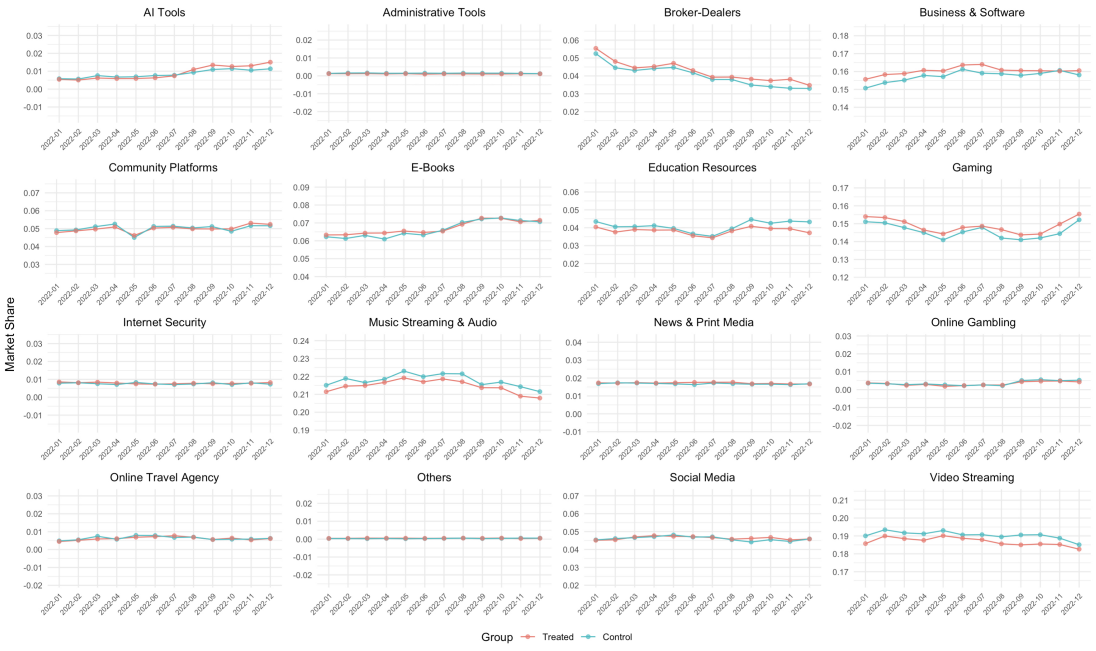


Figure A.3. Distribution of Months with Positive Spending across 16 Subcategories

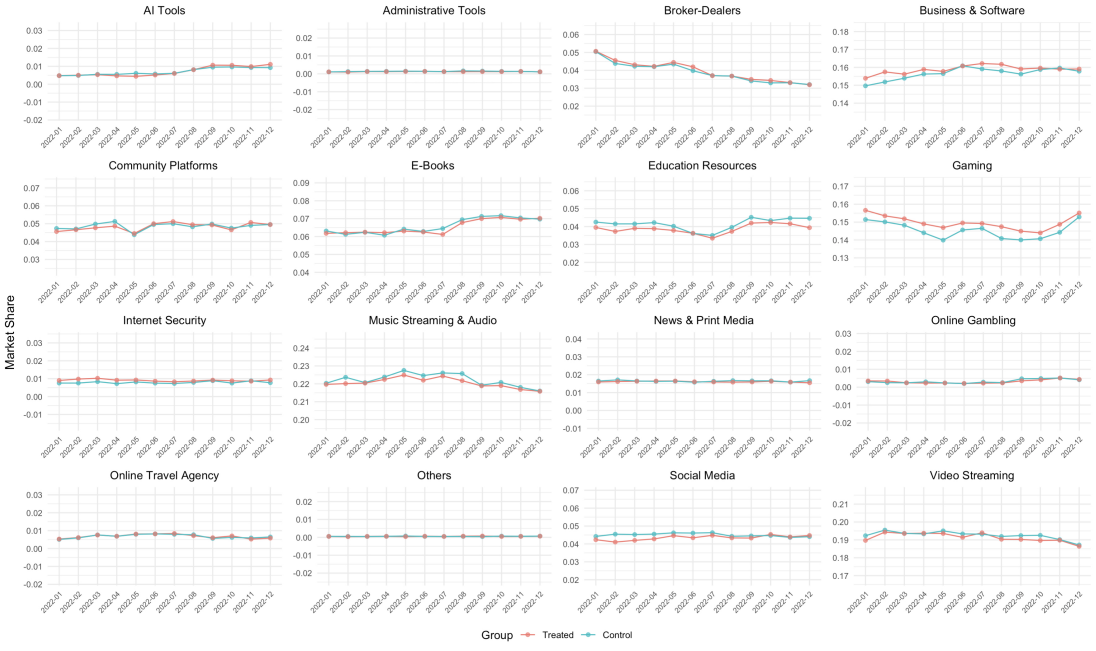
Figure A.3 shows the distribution of months with positive spending across 16 different subcategories for treated consumers. Each subcategory represents a different type of expenditure, and the histogram bars indicate the count of consumers who had positive spending in the respective number of months during 2022.

The distributions vary depending on the subcategory. For example, most cards did not pay for services related to Administrative Tools or Online Gambling. However, there are cards that used services for the entire year in categories such as Video Streaming and Music Streaming & audio. This distribution is used for the coarsened exact matching process to define comparability between the treated and control groups.



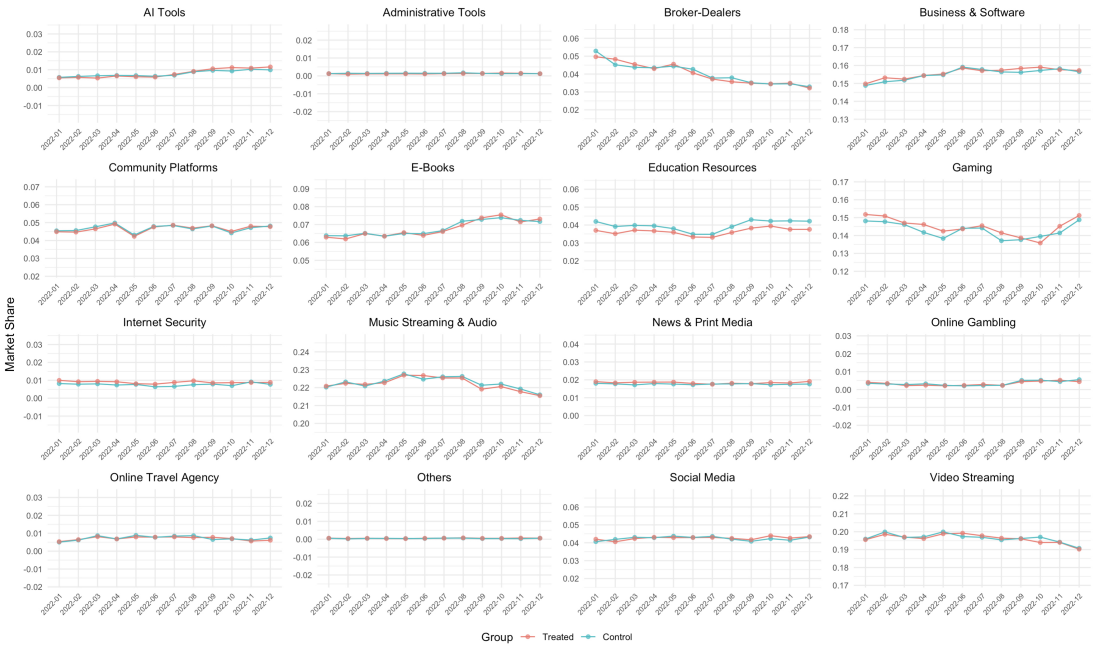
* *Note: Combined Panels*

Figure A.4. Pre-Treatment Trend in Matched Data with March Cohort



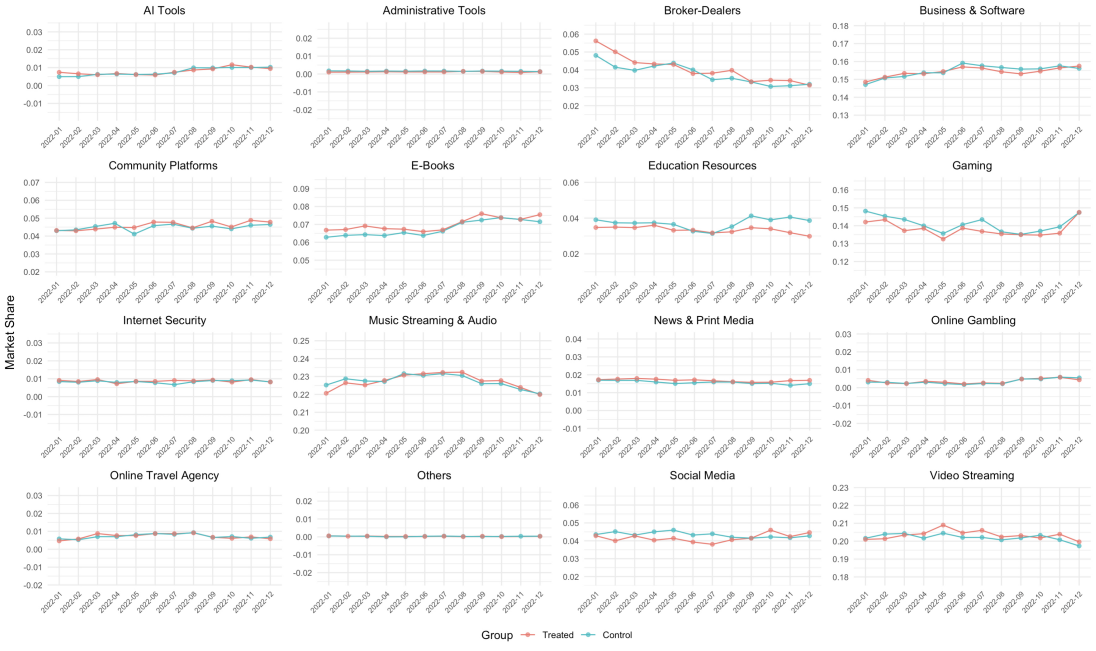
* *Note: Combined Panels*

Figure A.5. Pre-Treatment Trend in Matched Data with April Cohort



* *Note: Combined Panels*

Figure A.6. Pre-Treatment Trend in Matched Data with May Cohort



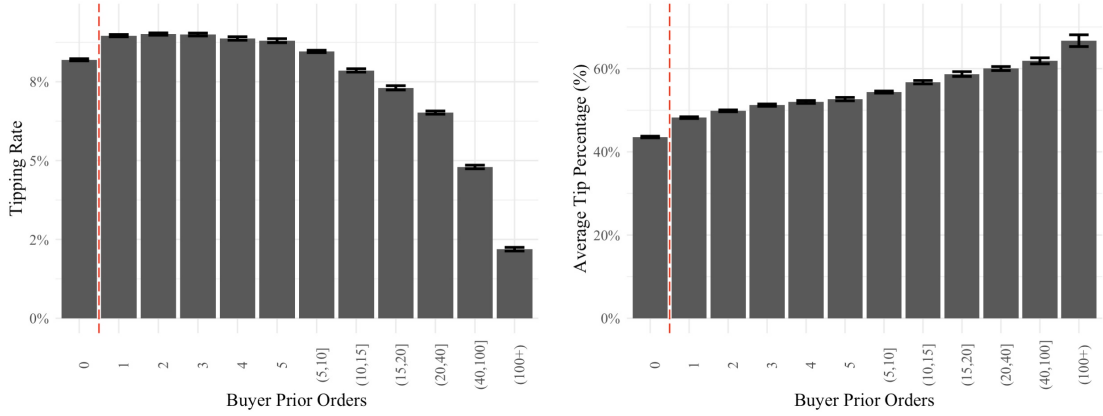
* *Note: Combined Panels*

Figure A.7. Pre-Treatment Trend in Matched Data with June Cohort

Appendix B

for Chapter 2

B.1 Additional Data Descriptives



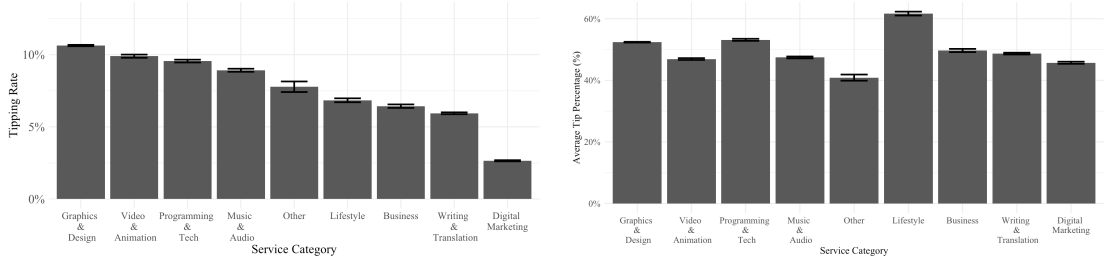
(a) Tipping Rate **(b)** Average Tip Percentage

Figure B.1. Tipping by Buyer Prior Orders

The variable Buyer Prior Orders counts buyer transactions going back to the first transaction in 2019. Figure B.1 classifies pre-treatment orders into comparably sized bins, showing that Tipping Rate exhibits an inverted "U" shape in buyer prior orders but is positively correlated with Average Tip Percentage.

B.2 Default Tips Conditions

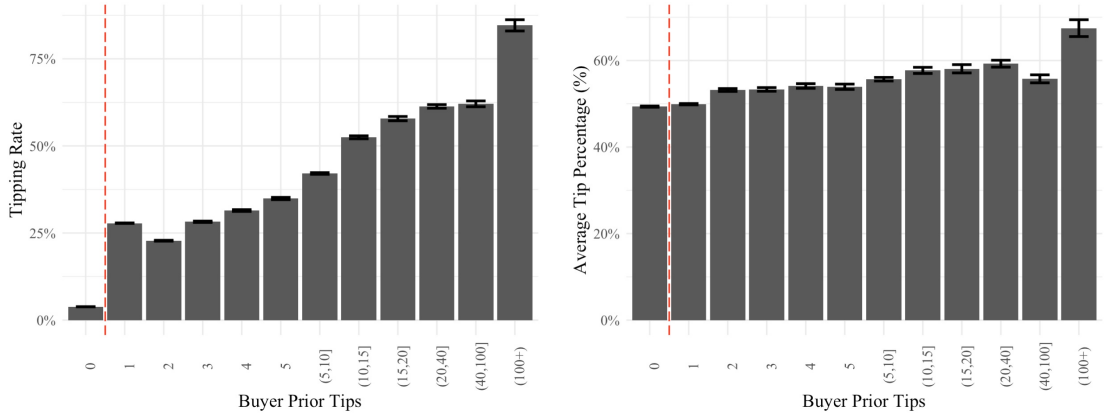
Figure B.7 shows how default tip usage changed with order prices under $25 in the pre-experiment data. Over 90% of consumers used default tips rather than custom tips for gigs under $25, with about 70% of those choosing the $5 default tip, about 20% choosing the $10



(a) **(b)**

Figure B.2. Tipping and Service Category

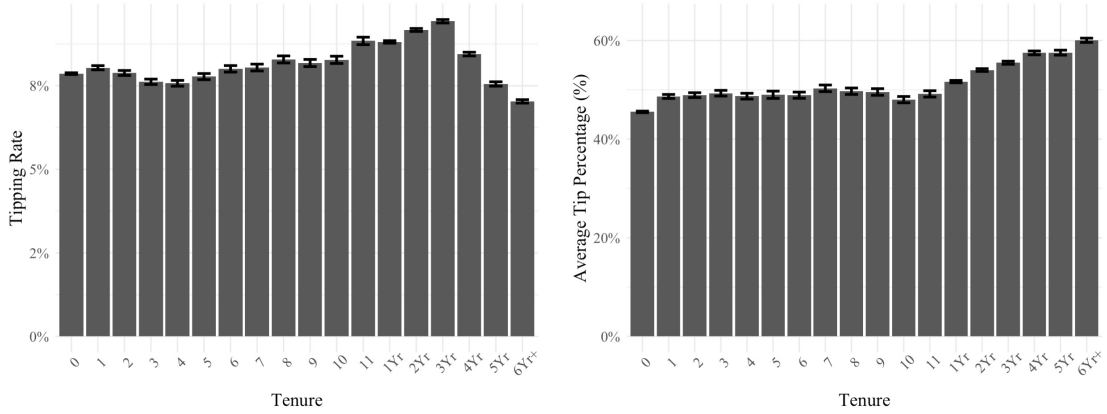
Figure B.2(a) shows that a nearly four-fold spread in Tipping Rate, ranging from 2.7% of Digital Marketing gigs up to 10.6% in the Graphics & Design category. Average Tip Percentage shows less variation across categories, relative to its mean, than tipping rate.



(a) Tipping Rate **(b)** Average Tip Percentage

Figure B.3. Tipping by Buyer Prior Tips

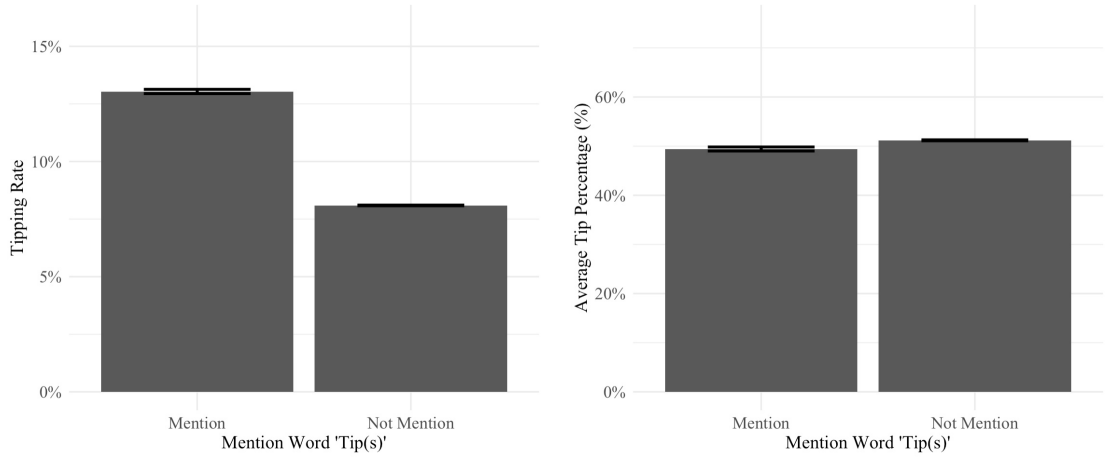
The variable Buyer Prior Tips counts the number of tips paid by each buyer before a transaction. Figure B.3(a) shows that buyers without any prior tips tipped on just 1.6% of orders, rising to 31.3% among buyers with one prior tip, up to 84.8% among buyers with 100 prior tips. The figure does not distinguish state dependence from heterogeneity–i.e., learned tipping behavior from individual tipping propensity. However, Average Tip Percentage varies much less with buyer prior tips than tipping rate.



(a) Tipping Rate **(b)** Average Tip Percentage

Figure B.4. Tipping by Buyer Tenure

Figure B.4 shows how tipping changes with the time elapsed since each buyer's first transaction date. Tipping generally increases with buyer tenure up to a peak at 3 years, then declines among older buyer cohorts. We believe that is because Fiverr used low prices to attract frugal buyers initially, then moved upmarket more recently. Average Tip Percentage is nearly flat among newer buyers but then rises with buyer tenure among older cohorts.



(a) Tipping Rate **(b)** Average Tip Percentage

Figure B.5. Tipping by Seller Mention "Tip" or "Tips"

Seller messages to buyers accompanied completed orders in 96.5% of transactions. 3.1% of seller messages mentioned the words "tip" or "tips." Figure B.5 shows that Tipping Rate increased with seller tip mentions from 8.1% to 13%. However, Average Tip Percentage was nearly the same regardless of seller tip mentions.

Table B.1. Variables and Summary Statistics in Pre-Treatment Transaction Data

	Mean	Standard Deviation
Buyer Characteristics X^B		
Buyer Prior Orders	10.51	42.89
Buyer Prior Ratings	9.03	31.22
Buyer Average Rating (Given)	4.96	0.15
Buyer Prior Tips	0.51	2.98
Buyer Tenure (Months)	20.97	25.21
User-defined Female	1.20%	
User-defined Male	3.11%	
Transaction Characteristics X^T		
Buyer-Seller Repeat Indicator	29.93%	
Count of Buyer-Seller Repeat Orders	1.78	9.21
Buyer Rating = 3	0.53%	
Buyer Rating = 4	3.34%	
Buyer Rating = 5	56.55%	
Price	$ 36.86	$ 82.29
Seller Tip Mention	2.98%	
Seller Characteristics X^S		
Seller Prior Orders	225.1	641.84
Seller Prior Ratings	120.90	337.99
Seller Average Rating (Received)	4.95	0.07

Note: Buyer Region, Seller Region, and Seller Category are also observed but excluded for brevity.

Table B.2. Descriptive Regression without Single Transaction Buyers and Sellers

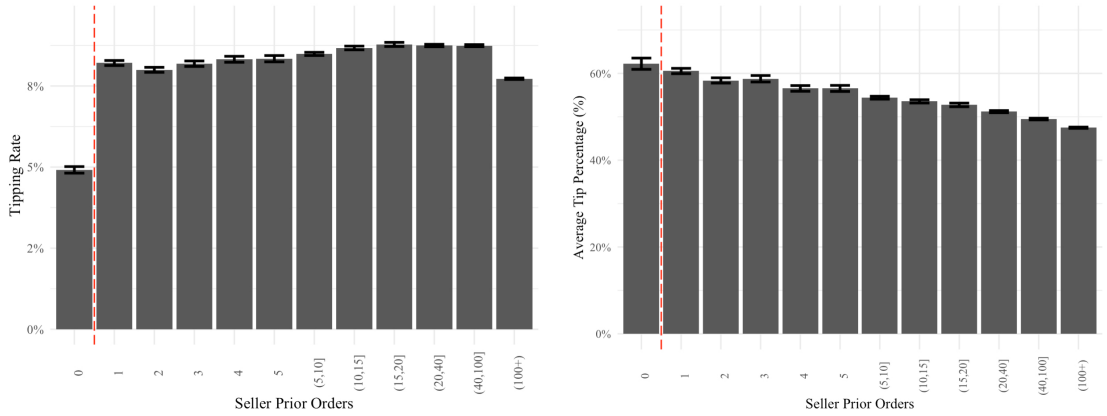
	(1)	(2)	(3)	(4)	(5)
Buyer Prior Orders	0.0001***	0.0001***		0.0001***	
	(0.00000)	(0.00000)		(0.00000)	
Buyer Prior Ratings	−0.009***	−0.007***		−0.008***	
	(0.001)	(0.001)		(0.001)	
Buyer Average Rating (Given)	0.002***	0.001***		0.001***	
	(0.0002)	(0.0002)		(0.0002)	
Buyer Prior Tips	0.019***	0.018***		0.018***	
	(0.00004)	(0.00004)		(0.00005)	
Buyer Tenure (Month)	0.00002***	−0.0001***		−0.0001***	
	(0.00001)	(0.00001)		(0.00001)	
User-defined Female	0.018***	0.011***		0.010***	
	(0.001)	(0.001)		(0.001)	
User-defined Male	−0.010***	−0.008***		−0.008***	
	(0.001)	(0.001)		(0.001)	
Buyer-Seller Repeat Indicator	−0.001***	−0.0003	0.001***	−0.002***	−0.002***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Count of Buyer-Seller Repeat Orders	−0.001***	−0.001***	0.0001***	−0.0005***	0.00003
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Buyer Rating = 3	0.030***	0.031***	0.019***	0.033***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
4	0.066***	0.068***	0.053***	0.067***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
5	0.127***	0.128***	0.117***	0.125***	0.114***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Price	0.0001***	0.0001***	−0.00002***	0.0001***	−0.00000
	(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)	(0.000)	(0.000)
Seller Tip Mention	0.045***	0.044***	0.038***	0.060***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Seller Prior Orders	−0.00000**	−0.00000	−0.00000***		
	(0.00000)	(0.00000)	(0.00000)		
Seller Prior Ratings	−0.00000	−0.00001***	−0.00000		
	(0.00000)	(0.00000)	(0.00000)		
Seller Average Rating (Received)	0.0003**	−0.0004***	−0.001***		
	(0.0001)	(0.0001)	(0.0001)		
Category FE	Y	Y	Y	Y	Y
Buyer Region FE		Y		Y	
Seller Region FE	Y	Y	Y		
Buyer FE			Y		Y
Seller FE				Y	Y
Observations	3,439,974	3,439,974	3,439,974	3,439,974	3,439,974
R^2	0.115	0.127	0.505	0.185	0.533
Adjusted R^2	0.115	0.127	0.398	0.155	0.407

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

Table B.3. Descriptive Regression with Single Transaction Buyers and Sellers

	(1)	(2)	(3)	(4)	(5)
Buyer Prior Orders	0.0001***	0.0001***		0.0001***	
	(0.00000)	(0.00000)		(0.00000)	
Buyer Prior Ratings	−0.009***	−0.007***		−0.008***	
	(0.001)	(0.001)		(0.001)	
Buyer Average Rating (Given)	0.002***	0.001***		0.001***	
	(0.0002)	(0.0002)		(0.0002)	
Buyer Prior Tips	0.019***	0.018***		0.018***	
	(0.00004)	(0.00004)		(0.00005)	
Buyer Tenure (Month)	0.00004***	−0.0001***		−0.0001***	
	(0.00001)	(0.00001)		(0.00001)	
User-defined Female	0.018***	0.012***		0.010***	
	(0.001)	(0.001)		(0.001)	
User-defined Male	−0.010***	−0.008***		−0.007***	
	(0.001)	(0.001)		(0.001)	
Buyer-Seller Repeat Indicator	−0.002***	−0.001***	0.001***	−0.003***	−0.002***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Count of Buyer-Seller Repeat Orders	−0.001***	−0.001***	0.0001***	−0.0005***	0.00003
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Buyer Rating = 3	0.030***	0.033***	0.020***	0.034***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
4	0.068***	0.071***	0.053***	0.070***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
5	0.130***	0.131***	0.117***	0.128***	0.114***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Price	0.0001***	0.0001***	−0.00002***	0.0001***	−0.00000
	(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)	(0.000)	(0.000)
Seller Tip Mention	0.047***	0.046***	0.038***	0.060***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Seller Prior Orders	−0.00000***	−0.00000***	−0.00000***		
	(0.00000)	(0.00000)	(0.00000)		
Seller Prior Ratings	0.00000	−0.00000	−0.00000		
	(0.00000)	(0.00000)	(0.00000)		
Seller Average Rating (Received)	0.001***	0.0001	−0.0003***		
	(0.0001)	(0.0001)	(0.0001)		
Category FE	Y	Y	Y	Y	Y
Buyer Region FE		Y		Y	
Seller Region FE	Y	Y	Y		
Buyer FE			Y		Y
Seller FE				Y	Y
Observations	4,134,928	4,134,928	4,134,928	4,134,928	4,134,928
R^2	0.108	0.121	0.582	0.177	0.610
Adjusted R^2	0.108	0.121	0.396	0.142	0.404

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



(a) Tipping Rate **(b)** Average Tip Percentage

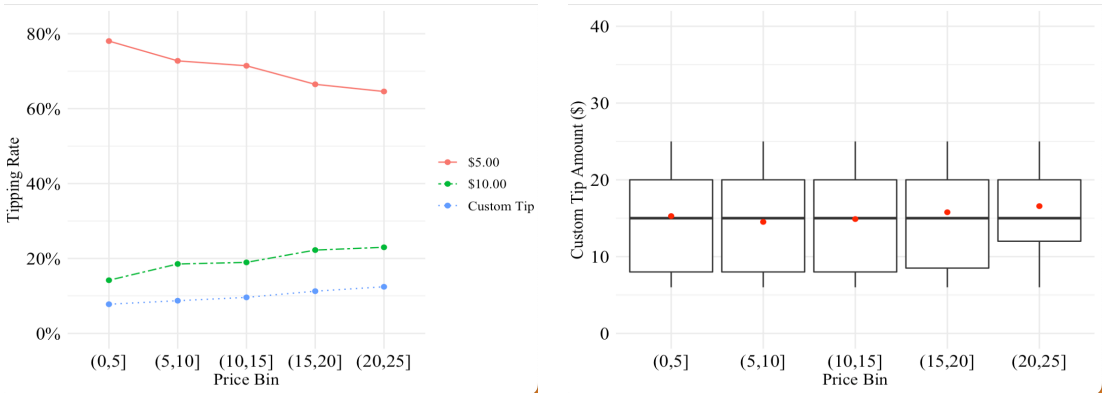
Figure B.6. Tipping By Seller Prior Orders

Figure B.6 shows how tipping related to seller experiences, as measured by seller prior orders delivered since their first transaction in 2019. Buyers tipped first-time sellers on just 4.9% of transactions, rising to 8.2% of transactions by one-time sellers. Average Tip Percentage declines slightly with seller prior orders, possibly due to both variables' correlation with price.

default tip, and the majority of custom tips entered near $15. Therefore, the company viewed adding a $15 default tip option as a natural way to lower consumer costs of entering a larger tip on a low-priced order.

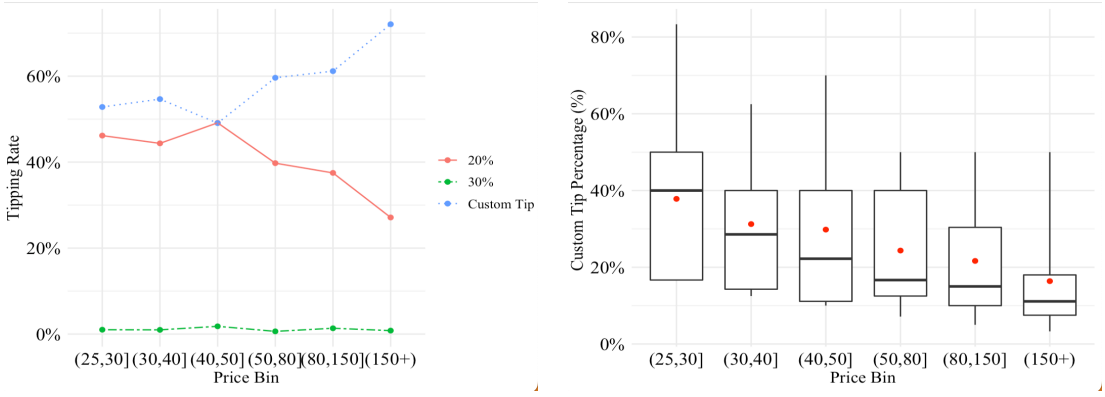
Figure B.8 shows how default tip usage changed with order prices over $25 in the pre-experiment data. About half of tipping consumers chose the 20% default tip, with the other half entering a custom dollar amount, and fewer than 2% of tipping consumers choosing the 30% default tip. Therefore, the platform decided to eliminate the nearly-unused 30% default tip, and it decided to replace it with 15% and 25% default tips. The platform also decided to increase the threshold between dollar and percentage default tips from $25 to $35, in part due to inflation since the original tip menu had been chosen and in part because of the new $15 default tip included at low prices.

We expect that default tips are typically likely to affect behavior, but two factors could limit how much they matter. First, the specific changes in default tips predict how consumers will respond. Imagine we replaced a popular $5 default with a $500 default; the changes would



(a) **(b)**

Figure B.7. Tipping Behavior Analysis by Price Bin: Dollar Default Option



(a) **(b)**

Figure B.8. Tipping Behavior Analysis by Price Bin: Percentage Default Option

be large because virtually no one would use the new $500 default tip, unless they either chose it by mistake or simply failed to notice the change. On the other hand, suppose we replaced a seldom-used 30% default with a 30.1% default; it seems unlikely that a minor change to a seldom-chosen option would have much effect on behavior. Second, the evidence of Chandar et al. (2019), in a similarly infrequently tipped context, showed that default tip manipulations had only limited effects on Uber tips.

In this specific case, because of the testing system's constraint to four treatment arms during the polite field experiment, the platform decided to change the default tips in two of the four conditions, as shown in Figure B.9. The default and Implicit Reciprocity conditions included the original default tips: ($5, $10, Custom) for prices below $25 and (20%, 30%, Custom) for prices above $25. The Reciprocity and Norms conditions included the new default tips of ($5, $10, $15) for prices below $35 and (15%, 20%, 25%) for prices above $35.

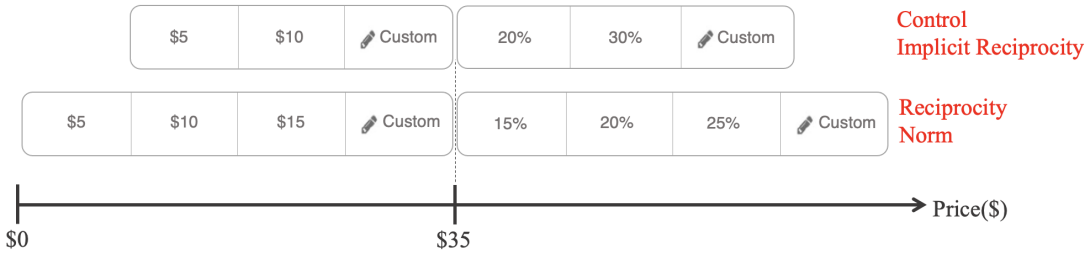


Figure B.9. Default Tip Manipulation

The confounding of tip messages and default tips muddles the interpretation of treatment conditions. Fortunately, we are able to exploit two post-test platform design changes to identify the default tips effect from the message effects separately. We also notice large differences between Norms and Reciprocity condition results, despite their common default tips conditions, which are hard to explain theoretically without main effects of treatment messages. Additionally, Fiverr runs a second field experiment on mobile apps without the confounding of tip messages and default tips, allowing for a clearer analysis.

We want to note several important caveats here. An ideal test design would independently manipulate tip message and default tips. The specific default tips effects we find depend on

the specific changes that were tested, which as we explain above, are unlikely to be very large. We refer to experimental treatments by their tip message motivations as shorthand, but we do not intend to diminish the default tips change in two conditions, which we explore further in the post-test analyses and the second field experiment to distinguish them from tip message treatment effects.

B.3 Field Experiment Mechanism Evidence

In this section, we further explore the potential moderating effect of Norms message through two field experiments to understand why this message significantly increases tipping while the two Reciprocity messages do not. We interact treatment effects with potential moderators tied to tipping motivations, informed by both descriptive analysis and relevant literature[1]. The findings below shed some light on when and why 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 in the experimental data.

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

1. *Regional tipping culture: North America vs. other regions*: Existing literature and Key Fact I showed that buyer region tipping cultural norms affect tipping much more than seller region. Given our main analysis showed Norms treatment message increases tipping, we hypothesize that North American buyers may tip more when presented with Norms message.

2. *5-star Satisfaction rating*: Bodvarsson and Gibson (1990) and key fact II indicate that satisfaction levels impact tipping. Buyers are more likely to tip when they rate a service

[1]Although there are numerous possible moderators such as Category, Price, we have chosen three moderators related to tipping motivations the most and have the most significant impact on tipping behavior.

with a 5-star rating. Thus, we hypothesize that recently satisfied buyers might respond more to tipping messages, examined by interacting the message treatment with 5-star buyer satisfaction ratings.

3. *Number of transactions between same buyer and same seller*: In traditional contexts, repeat buyers tend to tip more frequently than to get a better future service. However, the correlation between the frequency of patronage and tipping might be elusive in purely digital services with no direct buyer-seller interaction. We hypothesize that frequent buyers may be more likely to tip, especially if they've interacted with specific sellers multiple times. We examine this by interacting the message treatment with a binary indicator denoting prior interactions between the same buyer and seller.

Table B.4 shows how treatment effects vary with the three moderators for new buyers using services via web. Firstly, while the interactions between North American buyers and message treatments are not statistically significant, Norms main effect is positive. This suggests that local tipping culture doesn't primarily influence how buyers respond to treatment messages for online transactions. Secondly, the interaction between 5-star ratings and Norms treatment is positive and significant, indicating that buyer satisfaction is an important factor in predicting compliance with the tipping norm. Implicit Reciprocity also increased tipping among satisfied new buyers, but this effect was only significant at their first exposure and not in the larger pooled transaction sample. This highlights that new buyer satisfaction is predictive of tipping responses to treatment messages. Lastly, repeat interactions between buyers and sellers significantly boosted Norms effect on tipping during the third transaction. However, this effect was positive but not significant in the pooled transaction sample or when buyers transacted with the same seller a second time. This implies that Norms and Reciprocity messages don't strategically influence repeat buyers to tip more in web-based services.

We also check whether treatment effects on tip amount and tip percentage also vary with three moderators for web-based new buyers. Table B.5 and B.6 shows that the interaction

between 5-star ratings and Norm treatment is positive and significant, suggesting if buyers are satisfied with service quality, it also leads new buyers to not only tip more frequently but also tip a higher amount and percentage than those less satisfied with the service.

In both regional tipping culture, and the record of previous transactions with the same seller, a positive and significant effect of Norm message was observed on tip amount, but not significantly so on tip percentage. Norms increased tip percentage among North American new buyers, but only upon first exposure. Buyers who transacted with the same seller did not tip a higher percentage across all specifications.

Table B.7 show how the treatment effects change with the three moderators for app-based new users. In contrast to web-based new users in the first experiment, none of the three moderators increase the effect of Norm messages. However, the main Norm effects remain positive and significant. The findings might imply that user engagement through apps is not as high as it is on the web. This lower engagement may result in moderators not being the primary influence on buyers' responsiveness to messages. For example, when a buyer receives the final work from the seller through an app, accessing the quality of the final work might be challenging due to the smaller screen size. This difficulty may prevent potential moderators from enhancing Norms message effect [2].

Additionally, we check how treatment effects change depending on the three moderators among web-based and app-based repeat buyers treated in the second field experiment. Table B.8 and Table B.8 show that the three-way interactions between Norm treatment and the tipping moderators are not significant. Also, these three moderators did not increase the effect of Norm message on either tip amount or tip percentage in both web and app transactions.

[2]We also check how treatment effects on tip amount and tip percentage change varying with the three moderators, but we could not find any significant interaction effect

Table B.4. New Buyer Tipping Mechanism with First Field Experiment data

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*		
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Implicit_Reciprocity	0.007	0.009	−0.004	0.013	−0.005	−0.037	−0.004	0.002	−0.002	−0.011	0.002
	(0.008)	(0.012)	(0.011)	(0.016)	(0.008)	(0.023)	(0.010)	(0.014)	(0.008)	(0.012)	(0.016)
Reciprocity	−0.001	−0.011	−0.006	0.023	−0.008	−0.017	−0.011	−0.003	0.002	0.003	0.013
	(0.007)	(0.012)	(0.011)	(0.016)	(0.007)	(0.025)	(0.009)	(0.013)	(0.008)	(0.012)	(0.017)
Norms	0.032***	0.051***	0.020	0.036**	0.010	−0.009	0.018	0.007	0.041***	0.032**	0.009
	(0.009)	(0.014)	(0.012)	(0.017)	(0.008)	(0.025)	(0.012)	(0.014)	(0.009)	(0.013)	(0.017)
Implicit_Reciprocity:I_North_America	−0.017	0.003	−0.010	−0.021							
	(0.016)	(0.022)	(0.020)	(0.028)							
Reciprocity:I_North_America	0.009	0.031	0.013	−0.025							
	(0.016)	(0.022)	(0.020)	(0.029)							
Norms:I_North_America	0.032	0.039	0.042	0.0002							
	(0.018)	(0.024)	(0.022)	(0.030)							
Implicit_Reciprocity:I_5stars					0.005	0.055**	−0.008	0.0001			
					(0.012)	(0.026)	(0.018)	(0.025)			
Reciprocity:I_5stars					0.015	0.025	0.018	0.021			
					(0.012)	(0.027)	(0.017)	(0.025)			
Norms:I_5stars					0.052***	0.091***	0.035	0.047			
					(0.013)	(0.029)	(0.020)	(0.026)			
Implicit_Reciprocity:I_previous									−0.001	0.005	0.001
									(0.014)	(0.022)	(0.032)
Reciprocity:I_previous									0.004	−0.007	−0.007
									(0.014)	(0.022)	(0.032)
Norms:I_previous									0.029	0.027	0.088**
									(0.015)	(0.024)	(0.035)
Buyer Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller Char.	Y	Y	Y	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	25,215	7,258	3,545
Number of Buyers	7,880	7,880	7,258	3,545	7,880	7,880	7,258	3,545	7,880	7,258	3,545
R^2	0.092	0.077	0.102	0.102	0.092	0.077	0.102	0.102	0.091	0.101	0.105
Adjusted R^2	0.090	0.073	0.097	0.092	0.090	0.073	0.097	0.092	0.090	0.096	0.094

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

Table B.5. New Buyer Tip Amount Mechanism with First Field Experiment data

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*		
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Implicit_Reciprocity	−0.053	−0.089	−0.208	0.157	−0.065	−0.072	−0.104	−0.102	−0.104	−0.251	−0.013
	(0.122)	(0.176)	(0.186)	(0.224)	(0.120)	(0.320)	(0.173)	(0.228)	(0.117)	(0.191)	(0.257)
Reciprocity	−0.043	−0.247	0.058	0.347	−0.171	−0.194	−0.133	−0.186	−0.087	−0.107	0.124
	(0.145)	(0.165)	(0.284)	(0.270)	(0.099)	(0.256)	(0.139)	(0.182)	(0.111)	(0.188)	(0.248)
Norms	0.163	0.347	0.005	0.388	−0.038	−0.136	−0.157	0.048	0.425***	0.295	0.012
	(0.128)	(0.193)	(0.189)	(0.231)	(0.114)	(0.269)	(0.180)	(0.266)	(0.131)	(0.225)	(0.241)
Implicit_Reciprocity:I_North_America	−0.040	0.215	−0.056	−0.142							
	(0.216)	(0.298)	(0.310)	(0.432)							
Reciprocity:I_North_America	0.167	0.458	−0.050	−0.082							
	(0.230)	(0.281)	(0.386)	(0.472)							
Norms:I_North_America	0.800***	0.984***	0.914**	0.311							
	(0.262)	(0.370)	(0.403)	(0.472)							
Implicit_Reciprocity:I_5stars					−0.017	0.103	−0.215	0.286			
					(0.169)	(0.360)	(0.280)	(0.393)			
Reciprocity:I_5stars					0.290	0.204	0.251	0.762			
					(0.175)	(0.300)	(0.317)	(0.401)			
Norms:I_5stars					0.834***	1.118***	0.964***	0.779			
					(0.196)	(0.344)	(0.356)	(0.445)			
Implicit_Reciprocity:I_previous									0.125	0.031	0.272
									(0.219)	(0.319)	(0.498)
Reciprocity:I_previous									0.560	0.384	0.544
									(0.306)	(0.460)	(0.553)
Norms:I_previous									0.634**	0.506	1.670***
									(0.285)	(0.468)	(0.617)
Buyer Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller Char.	Y	Y	Y	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	25,215	7,258	3,545
Number of Buyers	7,880	7,880	7,258	3,545	7,880	7,880	7,258	3,545	7,880	7,258	3,545
R^2	0.095	0.114	0.116	0.094	0.094	0.114	0.116	0.094	0.094	0.115	0.096
Adjusted R^2	0.093	0.111	0.111	0.083	0.093	0.110	0.111	0.084	0.092	0.110	0.086

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

Table B.6. New Buyer Tip Percentage Mechanism with First Field Experiment data

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*		
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Implicit_Reciprocity	−0.146	0.358	−0.693	0.166	0.516	2.040	−0.096	0.961	0.251	0.005	−1.071
	(0.557)	(0.888)	(0.966)	(1.105)	(0.721)	(2.295)	(0.545)	(2.570)	(0.534)	(1.000)	(1.500)
Reciprocity	−0.248	−0.510	−0.395	0.775	−0.420	−0.264	−0.330	−2.227	−0.098	−0.492	−1.378
	(0.560)	(0.791)	(1.026)	(1.233)	(0.535)	(1.057)	(0.507)	(2.030)	(0.485)	(0.816)	(1.426)
Norms	1.260	1.461	0.929	1.747	0.444	0.488	1.913	−1.458	1.961***	1.436	−0.968
	(0.649)	(0.913)	(1.058)	(1.296)	(0.641)	(1.152)	(1.020)	(1.963)	(0.555)	(0.929)	(1.446)
Implicit_Reciprocity:I_North_America	0.851	1.784	2.196	−0.779							
	(1.109)	(1.481)	(1.550)	(2.626)							
Reciprocity:I_North_America	0.217	1.453	0.765	−4.146							
	(0.971)	(1.344)	(1.390)	(2.363)							
Norms:I_North_America	1.780	3.526**	2.842	−2.812							
	(1.133)	(1.596)	(1.558)	(2.376)							
Implicit_Reciprocity:I_5stars					−0.341	−0.912	0.772	−1.824			
					(0.923)	(2.451)	(1.293)	(3.048)			
Reciprocity:I_5stars					0.377	0.555	0.481	1.451			
					(0.783)	(1.299)	(1.138)	(2.574)			
Norms:I_5stars					2.364***	3.137**	0.689	2.814			
					(0.892)	(1.463)	(1.451)	(2.541)			
Implicit_Reciprocity:I_previous									0.109	1.099	2.564
									(1.225)	(1.605)	(2.975)
Reciprocity:I_previous									−0.206	1.318	0.303
									(1.057)	(1.477)	(2.493)
Norms:I_previous									0.876	2.685	4.009
									(1.176)	(1.777)	(2.598)
Buyer Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller Char.	Y	Y	Y	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	25,215	7,258	3,545
Number of Buyers	7,880	7,880	7,258	3,545	7,880	7,880	7,258	3,545	7,880	7,258	3,545
R^2	0.039	0.038	0.045	0.047	0.040	0.038	0.045	0.047	0.039	0.045	0.047
Adjusted R^2	0.038	0.034	0.040	0.036	0.038	0.034	0.039	0.036	0.038	0.040	0.036

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

Table B.7. New Buyer Tipping Mechanism with Second Field Experiment data

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*		
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Norms	0.013***	0.018***	0.002	0.012	0.009***	0.012***	0.010**	0.002	0.016***	0.012	0.021**
	(0.005)	(0.005)	(0.008)	(0.012)	(0.003)	(0.004)	(0.005)	(0.006)	(0.004)	(0.007)	(0.010)
Norms:I_North_America	0.005	−0.001	0.019	0.011							
	(0.007)	(0.008)	(0.013)	(0.018)							
Norms:I_5stars					0.011	0.010	0.007	0.029			
					(0.006)	(0.007)	(0.012)	(0.017)			
Norms:I_previous									0.006	0.024	−0.017
									(0.013)	(0.021)	(0.027)
Buyer Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	36,073	19,749	7,216	3,402	36,073	19,749	7,216	3,402	36,073	7,216	3,402
Number of Buyers	19,749	19,749	7,216	3,402	19,749	19,749	7,216	3,402	19,749	7,216	3,402
R^2	0.076	0.075	0.079	0.073	0.076	0.075	0.078	0.073	0.076	0.079	0.073
Adjusted R^2	0.075	0.074	0.074	0.063	0.076	0.074	0.074	0.064	0.075	0.074	0.063

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

Table B.8. Repeat Buyer Tipping Mechanism with First Field Experiment data

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*			
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	First	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Implicit_Reciprocity:Post	−0.001	−0.00001	−0.010	−0.002	−0.003	−0.001	−0.007	0.004	0.004	0.008	−0.009	−0.001
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.014)	(0.006)	(0.007)	(0.003)	(0.006)	(0.006)	(0.007)
Reciprocity:Post	0.005	0.005	0.006	0.002	−0.002	0.0005	−0.007	−0.004	0.002	−0.002	0.003	−0.004
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.014)	(0.006)	(0.007)	(0.003)	(0.006)	(0.006)	(0.007)
Norms:Post	0.005	0.017***	0.007	0.002	−0.001	0.002	−0.006	0.001	0.008**	0.028***	0.003	0.003
	(0.003)	(0.005)	(0.005)	(0.006)	(0.003)	(0.013)	(0.006)	(0.007)	(0.003)	(0.006)	(0.006)	(0.007)
Implicit_Reciprocity:I_North_America	−0.008	−0.008	−0.008	−0.006								
	(0.007)	(0.008)	(0.008)	(0.008)								
Reciprocity:I_North_America	−0.004	−0.001	−0.004	−0.005								
	(0.007)	(0.008)	(0.008)	(0.007)								
Norms:I_North_America	0.00002	−0.005	−0.003	0.005								
	(0.007)	(0.008)	(0.008)	(0.007)								
Implicit_Reciprocity:Post:I_North_America	0.001	0.004	0.006	−0.008								
	(0.004)	(0.008)	(0.008)	(0.010)								
Reciprocity:Post:I_North_America	−0.005	0.002	−0.007	−0.013								
	(0.005)	(0.008)	(0.008)	(0.010)								
Norms:Post:I_North_America	−0.0002	0.006	−0.004	−0.012								
	(0.004)	(0.008)	(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)				
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_previous									−0.002	−0.001	−0.002	0.001
									(0.002)	(0.002)	(0.002)	(0.002)
Reciprocity:I_previous									−0.002	−0.002	−0.002	0.001
									(0.002)	(0.002)	(0.002)	(0.002)
Norms:I_previous									0.001	0.0003	0.001	0.002
									(0.002)	(0.002)	(0.002)	(0.002)
Implicit_Reciprocity:Post:I_previous									−0.007	−0.009	0.005	−0.011
									(0.004)	(0.009)	(0.009)	(0.010)
Reciprocity:Post:I_previous									0.001	0.017	−0.002	−0.0002
									(0.004)	(0.009)	(0.009)	(0.010)
Norms:Post:I_previous									−0.006	−0.015	0.004	−0.016
									(0.004)	(0.009)	(0.009)	(0.010)
Buyer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	760,134	638,195	581,650	464,294	760,134	638,195	581,650	464,294	760,134	638,195	581,650	464,294
Number of buyers	40,823	40,823	34,871	22,226	40,823	40,823	34,871	22,226	40,823	40,823	34,871	22,226
R^2	0.544	0.562	0.562	0.569	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	0.453	0.462	0.461	0.469

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

Table B.9. Repeat Buyer Tipping Mechanism with Second Field Experiment

	Regional Tipping Culture				*5-star Satisfaction Rating*				*Previous Transaction*			
	Pool	First	Second	Third	Pool	First	Second	Third	Pool	First	Second	Third
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Norms:Post	0.010***	0.015***	0.009	0.004	0.007**	0.007	0.007	0.005	0.011***	0.011***	0.008	0.003
	(0.004)	(0.005)	(0.007)	(0.009)	(0.003)	(0.004)	(0.006)	(0.007)	(0.003)	(0.004)	(0.005)	(0.007)
Post	−0.011	−0.012**	−0.020***	0.009	−0.005	−0.011	−0.003	−0.010	−0.019***	−0.024***	−0.026***	−0.024**
	(0.007)	(0.006)	(0.007)	(0.019)	(0.004)	(0.009)	(0.012)	(0.009)	(0.006)	(0.009)	(0.007)	(0.009)
Norms:I_North_America	0.002	0.001	0.004	0.004								
	(0.004)	(0.004)	(0.004)	(0.004)								
Post:I_North_America	−0.010***	−0.005	−0.008	−0.005								
	(0.004)	(0.005)	(0.007)	(0.009)								
Norms:Post:I_North_America	0.001	−0.006	−0.0005	−0.002								
	(0.005)	(0.007)	(0.009)	(0.012)								
Norms:I_5stars					−0.003	−0.002	−0.003	−0.003				
					(0.004)	(0.004)	(0.004)	(0.004)				
Post:I_5stars					−0.026***	−0.021***	−0.029***	−0.023**				
					(0.004)	(0.005)	(0.007)	(0.009)				
Norms:Post:I_5stars					0.007	0.007	0.005	−0.003				
					(0.005)	(0.007)	(0.010)	(0.013)				
Norms:I_previous									−0.001	−0.002	−0.001	−0.001
									(0.002)	(0.002)	(0.002)	(0.002)
Post:I_previous									0.007	0.006	0.011	0.020
									(0.004)	(0.012)	(0.020)	(0.021)
Norms:Post:I_previous									−0.002	0.004	0.031	−0.011
									(0.006)	(0.017)	(0.027)	(0.028)
Buyer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	405,779	358,485	337,086	328,677	405,779	358,485	337,086	328,677	405,779	358,485	337,086	328,677
Number of Buyers	39,393	39,393	35,298	34,141	39,393	39,393	35,298	34,141	39,393	39,393	35,298	34,141
R^2	0.553	0.571	0.573	0.575	0.553	0.571	0.573	0.575	0.553	0.571	0.573	0.575
Adjusted R^2	0.369	0.382	0.386	0.388	0.370	0.382	0.386	0.388	0.369	0.382	0.386	0.388

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

B.4 Do Tips Affect Seller Behavior?

This section investigates the seller side of the market. First, we investigate the pre-treatment data collected before the first pilot experiment to estimate associations between tips and subsequent seller behavior[3]. 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 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, lagged seller average rating), transaction characteristics (order price, category fixed effects), buyer fixed effects, and seller fixed effects.

[3]For more accurate association estimates, we restrict our analysis to pre-treatment data preceding the first pilot experiment. If we base our analysis on the period before the second field experiment, it could risk including data potentially impacted by the treatment messages, obscuring the clear interaction we aim to see.

Table B.23 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 has an expected rating of 0.02 more stars, relative to no tip received on the previous transaction. 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 both the initial pilot experiment and the second 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 B.15 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.[4]

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 periods, then examine how sellers' first indirect treatment exposure predicts outcomes of the seller's second treated transaction. Figures B.17 and B.18 show that, overall, indirect seller treatments do not reliably predict changes in subsequent tips received, buyer satisfaction rating, seller tip mention, or price in both the first and second sample periods. We then restricted our attention to sellers who first interacted with new buyers during both treatment periods, given our previous finding that Norms treatment effect was larger for new buyers. Figure B.19 shows no significant effects of indirect treatments on sellers' subsequent transaction characteristics or tips received during the first sample period.[5] However, during the second sample period, Figure B.20

[4]Figure B.16 shows no significant differences if we restrict attention to sellers who first interacted with new buyers.

[5]We also examined new sellers' interactions with all treated buyers and new sellers who first interacted only with treated new buyers. Neither group showed significant differences across indirect treatment groups. Additionally, we checked for interactions between seller indirect treatments and indicators for North American seller region

shows that, although indirect Norm treatments generally had no significant impact on subsequent behaviors or tips received, sellers indirectly treated under Norms condition tend to offer higher lead prices compared to those not indirectly treated.

B.5 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 Norms to tip."

and 5-star buyer satisfaction on the first treated transaction. Once again, we found no evidence that indirect seller treatments influenced seller behavior.

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 B.10 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 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 B.11 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 for both Implicit Reciprocity and Reciprocity tip messages regardless of the seller's name, with "Social Proof" being the dominant interpretation of 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.

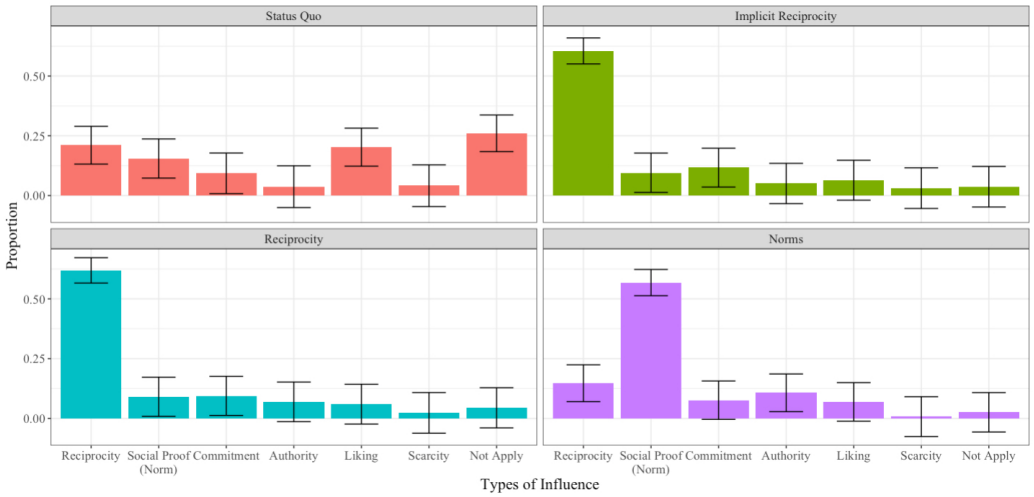


Figure B.10. Perception for Each Tip Message

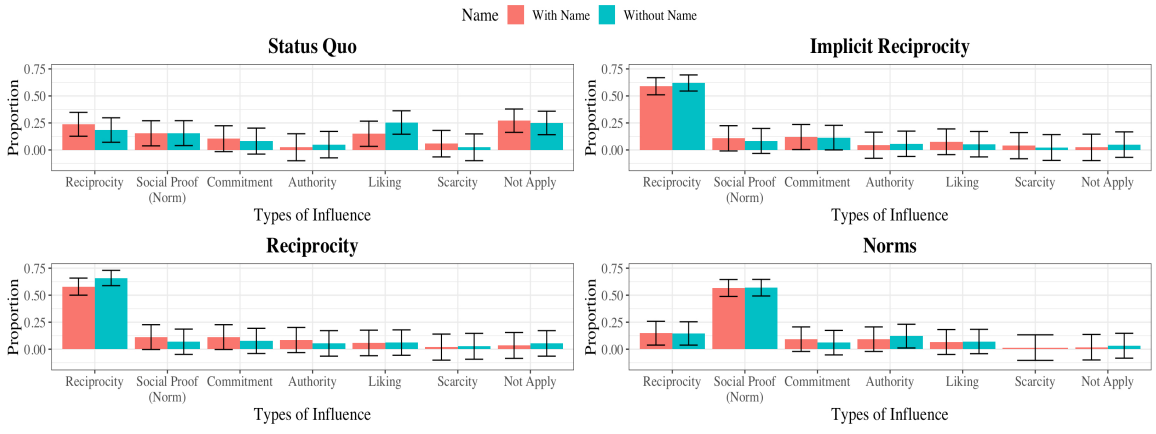


Figure B.11. Interaction between Dominant Perception and Seller Name

B.6 Additional Tables and Figures

Table B.10. First Pilot Experiment Randomization Check - New Buyers

Variable	Control	Implicit Reciprocity	Reciprocity	Norms	P-value
Buyer Region					
Africa	0.04	0.05	0.04	0.03	0.0579
Asia	0.13	0.12	0.11	0.11	0.3441
Europe	0.17	0.17	0.18	0.18	0.9291
Latin America & Caribbean	0.03	0.02	0.02	0.02	0.7022
North America	0.50	0.50	0.51	0.50	0.9838
Oceania	0.04	0.04	0.04	0.04	0.4962

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

Table B.11. First Pilot Experiment Randomization Check - Repeat Buyers

Variable	Control	Implicit Reciprocity	Reciprocity	Norms	P-value
Buyer Region					
Africa	0.03	0.03	0.03	0.03	0.4057
Asia	0.11	0.11	0.11	0.11	0.3448
Europe	0.16	0.15	0.15	0.16	0.7984
Latin America & Caribbean	0.03	0.03	0.03	0.03	0.922
North America	0.55	0.54	0.55	0.54	0.437
Oceania	0.04	0.04	0.04	0.04	0.917
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	14.44	15.24	0.3877

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

Table B.12. Second Experiment Randomization Check - New Buyers

Variable	Control	Norms	P-value
Buyer Region			
Africa	0.01	0.01	0.1397
Asia	0.07	0.07	0.3766
Europe	0.12	0.12	0.904
Latin America & Caribbean	0.02	0.02	0.3037
North America	0.64	0.65	0.5689
Oceania	0.04	0.04	0.4022

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

Table B.13. Second Experiment Randomization Check - Repeat Buyers

Variable	Control	Norms	P-value
Buyer Region			
Africa	0.01	0.01	0.590
Asia	0.07	0.07	0.998
Europe	0.11	0.11	0.804
Latin America & Caribbean	0.02	0.02	0.794
North America	0.66	0.66	0.798
Oceania	0.04	0.04	0.498
Buyer Characteristics			
Buyer Tenure	26.16	25.81	0.1421
Pre-Treatment Orders	2.30	2.32	0.5185
Pre-Treatment Expenditure	83.04	83.64	0.7399
Pre-Treatment Tip Frequency	0.18	0.20	0.0000
Pre-Treatment Tip Expenditure	1.81	1.98	0.0269

Note: The table displays characteristics of repeat buyers, including those who were treated in the first experiment, segregated by treatment group. It also presents Chi-square tests validating the null hypothesis of random assignment.

Table B.14. New Buyers Tip Amount Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity	−0.050	0.019	−0.274	0.104
	(0.098)	(0.149)	(0.149)	(0.201)
Reciprocity	0.042	−0.016	0.006	0.261
	(0.107)	(0.140)	(0.189)	(0.215)
Norms	0.423***	0.843***	0.193	0.216
	(0.117)	(0.184)	(0.186)	(0.223)
Buyer Char.	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category FE	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545
R^2	0.122	0.113	0.160	0.211
Adjusted R^2	0.120	0.110	0.156	0.202
Note:				**p<0.05; ***p<0.01

Table B.15. New Buyers Tip Percentage Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity	0.396	1.256	0.251	−0.139
	(0.491)	(0.733)	(0.758)	(1.221)
Reciprocity	−0.130	0.221	−0.125	−1.490
	(0.421)	(0.676)	(0.682)	(1.094)
Norms	1.469***	3.238***	1.183	−1.159
	(0.484)	(0.804)	(0.740)	(1.178)
Buyer Char.	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category FE	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y
Observations	25,215	7,880	7,258	3,545
R^2	0.076	0.038	0.104	0.152
Adjusted R^2	0.074	0.034	0.099	0.143
Note:				**p<0.05; ***p<0.01

Table B.16. Repeat Buyers Tip Amount Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity:Post	0.102	0.168	0.246	−0.237
	(0.104)	(0.244)	(0.192)	(0.191)
Reciprocity:Post	−0.029	0.007	−0.055	−0.301**
	(0.052)	(0.090)	(0.094)	(0.147)
Norms:Post	−0.012	0.075	0.118	−0.284
	(0.054)	(0.096)	(0.130)	(0.146)
Post	0.019	0.013	0.009	0.171
	(0.041)	(0.065)	(0.078)	(0.137)
Buyer FE	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category 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	0.482	0.498	0.518	0.521
Adjusted R^2	0.377	0.383	0.407	0.410
Note:				**p<0.05; ***p<0.01

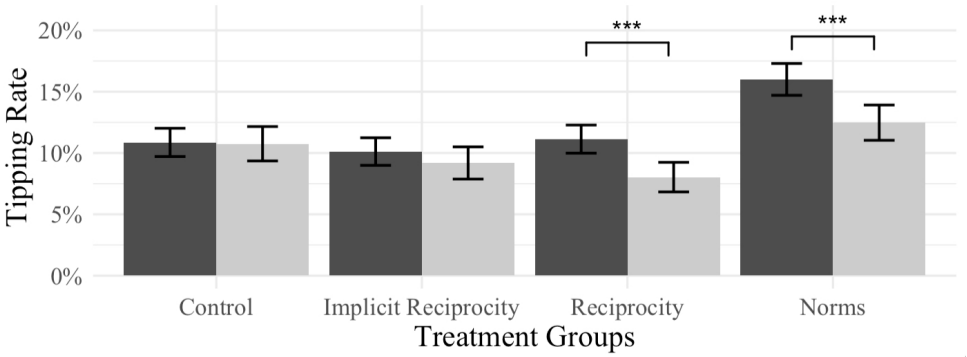


Figure B.12. Post-Treatment Effects on New Buyers

This figure shows the mean difference in new buyer Tipping rate between the test and post-test periods across treatment messages. Removing Norms treatment message without changing the default tips reduced new buyers' tipping during the post-test period from 16% to 12.5%. Removing the Reciprocity treatment message without changing the default tips reduced new buyers' tipping during the post-test period from 11.1% down to 8%.

Table B.17. Repeat Buyers Tip Percentage Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Implicit_Reciprocity:Post	−0.245	−0.242	−0.426	−1.209***
	(0.202)	(0.379)	(0.368)	(0.455)
Reciprocity:Post	−0.077	−0.166	0.180	−1.553***
	(0.202)	(0.372)	(0.371)	(0.448)
Norms:Post	0.062	0.509	0.331	−0.860
	(0.202)	(0.382)	(0.360)	(0.461)
Post	0.102	0.285	0.155	0.708**
	(0.146)	(0.278)	(0.258)	(0.344)
Buyer FE	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category 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	0.430	0.449	0.456	0.459
Adjusted R^2	0.316	0.322	0.330	0.333
Note:				**p<0.05; ***p<0.01

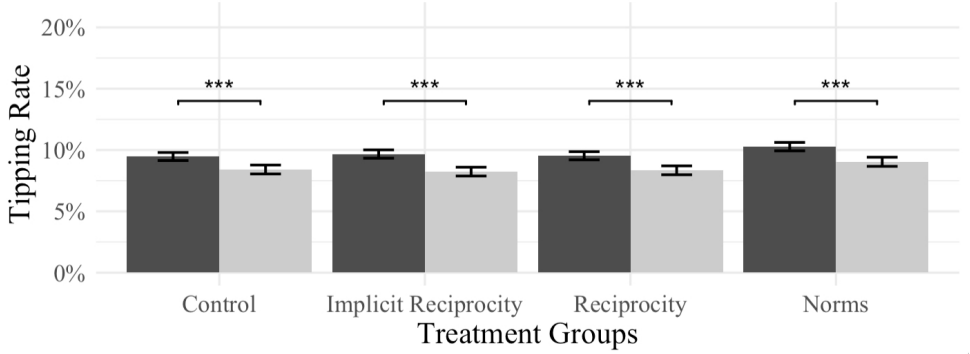


Figure B.13. Post-Treatment Effects on Repeat Buyers

This figure shows the mean difference in repeat buyer Tipping rate between the test and post-test periods across treatment messages. Removing Norms treatment message without changing the default tips reduced repeat buyers' tipping during the post-test period from 9.5% to 8.3%. Removing the Reciprocity treatment message without changing the default tips reduced repeat buyers' tipping during the post-test period from 11.9% down to 10.8%.

Table B.18. New Buyers Tip Amount Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Norms	0.161***	0.120**	0.203**	0.260
	(0.050)	(0.059)	(0.097)	(0.135)
Buyer Char.	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category FE	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y
Observations	33,810	18,599	6,737	3,174
R^2	0.073	0.079	0.078	0.089
Adjusted R^2	0.072	0.078	0.074	0.080
Note:				**p<0.05; ***p<0.01

Table B.19. New Buyers Tip Percentage Regression

	Pool	First	Second	Third
	(1)	(2)	(3)	(4)
Norms	0.791***	0.766***	0.752	1.043
	(0.215)	(0.241)	(0.450)	(0.612)
Buyer Char.	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y
Category FE	Y	Y	Y	Y
Seller Char.	Y	Y	Y	Y
Observations	33,810	18,599	6,737	3,174
R^2	0.041	0.040	0.046	0.041
Adjusted R^2	0.040	0.039	0.041	0.031
Note:				**p<0.05; ***p<0.01

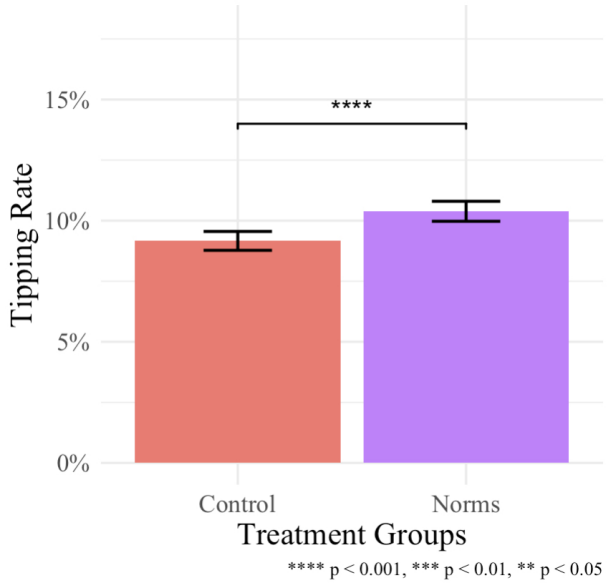


Figure B.14. Treatment Effects on 42,153 Repeat Buyers' First Treated Transactions

B.14 displays mean difference in 42,153 repeat buyer Tipping Rate by treatment groups in the first treated app transactions. Repeat buyers who received Norms treatment tipped about 14.3% more often than Control (9.1% vs. 10.4% $p< .0001$)

Table B.20. All Repeat Buyers Tipping and Repeat Transaction Regression

	Pool (1)	First (2)	Second (3)	Third (4)	Second Order Exists (5)	Third Order Exists (6)
Norms:Post	0.008***	0.009***	0.007	−0.0004	−0.005	0.005
	(0.002)	(0.003)	(0.004)	(0.006)	(0.005)	(0.007)
Post	−0.021***	−0.019***	−0.026***	−0.022***	−0.458***	−0.378***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Buyer FE	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y
Observations	553,656	498,592	476,295	467,341	498,592	476,295
Number of Buyers	42,153	42,153	38,058	36,901	42,153	38,058
R^2	0.529	0.544	0.546	0.547	0.772	0.749
Adjusted R^2	0.382	0.395	0.398	0.400	0.697	0.668
Note:						**p<0.05; ***p<0.01

In this analysis, we focus on transactions from all treated repeat buyers(N = 42,153), which includes those who were treated twice in both the first and second experiments. Notable, the significance of the estimations remains unchanged.

Table B.21. Frequent Repeat Buyers Tipping and Repeat Transaction Regression

	Pool (1)	First (2)	Second (3)	Third (4)	Second Order Exists (5)	Third Order Exists (6)
Norms:Post	0.008^{***}	0.011^{***}	0.008	0.0001	-0.450^{***}	-0.378^{***}
	(0.003)	(0.004)	(0.006)	(0.008)	(0.005)	(0.007)
Post	-0.023^{***}	-0.020^{***}	-0.026^{***}	-0.024^{***}	-0.005	0.011
	(0.002)	(0.003)	(0.004)	(0.006)	(0.007)	(0.009)
Buyer FE	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y
Observations	317,504	283,399	179,218	119,608	283,399	179,218
Number of Buyers	23,879	23,879	12,110	6,815	23,879	12,110
R^2	0.553	0.569	0.589	0.612	0.744	0.772
Adjusted R^2	0.367	0.381	0.377	0.378	0.632	0.655

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

Table B.22. Repeat Buyers Tip Amount and Tip Percentage Regression

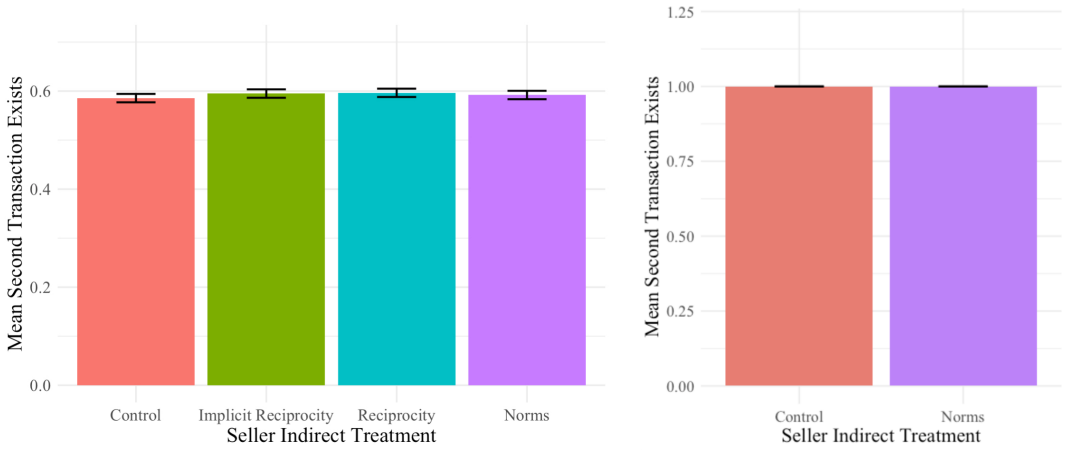
	Pool (1)	First (2)	Second (3)	Third (4)	Pool (5)	First (6)	Second (7)	Third (8)
Norms:Post	0.026	-0.001	-0.025	-0.024	0.215	0.204	0.460	-0.475
	(0.043)	(0.060)	(0.078)	(0.105)	(1.077)	(1.264)	(1.385)	(0.470)
Post	-0.244^{***}	-0.204^{***}	-0.184^{***}	-0.233^{***}	-1.089^{***}	-0.931^{***}	-1.461^{***}	-0.682
	(0.031)	(0.044)	(0.056)	(0.081)	(0.126)	(0.164)	(0.243)	(0.352)
Buyer FE	Y	Y	Y	Y	Y	Y	Y	Y
Transaction Char.	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y
Seller FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	412,115	364,821	343,422	335,013	412,115	364,821	343,422	335,013
Number of Buyers	39,393	39,393	35,298	34,141	39,393	39,393	35,298	34,141
R^2	0.529	0.549	0.557	0.559	0.460	0.473	0.471	0.471
Adjusted R^2	0.341	0.356	0.367	0.371	0.243	0.247	0.245	0.245

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

Table B.23. 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

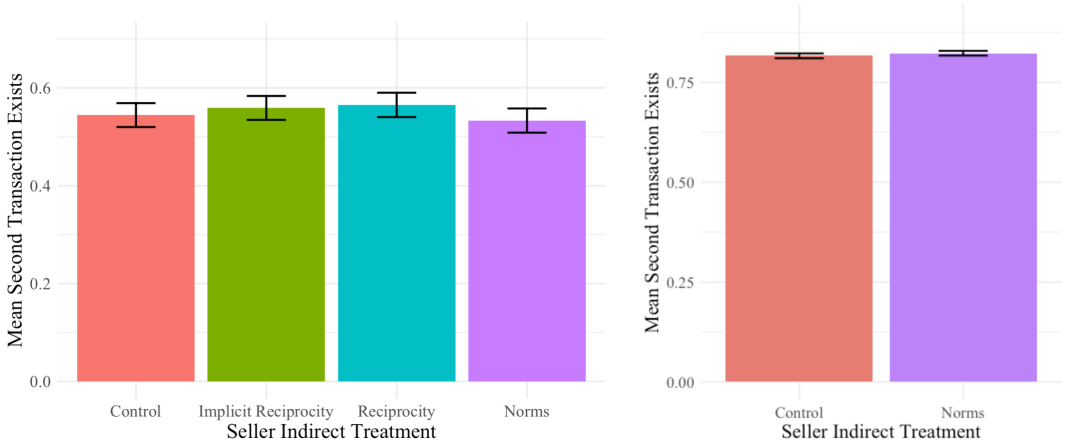
Note: **p<0.05; ***p<0.01



(a) First Pilot Experiment **(b)** Second Experiment

Figure B.15. Seller Repeat Platform Usage by Indirect Treatment Groups

These figures show the mean likelihood of a second transaction occurring within seller indirect treatment groups. Figure C5(a) considers the treatment group assigned to each seller's first matched buyer during the initial test period (N = 50,562). Figure C5(b) considers the treatment group assigned to each seller's first matched buyer during the subsequent test period (N = 20,510).



(a) First Pilot Experiment **(b)** Second Experiment

Figure B.16. Seller Repeat Platform Usage by Indirect Treatment Groups

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. Figure (a) takes into account sellers from the first test period (N = 6,268). Figure (b) involves sellers from the second test period (N = 11,834).

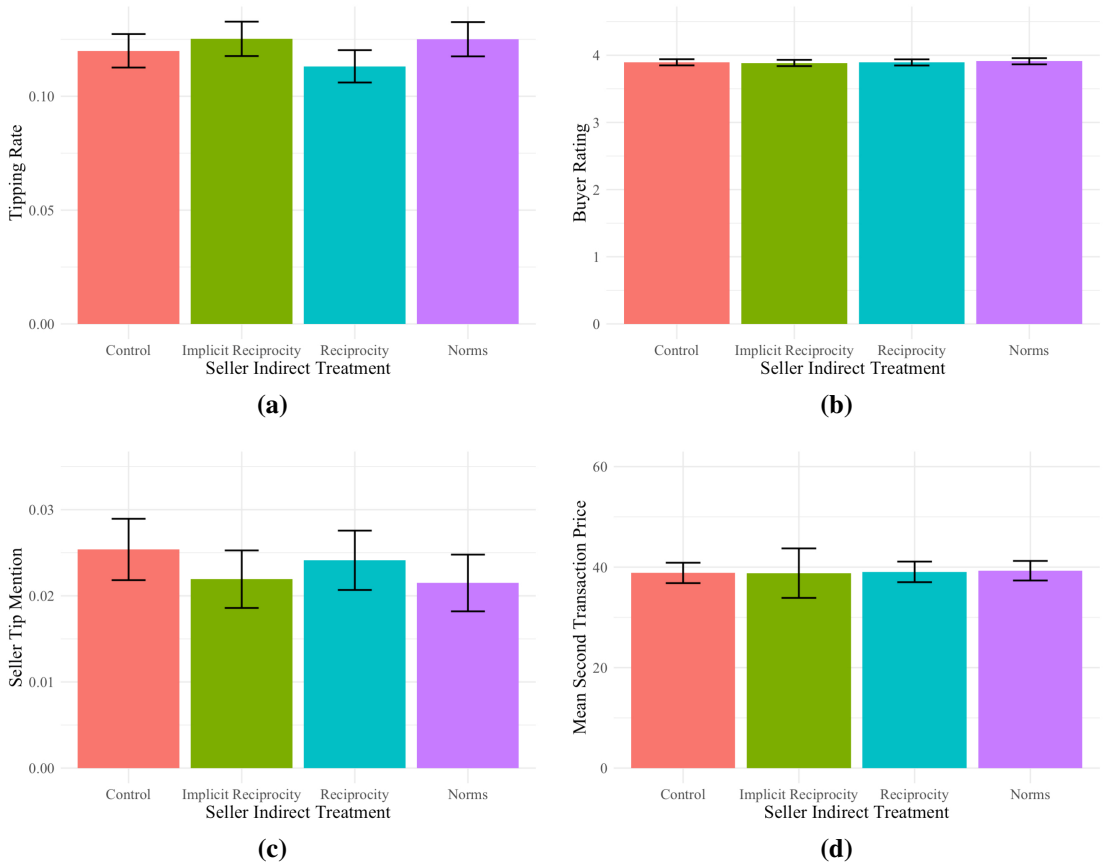


Figure B.17. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups

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

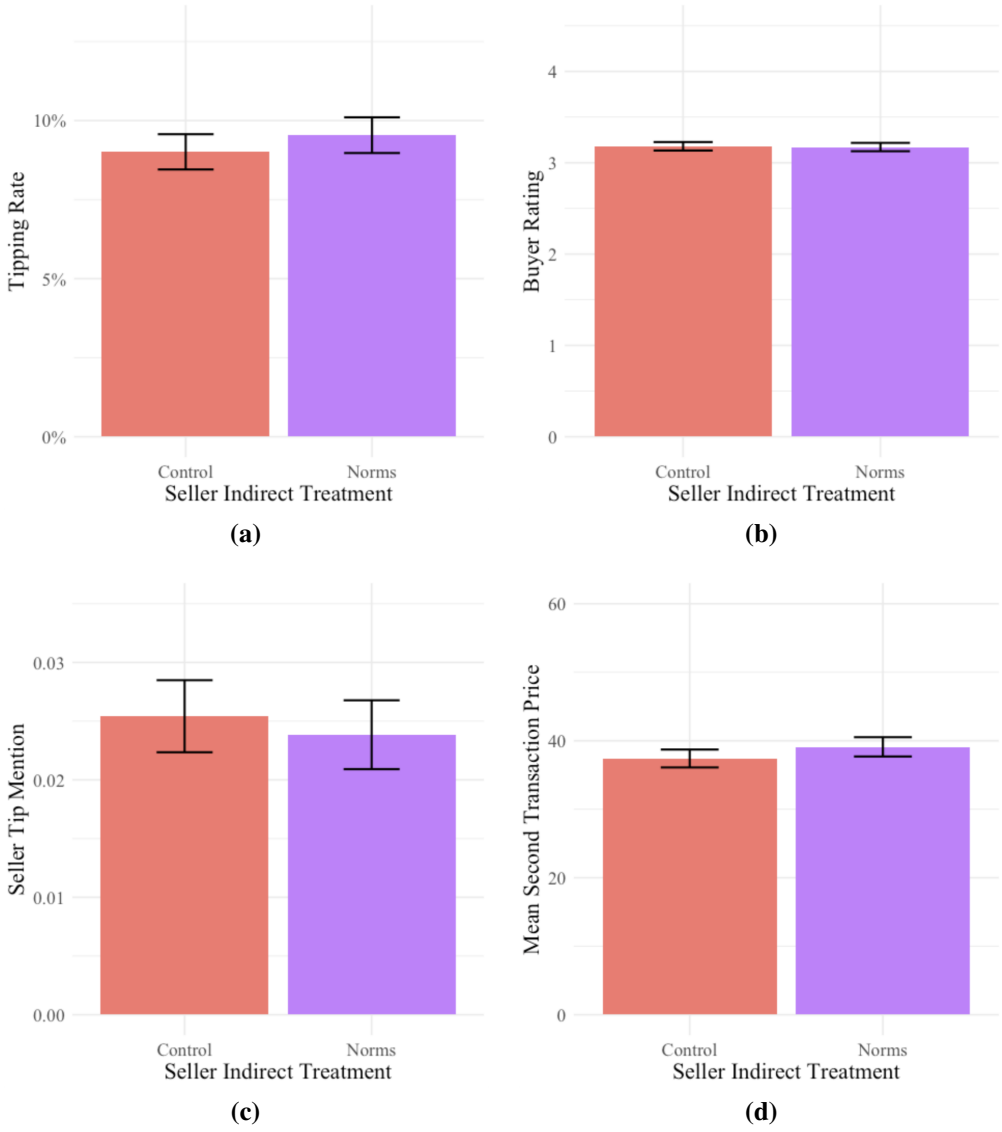


Figure B.18. Seller Subsequent Behaviors during Second Experiment by Indirect Treatment Groups

Figures show average outcomes of (a) tipping, (b) rating, (c) seller tip mention, and (d) price in all sellers' second transactions during the second field experiment period, differentiated by random treatments assigned to the first buyer each seller transacted with (N = 20,510).

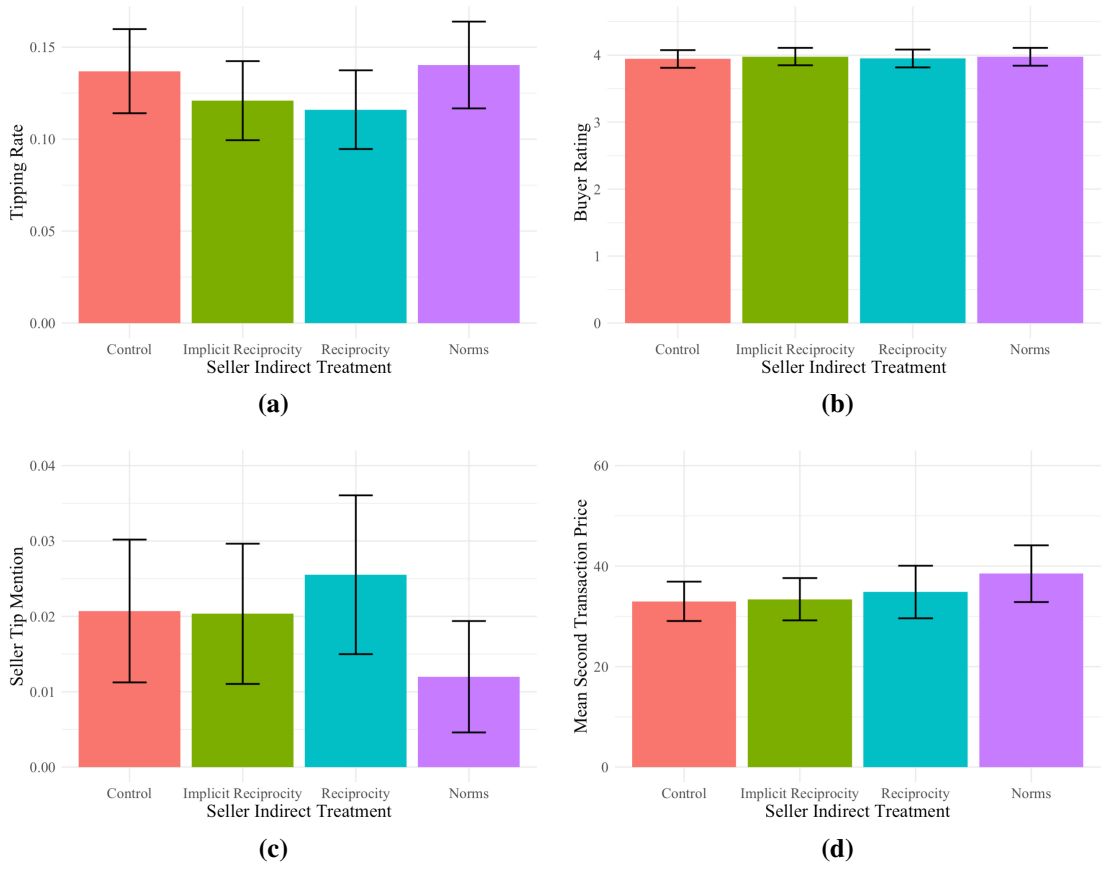


Figure B.19. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups - New Treated Buyers Only

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

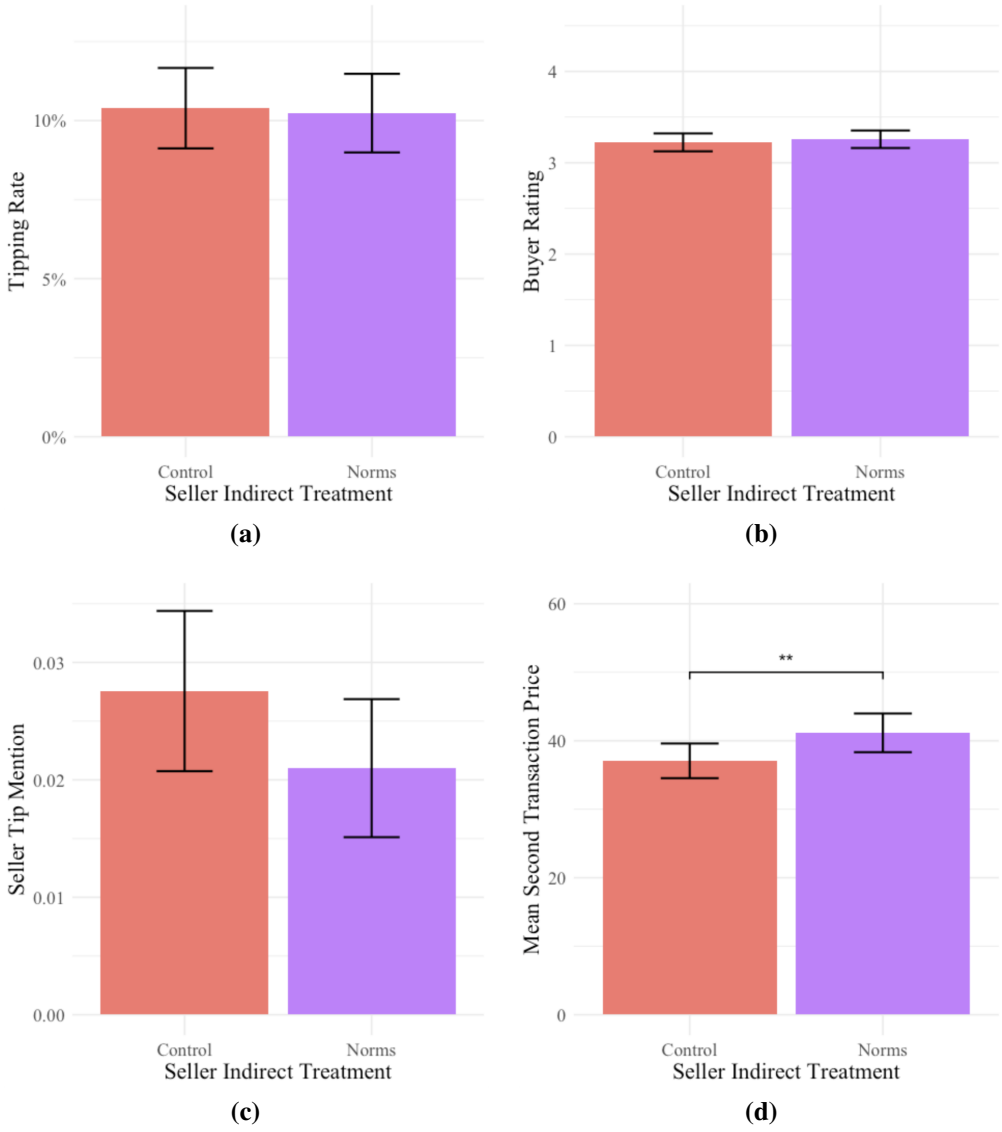
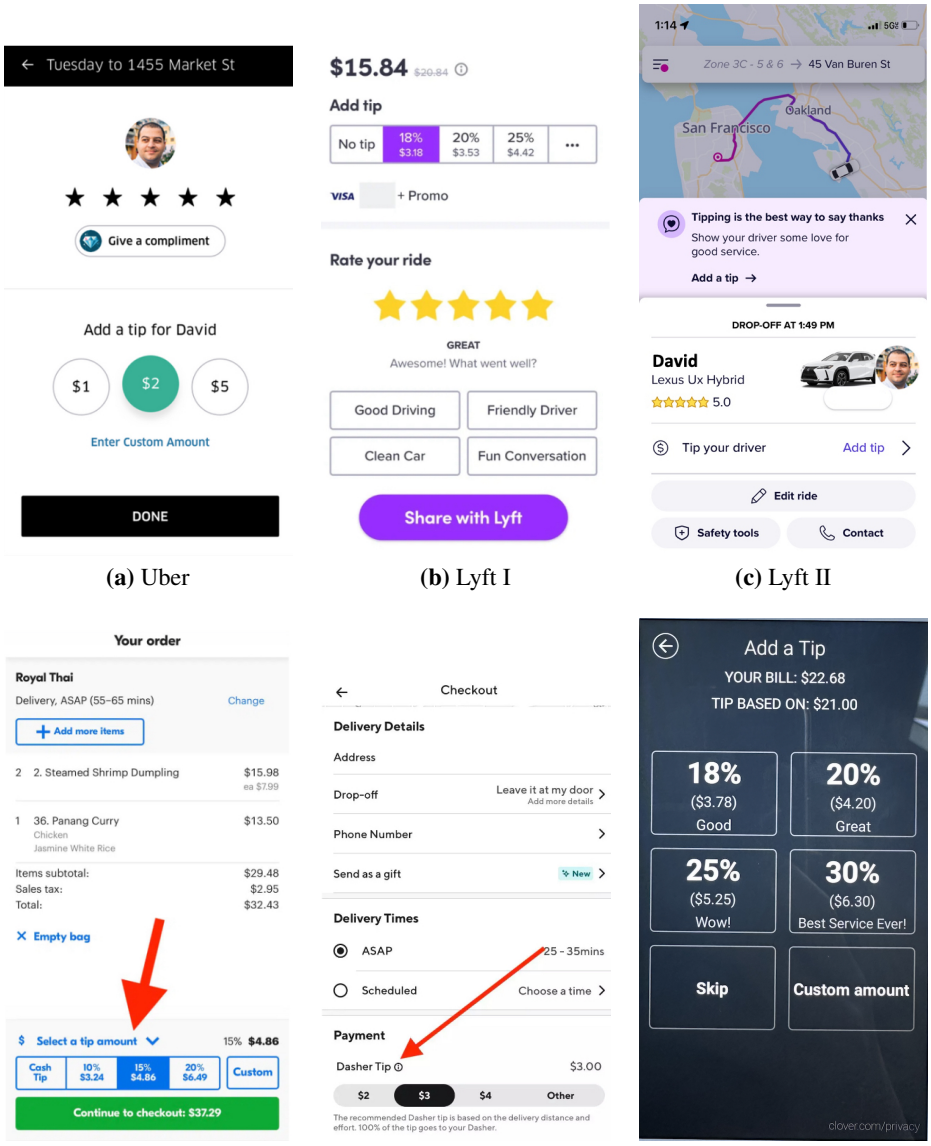


Figure B.20. Seller Subsequent Behaviors during First Pilot Experiment by Indirect Treatment Groups - New Treated Buyers Only

Figures show average outcomes of (a) tipping, (b) rating, (c) seller tip mention, and (d) price in all sellers' second transactions during the second field experiment period, differentiated by random treatments assigned to the first buyer each seller transacted with; filtered only to sellers whose first transaction was with a new buyer (N = 4,499).

B.7 Online Experiment Design



(a) Uber **(b)** Lyft I **(c)** Lyft II

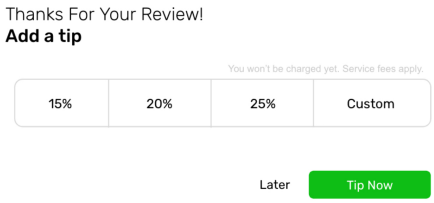
(d) Grubhub **(e)** DoorDash **(f)** Offline Store

Figure B.21. Actual Tip Window Example

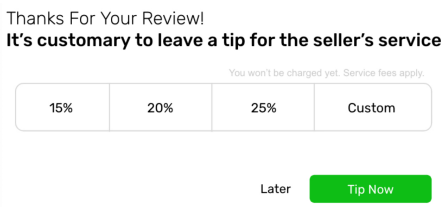
Each figure shows the actual tip window used by various online platforms and offline restaurants. As observed, each platform employs a unique combination of message and default tip options when requesting gratuities. Some platforms, like Lyft showcased in Figures (b) and (c), use a variety of messages to promote tipping. They transitioned from a simple "add tips" message to a more engaging prompt, "Tipping is the best way to say thanks." Furthermore, some venues have started to pair their default tip options with specific messages. This strategy can be seen in Figure (f), where default tip amounts are presented along with associated messages.

Table B.24. Message Manipulation

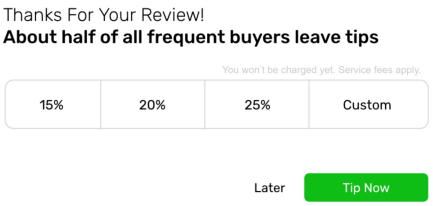
Context	Treatment	Type	Quantifiable	Tone	Subject	Message Example
Online Freelancer (e.g., Fiverr)	Treatment Messages	Descriptive	Yes		Buyer	About half of all frequent buyers leave tips
		Descriptive	Yes		Seller	About 75% of sellers receive tips
		Descriptive	No		Buyer	Many frequent buyers leave tips
		Descriptive	No		Seller	Many sellers receive tips
		Injunctive		Active	Buyer	Buyers should leave a tip
		Injunctive		Passive	Buyer	It's good to leave a tip
		Injunctive		Active	Seller	Most sellers believe that buyers should leave a tip
		Injunctive		Passive	Seller	Most sellers believe that it's good to leave a tip
	Control Messages	Uber Message				Add a tip
		Field Experiment				It's customary to leave a tip for the seller's service
Ride Sharing (e.g., Uber)	Treatment Messages	Descriptive	Yes		Buyer	Millions of riders leave tips
		Descriptive	Yes		Seller	Millions of drivers receive tips
		Descriptive	No		Buyer	Many riders leave tips
		Descriptive	No		Seller	Many drivers receive tips
		Injunctive		Active	Buyer	Riders should leave a tip
		Injunctive		Passive	Buyer	It's good to leave a tip
		Injunctive		Active	Seller	Most drivers believe that riders should leave a tip
		Injunctive		Passive	Seller	Most drivers believe that it's good to leave a tip
	Control Messages	Uber Message				Add a tip
		Field Experiment				It's customary to leave a tip for the rider's service
Food Delivery (e.g., DoorDash)	Treatment Messages	Descriptive	Yes		Buyer	About 95% of users leave tips
		Descriptive	Yes		Seller	About half of deliverers receive tips
		Descriptive	No		Buyer	Many users leave tips
		Descriptive	No		Seller	Many deliverers receive tips
		Injunctive		Active	Buyer	Users should leave a tip
		Injunctive		Passive	Buyer	It's good to leave a tip
		Injunctive		Active	Seller	Most deliverers believe that users should leave a tip
		Injunctive		Passive	Seller	Most deliverers believe that it's good to leave a tip
	Control Messages	Uber Message				Add a tip
		Field Experiment				It's customary to leave a tip for the deliverer's service



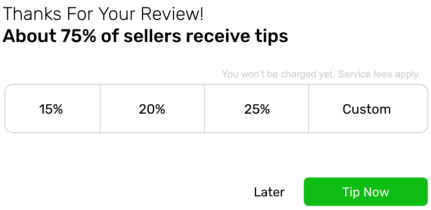
(a) Control 1 - Uber Message



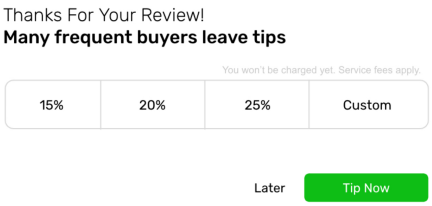
(b) Control 2 - Field Experiment Message



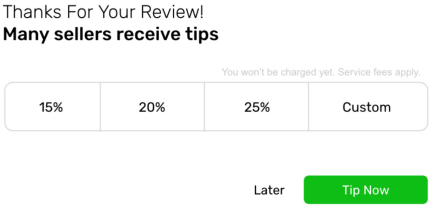
(c) Descriptive/ Quantifiable/ Buyer



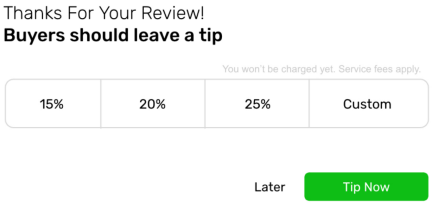
(d) Descriptive/ Quantifiable/ Seller



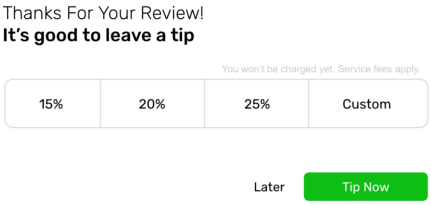
(e) Descriptive/ Not Quantifiable/ Buyer



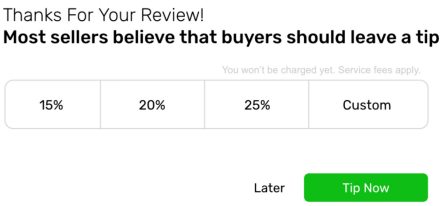
(f) Descriptive/ Not Quantifiable/ Seller



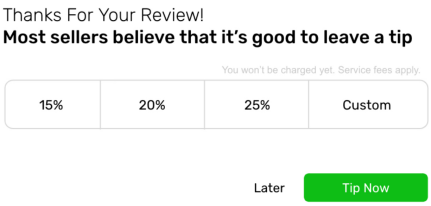
(g) Injunctive/ Active/ Buyer



(h) Injunctive/ Passive/ Buyer



(i) Injunctive/ Active/ Seller



(j) Injunctive/ Passive/ Seller

Figure B.22. Example of Fiverr Stimuli

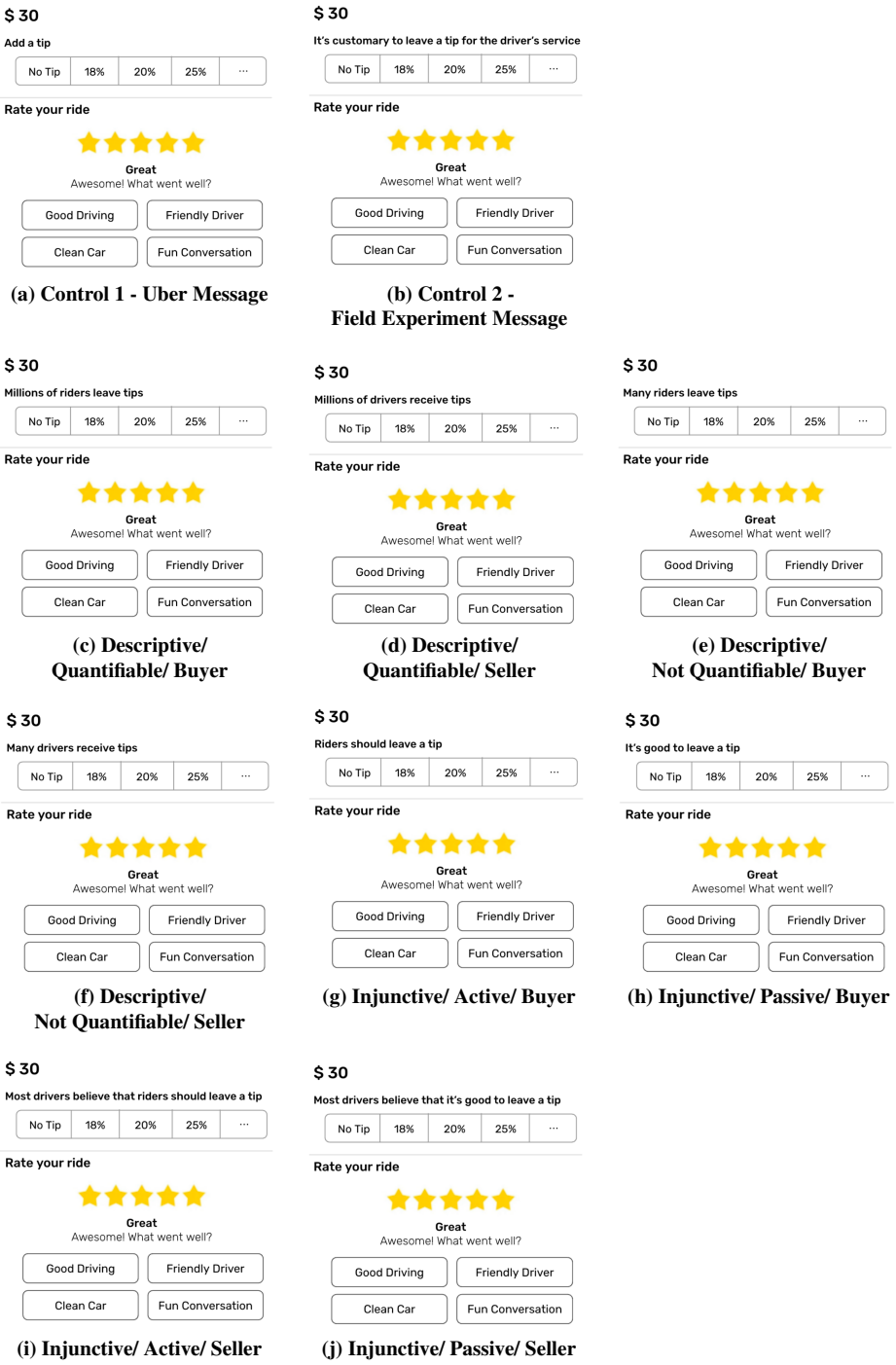
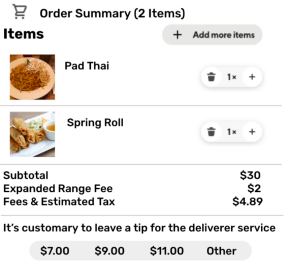
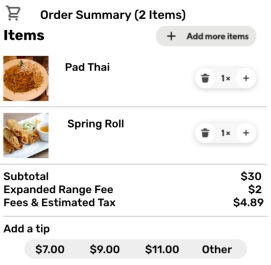


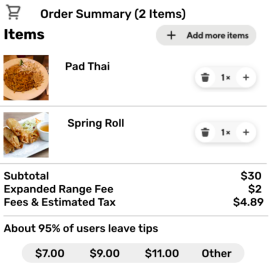
Figure B.23. Example of Ride-sharing Stimuli



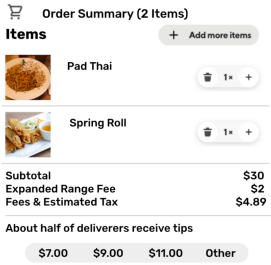
(a) Control 1 - Uber Message



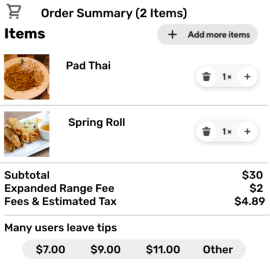
(b) Control 2 - Field Experiment Message



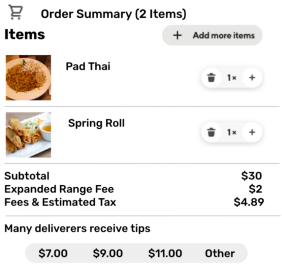
(c) Descriptive/ Quantifiable/ Buyer



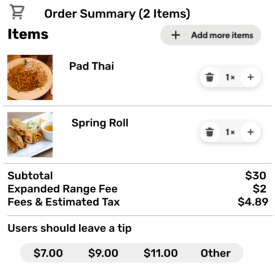
(d) Descriptive/ Quantifiable/ Seller



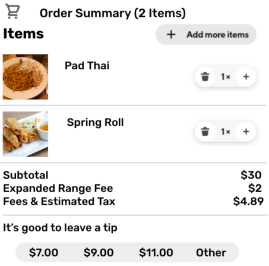
(e) Descriptive/ Not Quantifiable/ Buyer



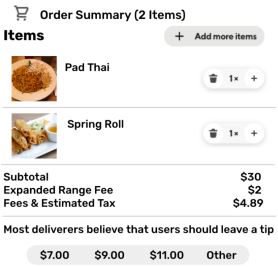
(f) Descriptive/ Not Quantifiable/ Seller



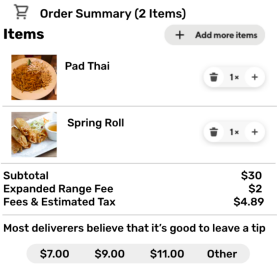
(g) Injunctive/ Active/ Buyer



(h) Injunctive/ Passive/ Buyer

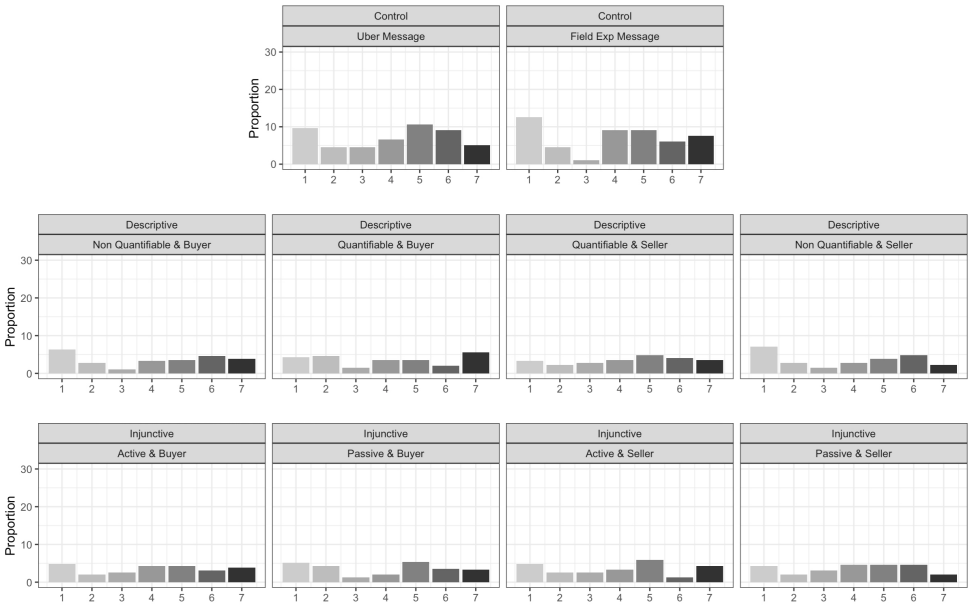


(i) Injunctive/ Active/ Seller

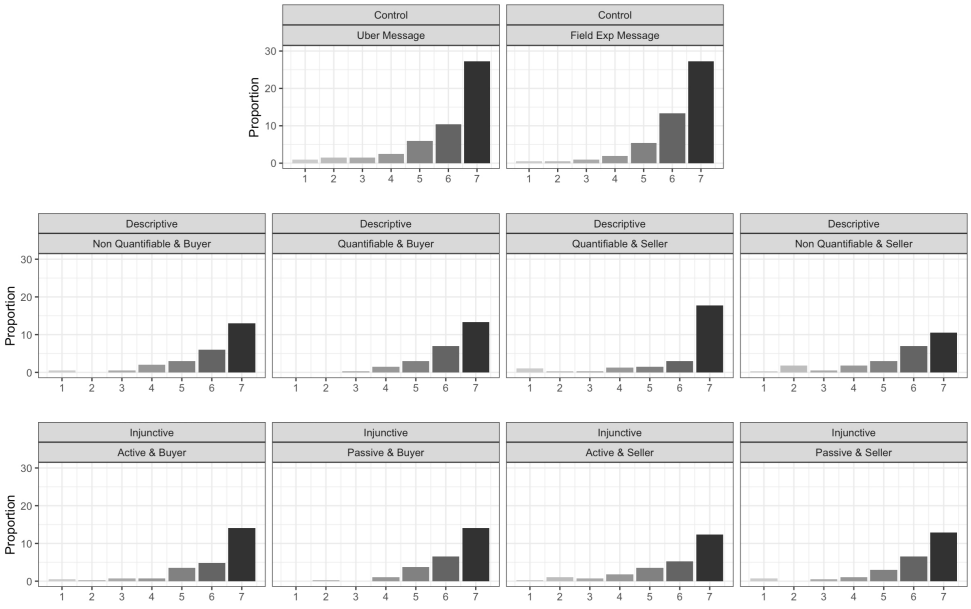


(j) Injunctive/ Passive/ Seller

Figure B.24. Example of Food Delivery Stimuli



(a) Fiverr

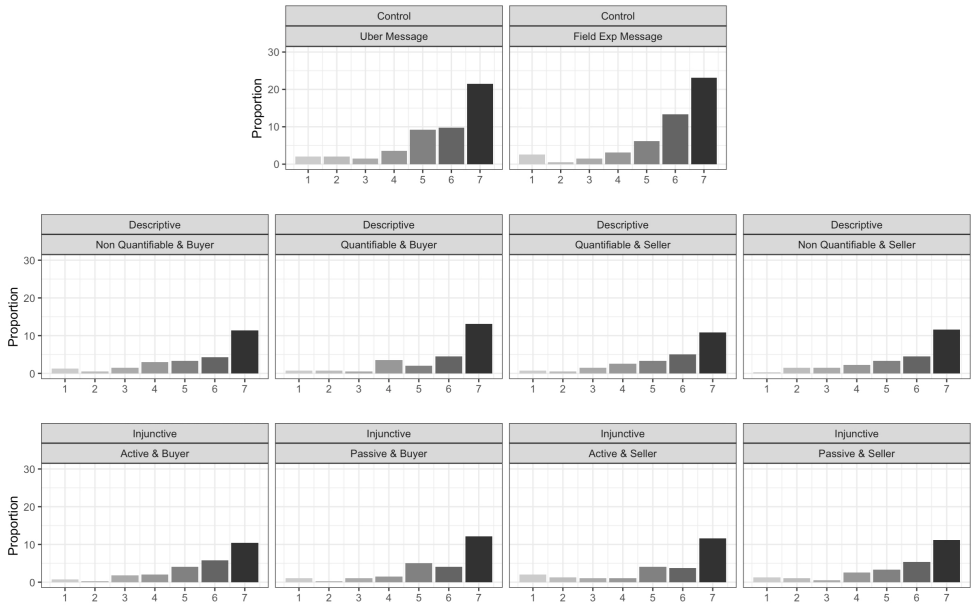


(b) Ride Sharing

Figure B.25. Frequency Graph (Part 1)

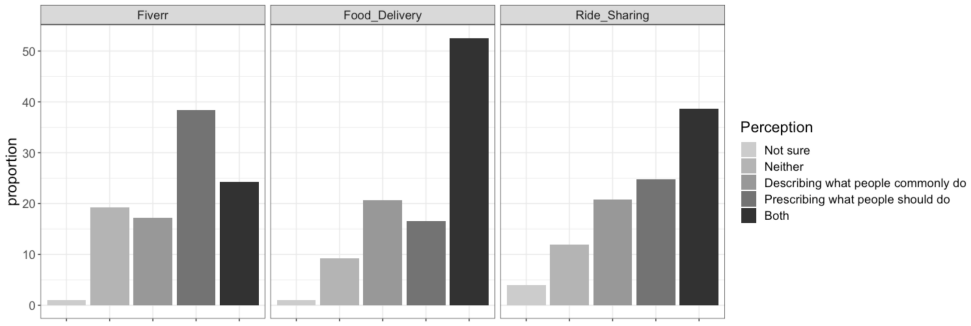
Table B.25. Hypothetical Scenario Example

Context	Hypothetical Scenario
Online Freelancer	You will see a hypothetical scenario where you hire a freelancer via online freelancer platforms like Upwork or Fiverr. Imagine that you've been working on your cover letter and resume, but you want to make sure they're as polished and professional as possible before submitting them to potential employers. After some research, you discover a freelancer on an online platform, Fiverr, specializing in editing cover letters and resumes. You decide to invest $50 in their services and hire them for their expertise. You send your current resume and cover letter along with a brief description of the types of jobs you're applying for and any specific concerns you have about your application materials. The freelancer responds quickly, assuring you that they'll have your revised documents back to you within two days. Once you receive their edited versions, you assess their work and feel satisfied with the improvements they made. Upon confirming your order, the platform prompts you to provide feedback on the freelancer's service, including the option to leave a tip on the next screen.
Ride Sharing	You will see a hypothetical scenario where you take a Uber or Lyft. Imagine you've just arrived at the airport after a long flight and are ready to head home. You request the ride via a ridesharing app like Uber or Lyft and select the ride option that costs $30. A few minutes later, your driver arrives and helps you load your luggage. During the ride, you sit back and relax. Upon arriving at your destination, the driver helps you unload your luggage. Following this, the ridesharing app prompts you to rate your driver's service, including the option to leave a tip on the next screen.
Food Delivery	You will see an example of ordering food through a food delivery app like DoorDash or Grubhub. Imagine you are feeling hungry but do not have the time to cook a meal. You open up the DoorDash app and start browsing through the different restaurant options available in your area. After a few minutes of scrolling, you settle on a Thai restaurant that has great reviews and offers delivery through DoorDash. You place your order, select your desired dishes, and add any special requests or modifications. The estimated delivery time is 25-35 minutes. Once you have confirmed your order, the delivery app will show a window summarizing your order including the option to leave a tip on the following screen.

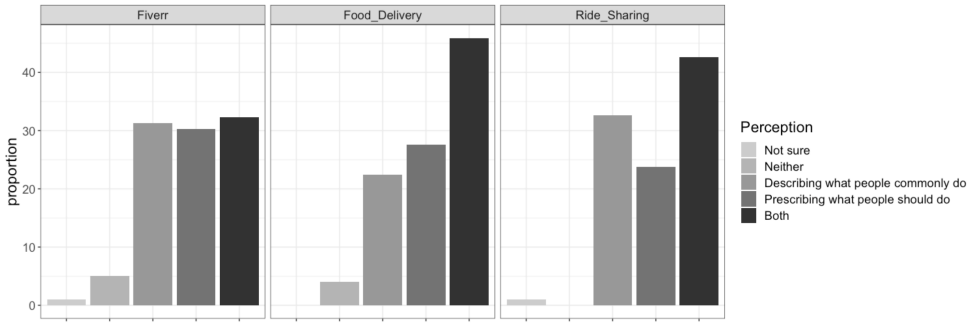


(c) Food Delivery

Figure B.25. Frequency Graph (continued)

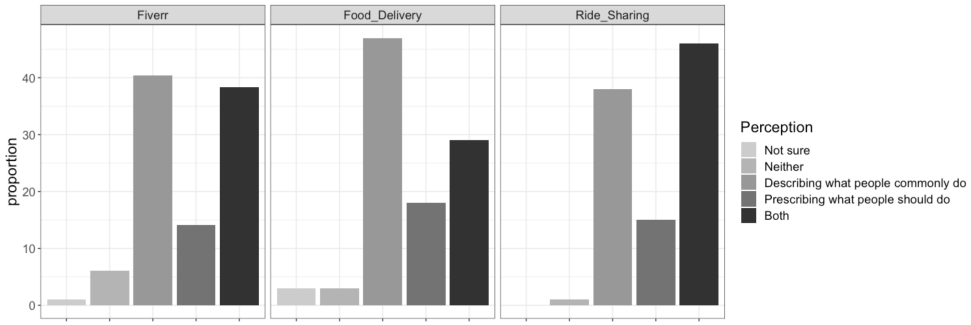


(a) Control1 - Uber Message

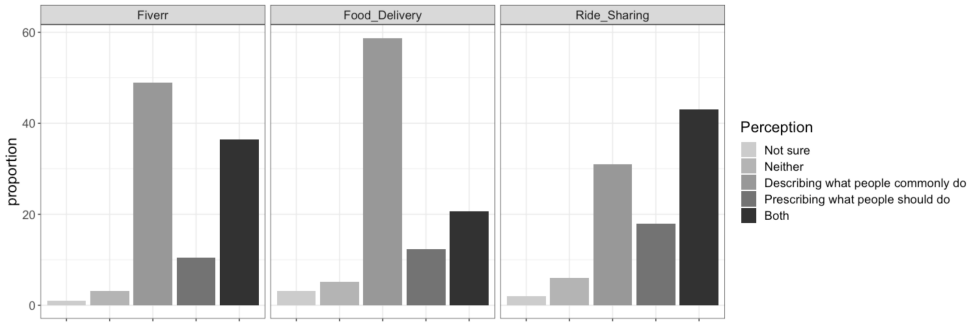


(b) Control2 - Field experiment Message

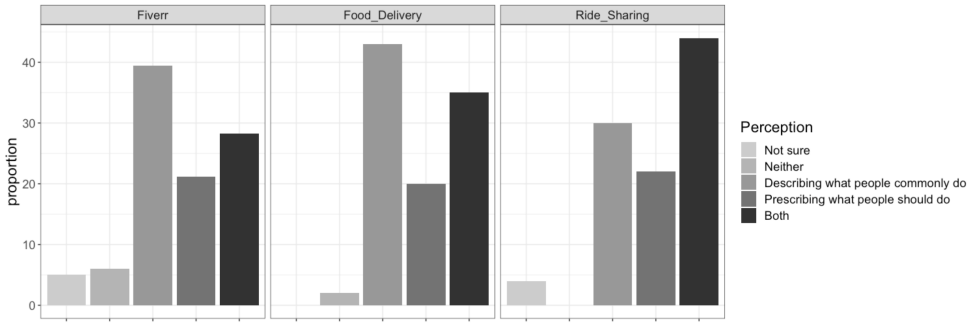
Figure B.26. Perception about Control Messages



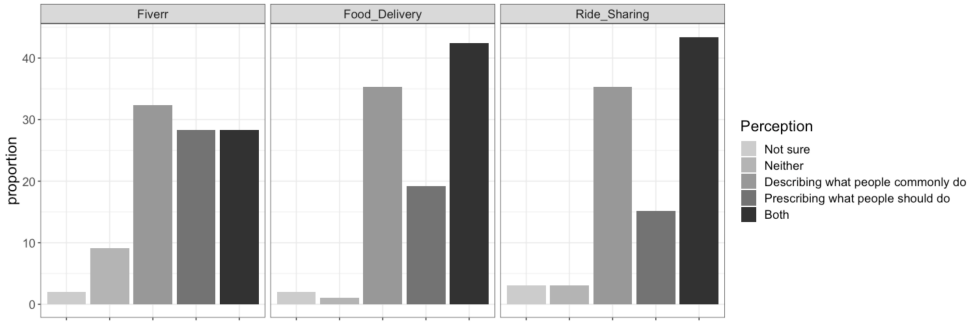
(a) Descriptive/ Quantifiable/ Buyer



(b) Descriptive/ Quantifiable/ Seller

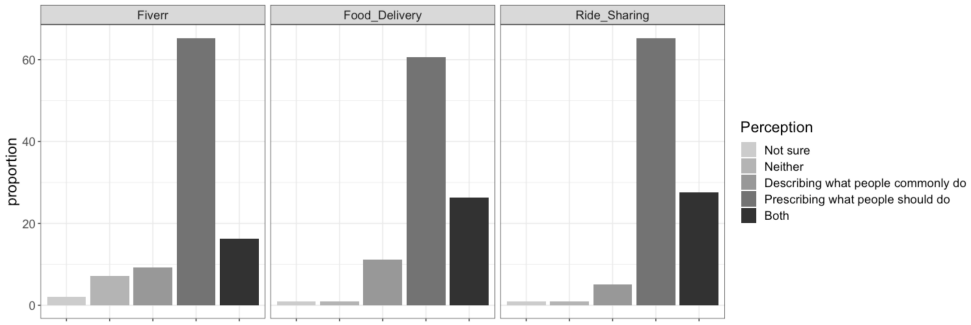


(c) Descriptive/Non-quantifiable/ Buyer

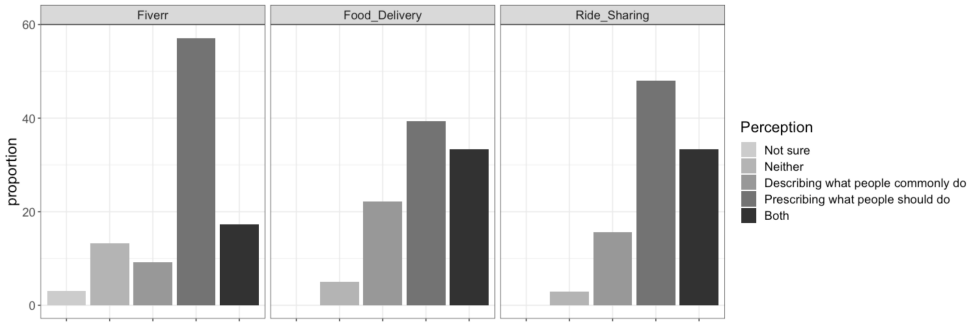


(d) Descriptive/Non-quantifiable/ Seller

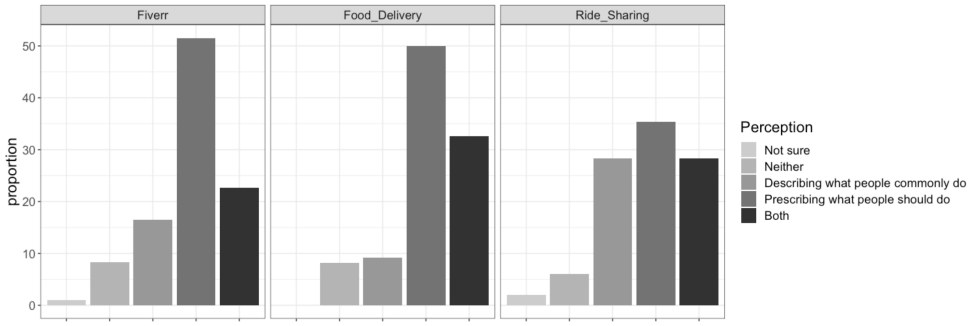
Figure B.27. Perception about Descriptive Messages



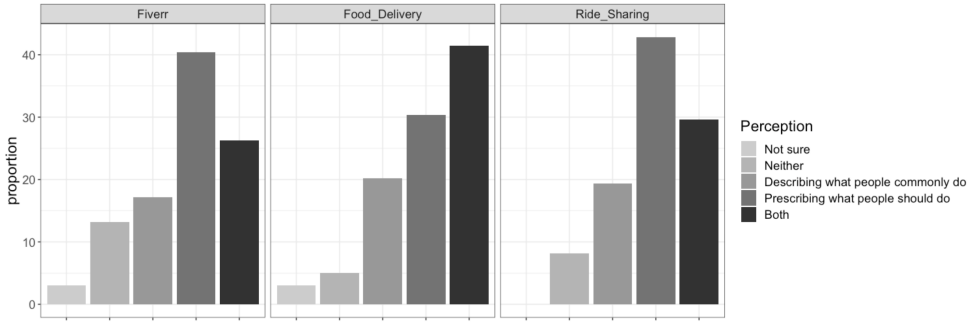
(a) Injunctive/ Active/ Buyer



(b) Injunctive/ Passive/ Buyer



(c) Injunctive/ Active/ Seller



(d) Injunctive/ Passive/ Seller

Figure B.28. Perception about Injunctive Messages

Bibliography

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

Allcott, H., Gentzkow, M., and Song, L. (2022). Digital Addiction. *American Economic Review*, 112(7):2424–2463.

Anderson, M., Mcclain, C., Faverio, M., and Gelles-Watnick, R. (2021). The State of Gig Work in 2021.

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

Azar, O. H. (2011). Business strategy and the social norm of tipping. *Journal of Economic Psychology*, 32(3):515–525.

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

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social Preferences and the Response to Incentives: Evidence from Personnel Data*. *Quarterly Journal of Economics*, 120(3):917–962.

Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). Cem: Coarsened Exact Matching in Stata. *The Stata Journal: Promoting communications on statistics and Stata*, 9(4):524–546.

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

Brynjolfsson, E., Li, D., and Raymond, L. R. (2023). Generative AI at Work.

Burtch, G., Lee, D., and Chen, Z. (2023). The Consequences of Generative AI for UGC and Online Community Engagement.

Chandar, B., Gneezy, U., List, J. A., and Muir, I. (2019). The Drivers of Social Preferences: Evidence from a Nationwide Tipping Field Experiment.

Chen, J., Xu, A. J., Rodas, M. A., and Liu, X. (2023). Order Matters: Rating Service Professionals First Reduces Tipping Amount. *Journal of Marketing*, 87:81–96.

Cialdini, R. B. (2009). *Influence: science and practice*. Pearson Education, Boston, 5th ed edition. OCLC: ocn227205869.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6):1015.

Cialdini, R. B. and Trost, M. R. (1998). Social influence: Social norms, conformity and compliance.

Conlin, M., Lynn, M., and Ted, O. (2003). The norm of restaurant tipping. *Journal of Economic Behavior Organization*, 52:297–321.

del Rio-Chanona, M., Laurentsyeva, N., and Wachs, J. (2023). Are Large Language Models a Threat to Digital Public Goods? Evidence from Activity on Stack Overflow. arXiv:2307.07367 [cs].

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., and Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality.

Deutsch, M. and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology*, 51(3):629.

Donkor, K. (2021). The Economic Value of Norm Conformity and Menu Opt-Out Costs. *SSRN Electronic Journal*.

Duhaime, E. P. and Woessner, Z. W. (2019). Explaining the decline of tipping norms in the gig economy. *Journal of Managerial Psychology*, 34(4):233–245.

Eisfeldt, A. L., Schubert, G., and Zhang, M. B. (2023). Generative AI and Firm Values.

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

Hadero, H. (2023). Is tipping getting out of control? many consumers say yes.

Haggag, K. and Paci, G. (2014). Default tips. *AMERICAN ECONOMIC JOURNAL: APPLIED ECONOMICS*, 6:1–19.

Hanbury, M. (2019). Instacart workers are fighting back against a policy change they say drastically cut their wages. *Business Insider*.

Hu, K. (2023). ChatGPT sets record for fastest-growing user base - analyst note. *Reuters*.

Hui, X., Reshef, O., and Zhou, L. (2023). The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market.

Kerr, P. M. and Domazlicky, B. R. (2009). Tipping and service quality: results from a large database. *Applied Economics Letters*, 16:1505–1510.

Kerschbaumer, K. (2005). Dvrs not so hot? espn study shows they're not for everyone. *Broadcasting and Cable*, page 26.

Li, P., Castelo, N., Katona, Z., and Sarvary, M. (2024). Frontiers: Determining the Validity of Large Language Models for Automated Perceptual Analysis. *Marketing Science*, 43(2):254–266.

Lu, S., Yao, D., Chen, X., and 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.

Lynn, M. (2006). Tipping in restaurants and around the world. In Altman, M., editor, *Handbook of contemporary behavioral economics: Foundations and developments*, page 626–643. M.E. Sharpe Publishers.

Lynn, M. (2015). Service gratuities and tipping: A motivational framework. *Journal of Economic Psychology*, 46.

Lynn, M. (2016). Motivations for tipping: How they differ across more and less frequently tipped services. *Journal of Behavioral and Experimental Economics*, 65:38.

Lynn, M. (2017). Should us restaurants abandon tipping? a review of the issues and evidence. *Psychosociological issues in human resource management*, 5:120–159.

Lynn, M. (2021). The effects of injunctive and descriptive tipping norms on tipping behavior and motives. *Journal of Behavioral and Experimental Economics*, 95:101786.

Lynn, M. (2022). Do tip percentages affect server job tenure, or vice versa?: Evidence from a panel dataset. *Compensation Benefits Review*.

Lynn, M., Jabbour, P., and Kim, W. G. (2012). Who uses tips as a reward for service and when? an examination of potential moderators of the service–tipping relationship. *Journal of Economic Psychology*, 33:90–103.

Lynn, M. and Sturman, M. (2010). Tipping and service quality: A within-subjects analysis. *Journal of Hospitality Tourism Research*, 34:269–275.

McClain, C. (2024). Americans' use of ChatGPT is ticking up, but few trust its election information.

Melnyk, V., Carrillat, F. A., and Melnyk, V. (2022). The influence of social norms on consumer behavior: A meta-analysis. *Journal of Marketing*, 86:98–120.

Miranda, D. and Main, K. (2023). U.s. digital tipping culture in 2023 – forbes advisor.

Newman, A. (2019). DoorDash Changes Tipping Model After Uproar From Customers. *The New York Times*.

Robinson, E., Fleming, A., and Higgs, S. (2014). Prompting healthier eating: testing the use of health and social norm based messages. *Health Psychology*, 33(9):1057.

Zipperer, B., McNicholas, C., Poydock, M., Schneider, D., and Harknett, K. (2022). National survey of gig workers paints a picture of poor working conditions, low pay.