Tip Nudging: A 3-Part Experimental Analysis of Influencing Tips Through Technological Default Menus

Tip Nudging:
A 3-Part Experimental Analysis of Influencing Tips through Technological Default Menus

A thesis presented

by

Dario Omid Zarrabian

To

Applied Mathematics

in partial fulfillment of the honors requirements

for the degree of

Bachelor of Arts

Harvard College

Cambridge, MA

March 29, 2019

Abstract

Restaurants and services are increasingly using technologically innovative point of sale iPad systems that offer default tipping amounts in their payment services. In this paper, I study the impact of these default tip suggestions through a behavioral economics lens. I first provide a meta-analysis of research on psychological tactics that waiters can use to increase tips. I then examine the effect that default suggestions have on consumer tipping using a series of experiments and a quasi-experiment. First, I show that higher default options anchor customers at higher tipping amounts. Second, I providence evidence that whiplash effects through reactance exist for anchoring customers at overly large amounts, creating negative downstream effects and feelings towards the restaurant. Third, using data from the field, I show that higher suggestions can induce higher tips in a fast-casual field setting, adding over 12% in tips to a restaurant in Cambridge, MA. The first two results are based on lab experiments and the last is based on a regression discontinuity design. Finally, further channels of research and implications are discussed.

Acknowledgments

I would like to deeply thank my thesis advisors, Professor Joshua Schwartzstein and Dr. Gregory Bruich for their incredible guidance and support throughout this research. This project and paper would not have been possible without all the constructive feedback, advice, and profound knowledge of the field. Thanks to Michael Norton, Alison Wood Brooks, and Leslie John for overseeing my work at Harvard Business School in Summer 2017. I am very grateful for all the help and feedback from Grant Donnelly, Oliver Hauser, Serena Hagerty, Ximena Garcia-Rada, and Annie Wilson, who all made my experiments possible thanks to their mentorship and guidance. Finally, thank you to Neema, Mom, and Dad for all the loving support over the years.

Contents

List of Figures

List of Tables

1 Introduction

A large body of work in behavioral economics and experimental psychology has highlighted the importance of default options in economic decisions. The effects of increasing defaults and anchoring customers at higher amounts have been documented in many different contexts, including organ donations, retirement savings in 401(k) accounts, charitable donations, and healthy eating options at Walt Disney World (Goswami & Urminsky, 2016; Peters et al., 2016; Johnson & Goldstein, 2003; Madrian & Shea, 2001). Evidence on the impact of defaults in customer tipping is more limited. Haggag and Paci (2014) applied the framework of anchoring through default menus to tipping in taxi cabs and provided evidence that riders were more likely to tip more with higher default menus provided, even though the choice set remained unrestricted. As tipping in the United States food industry alone amounted to $46.6 billion in 2011, small changes in consumer tipping behavior can have large-scale revenue and profitability impacts for restaurants and businesses, similar in magnitude to sales tax revenues (Azar, 2011).

This paper studies tipping within the restaurant industry, using experiments and quasi-experiments to shed light on the impact of recent technological innovations through point of sale systems, such as Clover, ShopKeep, Revel, and Square, on customer behavior. In particular, I examine how anchoring through default menus offered by suggested tips both at the beginning and end of meals in restaurants and food service settings affects the overall tip amounts, primarily on iPads but also through paper checks. I show new evidence that highlights restaurants' ability to nudge customers at fast-casual restaurants in a rapidly shifting food industry context, greatly affected by the recent innovation of technology and by the lack of standards in the tipping industry.

Restaurant owners typically provide the option of tipping so that they can reduce the hourly wages for their employees through a pay for performance structure. While the amounts tipped are influential factors, and even sometimes the majority of earnings for the 1.4 million U.S. food servers, the United States along with the rest of the world still has no normalized standard for tipping (Galante 2012; Martinez-Moncada, 2011; Seiter, Brownlee, & Sanders, 2011). Tips vary from completely optional to an expectation of 20% around the world, and even the United States has varying norms based on the industry (Kane, 2014). The differing expectations and lack of standards in the food industry have allowed waiters and managers to slightly shift consumer behavior in order to earn more money.

Cornell University's School of Hotel Administration professor and tipping expert Michael Lynn (1996) has conducted research on manipulations and tactics that waiters and waitresses can use to increase tips that span from introducing themselves by name and squatting down near eye level with the customer to writing "thank you" and gently touching customers while delivering the check. While these strategic implementations are performed by servers, restaurants can similarly affect human behavior and drive tipping amounts. However, there have been a limited number of studies examining the effect that suggestions have on tipping amounts, and even less existing research analyzing the effect of Brehm's reactance theory (1966) on tipping, which predicts that individuals take control and reassert themselves when their choices are eliminated or threatened.

The purpose of this paper is to quantify the impact of technological nudges from the higher-level perspective of companies on tips, complementing the psychological tactics that Lynn (1996) has researched. In order to test the methods and effect of anchoring and default options on tip suggestions, I design and implement two online experiments and conduct a quasi-experimental analysis of real world tipping behavior at a fast-casual restaurant in Cambridge, Massachusetts. I examine the impact that suggested tips have in traditional, old-fashioned recommendations at the end of paper bills through anchoring and focus on the effect of the more newly introduced iPad checks through default menu options. The first two studies were run through Amazon's Mechanical Turk platform to test the effect of shifting the recommendations for $20 and $40 checks, respectively. Since tipping is based on norm-driven behavior, suggested amounts included a control group with no suggestions provided, a minimum of 15%, and a maximum of 50% to observe participants' reactions.

In the randomized studies, the two tested scenes were suggestions with a write-in box and default options with buttons. Following the analysis of Haggag and Paci (2014), I call the tip suggestions that are presented to customers default options and default menus, since the choices provided by these companies are strongly effort advantaged and salient.

Results from the two randomized studies provide evidence that larger suggested amounts anchor participants at higher values and thus generate larger tips. In study one, the high suggested options provided amounts that had a 33.7% increase compared to the low options. Similarly, default options rather than provided fill-in boxes created a larger magnitude for the effect, where high suggestions drove tips significantly higher and low suggestions had a larger margnitude for the initial effect.

While revenue increased in the first online experiment, evidence of certain backlash effects was found in study two, as Brehm (1966) may predict. Participants stated that they were less likely to rate high-tip-asking restaurants as fair or appropriate, less likely to recommend the restaurant to a friend, and less likely to give a high review on Yelp; however, this group of participants still gave more money to the restaurant compared to those receiving lower suggestions. Therefore, customer satisfaction may decrease as a result of participants' negative feelings towards suggestions of unexpectedly high amounts of money on top of food costs and taxes.

A common critique of laboratory studies is that lab conditions cannot mimic real life scenarios. The third study addresses this limitation by using observational data from the field, where customers have real money on the line. I find that in a fast-casual setting with minimal services provided, customers are still anchored by tip suggestions. Specifically, one in three customers tip even though no tip is required, and the scene has historically been within a non-tipping industry. I exploit the discrete change at the $10 threshold from suggesting $1, $2, and $3 to suggesting $15%, 20%, and 25% (nearly equivalent to $1.50, $2.00, and $2.50) to show an addition of around $.07 on average per tip using an RD design, equivalent to a 12% increase over the average tip of $.58. While the means of the suggestions are the same, a small shift in the lowest suggestion has a positive effect on tips received.

This paper presents findings that, on top of waiters' psychological strategies to influence customers' experiences and tip donations, restaurants can alter customers' tips by anchoring their norm- and guilt-driven behavior through tipping suggestions. These results contribute to the behavioral economics literature. As governments are beginning to regulate tipping more seriously through the 2016 federally enacted Fair Labor Standards Act that helps control for worker wages, the research raises an important question: should default tipping options be regulated? In this paper, I focus on documenting empirical facts, and leave normative implications to policy makers and government officials. The evidence I provide through these experiments and quasi-experiments should be a key input for them in making these decisions.

In the next section, a background of the tipping industry is provided, including related studies on the default and anchoring effects and a meta-analysis of waiters' influencers of tip amounts. Next, I present the data, methods, and findings from studies one and two. I discuss limitations of scope that the results of the first two online studies. I then introduce the data and methods for study three, display summary statistics, and describe the identification strategy and results of the quasi-experiment. Finally, I summarize the results of the analysis and provide future directions of research.

2 Background and Motivation

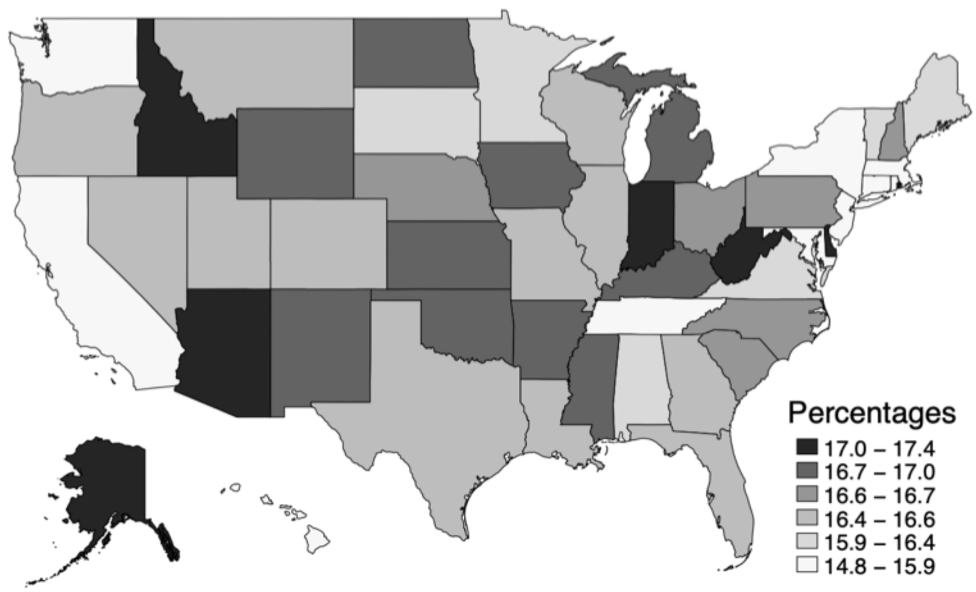
In the following section, I begin by discussing the current lack of standards in the restaurant tipping industry . Historical explanations for evolving social norms of tipping today are provided to ground a background in tipping history. Because of these varying expectations, waiters have been shown to positively manipulate customer tips, and an updated meta-analysis, adapted from research performed by Michael Lynn (2005), is displayed. Finally, I discuss the limited research performed in the field of tip suggestions, where managers and companies may have the ability to gain higher tips through the recommendations provided on checks.

2.1 The Tipping Industry's Lack of Standards

In the past few years, Americans spent around $20 per day on average on food, including about $3,100 per year on food outside of the house (Bloom, 2017; Bureau of Labor Statistics, 2017). While food can be purchased at different locations, such as grocery stores or farms, a big section of the industry today consists of spending at restaurants, especially within cities (Pickert & Lanman, 2018). Food and drink remain the biggest source of revenue for these restaurants, but an often-overlooked metric is tipping, traditionally at the end of sit-down meals. In the food industry alone, Azar (2011) estimated that tips amounted to nearly $47 billion in 2009, and this likely underestimated figure was expected to continually rise.

Even with these high figures, however, there remains no standard tip amount for service within the food industry. Internationally, a vast variation in averages and expectations of tips exists, with Brazil levying a 10% service charge, but expecting no tip, Egypt expecting a 5-10% tip, Japan serving as an almost entirely non-tipping society, and India customarily seeing 10-15% in tips (Galante 2012; Martinez-Moncada, 2011). The United States and Canada are the leaders in tip amounts, as most services for food expect 15-20% for tips. No law or legislature states the expected or required tip amounts, but as cultures have evolved, the United States' restaurant industry has continued to gain additional money after the meal compared to other countries; however, even within the United States, there are discrepancies in tipping. Figure 1 on the following page shows data aggregated by the point of sale system Square in 2014 of tips in the United States broken down by state (Michaels, 2018). Average tips can differ by up to 3.6% by state, leaving a situation where employees earn different tip amounts solely based on local standards.

Figure 1
Average Tip Percentage by State



Note: Figure 1 shows the average tip percentage by state in 2017 in the United States based on data collected by point of sale system Square. The state with the smallest amount tipped is Hawaii, shaded the lightest, at 14.8% for the average tip, and the state with the largest percentage is Idaho, shaded the darkest, at 17.4% of the order per tip.
Source: Michaels, 2018

2.2 Tipping Norms: Past and Present

As the trends above show, the United States has high tips and very different norms and standards around tipping that can partially be explained by the historical and evolutionary norms of tipping, as performed in the following section. Most historians attribute the origin of tipping to sixteenth and seventeenth century England, where urns were placed in both pubs and coffee houses with the label T.I.P., standing for "To Insure Promptitude" (Brenner, 2001). Visitors to private houses in England in the 1700s gave sums of money, called vails, for additional services from servants (Azar, 2004). By 1760, the expectation was that any service from servants was worthy of a tip. While there was discussion to abolish vails in 1764, no agreement was achieved, and guests in the early twentieth century in England were expected to give around $100 for a one-week stay.

Tipping did not begin in the United States until after the Civil War due to the lack of a servant class (Segrave, 1998). Wealthy Americans, who had traveled to Europe and saw their customs, wanted to exude their own advanced culture in the states. In the late 1800s, tipping was more officially established and media, including etiquette books, movies, and newspapers described the custom. In 1910, 5 million workers, or 10% of the labor force in the United States, received tips estimated to total $200M-$500M a year (Azar, 2004).

However, there still were many hold-outs for service tipping. Providers of services punished known non-tippers by marking their luggage with chalk so that other hotels do not serve them, providing slow service for customers who returned, and even putting powder in the food of customers, which ultimately led to the arrest of waiters (Azar, 2004).

In 1895, expected tips were already separated by location, as they were 5% on average in Europe and 10% on average in the states, even though U.S. waiters earned much higher incomes (Segrave, 1998). The large tips in the U.S. allowed some employers to charge a rent privilege or working fee for waiters who received high tips, to take all the tips from employees, or to give employees no hourly wage so that tips were their only source of compensation. No legislature or national regulation was in place to control for wages at the time. However, the argument for abolishing tipping began in 1764, and murmurs have existed since. Many debates and questions internationally about the standard around tipping have been raised: Should tipping be abolished? Should the U.S. follow other countries institute mandatory service charges? Should the U.S. have some sort of institutionalized, uniform requirements?

There are no current answers to these questions in the U.S., as the division by state in Figure 1 highlights. Without any answer, customers are left questioning why, how, and when to tip. Azar (2004) identified three main reasons that people have tipped throughout history: selfish economic rationale, social norms, and altruistic feelings. He argues that when no social norms existed at the start of history, customers tipped with altruistic motivations to show gratitude to workers who provided services beyond what was asked.

However, tip motivations have changed over time. While the price of a meal already includes service costs, tip compensation is now given to a worker for conventional services. Tipping is no longer an optional, altruistic idea. The main motivation now deals with social norms and avoiding embarrassment, at least in restaurant settings (Azar, 2004; Conlin, Lynn, & O'Donoghue, 2003; Segrave, 1998). However, benefits do exist for all parties: servers receive more money in a socially efficient world, managers bring in higher total revenue and can keep more money for themselves by slightly reducing workers' wages, consumers receive better service through servers' attempt to gain more tips, and society improves through the tit for tat exchange of more money for what is supposedly better service (Conlin et al., 2003). Similarly, customers with higher incomes can add higher tips in order to slightly reduce income gaps between less fortunate waiters and the upper class.

The government began to regulate tips recently due to the historical ability of managers to limit wages and even take tips from employees. Under the federally enacted Fair Labor Standards Act, employers can now pay employees who receive at least \$30 in tips a month as low as \$2.13 an hour, as long as the employer is also giving a tip credit of no more than \$5.12 an hour to achieve the minimum wage of \$7.25 (United States Department of Labor, 2016). Employees are required to maintain all their tips "except to the extent that they participate in a valid tip pooling or sharing arrangement" (United States Department of Labor, 2016). One can analyze the effect of tipping on wages and employment by applying the standard partial equilibrium tax incidence analysis from public finance. In that analysis, the labor supply curve would shift out leading to excess supply of workers at the initial (no tipping) wage level. To clear the market, employers could cut pretax wages until the market clears. As long as the elasticities of labor demand and supply are non-zero and finite, the economic incidence would be shared by both employers and employees. Unlike a tax, customers can decide how much to tip, and this is uncertain ex ante. Therefore, this analysis applies to the expected tipping behavior. Since a 1% increase in tipping would provide added industry revenue of over $466 million using the 2009 figure, any slight shift in the industry could be beneficial in helping waiters, waitresses, and restaurant owners (Azar, 2011).

While equilibria in tipping can be achieved through consumer giving and appropriate incidence between employer and employee, tipping fails when a social contract, needed to evaluate tipping efficiency and devise an appropriate compensation scheme, does not follow through. Conlin et al. (2003) assume a hypothetical service contract treats all parties as risk neutral and induces customers into paying more for higher service quality. They find through survey data that tips increase based on higher service quality, implying a basically efficient service contract. However, this contract isn't completely efficient as repetition of customer, day of week, age, group size, gender interactions, and the frequency of one's visits all impacted the amount tipped. In addition, waiters' roles in altering tips are highlighted through a meta-analysis in the next section.

In addition to the transformation of the purpose and motivations of tipping over the past centuries, the methods and industries of tipping have vastly changed in the past five years. The continuous introduction of tips and appreciation of services among multiple different platforms has also provided a disruption of the industry, and customer behavior will continue to adapt in an unpredictable, but significant manner. For example, the institution of the Square application and payment method has made tipping more implementable in a variety of locations (Swan, 2018). The application has also begun to gain tips in historically non-tipping industries, such as coffee shops and fast-casual restaurants. As more customers continue to pay with cards and mobile apps instead of cash, tipping is expected to increase due to the distance from a feeling of giving money (Prelec & Simester, 2001).

The market of mobile payment apps, which includes Square, Google Pay, Apply Pay, and Venmo, is expected to grow at a CAGR greater than 30% to 2023 to a valuation of over $4 trillion (Swan, 2018; Mobile Payment Market, 2017). The scale and spread of money transferred through mobile payment systems continues to grow along with introduction of other new systems, such as Clover, ShopKeep, and Revel in addition to Square and other point of sale technology. While these applications make money more easily transferable than ever, they can confuse customers even more about tipping expectations and norms, providing an opportunity for this paper to analyze how people act in these uncertain scenarios (Swan, 2018). Coupling uncertain consumer behavior due to a lack of industry standards with progress in technological advancement that makes tipping more salient can produce major growth in the tipping industry.

2.3 The Role of Waiters in Altering Tips

Even before the addition of advanced point of sale systems into shops and restaurants, different psychological nudges and manipulations by waiters have been shown to positively impact the consumer tips. The meta-analysis in Table 1 is taken and updated from research performed by tipping expert Michael Lynn (2005), a professor at the Cornell School of Hotel Administration. The meta-analysis provides some research performed over the past 45 years in restaurants. While these implementations may have found statistically significant results in their respective time frames, the papers are limited in their external validity. Nonetheless, evidence shows that waiters may be able to influence tips through slight changes in behavior. The following magnitudes show percentage increases rather than more incremental percentage changes.

Table 1
Meta-Analysis of Waiter Implementations to Increase Tips

Action	Description	Percentage Increase
Verbal		
Calling Customer by Name	Addressing the customer by name when returning the check as opposed to “Mr.,” “Mrs.,” “sir,” or “ma’am” has been shown to increase tips (Seiter & Weger, 2013).	10-19%
Introducing Self by Name	If the server begins the meal with an introduction by name, research has shown that amounts of people tipping in addition to average percentages tipped both increase (Garrity & Degelman, 1990).	53%
Repeating Customers’ Orders	Repeating customers’ orders verbatim through mimicry rather than paraphrasing or simply acknowledging them increased tips in a Netherlands field study (Van Baaren, 2005).	69%
Check- Related		
Giving Candy	Offering candy, and specifically a large amount of candy, increased tips in a New York restaurant (Strohmetz et al., 2002).	18%
Drawing a Picture of the Sun	Drawing on the effect of smiley faces and good weather, Gueguen and Legoherel (2000) showed that drawing a sun on the bottom of customers’ checks similarly increased customer tips.	37%
Telling a Joke to the Customer	Bringing a simple joke card along with the check increases both the percentage of customers tipping and the average percentage per tip (Gueguen, 2002).	40%
Drawing a Happy Face	Rind & Bordia (1996) obtained evidence showing that female servers could raise tips by drawing a smiley face on the back of the check unlike male counterparts.	19% (females)
Writing “Thank You” on the Check	Rind & Bordia (1995) randomized an experiment where the server wrote (a) nothing, (b) “Thank you,” or (c) “Thank you” and the credit card owner’s name for 51 dining parties. Writing “Thank you” alone was shown to positively impact tips.	11%
Physical Appearance		
Touching the Customer	Lynn, Le, & Sherwyn (1998) found waiters influenced tips when physically interacting with customers who asked for the check with a brief, 2-second touch, and a prolonged, 4-second touch.	29%
Smiling	Giving customers a maximal rather than minimal smile influenced customers’ tips and smiles at the end of the experience (Tidd & Lockard, 1978).	140%
Squatting Down	Waiters and waitresses receive more tips on average when squatting down near the table due to the perceived psychological closeness effects (Lynn & Mynier, 1993).	24-28%

Note: Table 1 is a meta-analysis of the main literature around increasing tipping, focusing on waiters’ ability to psychologically influence customer tipping. However, these studies have mostly been performed in singular locations and show the percent change rather than the percent increase. The analysis is modified from the analysis performed by Michael Lynn over 10 years ago.
Source: Lynn, 2005

Most of the existing literature around altering tips revolves around waiters' ability to influence tips. However, other behaviors, such as providing good service due to quick turnaround times, working more strategic shifts, obtaining customers who tip higher amounts, and receiving higher total bills are omitted from the meta-analysis either due to a lack of research or one-off studies that are unrelated to waiters' locus of control.

This meta-analysis shows that tipping can be influenced through subtle tactics. However, a less studied field is the impact that companies and restaurant owners can have in helping their employees receive higher tips, specifically through tip suggestions. The next section and following analysis of tip suggestions are based on two mechanisms: the anchoring effect and default options. These two mechanisms have been introduced in other fields and complement each other, as described below.

2.4 Company Manipulations Through Default Options and the Anchoring Effect

The anchoring effect introduced by psychologists Tversky and Kahneman is based around shaping others' behavior (1974). This psychological tactic nudges people to be impacted by their initial or starting values, so setting those values externally can impact others' decisions. Tversky and Kahneman (1974) found that after spinning a wheel with a number from 0 to 100, individuals that received a 10 predicted that there are 25 African nations in the UN on average, and people that received a 65 predicted 45 African nations on average. This random spin and the true number of African nations evidently have no correlation, but even with financial incentives involved, participants did not improve their accuracy. Other general fields with anchoring include influencing the boiling point on Mt. Everest and the year that George Washington was elected President, increasing how much people spend on wine and books, and gaining more money in charitable donations by asking for $400 rather than $5 (Ariely, Loewenstein, & Prelec, 2003; Epley & Gilovich, 2001; Tversky & Kahneman, 1974).

The literature of anchors' effect on consumer tip behavior is limited and slightly conflicting. Rind and Strohmetz (2001) ran a randomized field study with 120 sit-down meals that used the anchoring effect applied in the tipping field. In this experiment, a server gave either (a) a tip card with 19 rows that had guidelines of "15% for adequate service", "20% for better-than-average service", and "25% for outstanding service" for increments between $10 and $100 or (b) no tip card to customers paying for the check. They found that these "tip cards" had no impact on the mean of the amount tipped. However, the suggestions reduced the variability of tips in this study.

A follow-up online study performed by Karniouchina, Mishra, and Verma (2008) examined tip guidelines with differing service quality in a hypothetical scenario. The guidelines were either (a) a control with nothing provided, (b) an educational statement, such as "Quality service is customarily acknowledged by a gratuity of 15% to 20%," or (c) a calculation assistant that displayed the 15% and 20% values with conversion rates. The authors found a significant effect of providing suggestions, with calculation-assistance being the most significant for increasing tips.

While Rind and Strohmetz (2001) found that helpful suggestions had no impact on means and Karniouchina et al. (2008) found that calculation assistance positively impacted tips in an online scenario, Seiter et al. (2011) sought to implement the calculation assistance in the field. 113 participants received a check either with no suggestions on the bottom or 15% and 20% calculations of the check from the printing software. Customers who had the calculation assistance tipped 2.5% higher than those who did not, equivalent to a 15% increase. Calculation assistance may positively affect tips received due to anchoring customers at a higher amount unlike a guideline helper, which explains the conflicting results that Rind and Strohmetz (2001) provided.

Related to the anchoring effect is the idea of defaults: pre-set or easy-to-choose options. By showing easily selectable options, people have demonstrated automatic process thinking by choosing these options more often than spending extra time for more optimal solutions. By providing opt-out rather than opt-in policies for being an organ donor, the percentage of organ donors was around 90% rather than 10% (Davidai, Gilovich, & Ross, 2012; Johnson & Goldstein 2003). Studies have similarly shown that automatic enrollment can significantly increase 401K participation (Madrian and Shea, 2001; Choi, Laibson, Madrian, & Metrick, 2004). In both scenarios, people were not motivated enough to shift out of their already established positions.

As tipping has evolved over the past years, default options have been integrated in taxi cabs, replacing the more typical cash payments. A field study performed by Haggag and Paci (2014) examined quasi-random taxi cab trips from the airport with two companies who had different tipping recommendations: one provided 15%, 20%, and 25% options in all circumstances, and the other asked for \$2, \$3, or \$4 for rides under \$15 and 20%, 25%, or 30% for rides that are more expensive. They used a regression discontinuity design to show that higher suggestions in terms of percentages drove overall tips.

While tip suggestions may not seem entirely equivalent to defaults, Haggag and Paci (2014) argue that the suggestions offered are salient and are strongly effort advantaged. Therefore, instead of using terms such as "enhanced active choice" options, they call these suggestions default options. This paper uses the term default options and menus to emphasize that customers need to deliberately and actively avoid the salient options that are provided in order to enter personal amounts on sometimes even separate screens.

Haggag and Paci (2014) similarly show that the higher suggestions increased customer likelihood of tipping \$0 by about 50%. A potential explanation for this increase is Brehm's psychological reactance theory (1966). He argues that when individuals' freedom or choices are eliminated or threatened, they attempt to reassert themselves and take control, as external threats cause an unpleasant heightened motivational state that induces cognitive and behavioral assertion to re-establish freedom. In this situation, taxi riders may feel manipulated into tipping amounts that are too high and they thus react adversely by attempting to punish the cabs.

Reactance theory has been analyzed in a variety of situations and fields, including health, education, income, and politics. In tipping, Rind and Strohmetz (2001) find no evidence of reactance theory with the helping suggestions, while Haggag and Paci (2014) see the percentage of customers tipping $0 go from 1.7% to 2.8% with higher suggestions. However, very little other research exists about the trade-off between higher tips due to anchoring and lower tips due to reactance, and this paper seeks to provide initial evidence in showing the effect of unnecessarily high tips.

While the above findings have shown that anchoring through defaults can impact taxis, the added value of default options has not been discussed in the restaurant industry, potentially due to the very recent addition of technology for payment in the previous decade. Tipping has historically been for services provided within different industries. The Bureau of Labor Statistics identifies employees of these sectors that are most likely to be tipped, including restaurants, bars, casinos, hotels, beauty salons, valet, ride-sharing platforms, and other related services (Simpson, 1997). Today, more and more services are beginning to be tipped than before. Coffee shops, takeout restaurants, ice cream shops, and hotels, have historically not requested or been accustomed to tips until recent years. The shifting expectations and higher standards among employees and companies create the potential to drive the industry even further than its $47 billion in only the food tipping industry (Azar, 2011). While certain companies and platforms are expecting sizeable increases in their earnings, customers continue to remain confused about expectations and standards around tipping.

In the next section, I discuss the methodology of the 3-part experimental analysis and results from the analysis of changing recommendations for suggested tips.

3 Experiments

In this paper, I conduct two lab experiments and study a natural experiment using observational data to shed light on the effect that tip suggestions have on consumer tipping and behavior. In the following sections, I describe the data and methods used for study one on suggested tip amounts, study two on downstream effects from tipping, and quasi-experimental study three analyzing observational field data. Studies one and two were run through Amazon's Mechanical Turk platform to examine if an initial effect exists for increasing tips. While randomized data from these separate between subjects' experiments is given, limitations through mTurk do exist and are discussed in order to motivate study three. Study three is a quasi-experiment using a regression discontinuity design with data from a restaurant in Cambridge, MA. The purpose of study three is to estimate the causal effect of suggested tipping amounts in a legitimate and authentic field situation.

3.1 Study One: Anchors, Defaults, and Tipping

3.1.1 Data and Methods

To measure if increasing suggested tipping amounts has an initial effect on the amount given, I recruited participants and rewarded them $.20 for completing an approximately two-minute online study. 815 participants responded to this survey, but 21 data points are excluded in the following analysis, as three quit the survey before completing the opening questions, and 18 others provided tipping amounts that were unrealistic. I define "realistic" tipping amounts as percentages between zero and fifty of the full check's value, but these 18 provided unrealistic values that ranged from 75% to 500% of the bill and may be attributed to a mis-interpretation of the study or a lack of attention. These unrealistic amounts were spread out across the 4 different conditions described below, so excluding them did not impact the overall effect of the analysis. The following data analyzes the 794 participants who remained (51.2% male, $M_{age} = 34.9$, $\sigma_{age} = 11.2$).

Participants were told that they had just sat down alone for a one-hour meal with good service. They were asked how much they would tip on a $20 check based on one of the eight randomly-assigned conditions. The study followed a 2x4 structure, where both the type of check and the suggestions varied. The two possibilities for the type of check were paper and iPad. Similar to a restaurant, paper checks offer suggestions in writing, but there is a line where the participant has to write in their tip amount, whether or not it mirrors a suggestion. For the iPad check, customers are given four boxes: three of which are default values that can be clicked, and one of which is a box labeled "other," where the customer can alternatively write any amount (See Appendix A). By pressing one of the three default values, the customer is automatically committing to tipping that amount, similar to taxis and point of sale systems (see Appendix C). In both scenarios, the customers receive three suggestions if they are not in the control group, but still have the full choice set to tip any realistic amount.

The four possibilities for the tip suggestions were control, low, middle, and high. On bills, three suggested values are typically provided in restaurants, shops, and taxis; some attribute these three never-defined trifold values to bad service, adequate service, and great service, but the lack of industry standards has never officially defined these three (Rind & Strohmetz, 2001). The mTurk study aimed to mimic these options by providing three suggestions for each of the low, middle, and high conditions. The control condition had no suggested tipping amounts, so customers were not influenced by any anchors in the study. The low condition had the following suggested tip amounts: 15% = $3.00, 18% = $3.60, and 20% = $4.00. The middle condition provided 20% = $4.00, 25% = $5.00, and 30% = $6.00 for its suggested tips. Finally, the high condition provided 30% = $6.00, 40% = $8.00, and 50% = $10.00 for its suggestions. Each participant was randomly placed in both a type of check (paper vs. iPad) and suggestion (control, low, middle, or high) condition, giving eight possible options.

After participants entered their tipping amounts, they were asked a series of follow-up questions. They were asked to describe what influences how much they tip and to what extent they feel that they are influenced by suggested tipping amounts on a scale from one to seven. If they were not in the control condition, they were asked to what extent they feel as if they should tip at least the lowest amount that they were given. Finally, demographic questions about age, gender, ethnicity, and household income were asked.

I had two hypotheses for study one. First, increasing the suggested tipping amounts anchors customers at higher numbers and thus increases actual tipping amounts, as research has evidenced for purchasing decisions, charitable donations, and general knowledge (Ariely et al., 2003; Cialdini & Schroeder, 1976; Tversky & Kahneman, 1974). Second, providing defaults will cluster and funnel participants around these options, despite not limiting the choice set.

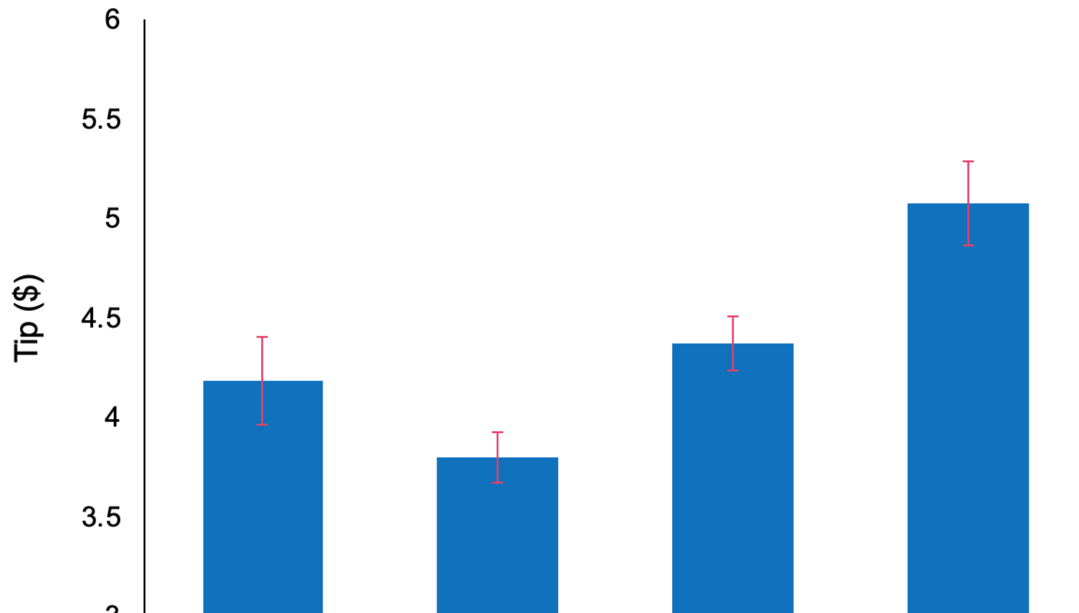
3.1.2 Results and Analysis

To analyze hypothesis one about the effect of higher suggested amounts, Figure 2 analyzes the difference between the different suggestion levels by pooling the iPad and paper conditions together.

Results from study one show that anchoring is prevalent and has at least some effect in driving consumer spending. On average, a 33% increase exists between providing low suggestions and high tip suggestions, pooling the scenes together. The low condition, which is typically displayed in restaurants and as default tips in some locations, has an even lower average than the control group. Figure 3 splits the sample by scene (iPad results represented by the dark bars and paper check by the light bars), with dollars as the y-axis and equivalent percentages labeled above each bar.

Figure 3 shows a trend that lower suggestion amounts may be driving slightly lower tips, while higher suggestions drive higher amounts. Broken down by scene, the iPad appears to have a larger magnitude compared to the control on driving the lower suggestions to smaller values and a highly statistically significant impact on driving higher suggestions to

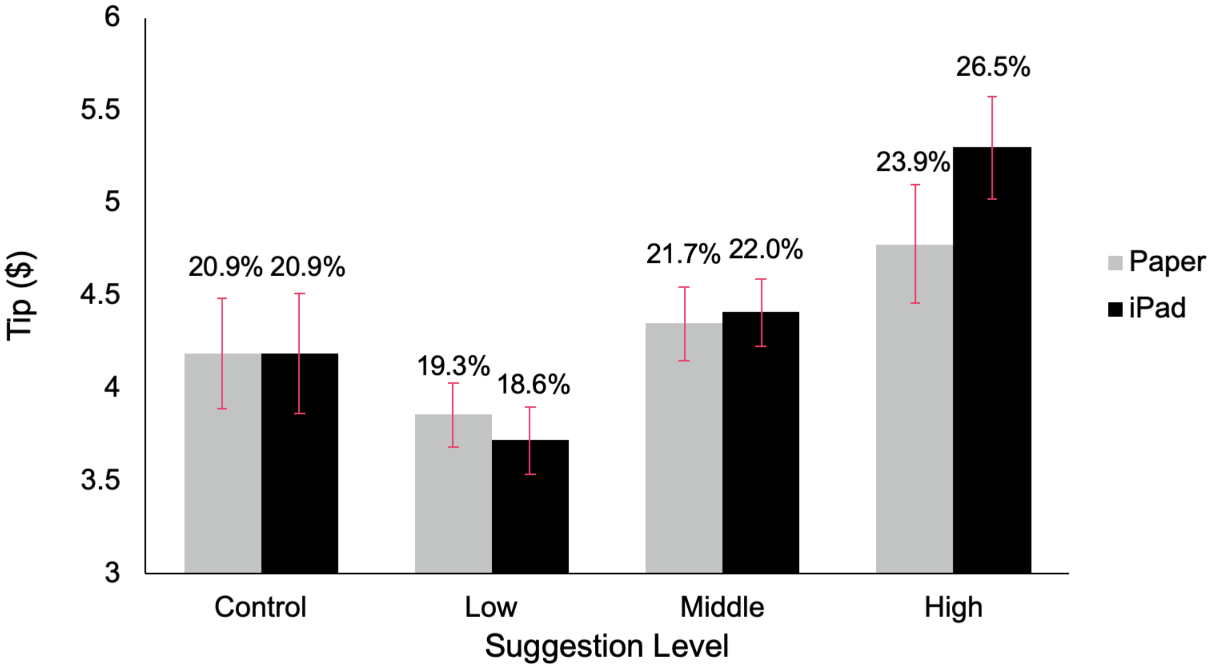
Figure 2
Tipping Amounts by Suggestion, Pooled by Scene



	Control	Low	Middle	High
Tip (%)	20.9%	19.0%	21.9%	25.4%
Tip ($)	$4.19	$3.80	$4.38	$5.08

Note: Figure 2 shows the average tip in dollars on the graph with percentage equivalents underneath, broken down by the suggestions in study one. Study one asks participants how much they would tip for a one hour-sit down meal that costs $20 with the following four treatment suggestions: Control: no suggestions; Low: 15% = $3.00, 18% = $3.60, 20% = $4.00; Middle: 20% = $4.00, 25% = $5.00, 30% = $6.00; High: 30% = $6.00, 40% = $8.00, 50% = $10.00. Customers are randomly placed into one of these four treatments. 95% confidence intervals are reported for the tipping amounts.

Figure 3
Tipping Amounts by Suggestion, Scene



Note: Figure 3 shows the average tip in dollars and percentage for each condition and scene. Study one asks participants how much they would tip for a one hour-sit down meal that costs $20 with the following four treatment suggestions: Control: no suggestions; Low: 15% = $3.00, 18% = $3.60, 20% = $4.00; Middle: 20% = $4.00, 25% = $5.00, 30% = $6.00; High: 30% = $6.00, 40% = $8.00, 50% = $10.00. 95% confidence intervals are reported for the tipping amounts. Initial graphical evidence shows that higher suggestions cause higher tips and vice versa, with a stronger effect existing with default options even larger values. To analyze the statistical results of the anchoring and defaults, I run regressions 1-3 below to see if statistical differences exist.

Equations 1 and 2 below highlight the effect of different suggested amounts for the iPad and paper conditions, respectively. Equation 3 pools these two together in order to measure the statistical effect that higher suggestions have, regardless of scene.

$$Tip\_iPad_i = \beta_0 + \beta_1 iPad_i * low_i + \beta_2 iPad_i * middle_i + \beta_3 iPad_i * high_i + \epsilon_i \quad (1)$$

$$Tip\_paper_i = \beta_0 + \beta_1 paper_i * low_i + \beta_2 paper_i * middle_i + \beta_3 paper_i * high_i + \epsilon_i \quad (2)$$

$$Tip\_pooled_i = \beta_0 + \beta_1 low_i + \beta_2 middle_i + \beta_3 high_i + \epsilon_i \quad (3)$$

The output in Table 2 shows that the high condition has a significant effect in both scenes, with a 3 percentage point increase in tipping in the paper condition and 5.6 percentage point increase in tipping on the $20 iPad check. Similarly, the low condition actually reduces tipping amounts for this small check by 1.6 percentage points on average for the paper condition and 2.3 percentage points for the iPad condition. The finding from earlier that the iPad condition has greater magnitudes in both directions is further supported. There is no statistically significant evidence that providing the middle suggestions of 20%, 25%, and 30% in this study alters consumer tipping, relative to the control of not providing recommendations. However, despite this minimal increase, study one provides strong evidence that different suggestions anchor individuals at different values to increase tipping, at least in this online setting, which is limited by the lack of financial incentives and the small value of the $20 check.

Table 3 analyzes the statistical differences between all the different potential conditions of tipping suggestions, pooled by scene. Results from Welch's unpaired t-test between the different pooled scenes are displayed. The analysis shows that every suggestion level has a significant difference from all others, except for that between the control and the middle condition.

Table 2
Suggestions' Impact on Final Tip by Scene

	(1)	(2)	(3)
VARIABLES	Tip ($)	Tip ($)	Tip ($)
Scene	iPad	paper	pooled
High	1.114***	0.604***	0.892***
	(0.216)	(0.224)	(0.156)
Middle	0.216	0.163	0.189
	(0.188)	(0.183)	(0.131)
Low	-0.465**	-0.327*	-0.388***
	(0.189)	(0.176)	(0.129)
Control Mean	4.189***	4.185***	4.187***
	(0.163)	(0.152)	(0.112)
Observations	403	391	794
R-squared	0.162	0.066	0.114

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 2 shows three separate regressions based on the suggestions compared to the control. The dependent variable in all three conditions is the tip amount, but the subset differs. Regression one analyzes the effect on the iPad scene with default options, regression two analyzes the effect on the paper scene with tip suggestions, and regression three looks at two scenes pooled together. The independent variables are indicators for condition type, and the regressions look at the effect on the tip based on the condition.

Table 3
T-tests Between Suggestion Levels

VARIABLES		(1) Difference	(2) Std. Error	(3) T	(4) P-Value	(5) DF	(6) 95% Confidence Interval	
Control	Low	0.388***	0.130	2.98	0.003	395	0.132	0.645
	Middle	-0.189	0.132	-1.44	0.151	400	-0.448	0.070
	High	-0.892***	0.156	-5.72	0.000	397	-1.198	-0.585
Low	Control	-0.388***	0.130	-2.98	0.003	395	-0.645	-0.132
	Middle	-0.578***	0.095	-6.11	0.000	393	-0.764	-0.392
	High	-1.280***	0.127	-10.11	0.000	390	-1.529	-1.031
Middle	Control	0.189	0.132	1.44	0.151	400	-0.070	0.448
	Low	0.578***	0.095	6.11	0.000	393	0.392	0.764
	High	-0.702***	0.128	-5.48	0.000	395	-0.954	-0.450
High	Control	0.892***	0.156	5.72	0.000	397	0.585	1.198
	Low	1.280***	0.127	10.11	0.000	390	1.031	1.529
	Middle	0.702***	0.128	5.48	0.000	395	0.450	0.954

*** p<0.01, ** p<0.05, * p<0.1

Note: Table 3 shows Welch's unpaired t-test for unequal variance with the differences between all the potential combinations of the suggestions, pooled with 794 total data points from the iPad scene and paper scene. Welch's test is used to compare the between subjects groups, who were each only shown one condition. The difference is computed by subtracting the average value of the second variable from the average value in the first column's variable. The standard error, t-statistic, p-value, and degrees of freedom are displayed along with the statistical significance of the difference.

While the distinctions between the different pooled conditions are shown above, differences also exist between the same suggestions in the different two scenes. Table 4 compares the effect of providing the same suggestions in the paper check versus the iPad check. While some marginal differences exist in the amount given based on the check provided, the only statistically significant difference is the high suggestion, providing evidence that in this online setting, default options do indeed have an even larger effect than paper suggestions for very large amounts. However, there is no significant evidence that low iPad suggestions drive tips significantly lower than low paper suggestions, even though the low iPad suggestion has statistically stronger evidence against the control condition than the low paper suggestion.

In order to analyze the second hypothesis that default menus and buttons funnel tipping behavior, I focus primarily on the iPad scene, as default options and choices are not provided in the paper condition. Regression 1 in Table 2 shows that the low amounts decreased average tips by 2.3 percentage points and increased high amounts by 5.6 percentage points. In order to further validate the idea that default options can shift consumer decisions, I examine the choices of individuals in each condition in Figure 4. It is important to note that the choice sets in these decisions are not limited, as participants have the ability to choose any amount from $0 to $15, even with the three provided options.

As the graph in Figure 4 shows, 32% of individuals in the low condition tip the option of $3, 20% of individuals pick the middle option of $3.60, 41% pick the high option of $4, and 7% pick some "other" manually-entered amount. However, if the defaults are switched to the middle suggestions, the number of people tipping $4 goes up dramatically. While in the low condition, 41% of people chose the $4 default option, 56% of participants in the middle condition chose the same $4 option. In the low condition, only 7% of individuals tipped more than 20% ($4.00), but in the middle condition, 37% of participants tipped over this same $4.00 threshold.

The most drastic shift happens between the middle and high conditions, even though the high options of up to 50% are rarely offered in field settings (Rind & Strohmetz, 2001).

Table 4
Tip Differences Due to Scene

	(1) Paper			(2) iPad			
VARIABLE	Mean	SD	N	Mean	SD	N	T-test
Control	4.19	1.49	96	4.19	1.68	106	0.017
Low	3.86	0.92	109	3.72	0.86	86	-1.03
Middle	4.35	1.02	100	4.41	0.93	100	0.417
High	4.79	1.52	86	5.30	1.49	111	2.38***

*** p<0.01, ** p<0.05, * p<0.1

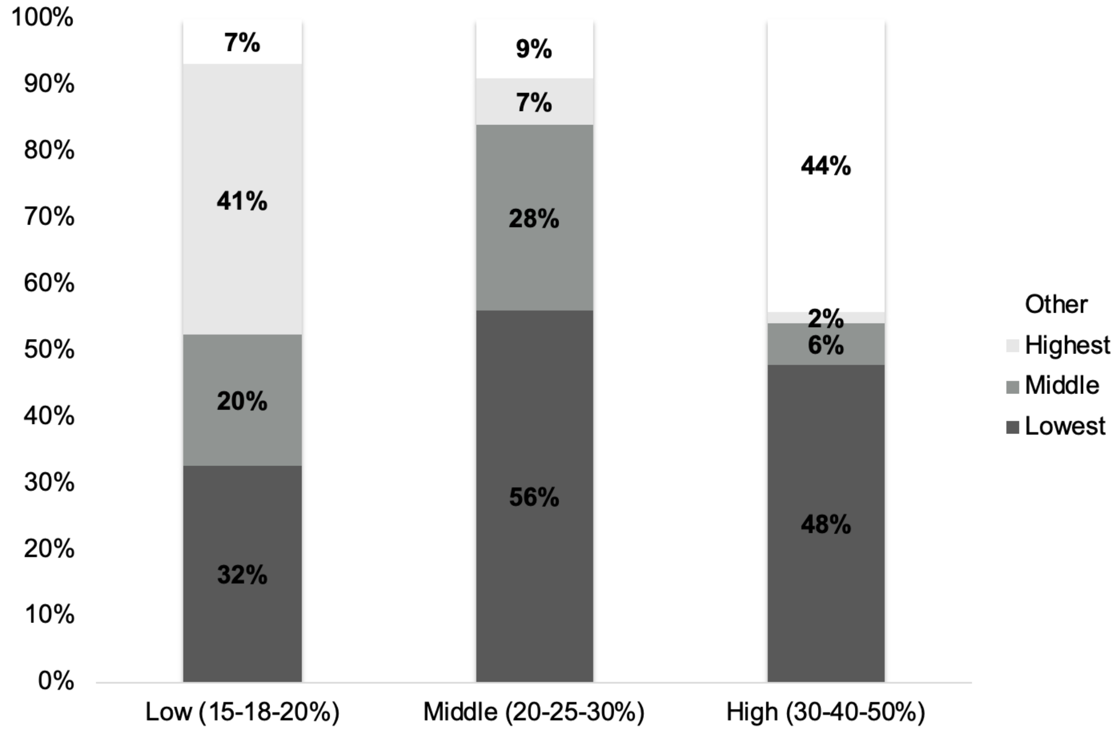
Note: Table 4 shows results for Welch's t-test for unequal variances and unpaired data between average tips in the paper and iPad scene for each condition. Summary statistics for each condition and scene combination are shown along with the t-statistic in the far right column, comparing the difference between the two scenes for each of the four given conditions.

In the middle condition, only 7% of respondents picked the default option of 30% ($6.00), but respondents from the high condition picked this very same value 48% of the time. This 41 percentage point increase can be explained by the changing set of default options. This shift explains the difference in Figure 3 between bars in the same condition but different scenes; low defaults are chosen often and drive lower amounts, but high defaults are chosen at a rate that is sufficient to drive high tips. The interaction between higher anchoring amounts and providing default options appears to be the ideal method of increasing the overall tips from a pure one-off revenue perspective.

As the high condition suggests, there also seems to be some sort of reverse effect where if the tip suggestions become too high, more people will select the "other" option. A potential interpretation for this effect is that there is some cost to entering a custom tip amount. Past a certain point, the suggested tip amount is so high that the optimal method is to incur that cost. In this study's example, 44% of participants in the 30-40-50% condition selected some other option relative to the default amounts. Compared to this 44%, the first two conditions both had under 10% of participants select some other option than the defaults, despite maintaining a full choice set.

This reverse effect could be attributed to Brehm's psychological reactance theory

Figure 4
Default Option Choices



Note: Figure 4 analyzes the choices of default options within the iPad scene, broken down by the tip suggestions. Column one shows the percentage of participants from the condition choosing 15% on the bottom, 18% right above it, 20% second from the top, and other at the top. Column two shows the percent choosing 20%, 25%, 30%, and other, respectively. The right column shows the percentage choosing 30%, 40%, 50%, and other, respectively.

(1966). Since customers may feel that their freedom and choice sets are being threatened through manipulations of unexpectedly large, dis-engaging suggestions, they switch their automatic, system one thinking to a more rational system two process. Therefore, the high amounts may raise internal feelings within participants that cause them to react and reassert dominance by selecting their own amount over a default.

A potential explanation for the reactance that individuals are experiencing is the time spent thinking or reacting about how much to tip. Table 5 analyzes the difference between the amount of time spent by participants on each tipping screen. The low and middle scenes have significantly lower time spent on the iPad screen, which could be attributed to customer system one decision making in simply pressing a default option. This significance goes away in the high condition, however, where customers react in a more systematic, rational way by entering their own tipping amount.

Table 5
Average Time Spent on Tipping Screen

	(1) Paper			**(2) iPad**			
VARIABLE	Mean	SD	N	Mean	SD	N	T-test
Control	12.03	8.79	96	14.15	8.39	106	1.753
Low	14.22	9.24	109	11.71	7.77	86	-2.025**
Middle	17.62	10.34	100	12.19	9.46	100	-3.872***
High	21.63	12.92	86	20.88	15.82	111	-0.357

*** p<0.01, ** p<0.05

Note: Table 5 shows results for Welch's t-test for unequal variances for the time spent in seconds on the tip screen between the paper and iPad scene for each condition. Brief summary statistics for each condition and scene combination are shown in the first six columns along with the t-statistic in the far right column, comparing the difference between the time spent by participants on the two scenes for each of the four given conditions, with significant values noted.

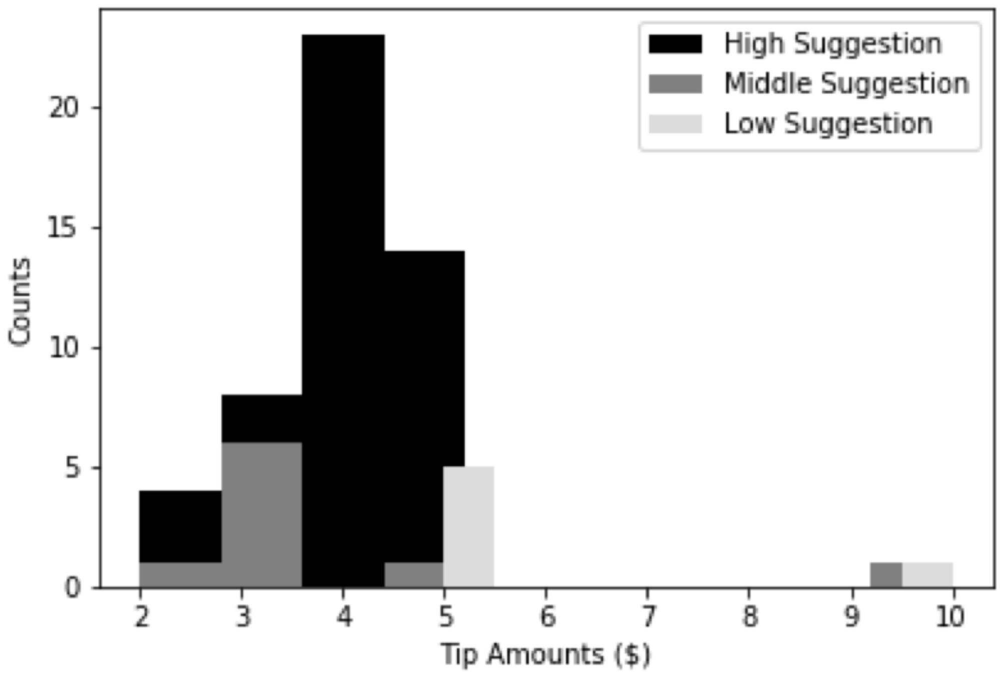
A steady trend exists for increasing means of time spent with increasing tip suggestions. The iPad scene has a slight drop in the time spent for the low and middle conditions compared to the control, but then a major increase in the time on the high condition, which provides evidence for system one decision making with defaults until reactance takes over.

While reactance exists in this study, the results are not as detrimental to pure, short-term revenue as may seem, considering all else equal. Figure 5 examines individual's choices, using 10 distinct bins for each suggestion, that shows the spread of choices for people who selected the other option in the iPad scene for the low, middle, and high scenes. As the graph displays, a much higher quantity of participants selected this option for the high condition. Similarly, customer reactance in the high condition is shown as none of the 49 participants who selected the "other" amount tip more than 25%, even though the 30-40-50% suggestions are provided. 10 out of these 44 participants select to tip between 10% and 15% as well, showing that some people tip less than usual when they are asked to tip too much.

However, the average tip for the 49 participants selecting another amount was still \$3.97, equivalent to 19.9% of the \$20 check. Therefore, while the high default options drove participants to tip alternate amounts, they weren't tipping significantly differently from their no-suggestion, standard values. Coupling this with the fact that 56% of participants tipped \$6, \$8, or even \$10 on a \$20 check explains the increase in revenue from these results.

Study one highlights the initial effect that anchoring can increase tipping in this online, hypothetical setting. This pilot study showed that recommendations on paper increased tipping by 23.8% and adding the default options along with higher anchors increased tipping by 42.5%. The magnitudes of these increases are likely upward-biased and quite large for a nearly \$50 billion industry and should be examined further in field settings, where financial incentives are tied and if a true effect exists, it would most likely be smaller as discussed in the limitations section (Azar, 2011). However, while the direct revenue may increase, people may feel disrespected by being asked for high amounts and in turn, violate the norms of reciprocity that are assumed in socially efficient tipping contracts (Conlin et al., 2003). Despite the difference in general tipping in the entire dataset, no statistically significant differences were observed based on any demographic information, including gender, age, race, and income in this study. In the next study, I examine more potential psychological mechanisms driving different reactant effects and provide a slightly more nuanced approach and

Figure 5
Histogram of "Other" Choices by Suggestion Level: Examining Reactance Theory



Note: Figure 5 plots a histogram of the choices of individuals who were in the iPad condition and selected the "other" option. The lightest group shows the choices (in $) for those who avoided the defaults of $3, $3.60, and $4, the middle group shows the number of participants who chose something other than the $4, $5, and $6 defaults, and the dark group plots the counts of participants who chose numbers other than $6, $8, and $10.

reasoning for not increasing tip suggestions.

3.2 Study Two: Downstream Whiplash Effects

3.2.1 Data and Methods

Study two was designed to provide more insights through minor modifications of study one in addition to providing some more context for potential downstream effects. In this two-minute study on Amazon Mechanical Turk, a different set of participants were recruited and paid $.20 for completion of a survey asking about their tipping habits. They were asked how much they would tip on a $40 check after eating an hour-long meal with good service in exactly the same format as the previous study. Participants were randomly assigned to one of four conditions on the iPad scene from study one. However, participants were only shown percentage suggestions without the calculation assistance of dollar equivalents. While most restaurants, taxis, and iPads relay the dollar equivalents, they are excluded to observe if the calculation assistance that was provided in study one is more effective in gaining high tips (Karniouchina et. al, 2008). The control condition again provided no suggested amounts and asked participants to enter how much they would tip. The low (15%, 18%, 20%), middle (20%, 25%, 30%), and high (30%, 40%, 50%) conditions all retained the same percentages as the previous study. Therefore, study two aimed to see if there was any change in the percentage tipped on the iPad condition that was found in study one through two alterations: changing the bill's order size ($40) and leaving dollar equivalents out of the equation.

After participants entered the amount that they wanted to tip, a few follow up questions were asked. Participants were asked to answer how fair, appropriate, and greedy the restaurant is on a scale of 1-7 without any explicit reference to the tipping amount. Afterwards, a few downstream effect questions were relayed. Participants were asked what Yelp review they would give the restaurant on a scale of 1-5. In addition, they were asked on scales of 1-7 how likely they were to return to the restaurant and how likely they were to recommend the restaurant to a friend. The order of these follow up questions was randomized and succeeded by some questions related to normal tipping. Participants were asked how often they eat out in restaurants on a scale from "More than once a day" to a "Few times a year," what they believe the standard tipping percentage is, and how much they spend on average for a nice dinner for two in a random order. Finally, participants completed demographic questions about age, race, gender, and household income.

The hypotheses in study two added on those provided in study one. Again, higher options, and specifically default options, were expected to increase tipping among customers. However, the higher value of the check was expected to reduce the overall value of the tips due to the higher absolute costs of giving money for a larger bill. Finally, in examining the whiplash effect, higher suggested tips were expected to make customers less likely to enjoy and recommend the experience.

415 participants filled out the survey, but three opened the survey and did not make it to the tip screen, so their results were discarded. The data described below thus provides information on the remaining 412 participants (43.9% male, $M_{age} = 34.0$, $\sigma_{age} = 11.1$), who all tipped between 0% and 40% of the $40 bill.

3.2.2 Results and Analysis

In this slightly altered context, participants provide mostly similar, but slightly modified tip submission behavior as Figure 6 shows. The direction of the behavior remains the same; higher suggested amounts warrant higher tips by customers in the online platform. However, in all 4 conditions, the percentage tipped in the $40 iPad check with only percentages is lower than the previous study's $20 iPad check with percentages and dollar equivalents provided. While a few potential explanations exist, the control group has the largest difference between studies, providing evidence that the bill size may be most likely the driving factor, as no tip suggestions nor educational assistance were provided in the controls. Tipping may follow a logarithmic curve, where the higher the bill is, the steadier the percentage becomes. Another potential explanation is that when people see percentages, they don't consider the dollar equivalents and thus are more likely to select the more standard 15-20%. Finally, a low sample size could limit the results. A mixture of these hypotheses could be valid and should be tested in further scenarios in order to isolate the effects.

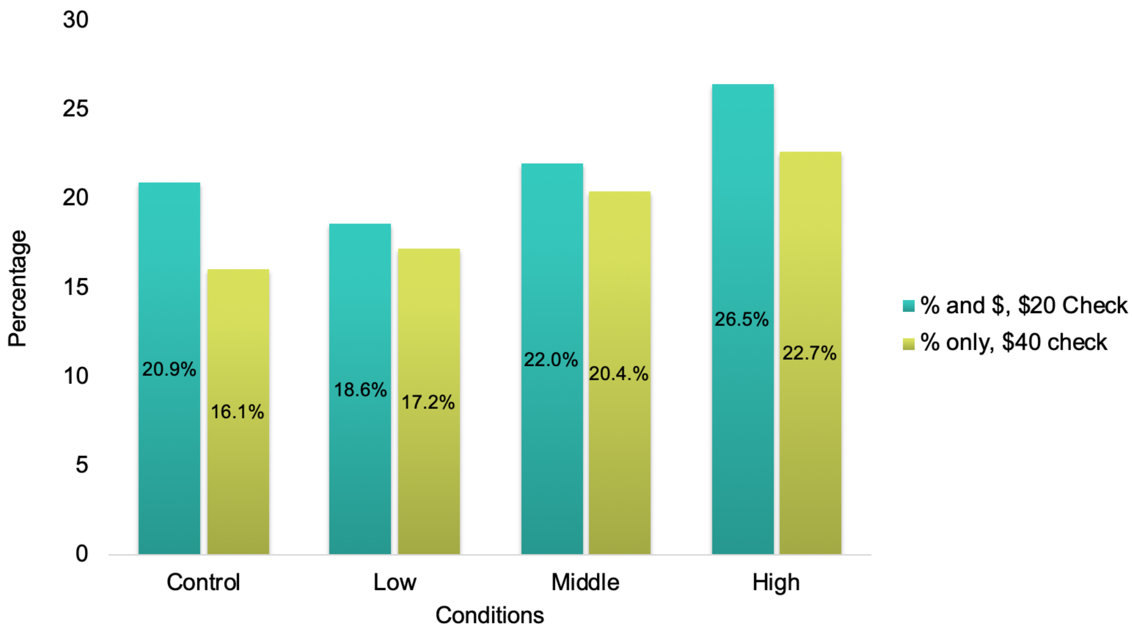
The other difference between the results in study one and study two is that the low percentage becomes insignificant relative to the control, while the middle group becomes significantly higher. Again, the results are most likely due to one of the two main shifts in design, showing the variation in tipping based on different circumstances. However, the overall effect of higher percentages driving higher tipping amounts remains true, as the middle and high suggestions had statistically higher values than the control groups in both circumstances.

Putting the results of the percentages in context, the average tipping amount for the 406 respondents who answered the open response question, "What do you believe the standard tipping percentage in restaurants is?" was 16.9%, which is right in between the control and low conditions. However, customers' ability to be influenced is shown is in the following chart and table, which divides up the standard tipping percentage based on which condition participants were randomly placed into.

The key finding through this analysis is that consumers do not know what their own standard tip percentage is. They are influenced by the suggestions shown previously in the study. The most common responses were 15% and 20%, but participants were more likely to be anchored closer to 20% after receiving higher suggestions earlier in the survey. The lack of a common standard makes final tip amounts variable in this online setting and makes anchoring with default options that much more effective in altering consumer behavior. The middle and high conditions even provide some evidence that participants may tip more in the future based on the prior situation.

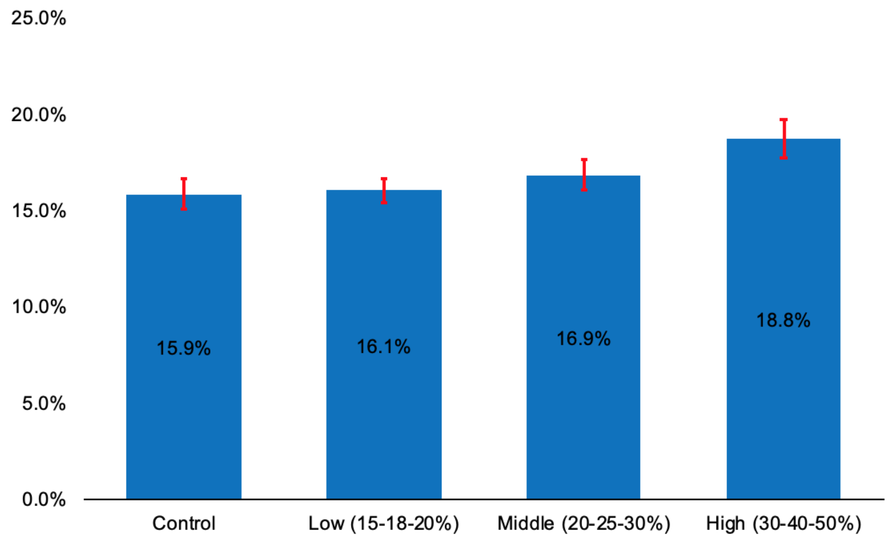
In Figure 8, the major difference between this study and Figure 4 from the previous experiment in terms of default choices relies on the middle suggestion group. The participants

Figure 6
Average Tip Percentage by Condition for Studies 1 and 2



Note: Figure 6 plots the average percentage tipped for each condition in study one and study two. The bars on the left of each duad represent the results from study one, where calculation assistance was provided in terms of percentages and equivalent dollars for a $20 bill, while the four bars on the right of each duad represent the results of study two, which show only percentage suggestions on a $40 check in study two. The remaining conditions were the same in both studies.

Figure 7
Participants' Belief in Standard Percentage



Note: Figure 7 shows the difference in what participants believe the standard average tipping percentage is after they had participated in the experiment and entered their tipping amounts for the check. The previous suggestions provided to participants appear to affect their beliefs about what the standard percentage is, based on the condition they were randomly placed into.

Table 6
Difference in Participants' Belief in Standard Percentage

VARIABLES	standard
high	2.789***
	(0.657)
middle	1.021*
	(0.569)
low	0.206
	(0.522)
Constant	15.89***
	(0.411)
Observations	406
R-squared	0.065

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

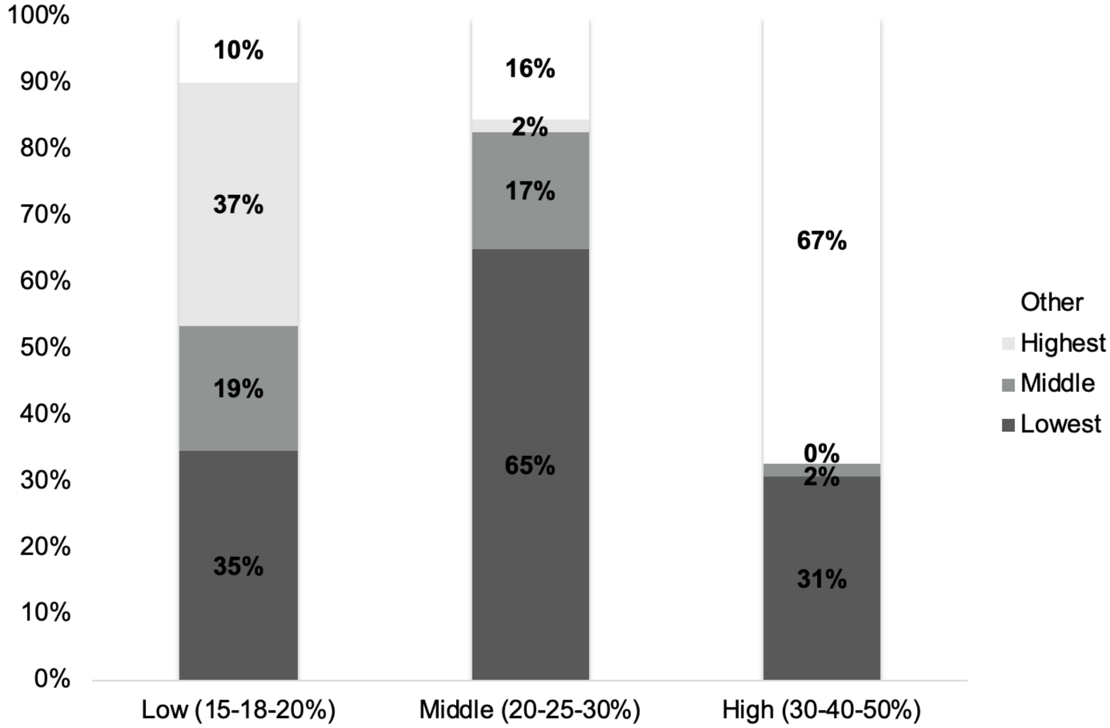
Note: Table 6 shows a regression based on the analysis in Figure 7 that shows the difference from the control value for each suggestion, with statistically significant differences displayed and robust standard errors shown in parentheses.

in this given condition were more likely to select 20% or "other" over the 25% or 30% value. Since randomization occurred and this shift is probably not due to different consumer preferences, the $40 check or the lack of calculation assistance may drive participants to select a more average amount. Similarly, a low sample size may impact these results. Even with this finding, higher tip suggestions were shown to drive higher tip amounts.

Reactance theory still holds true in this example, but in a slightly different manner. At a certain point, rationality overcomes automatic, system one behavior and the majority of participants select the "other" option rather than choosing one of the listed defaults. However, in this example, the amounts are not vastly different from what they pick in a lower treatment. Out of the 104 participants who were in the high suggestion condition, 67% selected "Other" instead of the three default options of 30, 40, or 50%, and these 70 individuals tipped an average of 18.9%. Using Welch's t-test with unequal variances, there is a significant difference in the percentages for the control group (M=16.1, SD=6.3) and the high condition group that selected the other button (M=18.9 SD=3.4) option ($t(164)= 3.74$, $p < 0.001$), even though the individuals were filling in their own desired amounts in both scenarios. Therefore, high tip suggestions seem to be driving higher tips for the interaction between two reasons: a percentage of individuals are nudged into selecting the incredibly high defaults and while a majority actually avoids these defaults and chooses the "other" option, these selections are in line with low tip suggestions or even higher than no tip suggestions. This implies that the high default suggestions may still have anchoring effects on final amounts even when not chosen.

However, while this study looked at tipping tendencies without financial incentives, there also may be some downstream effects from instituting these high amounts. As follow-up questions, the survey asked participants about their Yelp reviews (1-5), recommendations (1-7), and likeliness to recommend (1-7). While the previous analysis showed that high tipping amounts were more likely to drive profits in this online setting, there are many other potential downstream negative effects that could arise from asking participants for higher

Figure 8
Default Option Choices, Study Two



Note: Figure 8 performs the same analysis as Figure 4 in study one and analyzes the choices of default options within the iPad scene, broken down by the tip suggestions. Column one shows the percentage of participants from the condition choosing 15% on the bottom, 18% right above it, 20% second from the top, and other at the top. Column two shows the percent choosing 20%, 25%, 30%, and other. The right column shows the percentage choosing 30%, 40%, 50%, and other, showing slight differences in percentages choosing each value between study one and two based on design.

amounts. These whiplash effects, which stem from reactance theory, predict individuals to potentially want to punish places who aim to rip them off.

In this online platform, evidence of whiplash exists, especially in the high and middle suggestion conditions. As Table 7 shows, high tip suggestions cause people to feel somewhat offended and think lower of the restaurant on different characteristics, such as how greedy the restaurant is, how likely customers are to recommend the restaurant to a friend, how likely customers are to return, how highly they would rate a Yelp review, and how fairly participants believe the restaurant treats their employees. Table 7 is a regression modeling Z-Scores, converting the raw data into number of standard deviations away from the mean, in order to see the effect that moving one standard deviation has on outcome factors.

Having customers who are less likely to recommend the restaurant or return for another meal may actually lower total profits, even considering the increase in tipping. More research and analysis would need to be done at an economic level to understand the short-term gains versus the long-term trade-offs. Similar to study one, no significant effects were found for groups from different demographic backgrounds, and further analysis may be necessary to analyze different groups.

3.2.3 Limitations of Studies 1 and 2 and Motivations for Study 3

Studies one and two show some relevant initial effects in an online setting: higher tip suggestions anchor participants at higher tip amounts, default options can have a more significant effect than pure tip suggestions, and reactance along with whiplash effects may exist from recommending tipping at too high of a value. Study three adds some field evidence that different suggestions have small, but significant effects through a minor change in a historically non-tipping fast-casual business setting.

Studies one and two were intended to show the initial effects rather than absolute significance through strong generalizations about exactly how individuals tip or the state of the field on the whole. The purpose of using Amazon's Mechanical Turk platform was to

Table 7
Downstream Effects from Tipping Amounts, Z-Score Normalized Results

VARIABLES	(1) greedy	(2) fairly	(3) recommend	(4) return	(5) yelp
high	1.173***	-0.398***	-0.784***	-0.490***	-0.783***
	(0.123)	(0.142)	(0.140)	(0.146)	(0.131)
middle	0.639***	-0.0998	-0.298**	-0.349***	-0.386***
	(0.116)	(0.128)	(0.134)	(0.129)	(0.129)
low	0.295**	-0.104	-0.0399	-0.0246	-0.142
	(0.117)	(0.129)	(0.128)	(0.118)	(0.134)
Constant	-0.531***	0.152*	0.286***	0.219**	0.333***
	(0.0725)	(0.0866)	(0.0955)	(0.0850)	(0.0888)
Observations	410	410	407	407	407
R-squared	0.192	0.022	0.099	0.044	0.089

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7 runs five regressions with the dependent variables of greediness, fairness, recommendations, likelihood to return, and yelp review. The first four dependent variables were tested on scales of 1-7, while Yelp reviews were tested on scales of 1-5, as is customary. The independent manipulations are the suggestion values. The values for the tips are normalized, so the above shows the Z-scores and number of standard deviations away from the mean each of the values are, rather than pure 1-7 and 1-5 outputs. Robust standard errors are shown in parentheses.

examine an initial effect that anchoring values and default options can influence consumer behavior and spending. However, some potential limitations have been documented around the participants who engage in mTurk, generalizability using the analysis, and data quality (Stritch, Pedersen, & Taggart, 2017). Therefore, studies one and two are exploratory analyses that show initial effects may exist in the field with more evidence needed to support the findings.

The design of the online surveys also had three limitations. The main limitation is the lack of financial ties for giving small or large sums of money as tips. People are more familiar with the cost associated with spending $5 on a tip in real situations rather than entering the amount in a hypothetical online scenario. Second, people consider more than the suggestions when deciding how much to tip. The full experience within a restaurant includes the waiter-customer interaction, delivery times, food quality, cost of food, cleanliness, and overall happiness with the experience. While online and lab experiments isolate the tipping suggestions and reduce the effect of other variables, the external validity of the real-life situation is increased in a field experiment with all other factors considered. Therefore, coupling a field analysis with the online experiments adds more externally valid results on the mechanisms from the online scenarios. Finally, while $20 and $40 were simply chosen values for simplicity based on TripAdvisor's average price for a meal for two, generalizations cannot be made about all prices and the effect of the treatment (Mulcahy, 2013).

The purpose of the field study is to address these three limitations by tying in financial incentives, providing the full experience, and naturally altering prices based on consumer spending through an observational, quasi-experiment in the field. However, the specificity of a fast-casual setting and its specific application may limit generalizability so that this effect may not remain true at every restaurant that has customers pay through the iPad. Similarly, customers were spending $11 on the average order, while most other restaurants have larger overall bills that could potentially show more significant results if the psychological mechanisms driving the change in tips hold true in those settings. Finally, the first two studies devised scenarios with one-hour sit-down meals, which differs from the fast-casual business model. Even with these considerations, study three aimed to provide more evidence in a field situation that tip suggestions can impact consumer spending.

3.3 Study Three: Field Data Analysis

The following section describes a field analysis of tipping for a fast-casual restaurant in Cambridge, MA. In the following section, I describe the establishment, its location, and the products it sells. I then describe the dataset that is analyzed and provide some summary statistics about consumer behavior and tipping. Finally, I exploit a discontinuity in the rule used by the restaurant for its tipping suggestions using a regression discontinuity design.

3.3.1 Data Description

This restaurant is defined as "fast-casual," since it does not offer full table service, but provides a quality of food that is typically labeled as better than fast food. At this restaurant, customers order their food and pay while waiting nearby for their meal to be called and left on the counter. While the only service is the creation of their order, customers have the option of tipping before their food is given to them, unlike in more traditional, sit-down settings, where tipping occurs after the meal. The average meal item costs around 8-9$, and most menu items fall in this range.

Data was obtained from 7,923 transactions over 32 days from mid-October to mid-November 2018. The following analysis excludes 30 entries that had sales of $0 total, refunds with negative dollar amounts, and gift cards in order to isolate tipping amount. Financially sensitive information is withheld at the restaurant's request. The dataset included information including the order amount, tax added, tip amount in dollars, order description, payment method, and time of day.

I also define indicator variables for repeat customers and Harvard undergraduate students. Other demographic variables such as race, age, or income demographics are not available. A key feature of the data is the first and last name of the ordering customer in each transition. To define repeat customers, I use first and last name to define a unique customer identifier. In order to define a dummy variable for Harvard undergraduate, I blindly scraped the online directory of Harvard undergraduates and merged the resulting data with order names. While students may not have exactly the same names online and in their order information, 15.3% of the transactions came from names that were in the Harvard directory. Since I was interested in general trends and not specific names or other identifiable information, I assume that these encompass all Harvard undergraduates in the data set and do not probe the data any further. After defining repeat customer and Harvard student variables, I drop the first and last name from the data and conduct the analysis on the de-identified data set.

The restaurant uses a pre-loaded tip suggestion option on their POS system, where every customer in the 7,893 transactions is offered three default tipping amounts, a "no tip" option, and an "enter custom" option. An example tip screen is provided in Appendix C from the company's point of sale system. For total orders under $10, which included the price of the meal, tax, and discounts, the three default suggestions are $1, $2, and $3. For total orders over $10, the three suggestions are 15%, 20%, and 25% of the bill, with dollar equivalents portrayed under each option.

In the following analysis, tip percentages are often referenced, even though exact values of percentages are not given in the data set. Tip percentages were calculated by dividing the tip value by the tax plus pre-tip order value. Since tips were provided in dollars, rounding errors may occasionally occur in the dataset in cases where participants selected the defaults of 15%, 20%, and 25%. Therefore, the following analysis includes rounded percentages that round any number within .02% of a whole number to that whole number. Since the number of rounded percentages within .02% above and below the threshold do not have a statistically significant difference and since the percentages are so small relative to the average order of $10.91, I assume that the results outside of default choice analysis are not significantly affected by this data intervention.

3.3.2 Descriptive Statistics

Table 8 shows the summary statistics based on different orders and customers, divided at the threshold of $10. A higher percentage of individuals tip for orders greater than $10, and they also give more on average, conditional on tipping something.

Because the orders over $10 are receiving a higher percentage of tips, and higher tip amounts from those giving, the average tip percentage has a 48% increase per transaction in the high condition relative to the low one. In Sections 3.3.3 and 3.3.4, I use a regression discontinuity to test if this is a causal effect based on the tipping suggestions provided, or if there are differences between the groups of individuals, such as income levels, that are driving the effect at the $10 threshold.

Repeat customers have similar average tipping amounts as one-time customers, but students tip at much lower rates than non-undergraduates. Students who tip give similar amounts as non-students, but only about half as many students give tips compared the rest of the customers in the data set. Lower student income levels, an understanding of the optionality of tipping at fast casual locations, or a mixture of both may be driving this lower percentage.

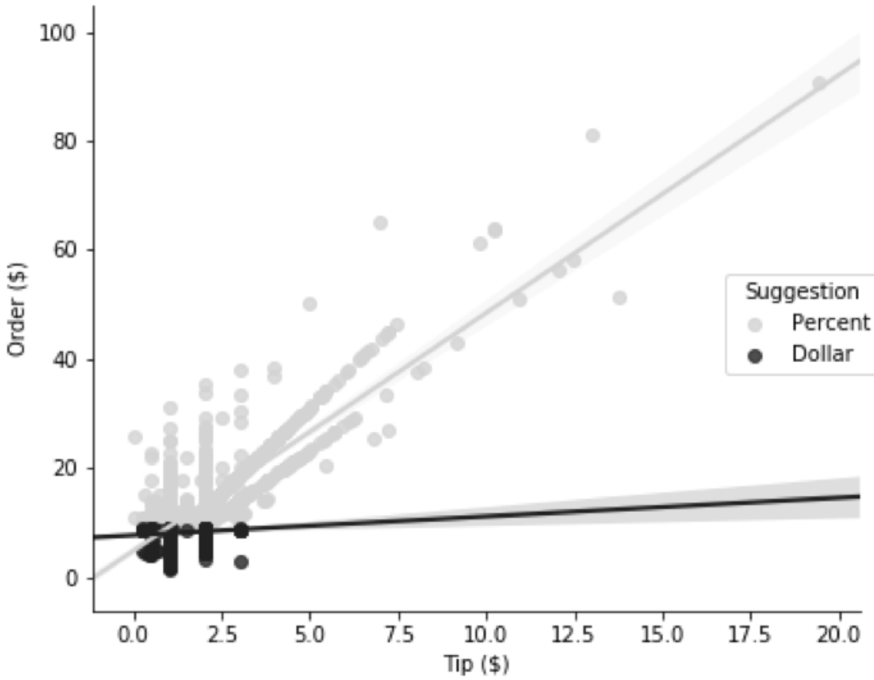
Because of the difference in type of suggestions in terms of dollar versus percent, the data also shows different results for how tips scale with higher orders. In both scenarios around the cutoff of $10, there is a positive correlation depicted in Figure 9. However, the coefficient around the correlations are very different: the high order has a correlation of .848, while the low order has a correlation of .086.

Table 8
Summary Statistics by Order Amount, Type of Customer

VARIABLES	(1) Number of Orders	(2) Average Order ($)	(3) People Tipping (%)	(4) Average Tip Size (%)	(5) Tip Per Transaction (%)
High Order	**3462**	**14.91**	**40.18**	**14.38**	**6.15**
Students	489	13.60	18.61	14.37	2.80
Non-Students	2973	15.12	43.73	14.38	6.65
Repeat Customers	1778	14.23	38.13	14.06	5.59
One-time Customers	1684	15.62	42.34	14.64	6.69
Low Order	**4431**	**7.80**	**28.84**	**12.86**	**4.14**
Students	721	7.72	17.61	13.00	2.47
Non-Students	3710	7.82	31.02	12.85	4.46
Repeat Customers	2139	8.08	31.51	12.45	4.20
One-time Customers	2292	7.54	26.35	13.32	4.07
All Orders	**7893**	**10.92**	**33.81**	**13.87**	**5.34**

Note: Table 8 shows descriptive statistics for orders broken down by students found through a blind algorithm and repeat customers, whose names were converted to customer IDs and analyzed. High orders have a total order plus tax value greater than or equal to $10, and low orders are below $10. The bottom line summarizes statistics for the entire dataset.

Figure 9
Tip Order Versus Amount



Note: Figure 9 scatters the order amount in dollars against the tip amount in dollars. The dark scatter is representative of the low orders, while the light scatter represents the high orders. Least square regression lines for the two different scatters are plotted with 95% confidence intervals fading around them.

The difference in these two correlations can be explained by the scaling of the suggested amounts. Since the under \$10 condition does not change its suggestion in dollars with higher orders, the percentage will continue to go down as orders get larger if customers are not choosing to enter their own amounts. The tip amounts for orders under \$10 are clustered at 3 or 4 different values, with the defaults of \$1, \$2, and \$3 selected at high rates. The scatters right above and below the regression line for the percent suggestion that appear to be almost perfectly correlated have slopes that equate to the 15% and 20% suggestions, so it appears as if customers are likely to select these options.

This choice of default options is examined in Figures 10 and 11, which show the percentage of people tipping and the default options chosen in the two scenarios. While different proportions of customers tip in the two scenarios, an overwhelming percentage of tips are the lowest default selection. In the low order condition, only 6.1% of people decide to tip any amount other than the three options provided, highlighting the willingness of customers to pick the presented options.

In both examples, while a low percentage chose "Other," these individuals are tipping about half the lowest suggestion on average, equivalent to $.50 for low orders and 7.74% in the higher condition. As study two would predict, a slight negative emotional reactance also may exist here, as more people are selecting some other option when the restaurant asks for higher dollar amounts in large orders. In this field data, the reactance happens at a lower suggested value than in the hypothetical online scenario, potentially because customers are unaccustomed to tipping in fast-casual restaurants outside of change in tip jars or boxes. However, even with the larger percentage selecting "other," revenues in percentage for the over $10 condition are larger, which suggests that reactance is not enough to overcome the added revenue of percentage tip suggestions that scale in proportion to orders. The extent to which this relationship is causal is explored in the following section.

3.3.3 Methods for Quasi-Experiment

In order to identify a casual effect between the tip suggestions provided, I exploit the discontinuity in the tipping options that occurs at transactions of $10. The tip suggestions at the discontinuity are $1, $2, and $3 on the lower end, and 15%, 20%, and 25% on the higher end, which provide dollar equivalents of $1.50, $2.00, and $2.50. These two separate suggestion types have the same means between the three amounts, but the lower end has a larger standard deviation. In the following analysis, therefore, very similar suggestions are provided in the two examples.

The key identification assumption is that all other determinants of tipping evolve smoothly across the $10 threshold. I validate this identification graphically. First, I perform the McCrary test of a manipulation of the running variable across the threshold to observe any differences in Figure 12.

Figure 10
Default Option Selection for Orders Under $10

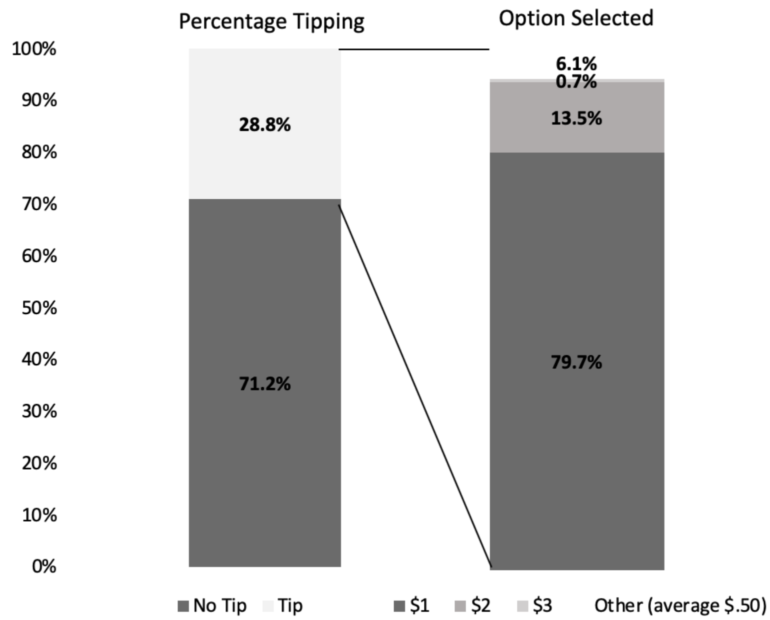
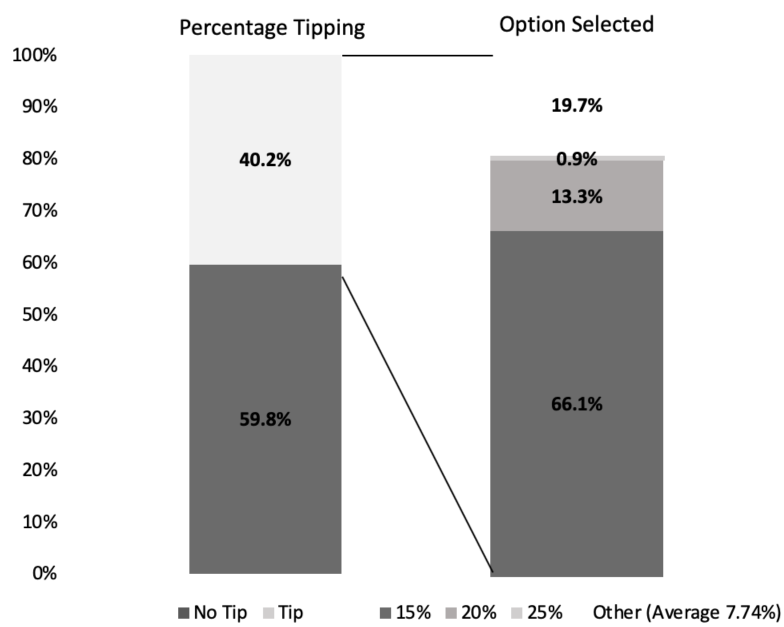


Figure 11
Default Option Selection for Orders Over $10



Note: Figures 10 and 11 show the percentage of customers tipping a value greater than 0$ on the left and of those customers tipping, the default options selected on the right for orders that are under $10 and over $10, respectively. The majority of participants who tipped chose the lowest suggested amount of $1 and 15%, respectively.

In Figure 12, the running variable does not appear smooth across the $10 threshold. However, this result is most likely due to reasons other than customers attempting to avoid specific tip amounts. For example, nearly one-third of the customers in the entire dataset had an order of exactly $9.47. In fact, deletion of all values of $9.47 in Figure 13 shows a much different cross along the discontinuity.

Similar to grade point averages for students, these orders are discrete variables due to the limited number of order options; therefore, as Zimmerman (2014) performs with his GPA analysis, I conclude that this spike underneath the threshold is most likely due to a common menu choice that is limited mathematically, rather than an intentional avoidance of the different suggested tip amounts. This conclusion is based off the fact that people are primarily basing their orders on food choice and potentially the associated cost, rather than the tip suggestions provided, which are even undetectable to customers who have not visited this specific restaurant before. The only potentially limiting factor at $10 could be a lack of cash on hand, but over 90% of the dataset used a credit, debit, or gift card to pay. Since most customers will not make an order decision over a couple cents at the discontinuity, I conclude that the McCrary test does not need to be flawlessly smooth.

On top of the McCrary test, I examine other variables that may show similarities in customers on both sides of the threshold. The graphs in Figures 14 , 15, and 16 provide some initial evidence that there are minimal-to-no differences between the two groups divided by the $10 cutoff on a variety of factors, including fees as a proxy for revenue, time of day, and day of week. No race, age, income, or gender demographics are given in the dataset, which limits the ability to ensure that smoothness for individual characteristics exists along the curve.

Figure 12
Order Counts

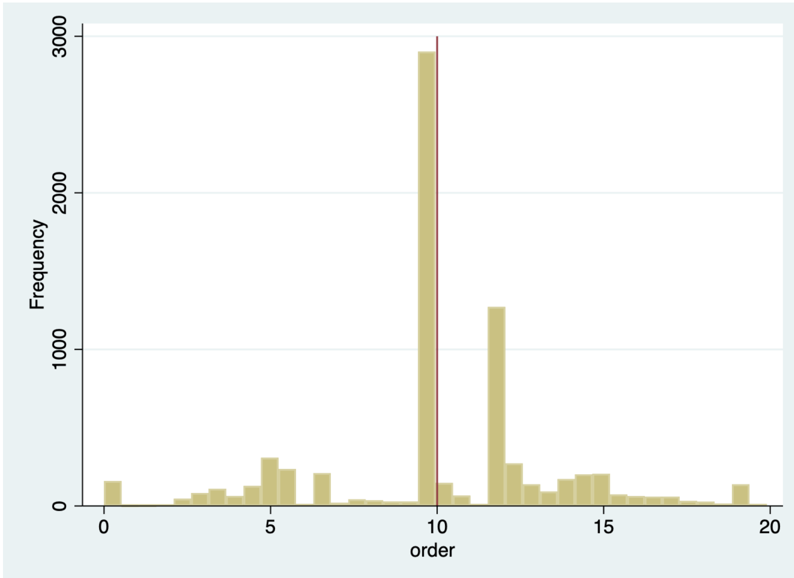
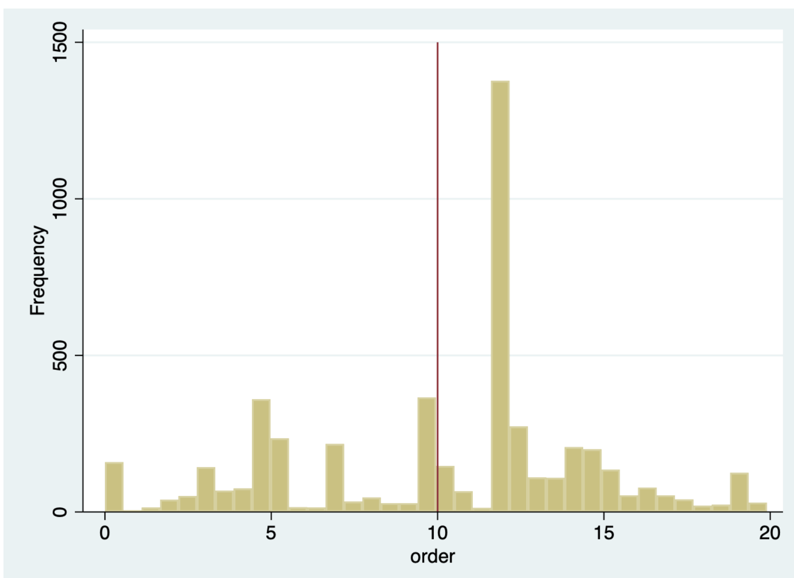
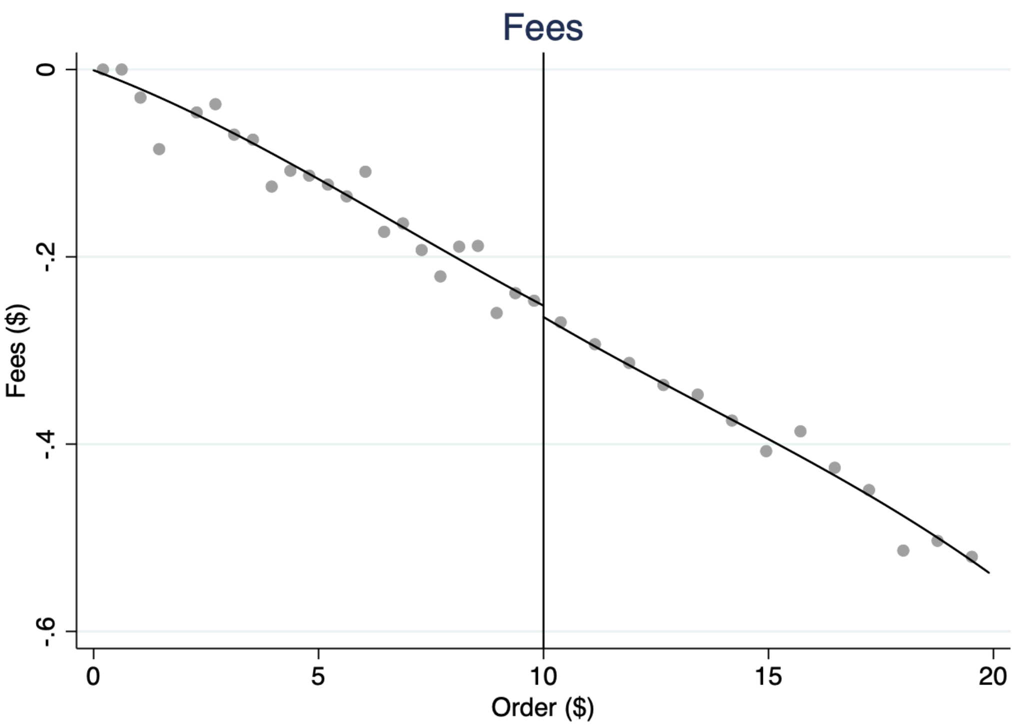


Figure 13
Order Counts Without, Excluding $9.47



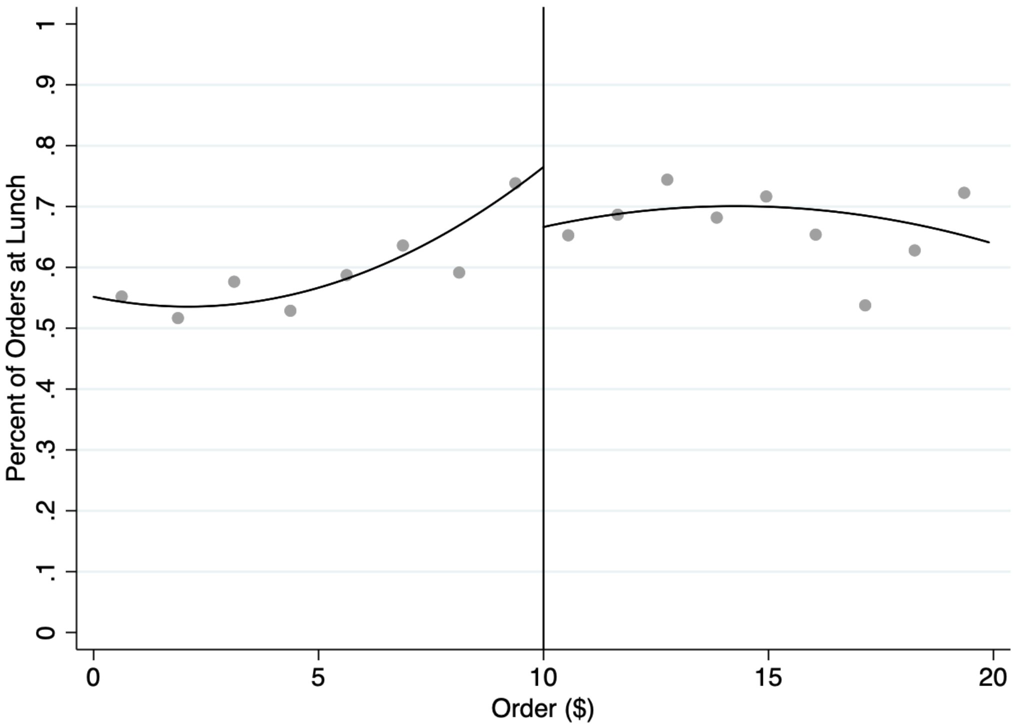
Note: Figures 12 and 13 show counts for McCrary tests that plot histograms of orders for the running variable. The percentage of orders within each bin is shown on the y-axis, and a line is drawn at the threshold of $10 to examine differences in smoothness across the threshold. Figure 12 plots all the orders in the dataset, while Figure 13 plots all the values, except orders of exactly $9.47.

Figure 14
Fees Sent to the Application as a Proxy for Revenue



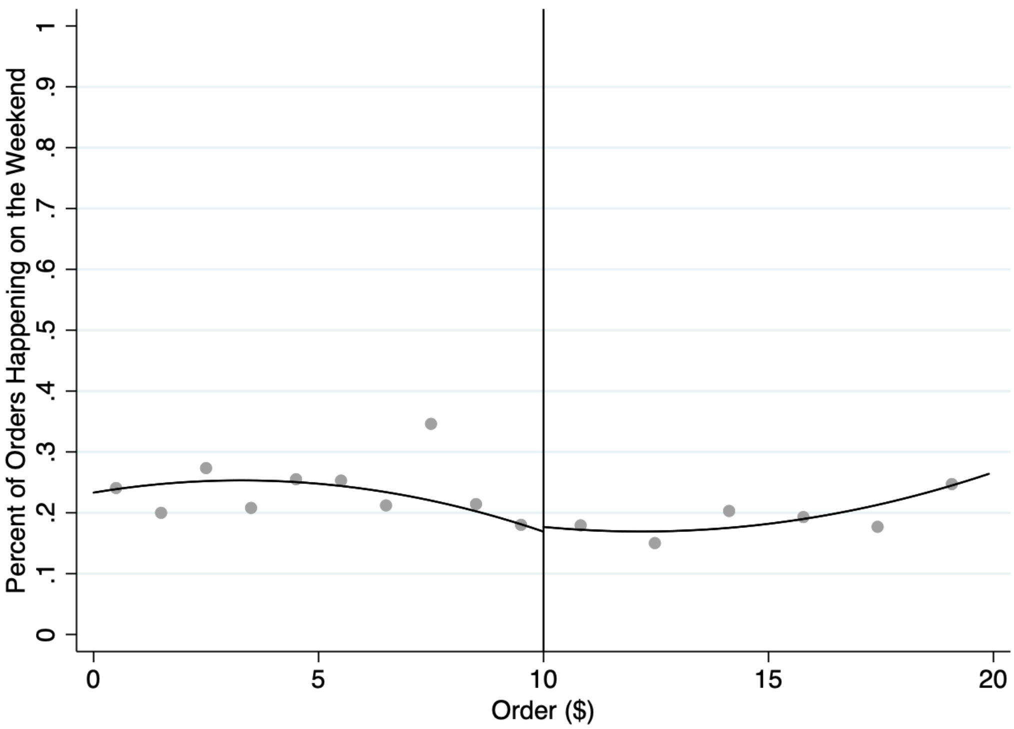
Note: Figure 14 plots the a third order polynomial for the fees that go to the application against the order in dollars. According to the manager, the fees serve as a good proxy for the discounts that customers are gaining, since a proportion of the final revenue that the store gains goes to the application.

Figure 15
Percentage Lunch Orders



Note: Figure 15 plots second order polynomials for the percentage of lunch orders. The restaurant has two segments during their day: lunch and dinner. Lunch spans the hours of 8AM and 4PM, while dinner spans from 4PM to 10PM. Staff shifts normally end for the lunch period and begin for the dinner period around 4PM, so this graph examines the smoothness along the $10 threshold.

Figure 16
Percentage Weekend Orders



Note: Figure 16 plots the percentage of orders for each bin that happen on Saturday and Sunday versus any other day during the week. A second order polynomial is used to fit curves on both sides of the $10 threshold.

Due to the assumptions underneath the McCrary test and the covariate graphs, I use the natural, quasi-randomness of the data at the threshold of \$10 to see the effect that crossing the threshold has on tipping amounts. In Equation 4, I aim to estimate this effect β_1 by using a control function f to reduce the noise.

$$Tip_i = \beta_0 + \beta_1 \mathbb{1}(Order_i \geq 10) + f(Order_i - 10) + \epsilon_i \tag{4}$$

To assess the robustness of my estimates, the control function in the equation above is analyzed through both linear and quadratic equations that plug in the order centered around the threshold of \$10. Mathematically, the identification assumption can be written as follows:

$$\lim_{x \to 10-} E[\epsilon | X + dX] = \lim_{x \to 10+} E[\epsilon | X + dX] \tag{5}$$

As the graphs of covariates in this section show, there are no significantly observable differences in characteristics. Similarly, since I assume that the McCrary test has no effect on self-selection, I do not have evidence that the two sides of Equation 5 are different, so I assume that any difference in tipping is due to the different tip suggestions.

Therefore, in order to test the significance of this change in Section 3.3.4, I regress with linear controls in Equation 6 and linear and quadratic controls in Equation 7, where the control functions emcompass all the terms after $Above_i$:

$$Tip_i = \beta_0 + \beta_1 Above_i + \beta_2 Order_i + \beta_3 Order_i * above_i + \epsilon_i \tag{6}$$

$$Tip_i = \beta_0 + \beta_1 Above_i + \beta_2 Order_i + \beta_3 Order_i * above_i + \beta_3 Order_i^2 + \beta_5 Order_i^2 * Above_i + \epsilon_i \tag{7}$$

In these equations, order is a centered variable that is equal to the the value of the order minus 10 to ensure that β_1 measures the discontinuity. $Order * above$ is an interaction between being above the cutoff threshold and the $Order$ values. $Order^2$ and $Order^2 * above$ have the same methods but are just squared quadratic controls.

3.3.4 Regression Discontinuity Results and Interpretation

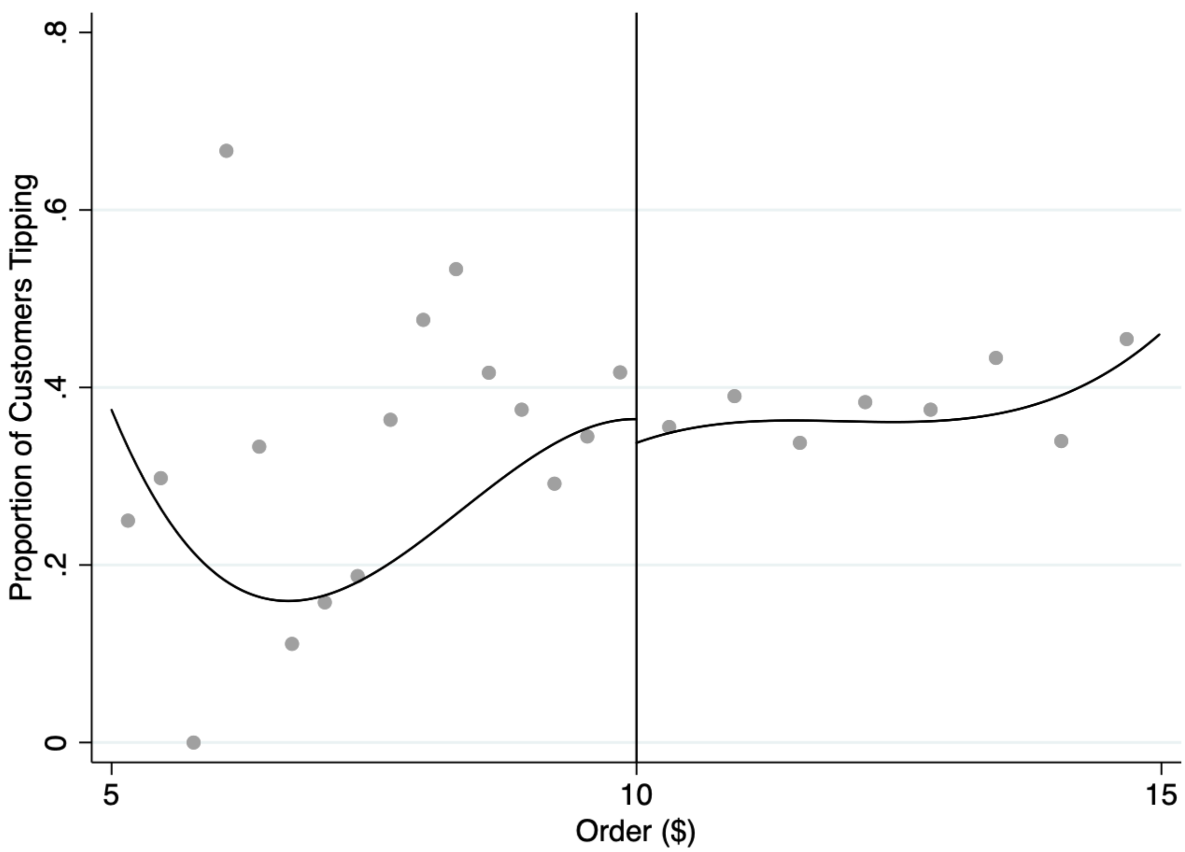
I start by reporting the RD estimates of the impact of the discontinuity at \$10 on the number of customers tipping any amount. This is the extensive margin. The summary statistics show that 11.3% more customers tip on orders over \$10 compared to orders under \$10. However, the discontinuity shows no significant difference between the percentages tipping in Figure 17.

There is no statistical evidence that providing \$1, \$2, and \$3 versus 15%, 20%, and 25% across a \$10 threshold has an effect on the percentage of people tipping, potentially due to the fact that very similar suggestions are provided in both scenarios. While there is no discontinuity evident in the general number of people tipping, the amount that people tip may change across the threshold due to the suggestions. The next section analyzes the amount given by people who actually tip. Figures 18 and 19 show small, but statistically significant effects in terms of both percentages and dollars at the cutoff of \$10.

Since the observable characteristics are essentially the same around the cutoff, there seems to be a causal effect of changing the suggestions. Unlike typical shifts in study designs, the natural shift here does not shift all three options higher or lower at the discontinuity.However, although the mean of the suggestions remain the same, the lowest suggestion specifically is most likely driving the overall tip amount, as it is picked in over 60% of scenarios. Therefore, the slight \$.50 increase in the suggestion from \$1.00 to 15%, or \$1.50, impacts consumers' mindsets in tipping. In order to test the significance of this change, the results from regressing with linear controls in Equation 6 and linear and quadratic controls in Equation 7 are examined.

For both dollars and percentages in Table 9, crossing the threshold provides a positive, but small effect for the \$10 check in all but one scenario. By providing the alternate

Figure 17
Percentage of Participants Tipping by Order Size



Note: Figure 17 shows the percent of individuals tipping a value greater than $0 for each binned order group. A value of 1.0 on the y-axis signifies that every individual in the binned group tipped something. A line is drawn to examine graphically if any discontinuity exists at the $10 threshold. Third order polynomials are used to fit the data to the left and right of the threshold.

Figure 18
Average Tip Amount ($) by Order Size

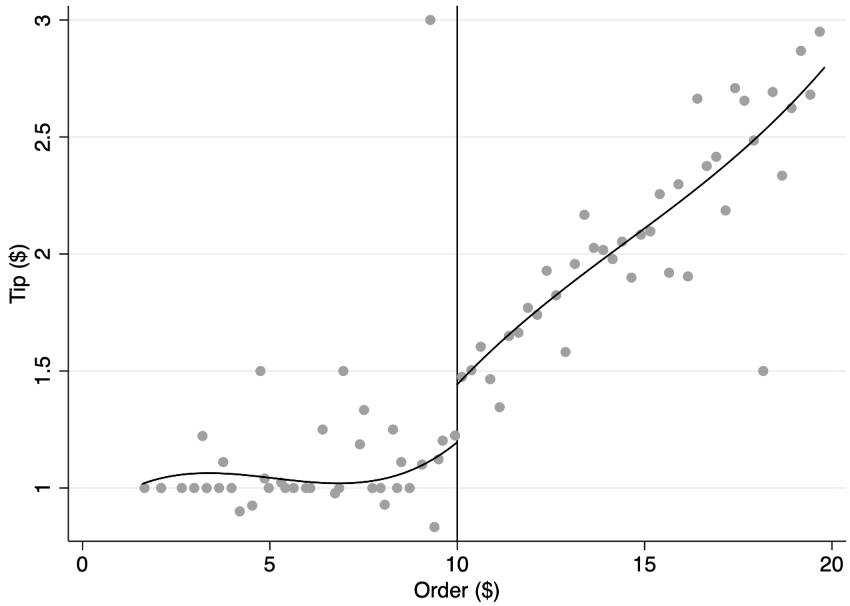
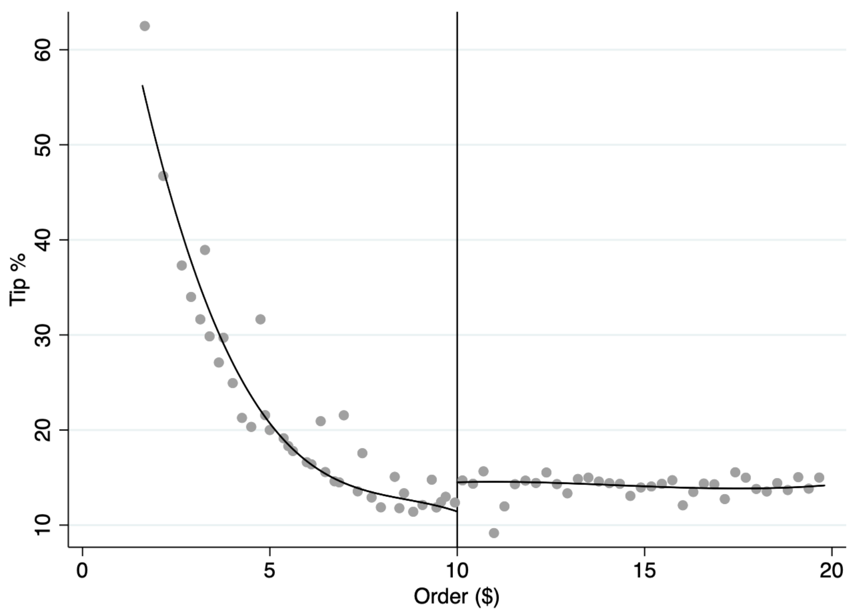


Figure 19
Average Tip Amount (%) by Order Size



Note: Figure 18 and 19 use binscatter to show the relationship between tips and the order amount. A third degree polynomial is used in both cases to model the trends of the tip in dollars for Figure 18 and in percent for Figure 19. A discontinuity line is drawn at the order threshold of $10 to examine graphical differences between orders below $10 and above $10.

suggestions, an approximately \$.07 or .7% increase per tip exists. While the magnitude of this increase seems small, the average tip for the average order of \$10.91 was just \$.58 cents, so an increase of \$.07 per order actually adds over 12% to the tip in every order.

At the threshold, many customers are shifting their tipping amounts from \$1.00 to 15%, or \$1.50. Assuming that some of these customers still prefer to tip \$1.00 without the nudge, they are giving up \$.50 more than their rationally economic self would typically give. Pressing the "enter custom" option, thinking about their choice, and then submitting should take no longer than 10 seconds avoid the default of \$1.50 and enter \$1.00. Performing a back-of-the-envelope calculation shows that a \$.50 increase for saving 10 seconds equates to an average personal evaluation of an individual's time at \$180/hour, well above the average individual's assessment.

In this data analysis, I find evidence that providing higher bottom suggestions for the tip amount through percentages causes individuals to tip more, even with the slight reactance of the increased high tip suggestion, due to a mixture of the anchoring and default effects. Although the top default option at the discontinuity was lower in high orders, less than 1% chose this amount regardless of order size, causing no sizeable impact on overall tips.

4 Further Discussion

In this paper, I provide experimental and quasi-experimental evidence that restaurants can influence their employees' earnings through increased tipping mechanisms by anchoring through default suggestions. While some reactance and whiplash effects may exist from customers' heightened valence and more rational thinking due to very high tip suggestions, suggesting higher tip recommendations, specifically at the lowest of three default choices for fast-casual and other historically non-tipping industries may warrant higher yield and magnitudes of tips.

Table 9
Analyzing the Discontinuity with Control Functions

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLE	Tip	Tip	Tip	Tip	Tip	Tip
UNITS	$	$	$	%	%	%
FUNCTION	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Above	0.0621*	0.0990***	0.0741***	0.731**	0.190	0.701***
	(0.0351)	(0.0237)	(0.0265)	(0.327)	(0.225)	(0.251)
Order	0.0448***	0.0432***	0.0502***	0.000963	0.179***	-0.316**
	(0.0100)	(0.00458)	(0.0153)	(0.0934)	(0.0434)	(0.145)
Order*above	0.0342**	0.0268***	0.0168	0.202	0.110*	0.633***
	(0.0154)	(0.00642)	(0.0160)	(0.143)	(0.0609)	(0.152)
Order^2			0.000937			-0.0660***
			(0.00195)			(0.0185)
Order^2* above			0.000756			0.0506***
			(0.00203)			(0.0192)
Constant	0.418***	0.417***	0.422***	4.144***	4.340***	4.013***
	(0.0161)	(0.0148)	(0.0177)	(0.150)	(0.141)	(0.168)
Bandwidth	5	10	10	5	10	10
Observations	6,084	7,490	7,490	6,084	7,490	7,490
R-squared	0.048	0.098	0.099	0.008	0.017	0.020

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 9 runs results for the regression discontinuity and finds statistically significant results for five of the six cases. Tips in dollars are the dependent variables for the first three columns, while columns 4-6 have tips as percentages as the dependent variable. Regressions 1, 2, 4, and 5 use a linear control function, while 3 and 6 use a quadratic control that is defined by 6 and 7, respectively. The analysis measures results from 0 to 20 for all the regressions except for columns one and four, where the bandwidth is 5. The variable *Above* measures the effect that the discontinuity has on tips.

A major question that this analysis raises is why people tip in industries that have historically received little-to-no tips. It may be useful to draw an analogy with behavior in the well-known Prisoner's Dilemma game. While in the short term or in one-off games, individuals may defect, models predict that consumers will evolve to either cooperate or act in the same way as their opponent in the long-run (Brede, 2013). However, repeat and non-repeat customers appear to have similar strategies with tipping. The model predicts that repeat customers will constantly be tipping because the restaurant provided a good enough overall experience for them to return for more, while one-time customers should feel no need to tip. The analysis from this paper shows that customers do not follow the same logic in the tipping realm.

The Prisoner's Dilemma uses the logic that individual actors are selfish in economic terms. An explanation for the violation of the common strategy is that there are enhanced societal norms that are pressuring individuals to act in non-economically selfish ways. One of these norms is that of reciprocal nature and altruistic actions, where customers want to reward positive service. However, in fast-casual restaurants, no service is being provided by the cashiers. A tipping of service in this scenario would then expand tipping to grocery stores, sports shops, and any other place with someone ringing up orders or providing a basic service.

Another explanation for the violation of the Prisoner's Dilemma strategy is the lack of standards. People are unsure about when they are supposed to tip, and therefore these suggestions, and especially the defaults, serve as great nudges that have a large effect.

While this paper introduces initial findings in two online studies and one quasi-experiment, more research can examine the specific mechanisms and effects on data in different industry contexts and settings. Different consumer behavior in alternate industries could provide important insights. Many coffee shops are installing similar default tips for very small orders, and examining these effects on revenue could provide insightful results. In the restaurant industry specifically, this paper opens further research about the threshold for tip suggestions. Specifically, what is the ideal suggestion menu that balances high anchoring and reactance theory with whiplash effects. Does an equilibrium exist where higher tips are given without customers feeling reactionary and not desiring to return?

In addition to restaurant managers, this paper raises questions about the tipping industry for policy makers. If fast-casual restaurants are gaining extra revenue through higher suggestions on iPad devices, sit-down restaurants may want to explore investing in technological upgrades instead of paper checks in addition to shifting suggestions. With the rapidly shifting tipping standards in restaurants, however, conversation around abolishing or standardizing tipping will continue to increase. Policy makers and governmental officials should analyze these results to consider creating a unified tipping framework or devising regulations for default suggestions. In the meantime, however, this new evidence shows that restaurants, specifically using newly introduced point of sale technology in fast-casual restaurants, may be able to influence customer tipping through slight nudging by providing altered suggested tipping amounts.

5 Appendices

5.1 Appendix A

Figure 20
Study One Tip Screen: Paper Scene, Middle Suggestions

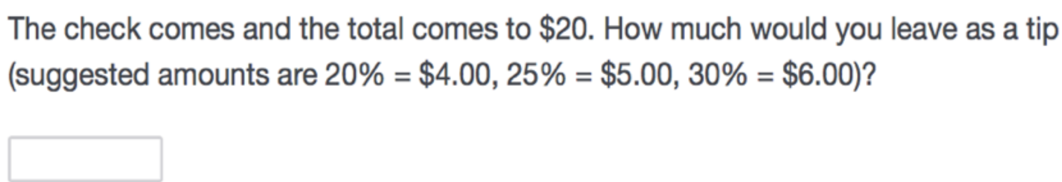
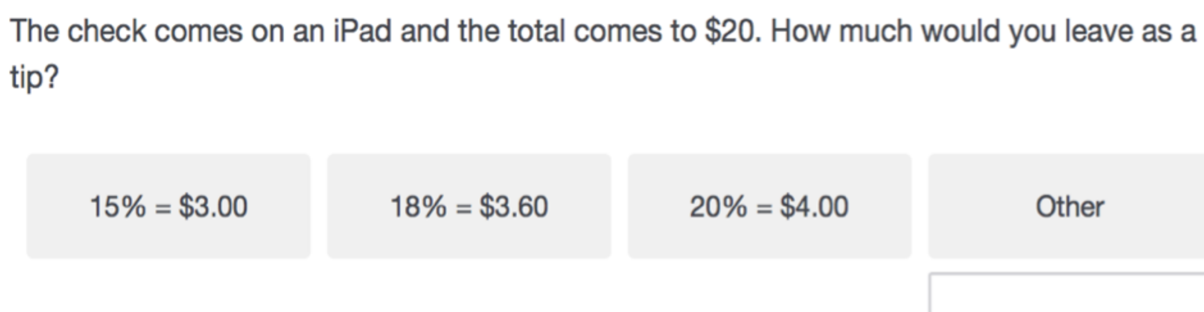


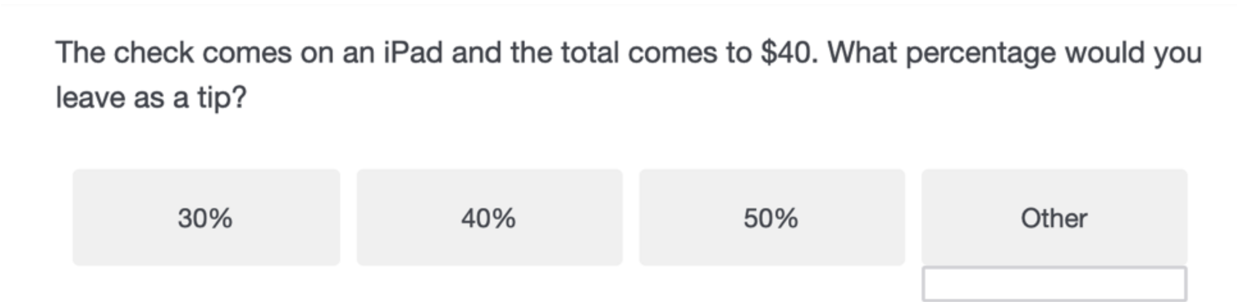
Figure 21
Study One Tip Screen: iPad Scene, Low Suggestions



Note: Figure 20 is an example of a one of 8 conditions of a tip screen in study one. Participants were placed in one of two scenes: iPad (as shown in Figure 20) or paper (as shown in Figure 21). In all 8 conditions, participants are told the check comes out to $20 after sitting down for a one-hour meal. Participants are randomly placed in one of four suggestion conditions in addition, spanning a control with no suggestion, low, middle, and high. Figure 20 displays the middle conditions of 20%, 25%, and 30%, while Figure 20 displays the middle conditions of 20%, 25%, and 30%, both with calculation assistance and dollar equivalents.

5.2 Appendix B

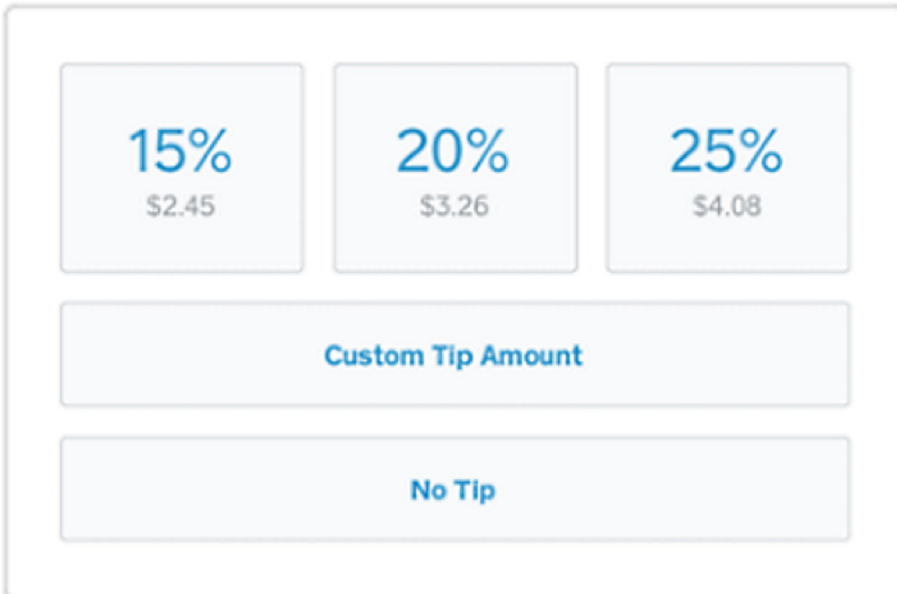
Figure 22
Study Two Tip Screen: High Suggestions



Note: Figure 22 is an example of a tip screen in study two. In all the 4 conditions, the line remains constant with the $40 bill after participants had previously been told that they had sat down for a one-hour meal. In this study, calculation assistance was not provided, and participants were just told to select one of the amounts of enter their own percentage under the "Other" box.

5.3 Appendix C

Figure 23
Field Tipping Screen Example



Note: Figure 23 is an example of a newly introduced technological screen that provides default options that can be pressed as tips in contrast to formerly used tip jars and write-in boxes. This example comes from the Square application, but similar technology is being used through different applications.

6 References

References

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent Arbitrariness": Stable Demand Curves without Stable Preferences. *The Quarterly Journal of Economics*, 118(1), 73-105.

Azar, O. H. (2004). The history of tipping—from sixteenth-century England to United States in the 1910s. *Journal of Socio-Economics*, 33(6), 745-764.

Azar, O. H. (2009). Incentives and service quality in the restaurant industry: The tipping–service puzzle. *Applied Economics*, 41(15), 1917-1927.

Azar, O. H. (2011). Business strategy and the social norm of tipping. *Journal of Economic Psychology*, 32(3), 515–525. https://doi.org/10.1016/j.joep.2011.03.018

Bloom, E. (2016, July 15). Square is guilting us into tipping basically everyone. Retrieved from https://qz.com/730315/squares-design-guilts-us-into-tipping-basically-everyone/

Bloom, E. (2017, September 27). See how your spending compares with that of the average American-and the US government. Retrieved from https://www.cnbc.com/2017/09/27/how-your-spending-compares-to-the-average-american-and-us-government.html

Brede, M. (2013). Short versus long term benefits and the evolution of cooperation in the prisoner's dilemma game. *PLoS One*, 8(2) doi:http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1371/journal.pone.0056016

Brehm, J. (1966). *A theory of psychological reactance (Social psychology)*. New York: Academic Press.

Brenner, M.L. (2001). *Tipping for Success: Secrets for How to Get In and Get Great Service*. Brenmark House, Sherman Oaks, CA.

Bureau of Labor Statistics. (2018, September 13). CONSUMER EXPENDITURES–2017. Retrieved from https://www.bls.gov/news.release/cesan.nr0.htm

Cialdini, R., Schroeder, D., & Lanzetta, John T. (1976). Increasing compliance by legitimizing paltry contributions: When even a penny helps. *Journal of Personality and Social Psychology*, 34(4), 599-604.

Conlin, Lynn, & O'Donoghue. (2003). The norm of restaurant tipping. *Journal of Economic Behavior and Organization*, 52(3), 297-321.

Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin*, 30(4), 447-460.

Galante, M. (2012, January 04). A Very Simple Guide to Tipping Etiquette Around The World. Retrieved from https://www.businessinsider.com/guide-to-tipping-around-the-world- 2012-1

Garrity, K., & Degelman, D. (1990). Effect of Server Introduction on Restaurant Tipping 1. *Journal of Applied Social Psychology*, 20(2), 168-172.

Goswami, I., & Urminsky, O. (2016). When Should the Ask Be a Nudge? The Effect of Default Amounts on Charitable Donations. *Journal of Marketing Research*, 53(5), 829-846.

Guéguen, N. (2002). The effects of a joke on tipping when it is delivered at the same time as the bill. *Journal of Applied Social Psychology*, 32(9), 1955-1963.

Guéguen, N., & Legoherel, P. (2000). Effect on Tipping of Barman Drawing a Sun on the Bottom of Customers' Checks. *Psychological Reports*, 87(1), 223-226.

Haggag, K., & Paci, G. (2014). Default Tips. *American Economic Journal: Applied Economics*, 6(3), 1-19.

Johnson, E., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338-1339.

Kane, L. (2014, April 07). Here's how much you should tip in every situation. Retrieved from https://www.businessinsider.com/what-you-should-tip-in-every-situation-2014-4

Karniouchina, E., Mishra, H., & Verma, R. (2008). Exploring consumer reactions to tipping guidelines: Implications for service quality. *Cornell Hospitality Report*, 8, 4–16.

Lynn, M., Le, J. & Sherwyn, D. (1998). Reach out and touch your customer. *Cornell Hotel and Restaurant Administration Quarterly*, 39, 60-65.

Lynn, M. (1996). Seven ways to increase servers' tips. *Cornell Hotel and Restaurant Administration Quarterly*,37(3), 5,24-5,29.

Lynn, M. (2005). Increasing Servers' Tips What Managers Can Do and Why They Should Do It. *Journal of Foodservice Business Research*, 8(4), 87–96. https://doi.org/10.1300/J369v08n04_07

Lynn, M., & Mynier, K. (1993). Effect of Server Posture on Restaurant Tipping 1. *Journal of Applied Social Psychology*, 23(8), 678-685.

Madrian, B., & Shea, D. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics*, 116(4), 1149-1187.

Martinez-Moncada, D. (2011, September 19). Tipping Etiquette Around the World [Infographic]. Retrieved from https://www.dailyinfographic.com/tipping-etiquette- around-the-world-infographic

Michaels, M. (2018, April 09). How much people tip on average across the US - and there's a big difference between red and blue states. Retrieved from https://www.businessinsider.com/how-much-americans-tip-every-state-ranked-2018-3

Mobile Payment Market by Mode of Transaction (SMS, NFC, and WAP), Type of Mobile Payment (Mobile Wallet/Bank Cards and Mobile Money) is anticipated to grow at a CAGR of 32% from 2017 to 2023. (2018). M2 Presswire, p. M2 Presswire, June 20, 2018.

Mulcahy, J. (2013, October 2). $48.56: What Can You Get For NYC's Average Meal Cost? Retrieved from https://www.zagat.com/b/48.56-what-can-you-get-for-nycs-average- meal-cost

Peters, J., Beck, J., Lande, J., Pan, Z., Cardel, M., Ayoob, K., & Hill, J. (2016). Using Healthy Defaults in Walt Disney World Restaurants to Improve Nutritional Choices. *Journal of the Association for Consumer Research*, 1(1), 92-103.

Pickert, R., & Lanman, S. (2018, August 15). Americans Are Spending Like Crazy at Restaurants. Retrieved from https://www.bloomberg.com/news/articles/2018-08- 15/americans-are-spending-like-crazy-at-restaurants-maybe-too-much

Prelec, & Simester. (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. Marketing Letters, 12(1), 5-12.

Rind, B., & Bordia, P. (1995). Effect of Server's "Thank You" and Personalization on Restaurant Tipping. *Journal of Applied Social Psychology*, 25(9), 745–751. https://doi.org/10.1111/j.1559-1816.1995.tb01772.x

Rind, B., & Bordia, P. (1996). Effect on Restaurant Tipping of Male and Female Servers Drawing a Happy, Smiling Face on the Backs of Customers' Checks. *Journal of Applied Social Psychology*, 26(3), 218–225. https://doi.org/10.1111/j.1559-1816.1996.tb01847.x

Rind, B., & Strohmetz, D. (2001). Effect on restaurant tipping of presenting customers with an interesting task and of reciprocity. *Journal of Applied Social Psychology*, 31, 1379-1384.

Segrave, K. (1998). *Tipping: An American social history of gratuities*. McFarland.

Seiter, J., Brownlee, G., & Sanders, M. (2011). Persuasion by Way of Example: Does Including Gratuity Guidelines on Customers' Checks Affect Restaurant Tipping Behavior? *Journal of Applied Social Psychology*, 41(1), 150-159.

Seiter, J., & Weger, H. (2013). Does a customer by any other name tip the same?: The effect of forms of address and customers' age on gratuities given to food servers in the United States. *Journal of Applied Social Psychology*,43(8), 1592-1598.

Shai Davidai, Thomas Gilovich, & Lee D. Ross. (2012). The meaning of default options for potential organ donors. *Proceedings of the National Academy of Sciences*, 109(38), 15201-15205.

Simpson, Hilery. (1997). Tips and excluded workers: The New Orleans test. (New Orleans, Louisiana). *Compensation and Working Conditions*, 2(2), 32-36.

Stritch, J., Pedersen, M., & Taggart, G. (2017). The Opportunities and Limitations of Using Mechanical Turk (MTURK) in Public Administration and Management Scholarship. *International Public Management Journal*, 20(3), 489-511.

Strohmetz, D., Rind, B., Fisher, R., & Lynn, M. (2002). Sweetening the till: The use of candy to increase restaurant tipping. *Journal of Applied Social Psychology*, 32(2), 300-309.

Swan, A. (2018, August 31). Proof That Square Is Completely Disrupting Payment Processing. Retrieved from https://www.forbes.com/sites/andyswan/2018/08/31/square-growth/#3d2c2b9a291d

Tidd, K., & Lockard, L. (1978). Monetary significance of the affiliative smile: A case for reciprocal altruism. *Bulletin of the Psychonomic Society*, 11(6), 344-346.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*,185(4157), 1124-1131. doi:10.1126/science.185.4157.1124

United States Department of Labor. (2016, September). Wage and Hour Division (WHD). Retrieved from https://www.dol.gov/whd/regs/compliance/hrg.htm

Van Baaren, R. B. (2005). The Parrot Effect: How to Increase Tip Size. *Cornell Hotel and Restaurant Administration Quarterly*, 46(1), 79–84. https://doi.org/10.1177/0010880404270062

Zimmerman, S. (2014). The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics*, 32(4), 711-754.