Default Tips

By Kareem Haggag and Giovanni Paci *

We examine the role of defaults in high-frequency, small-scale choices using unique data on over 13 million NYC taxi rides. Using a regression discontinuity design, we show that default tip suggestions have a large impact on tip amounts. These results are supported by a secondary analysis that uses the quasi-random assignment of customers to different cars to examine default effects on a wider range of fares. Finally, we highlight a potential cost of setting defaults too high, as a higher proportion of customers opt to leave no credit card tip when presented with the higher suggested amounts.

The large effects of default options on consumer choices have been documented in various high-stakes, but low-frequency contexts, ranging from organ donation to 401(k) contributions. Because defaults preserve freedom of choice, but nonetheless appear to strongly influence behavior, they have been of great interest to both policy makers and academics (*Nudge*, Thaler and Sunstein 2008). In contrast to the extant literature, we study the effects of defaults on a frequently encountered consumer choice: the decision of how much to tip a service provider.[1] By studying tipping, we demonstrate the ability of defaults to nudge behavior in a decision problem which agents have arguably encountered enough times to learn their optimal responses. In doing so, we also extend the literature by documenting a case in which default effects were exploited by a for-profit industry.

Our study introduces a unique data set that contains fare information for 170 million NYC taxi rides over the calendar year of 2009. Among these rides, we have tip information for the 38 million credit card transactions, from which we use a sample of more than 13 million rides to study tipping (all credit card transactions on rides with no tolls, taxes, or surcharges). At the end of each ride, customers who used credit cards were presented with a screen that provided them with the option to either type in a desired tip amount or to press one of three buttons with default tip suggestions. During the period of study, one of the credit card machine companies offered different tip suggestions depending on whether the fare

* Haggag: University of Chicago, Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637 (e-mail: kareem.haggag@chicagobooth.edu); Paci: Columbia University, 420 W. 118th St., Room 1022, New York NY 10027 (e-mail: gp2282@columbia.edu). We thank Richard Johns and Jason Gonzalez at the New York City Taxi and Limousine Commission for helpful discussions and data, Stephan Meier, Devin Pope, Bernard Salanie, and Eric Verhoogen for reading earlier drafts of this paper and providing excellent suggestions that greatly improved its quality, and Chanont Banternghansa, Seth Blumberg, Colin Camerer, Ray Fisman, Eric Johnson, Supreet Kaur, John List, Chenfei Lu, Suresh Naidu, Adam Osman, Kiki Pop-Eleches, Sebastien Turban and Miguel Urquiola for helpful comments and discussions at various stages of this project.

[1] Though difficult to precisely measure, Azar (2011) estimates total annual tipping in the US food industry alone at $47 billion, or approximately 0.3% of annual GDP.

was above or below \$15. For rides under \$15, tip suggestions were \$2, \$3, and \$4, while rides above \$15 were presented with 20%, 25%, and 30% tip suggestions and the corresponding dollar amounts. At the discontinuity, this shift represents an increase in the suggestion categories (low, medium, and high) of approximately \$1, \$0.75, and \$0.50. Importantly, the shift in suggestions did not change the choice set; customers were still free to key in any tip amount. Under the assumption that all ride characteristics that affect tips vary smoothly with the base amount, the difference at the discontinuity can be used to identify the causal effect of this particular increase in default suggestions on tipping. We find that this local treatment effect is an increase in tip amounts of approximately \$0.27 - \$0.30, which is more than a 10% increase in the average tip at that margin.

The design of this choice context is not typical of the default effects literature. In order to complete a transaction, customers had to enter a tip amount. As such, our study is closely related to the literature on *active choice*, a type of design in which customers are not provided with a no-action option, and are thus required to actively choose among a set of similarly presented options (e.g. Carroll et al. (2009) in the context of retirement savings). Keller et al. (2011) go a step further, presenting cases in which some options are advantaged by using favorable language, a design they describe as *enhanced active choice*. Our context goes even further in the direction of a default, with the featured amounts being strongly effort advantaged and highly salient. These features draw the influence of these options over choice closer to that of a default, and so we describe the buttons as *default tip suggestions*.

To examine the role of default suggestions across a larger range of fares, we present a second econometric strategy. We use the quasi-random assignment of passengers to taxi cabs at LaGuardia airport to compare across credit card machine companies. For rides above \$15, both companies provided percentage defaults; however, one company provided 15%, 20%, and 25% percent, while the other provided defaults of 20%, 25%, and 30% percent. The distribution of tips clearly reflects this shift, and again, we find that higher defaults are associated with higher average tip amounts, controlling for time-invariant driver characteristics.

Having demonstrated the benefits of higher default suggestions on the intensive margin of tipping, we next highlight a potential cost of setting defaults too high. First, in both the regression discontinuity design and the comparison across vendors, we find that the higher default suggestions reduce the probability of leaving a tip that corresponds to one of the default suggestions (24 and 7.8 percentage point reductions respectively).[2] More striking is the result that rides with the higher tip suggestions are over 50% more likely to receive a zero-valued tip than those with the lower suggestions (1.7 and 2.8 percentage point increases). Such customers may have been penalizing drivers for using tip defaults that are perceived as *unfairly* high. Similarly, the absence of the lower amount button may have been construed by customers as an attempt to manipulate their behavior which induced them to respond adversely, as described in social psychology by *reactance* theory (Brehm

[2]However, we also find that the average manually entered tip amount increases. Thus, it is not clear that those induced to leave a manual tip are leaving lower tips than they would with the lower suggestions.

1966).

Several factors may explain our observed default effects. Customers may be rationally inattentive, failing to compute their preferred tip due to the opportunity cost of time and/or the cognitive costs associated with that computation. For instance, customers may have difficulty converting between dollars and percentages, and the alternative measures may evoke different intuitions on the appropriate tip (Kahneman 2011, p.372).[3] Moreover, customers that are unfamiliar with the tipping norm may interpret the defaults as the socially endorsed norm. Both uninformed and informed customers may experience disutility from deviating from these options. Ultimately, we cannot disentangle these different possible mechanisms.

We build upon the broad literature on defaults. Default effects have been demonstrated across a wide variety of consumer choices. Most notably, Madrian and Shea (2001) and Choi, Laibson, and Madrian (2004) found large effects in retirement savings contributions, with Madrian and Shea (2001) finding a 50% increase in enrollment from switching from an opt-in to automatic enrollment default. In a similarly sparsely encountered consumer choice, Johnson and Goldstein (2003) and Abadie and Gay (2006) used cross-country analyses to suggest that presumed consent policies induce higher organ donation rates than opt-in policies. Johnson et al. (1993) studied a somewhat more frequently encountered type of decision problem, namely (car) insurance plan choice, while Johnson, Bellman, and Lohse (2002) studied default effects in the decision to accept email marketing. Our paper contributes by showing that default effects can persist in a similarly habitually encountered consumer choice, using a much larger naturalistic field data set. We also add to the literature by tracing out the response to higher defaults, including its limitations. Beshears et al. (2010) similarly study the limits to setting high defaults. They provide a case study of a firm that set the default contribution rate at 12%, a rate much higher than previously studied defaults in this area (2% - 6%) and one that the authors note was sub-optimal for all employees in the sample. They find that roughly 25% of employees remain at this default rate after 12 months of tenure, in comparison to the 60% adherence rate seen at firms in previous studies. In our study, we find that a substantial proportion are induced to opt out of the default when presented with the higher suggestions. We still find a higher average contribution despite this result; however, our analysis also highlights the emergence of a cost (zero-valued tips) that suggests a potential reduction in tips if defaults are set sufficiently high.

The paper proceeds as follows. Section I provides background on taxis and tipping and describes the data used. Section II presents our regression discontinuity results. Section III presents an analysis that compares across credit card machine companies.

[3]On its own, the difficulty of comparing across the measurements does not imply that we should find higher tip amounts for the percentage suggestions. One potential explanation of this pattern would rely on this computational difficulty interacting with a particular type of self-deception. If customers adhere to tip *percentage* norms, then dollar suggestions could result in less generous tips by lowering the cost of self-deception. For example, consider a customer that has a fare of $13 and adheres to a 25% tipping norm (i.e. a tip of $3.25). This customer may be able to convince herself that she is adhering to the norm by selecting the $3 option (rounding in the direction of her self-interest), whereas she could not ignore her deviation from the norm if explicitly presented with the 25% option.

Section IV concludes.

I. Institutional Context and Data

The data for our study were provided by the Taxi and Limousine Commission (TLC) of New York City. In May 2004, the TLC mandated that all taxi cabs be outfitted with a set of technological improvements, including the electronic collection and transmission of trip data and the introduction of equipment to accept credit cards.[4] These technological improvements also marked the introduction of a system that measured and saved the GPS coordinates of all pick-up and drop-off locations. Though mandated in 2004, the entire taxi fleet was not outfitted with the equipment until 2008.[5] Our data spans the entirety of 2009, covering all rides by licensed Yellow Cab drivers in NYC. Before describing the data, we first present details about the institutional context.

A. *Institutional Context*

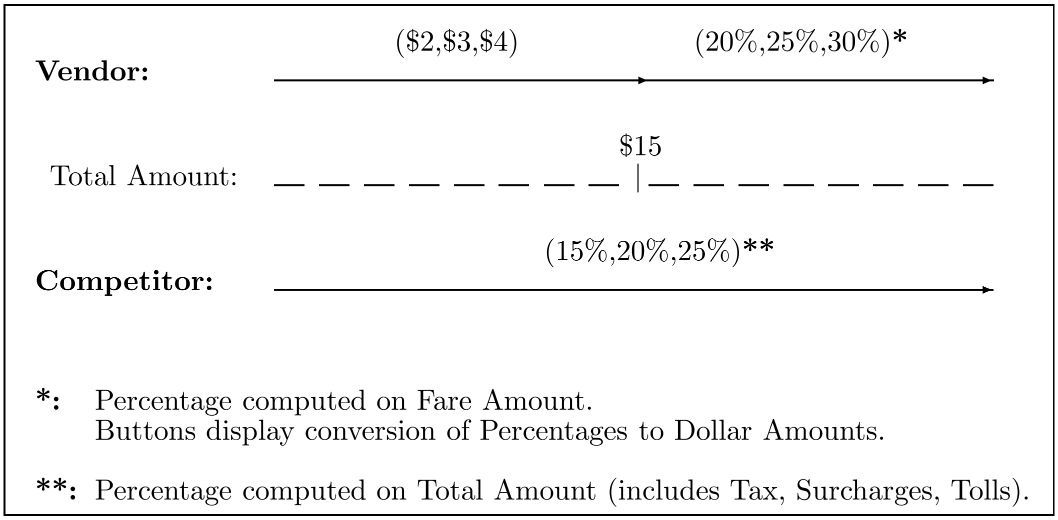
During the period of study, three companies were contracted to provide taxi cabs with credit card machines. The largest two, which we denote as "Vendor" and "Competitor", account for 49% and 45% of observations in the raw data respectively. Each taxi cab was equipped with its own Passenger Information Monitor (PIM) which would display advertisements and other viewing material during the ride. At the end of the ride, the PIM displayed a payment screen (see the Online Appendix for an example). Customers were presented with the base amount and had the option of keying in their own tip amount or using one of the suggested tip buttons. Each vendor was allowed discretion over how this page was displayed, and the two companies elected to offer different default buttons during the period we study.

There were three key ways in which the Vendor and Competitor differed with respect to tip suggestions. The Competitor offered three suggestions on all base amounts: 15%, 20%, or 25%. In contrast, the Vendor provided one set of suggestions ($2, $3, or $4) for all base amounts lower than $15, and another set of suggestions (20%, 25%, or 30%) for all base amounts above $15. The second difference is how the percentages were calculated. Though the Vendor used the base amount (i.e. the fare, toll, tax, and surcharge) to determine which set of suggestions to provide, the percentage tips were calculated on only the fare. The Competitor instead calculated percentages on the entire base amount. Thus, if the ride consisted of a $10 fare and a $10 toll, a customer that pushed the 20% button with the Vendor equipped cab would be paying $2 in tip, whereas a customer in a Competitor equipped cab would be paying $4. A third difference is that the Vendor displayed corresponding dollar amounts alongside the percentage tips, while the Competitor did not display these conversions. These differences are summarized in Figure 1.

[4]Source: http://www.nyc.gov/html/tlc/html/industry/taxicab_serv_enh.shtml.

[5]Source: "Despite some grumbling, however, the TLC is moving to install the devices in all cars by August 31." http://www.nysun.com/business/hot-tip-for-cabbies-credit-cards-boost-tips/72783.

Figure 1. : Default Tip Suggestions by Vendor and Competitor



Taxi meters determined the fare through a combination of time and distance measures. The standard city rate (Rate Code 1) charged customers $2.50 upon entry, and $0.40 for each additional unit.[6] One unit is defined as either (1) a 60 second interval in which the car is idle or driving less than 6 miles per hour or (2) 0.20 miles when the car is driving 6 miles per hour or faster. Fractional amounts are rounded up to the next unit. Riders were also subject to different sets of surcharges depending on the period of the year or the time of the day. To maintain comparability on either side of the Vendor default discontinuity, we limit our analysis to a period of time during which there were no taxes or surcharges and we exclude rides with tolls. This period includes January 1, 2009 to October 31, 2009 (i.e. prior to the introduction of a $0.50 MTA tax), and spans 6am to 4pm on Monday - Friday and 6am to 8pm on Saturday and Sunday (i.e. periods of time not subject to the $1.00 peak weekday or $0.50 night time surcharges). During this time period, the largest base amount to the left of the $15 discontinuity was $14.90, and the smallest base amount to the right of the discontinuity was $15.30.

[6] Other Rate Codes include:

- Rate Code 2 - Rides to and from JFK - Charged a flat rate of $45.
- Rate Code 3 - Rides to and from Newark Airport - Charged the standard rate in addition to a $15 surcharge.
- Rate Code 4 - Rides to Nassau or Westchester county - Charged the standard city rate while in city limits, and double the standard rate while in Nassau or Westchester county.
- Rate Code 5 - Rides outside NYC, excluding Nassau, Westchester, or Newark Airport - Charged flat rate (determined through negotiation between rider and driver).

Source: http://www.nyc.gov/html/tlc/html/passenger/taxicab_rate.shtml.

B. Data Description

Our preliminary data set included 170,896,479 observations. Though the TLC has its own private routine for removing electronic glitches, the provided data set still contained a number of possible electronic errors, including zero-valued distances/durations and surcharges that did not correspond to the appropriate schedule. We took a number of steps to clean the data, which we outline in greater detail in the Appendix. Our largest sample reductions were the removal of Cash payments (approximately three-quarters of the sample), and the removal of rides with taxes or surcharges. Our final dataset consists of 13,820,784 rides. For the majority of our estimates, we limit our sample to rides complete on cars equipped by the Vendor (7.28 million) and to fares between \$5 and \$25 (6.22 million observations).

The variables provided in the data are as follows: anonymized driver identifier, car identifier, credit card machine company, payment type, ride duration, ride distance, number of passengers, fare, surcharge, MTA tax, toll amount, and pick-up and drop-off latitude, longitude, and time. Because we do not have an indicator for whether the customer physically selected one of the default suggestion buttons, we needed to create this key variable. To do so, we make the assumption that all tip amounts that correspond to one of the relevant tip suggestions (e.g. \$2/\$3/\$4 for Vendor if the base amount is less than \$15) were selected from one of these buttons. We thus make the assumption that customers recognize this congruence and save the time of keying in this amount by pressing a single button.

For the purpose of computing heterogeneous treatment effects, we use data from the American Community Survey's 5 year estimates (2006 - 2010). This dataset provides census tract level summary statistics. We match these statistics to each pickup and drop-off location. To do so, we first assign each GPS coordinate to a census tract using a point-in-polygon operation in PostgreSQL (PostGIS). We then merge each pickup location and each drop-off location with the ACS census tract variables. We focus on one variable in particular: median household income. Table 1 provides summary statistics for the sample, split by Vendor and Competitor.

Table 1—: Summary Statistics by Ride

	(1) Competitor	(2) Vendor	(3) Difference [(1)-(2)]
Fare	9.690	9.813	-0.123***
	(5.479)	(5.668)	
Tip Amount	1.694	1.920	-0.227***
	(1.253)	(1.431)	
Tip as Percentage of Fare	18.266	21.463	-3.196***
	(9.254)	(16.602)	
Tip Corresponds to a Default Tip Option	0.556	0.495	0.061***
	(0.497)	(0.500)	
Ride Duration (Minutes; Dropoff Time - Pickup Time)	12.674	12.854	-0.180***
	(8.136)	(7.931)	
Ride Distance (Miles)	2.531	2.596	-0.065***
	(2.310)	(2.401)	
Zero Tip	0.020	0.029	-0.009***
	(0.139)	(0.167)	
High Choice [Pr(Select ‘High’ Default | Select a Default Option)]	0.129	0.037	0.092***
	(0.335)	(0.189)	
Med Choice [Pr(Select ‘Med’ Default | Select a Default Option)]	0.419	0.184	0.234***
	(0.493)	(0.388)	
Low Choice [Pr(Select ‘Low’ Default | Select a Default Option)]	0.452	0.778	-0.326***
	(0.498)	(0.415)	
Observations	6,542,807	7,277,977	13,820,784

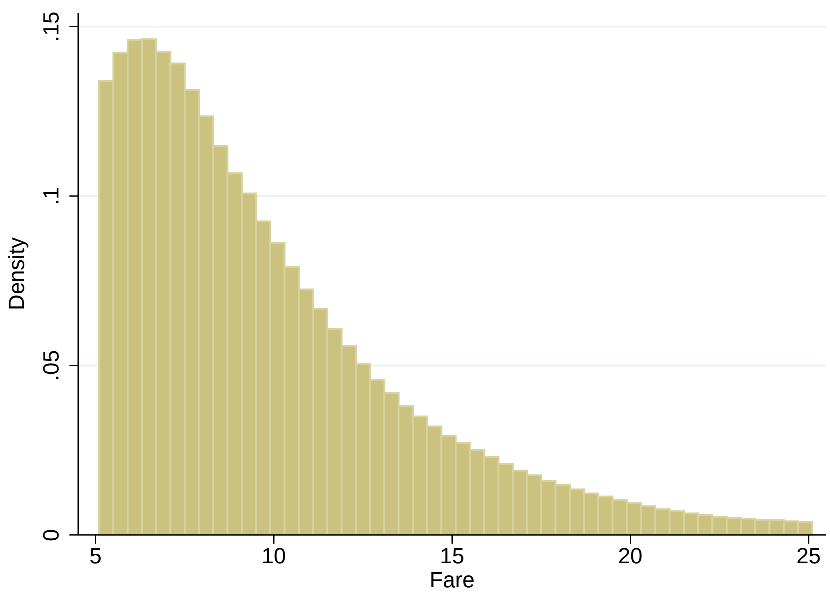
Note: Standard deviations are in parentheses. T-tests for the equality of columns 1 and 2 are rejected at the 1% level for all differences. The sample is limited to rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). High Choice (Vendor: \$4 or 30%, Competitor: 25%), Medium Choice (V: \$3 or 25%, C: 20%), and Low Choice (V: \$2 or 20%, C: 15%) estimates are conditional on using a default tip option.

II. Regression Discontinuity

A. Visual Evidence

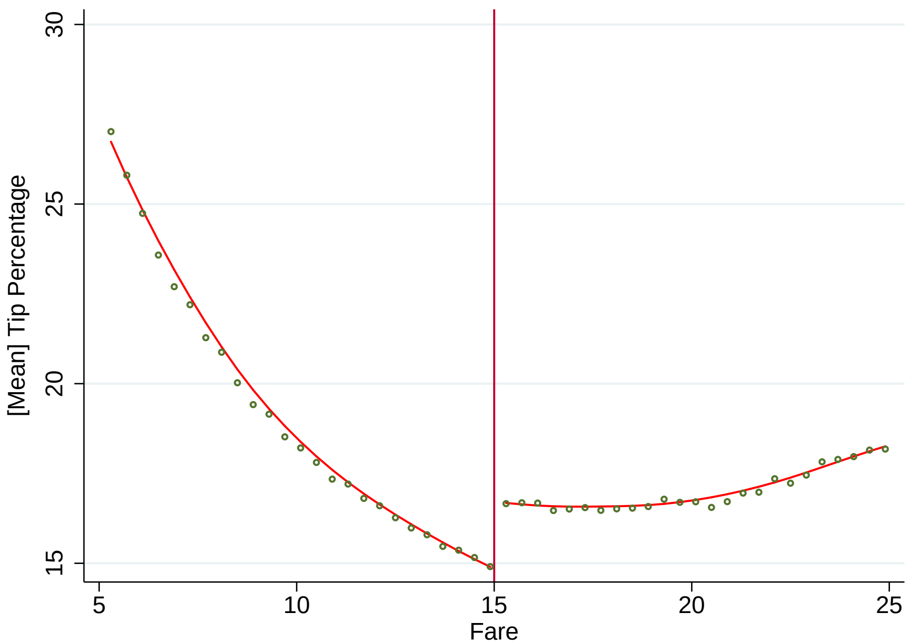
We start by presenting a visual analysis of the discontinuity. As a first test of the validity of the regression discontinuity approach, Figure 2 demonstrates that the density of the forcing variable is smooth. Next, we limit the forcing variable (fare) to be between \$5 and \$25 and calculate the mean tip percentage within each of the discrete fare amounts (\$0.40 increments). On each side of the discontinuity, we scatter plot these estimates and perform a lowess smoother separately on either side of the discontinuity. Figure 3 displays this plot for tip percentages on Vendor-equipped cabs, clearly demonstrating a jump at \$15. Finally, as a robustness test, Figure 4 repeats this graph for the Competitor, showing no jump at \$15.

Figure 2. : Histogram of Fares



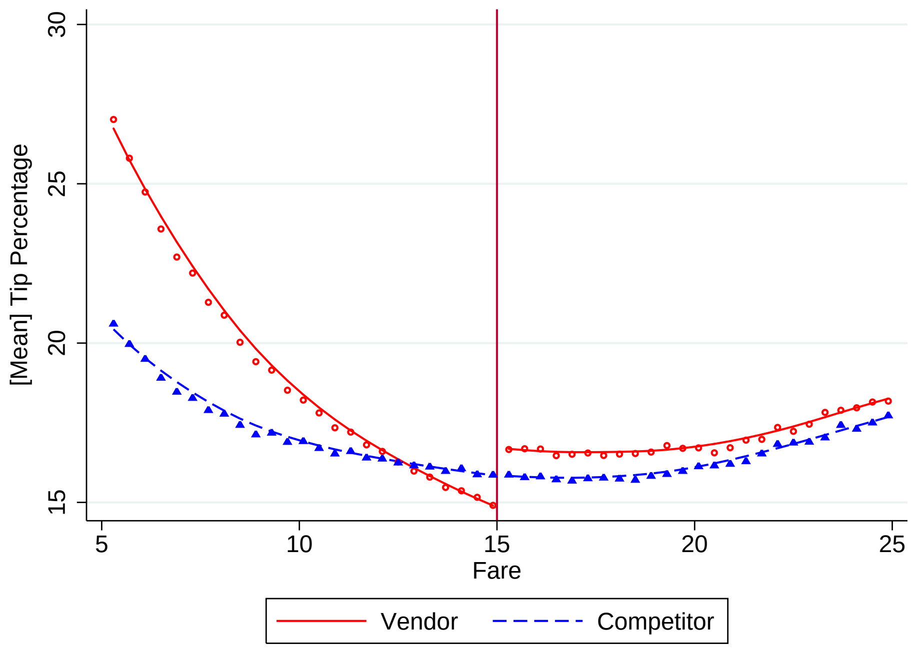
Note: The sample is limited to fares between \$5 and \$25 on **Vendor-equipped** cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). Fares are in \$0.40 intervals.

Figure 3. : Lowess Smoothed Mean Tip Percentages Within Each Discrete Fare Amount ($0.40 Intervals)



Note: Each dot is the average within a discrete fare amount ($0.40 intervals). Solid lines display the smoothed values from locally weighted regressions performed separately on either side of the discontinuity ($15). The sample is limited to fares between $5 and $25 on **Vendor-equipped** cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). Fares are in $0.40 intervals.

Figure 4. : Lowess Smoothed Mean Tip Percentages Within Each Discrete Fare Amount ($0.40 Intervals), for Vendor and Competitor



Note: Each dot is the average within a discrete fare amount ($0.40 intervals). Solid and dashed lines display the smoothed values from locally weighted regressions performed separately on either side of the discontinuity ($15). The sample is limited to fares between $5 and $25 on rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). Fares are in $0.40 intervals.

B. Regression

To supplement the visual evidence, we estimate a regression discontinuity model for the case of a forcing variable with discrete support. To the extent that the tip paid is smoothly related to the fare, observations at either side of the cutoff can be used to identify the causal effect of a change in the suggested amount. Following Lee and Card (2008), we estimate equation (1):

$$(1) \quad Y_r = \beta_0 \mathbb{1}(F_r \geq 15) + \beta_1 h\,(F_r - 15) + \beta_2 \mathbb{1}(F_r \geq 15) * g\,(F_r - 15) + \mathbf{X_r}\theta + u_r$$

Where Y_r is the tip amount in dollars, $\mathbb{1}(F_r \geq 15)$ is an indicator function that the fare is greater than or equal to \$15, $h\,(F_r - 15)$ and $g\,(F_r - 15)$ are polynomials in the fare, centered at \$15, on either side of the discontinuity, and $\mathbf{X_r}$ is a vector of fixed effects consisting of pick-up hour, day of week, pick-up borough, and drop-off borough indicators. Since our source of variation is at the fare value relative to the discontinuity, rather than the ride level, we follow Lee and Card (2008) and cluster our standard errors at the level of the forcing variable, thereby correcting our degrees of freedom and allowing for random specification errors due to the discrete bins. We estimate four specifications in Table 2, starting with a 2^{nd} order polynomial in the first column up to a 5^{th} order polynomial in the last column. Our local treatment effect is a \$0.27 to \$0.30 increase in tip amounts over a baseline level at the cut-off of \$2.22.

Table 2—: Regression Discontinuity Estimates of the Effect on Tip Amount

	Tip Amount			
	(1)	(2)	(3)	(4)
$\mathbb{1}(Fare_r \geq 15)$	0.292***	0.276***	0.275***	0.296***
	(0.004)	(0.006)	(0.008)	(0.010)
N	6,218,196	6,218,196	6,218,196	6,218,196
r2	0.207	0.207	0.207	0.207
DepVarMean	2.221	2.221	2.221	2.221

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Robust standard errors clustered at each fare value (\$0.40 intervals), in parentheses. Columns (1) - (4) present 2^{nd}-5^{th} order polynomials. $\mathbb{1}(Fare_r \geq 15)$ is an indicator function that the fare is greater than or equal to \$15. *DepVarMean* is the mean of the dependent variable on rides with fares of \$14.90. All specifications include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to fares between \$5 and \$25 on Vendor-equipped cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

C. Other Outcomes of Interest

Our primary outcome of interest, tip amount, is produced through movements along an extensive margin (using a default suggestion) and two intensive margins (amounts tipped either manually or through one of the suggestions). Table 3 presents results for a number of other variables.

Table 3—: Regression Discontinuity Estimates of the Effect on Alternative Outcomes

	(1) Tip Percent	(2) Default Tip	(3) Default Tip Amt	(4) Manual Tip Amt	(5) High Choice	(6) Med Choice	(7) Low Choice	(8) Zero Tip
$\mathbb{1}(Fare_r \geq 15)$	2.025***	-0.243***	0.714***	0.368***	-0.021***	-0.169***	0.190***	0.017***
	(0.038)	(0.002)	(0.003)	(0.011)	(0.001)	(0.002)	(0.003)	(0.001)
N	6,218,196	6,218,196	3,227,726	2,990,470	3,227,726	3,227,726	3,227,726	6,218,196
r2	0.097	0.058	0.568	0.122	0.015	0.029	0.038	0.008
DepVarMean	14.907	0.749	2.489	1.422	0.073	0.342	0.585	0.028

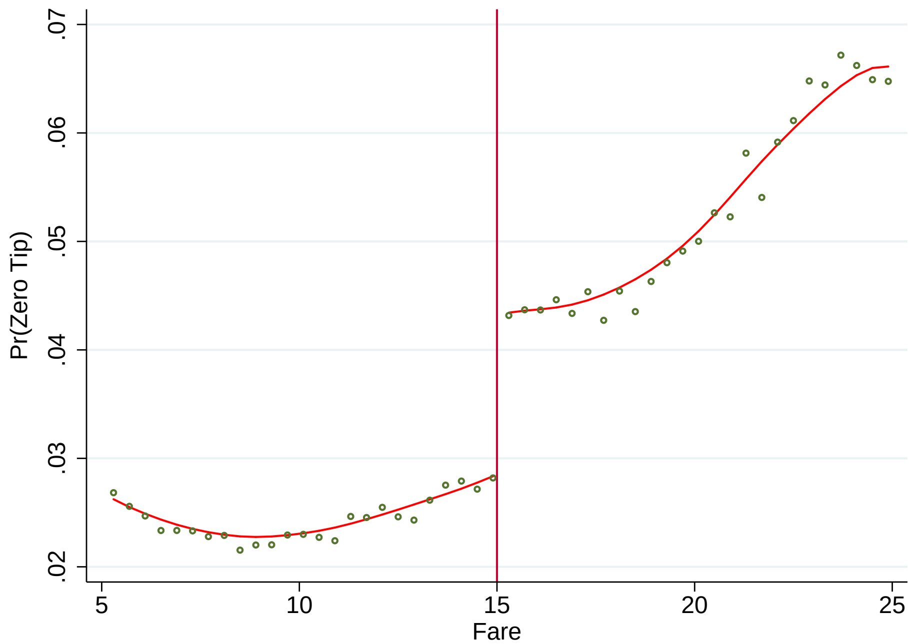
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Robust standard errors clustered at each fare value ($0.40 intervals), in parentheses. $\mathbb{1}(Fare_r \geq 15)$ is an indicator function that the fare is greater than or equal to $15. Columns 1 (*Tip Percent*), 3 (*Default Tip Amount*), and 4 (*Manual Tip Amount*) use continuous outcome variables, while columns 2 (*Default Tip*), 5 (*High Choice*), 6 (*Medium Choice*), 7 (*Low Choice*), and 8 (*Zero Tip*) use binary outcome variables. The dependent variable in column 2 (Default Tip) takes on value 1 if the customer selected one of the default tip suggestions (buttons). The dependent variable in column 3 is the tip amount conditional on using one of the default tip suggestions. The dependent variable in column 4 is the tip amount conditional on manually entering a tip using the keypad. The High Choice (Vendor: $4 or 30%, Competitor: 25%), Medium Choice (V: $3 or 25%, C: 20%), and Low Choice (V: $2 or 20%, C: 15%) variables are conditional on using a default tip option, and take on value 1 if the customer select the corresponding option. The dependent variable in column 8 (*Zero Tip*) takes on value 1 if the customer left zero credit card tip. *DepVarMean* is the mean of the dependent variable on rides with fares of $14.90. All specifications use 3$^{\text{rd}}$-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to fares between $5 and $25 on Vendor-equipped cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Notably, the higher tip suggestions induce a 24 percentage point reduction in the probability of using a default suggestion, and a shift in the composition of those that use default suggestions toward the low option. However, since the low option is approximately equal to the medium option to the left of the discontinuity, we see a net increase in the amount tipped by those that select a default option. There is also an increase in amount tipped manually, which reflects a change in the composition of the tippers, but may also reflect the influence of the higher suggestions on those that would tip manually when faced with either set of suggestions.

Column 8 shows perhaps the most interesting of these behavioral responses. The probability of leaving no credit card tip increases by 1.7 percentage points when customers are faced with the higher tip suggestions. Figure 5 repeats the visual analysis for this outcome variable. This negative response is consistent with the social psychology literature on "reactance". Within this framework, individuals react to a perceived reduction in their freedom of choice by doing the opposite of the intended

Figure 5. : Lowess Smoothed Mean of "Zero-Valued Tip" Within Each Discrete Fare Amount ($0.40 Intervals)



Note: Each dot is the average within a discrete fare amount ($0.40 intervals). Solid lines display the smoothed values from locally weighted regressions performed separately on either side of the discontinuity ($15). The sample is limited to fares between $5 and $25 on **Vendor-equipped** cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

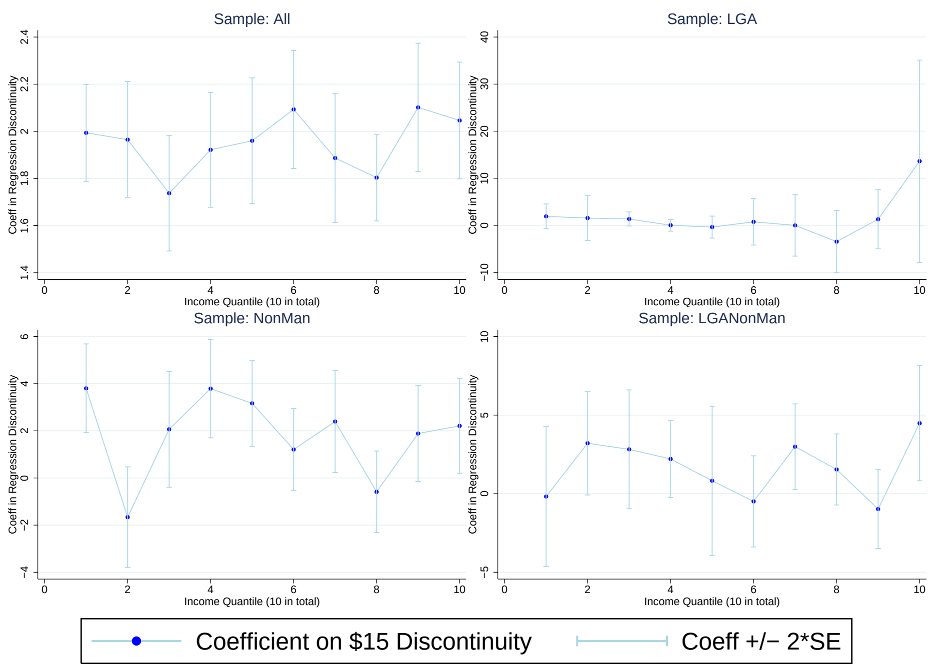
manipulation (Brehm 1966, 1989). Though customers' choice sets of tip amounts were not changed by the shift to higher suggestions, they may have still perceived the replacement of the lower button amounts as a threat to their behavioral freedom. The negative reaction also has some parallel in the vast literature on ultimatum games. Insofar as customers have some fixed notion of a "fair" tip, presenting the higher suggestions might have led them to punish the drivers by leaving a lower tip than would be provided in the absence of the "unfair" split. This result highlights a potential cost of setting defaults too high, although we cannot confirm whether this cost would persist in a context featuring homogeneous suggestions across vendors (e.g. all vendors currently offer the 20%/25%/30% suggestions). The backlash to high suggestions may hinge on the existence of a reference "fair tip" in a comparable market. Furthermore, without making strong assumptions, we cannot use this cost to trace out the set of optimal default suggestions. Finally, we cannot rule out an important possible confound. Since we do not observe cash tips, it may be possible that customers were induced to switch from paying both their tips and fares by credit to instead paying their fares by credit and their tips by cash.

D. Heterogeneity

We next examine heterogeneous treatment effects along proxied income. Because wealthier riders have a lower marginal utility of income, they have less incentive to be attentive to the shift in tip suggestions. Similarly, these customers will potentially have higher time costs. This reasoning suggests that higher income customers should (rationally) exhibit a larger default effect. Alternatively, wealthier individuals may have greater access to distraction-reducing devices, allowing them to deplete less attention during the day (Banerjee and Mullainathan 2008), and thus be less susceptible to default effects. A wealthier customer may also be more likely to take more taxi rides, and default effects possibly attenuate with experience (Lofgren et al. 2012). Ultimately, this exercise is theoretically ambiguous; however, it is an interesting source of potential heterogeneity that has been studied in other contexts. For example, Goldin and Homonoff (2012) found that low-income customers were more attentive to a low salience cigarette tax than were high income customers. In contrast, Beshears et al. (2010) found that 401(k) savings defaults had a greater influence on low-income employees than on high-income employees.

To proxy income, we use a variety of different sample groups. For the full sample, we proxy customer income by the average of the median income associated with the pick up location and drop-off location census tracts. These pick-up and drop-off locations would not be an accurate assessment of tourists or any other customers that do not start or end at their home addresses. To partially reduce this concern, we use a variety of other specifications that attempt to remove these non-representative customers. One set of specifications limit the sample to rides that both start and end outside Manhattan. Another set of specifications limit to rides that either start or end at LaGuardia (LGA) airport, proxying income by the median income in the pick-up (drop-off) location census tract if the ride ends (starts) at LGA. Finally, the most conservative set of specifications limits to rides that start at LaGuardia airport and end outside Manhattan. We then split these rides into ten income quantiles and run the regression discontinuity for each of these sub-samples. Figure 6 plots the coefficients from these regressions, finding no systematic pattern in income. Although it might be that the absence of a clear pattern is due to measurement error in the income variable, the constancy in the discontinuity suggests that default effects are similar for customers across the different proxied income brackets in our sample.

Figure 6. : Coefficients from Regression Discontinuity Estimates in Income Quantile Sub-Samples



Note: Robust standard errors clustered at each fare value ($0.40 intervals). All specifications use 3rd-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to fares between $5 and $25 on Vendor-equipped cab rides without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). All income measures correspond to American Community Survey 5-year (2006-2010) estimates of the census tract median income. The LGA sample is restricted to rides that either originate or end within the census tract associated with LaGuardia Airport. Income is proxied by the median income in the pick-up location census tract if the ride ends at LGA or by the income at the drop-off location if the ride starts at LGA. The NonMan sample is restricted to rides that both start and end outside Manhattan, and uses the average of the median income at the pick-up and drop-off census tracts. The LGANonMan sample corresponds to rides that start at LaGuardia Airport and end outside of Manhattan. The Average (Median ACS Census Tract) income within each quantile are as follows: **All**: 49,244 - 68,340 - 78,945 - 86,876 - 93,574 - 99,455 - 105,490 - 112,490 - 121,617 - 145,084.
NonMan: 26,636 - 38,634 - 45,384 - 49,978 - 53,906 - 57,856 - 63,013 - 71,318 - 82,629 - 106,607.
LGA: 27,775 - 43,948 - 52,174 - 58,943 - 69,240 - 81,628 - 92,334 - 102,506 - 114,219 - 138,099.
LGANonMan: 28,271 - 42,718 - 48,115 - 51,752 - 54,076 - 57,739 - 61,705 - 69,823 - 77,697 - 118,845.

III. Comparing Across Vendors

While our regression discontinuity design provides compelling identification, it is limited to a localized treatment effect. One way to expand upon our results would be to compare rides over which both credit card machine companies provided percentage default suggestions. For fares above $15, the Vendor provided suggestions of 20%, 25%, and 30%, while the competitor provided suggestions of 15%, 20%, and 25%. However, there are several potential differences in the matches of customers and drivers between the two companies (e.g. see the Online Appendix for a figure depicting the geospatial distribution of pick-up locations between the two companies). While we can control for the pick-up and drop-off location, there may be other unobservable differences in driver-customer match that affect the tip amounts. To address this challenge, we limit our analysis to rides that originate at LaGuardia airport.[7] Customers queue at lines that contain a mix of taxis equipped with both credit card machine companies. Panel A of Figure 7 shows that the distribution of fares is comparable across the two credit card companies when we limit the sample to rides that are above $15 and originate at LaGuardia.[8]

For the distribution of tip percentages, we limit the sample to fares with tip percentages less than or equal to 50% in order to provide greater visual clarity. Panel B of Figure 7 demonstrates the stark difference in the two distributions, with the higher set of defaults inducing a distribution that has significantly more density around its lowest option. The left tail of the figure is also larger for the Vendor; however, this effect is limited to zero-valued tips.

We provide a regression analysis of these effects in Table 4. To address concerns of any type of sorting between drivers and credit card machine companies, we also provide specifications with fixed effects for driver, pick-up hour, and borough of the drop-off location. These fixed effects specifications exploit the 7% of drivers in our sample that drove on cars equipped by both companies on rides from LaGuardia Airport, allowing us to identify the coefficient on "Vendor" while controlling for time-invariant driver characteristics.[9] However, it is important to note that we cannot control for the influence of possible additional differences between the two companies in the distribution of taxi cab characteristics, such as the display of the payment screens (e.g. the Vendor, unlike the Competitor, paired the percentage suggestions with their dollar conversions).

We find that the small difference in fare is significant at the 5% level in specifications that do not include fixed effects (Column 1), and insignificant in specifications that do include fixed effects (Column 2) – though these coefficients are not statistically different from each other. We find a significant increase in the tip

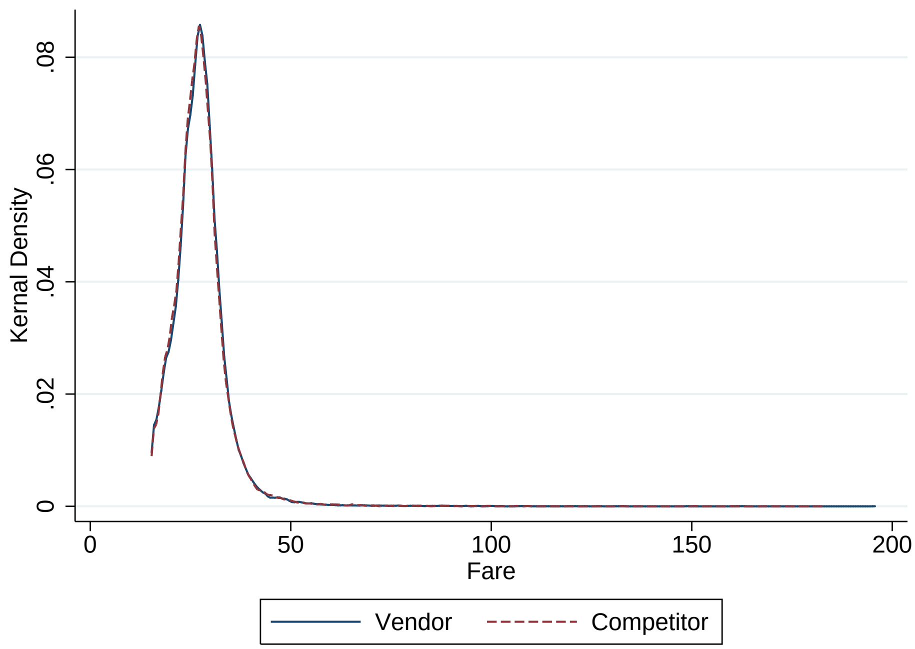
[7]We exclude JFK airport because the majority of rides use a $45 fixed fare, complicating our placebo test of equality in fares between vendors. In the Online Appendix, we repeat the analysis in this section pooling both LGA and JFK observations. Our estimates in that pooled sample are qualitatively similar.

[8]However, it should be noted that a simple t-test of fare across the vendor (27.47) and competitor (27.36) rejects equality ($p = .0089$).

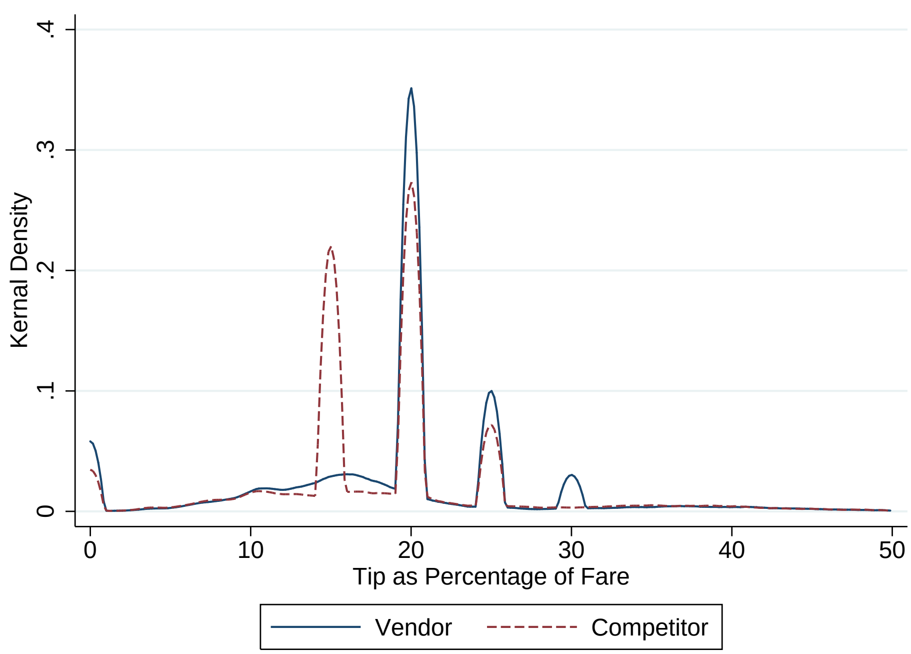
[9]There may be time-varying characteristics that present a threat to identification, such as motivation (e.g. drivers may choose to use Vendor-equipped cabs on days in which they are more motivated to earn income).

Figure 7. : Distribution of Fares (Panel A) and Tip Percentages (Panel B) across Vendor and Competitor Equipped Taxis

[Panel A: Fare]



[Panel B: Tip Percentage]



Note: The sample is limited to fares greater than $15 on cab rides that originated at the census tract associated with LaGuardia Airport, without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). Panel B is limited to rides with tip percentages less than 50%.

Table 4—: OLS - Comparison of Vendor (20%/25%/30%) and Competitor (15%/20%/25%)

	Fare		Tip Percent		Default Tip	
	(1)	(2)	(3)	(4)	(5)	(6)
Vendor	0.128**	0.176	0.675***	0.699***	-0.083***	-0.078***
	(0.056)	(0.146)	(0.099)	(0.259)	(0.004)	(0.013)
N	100,577	100,577	100,577	100,577	100,577	100,577
r2	0.000	0.295	0.001	0.265	0.007	0.222
MeanDepVar	27.047	27.047	18.652	18.652	0.634	0.634
Fixed Effects?		X		X		X
PVal_FEvsNoFE		0.721		0.917		0.661
	TipPercent0to10		Zero Tip		Tip25	
	(7)	(8)	(9)	(10)	(11)	(12)
Vendor	-0.003**	-0.016***	0.028***	0.028***	0.034***	0.037***
	(0.001)	(0.006)	(0.001)	(0.006)	(0.002)	(0.007)
N	100,577	100,577	100,577	100,577	100,577	100,577
r2	0.000	0.213	0.004	0.204	0.003	0.210
MeanDepVar	0.052	0.052	0.039	0.039	0.080	0.080
Fixed Effects?		X		X		X
PVal_FEvsNoFE		0.011		0.998		0.671

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Robust standard errors clustered at the driver level, in parentheses. Even columns include fixed effects for driver, pick-up hour, and drop-off borough. The dependent variable in columns 5 and 6 (*Default Tip*) takes on value 1 if the customer selected one of the default tip suggestions (buttons). The dependent variable in columns 7 and 8 (*Tip Percent >0 & <10*) takes on value 1 if the tip is greater than 0% and less than 10% of the fare. The dependent variable in columns 9 and 10 (*Zero Tip*) takes on value 1 if the customer left zero credit card tip. The dependent variable in columns 11 and 12 (*Tip Percent = 25*) takes on value 1 if the customer selected the 25% tip button. *DepVarMean* is the mean of the dependent variable in the control group (rides on Competitor-equipped cabs). *PVal_FEVsNoFE* is the p-value from a Chow Test for the equality of coefficients across even and odd columns. The sample is limited to fares greater than $15 on cab rides that originated at the census tract associated with LaGuardia Airport, without tolls, taxes, or surcharges (January 1, 2009 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

percentage, consistent with our regression discontinuity estimates, though the effect size is smaller in magnitude. We also find a reduction in the probability of using one of the suggested amounts, also consistent with Section II. Columns 7 and 8 show a small decrease in the probability of leaving a tip greater than 0% but less than 10%; however, columns 9 and 10 show a significant increase in the probability of leaving a zero-valued tip. As with the regression discontinuity analysis, we cannot rule out the possibility that customers presented with the higher options were simply switching to paying their tips in cash. Unlike the regression discontinuity analysis presented in Section II, this comparison suffers from the potentially confounding influence of other differences between the two companies. For example, these zero-valued tip entries could reflect data errors that were more likely to be produced by Vendor machines. In cleaning the data, we removed all zero-valued distance and ride duration observations; however, we found that there were slightly more of these distance errors associated with the Competitor (0.88% vs. 0.65%) and more ride duration errors associated with the Vendor (.65% vs. .05%). It is possible that some of the zero-valued tip percentage entries are residual electronic errors or tests, and that these tip errors are more concentrated in Vendor credit card machines. Despite this issue, the results are congruent with the regression discontinuity results for the zero tipping outcome.

Finally, Columns 11 and 12 present the proportion of customers who selected the 25% button, i.e. the "high" option for Competitor cabs and the "middle" option for Vendor cabs. We find that the proportion of customers who select the 25% button increases by 3.7 percentage points when it is the "middle" option, relative to when it is the "high" option. This result is suggestive of customers being influenced by a context effect similar to the "compromise effect" (Simonson 1989). A typical test of the compromise effect would compare the tendency to choose an option (e.g. 25%) between a choice set in which it is the high option out of two (e.g. 20% and 25%) against a choice set that adds a higher option (e.g. 20%, 25%, and 30%).[10] Our test differs in that the addition of a higher button (30%) comes with the removal of a low button (15%). Nonetheless, this result is suggestive of a compromise effect operating, even when the choice set of possible tip amounts if preserved and only the set of default tip suggestions is manipulated.

IV. Conclusion

Using an extensive dataset, we show that a small change in default tip suggestions has a significant effect on tipping amounts. Our data allow us to provide very clean identification in a large naturalistic field setting. We use a regression discontinuity

[10] One could also examine the removal of an "inferior" option from the choice set (e.g. the removal of the 15% button) on the propensity to choose what was previously the middle option (20%). We do not find this to be a particularly compelling test of the compromise effect because 15% is a highly relevant (frequently chosen) alternative, rather than a decoy (in contrast to the rarely chosen 30% option). Thus, in that manipulation, the default effect is the more relevant phenomenon and appears to dominate the compromise effect (Figure 7b shows that the propensity to choose 20% actually increases when it becomes the low option).

design to show that an upward shift in the set of suggestions induces higher average tip amounts, despite significantly reducing the probability that customers use one of the defaults. To analyze consumers' responses across sets of suggestions that were closely framed in terms of percentages, as well as to provide less localized treatment effects, we performed a secondary analysis of trips originating at LaGuardia airport. Exploiting these quasi-random driver-to-customer matches, we again find that higher default tip suggestions (20%/25%/30% vs. 15%/20%/25%) result in higher average tip amounts. This analysis also reveals a potential cost of setting defaults too high – customers are also more likely to leave no tip in response to the higher defaults.

The observed default effect may be attributed to three potentially complementary mechanisms. First, customers may be rationally inattentive if the cognitive effort or time costs are sufficiently high to justify the additional tip.[11] A second possible explanation is that the default serves an information transmission purpose, acting as a signal of the social norm to unfamiliar customers.[12] Finally, customers may experience disutility from deviating from the status quo, either due to social pressure or other forms of psychological resistance. Ultimately, our data do not allow us to cleanly parse the three proposed mechanisms.

As firms begin to use the insights of behavioral economics to inform their product design and promotion, our study suggests that default effects can be exploited even in habitually encountered consumer choices. However, there may be a backlash to defaults that exceed certain thresholds, and so firms and policy makers alike should be cognizant of this potential cost when optimally designing their defaults.

References

Abadie, Alberto, and Sebastien Gay. 2006. "The impact of presumed consent legislation on cadaveric organ donation: a cross-country study." *Journal of Health Economics*, 25(4): 599–620.

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

[11] The effects of these changes in suggestions on overall spending are relatively small even for regular NYC taxi users, though sizable for taxi drivers. To perform a back-of-the-envelope calculation, we use a question about the frequency of taxi use from a voluntary passenger survey administered in 2010 by the Taxi and Limousine Commission (http://www.nyc.gov/html/tlc/downloads/pdf/tot_survey_results_02_10_11.pdf). We approximate the average of these bucketed responses to ∼100 rides per year, and scale down by the proportion of rides paid by credit in 2009 (∼25%). Even with this selected sample of passengers, if we extrapolate from our RDD (∼$0.30) or Across Vendor (∼$0.20) point estimates, the change in overall spending is just $7.25/$5 for a passenger who spends over $1,000 a year on taxis. In contrast, for the median driver in the raw data with ∼5,000 rides per year (∼1,250 by credit), similar calculations produce estimates of $375/$250 increases in their annual incomes. We should stress that extrapolating from our local average treatment effects requires making several strong and unrealistic assumptions – these calculations are reported here solely to give the reader a rough sense of magnitude.

[12] A prime example of this social pressure mechanism is expressed in a New York Sun magazine article on the introduction of the credit card system: *"It forces you to tip," a Manhattan resident who recently tipped 15% on a $14 fare, Greg Mack, said. "What if you didn't enjoy the ride? It made me feel obligated."* (Source: http://www.nysun.com/business/hot-tip-for-cabbies-credit-cards-boost-tips/72783/).

Banerjee, Abhijit V, and Sendhil Mullainathan. 2008. "Limited Attention and Income Distribution." *American Economic Review*, 98(2): 489–493.

Beshears, John, James Choi, David Laibson, and Brigitte Madrian. 2010. "The Limitations of Defaults." *Prepared for the 12th Annual Joint Conference of the Retirement Research Consortium, Washington, D.C.*

Brehm, Jack W. 1966. *A Theory of Psychological Reactance.* New York:Academic Press.

Brehm, Jack W. 1989. "Psychological reactance: theory and applications." In *Advances in Consumer Research, 16.* , ed. Thomas K. Srull, 72–75. Provo, UT.

Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. 2009. "Optimal defaults and active decisions." *Quarterly Journal of Economics*, 124(4): 1639–1674.

Choi, James, David Laibson, and Brigitte Madrian. 2004. "Plan design and 401(k) savings outcomes." *National Tax Journal*, LVII(2): 275–298.

Goldin, Jacob, and Tatiana Homonoff. 2012. "Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity." *American Economic Journal: Economic Policy*, 5(1): 302–36.

Johnson, Eric, and Daniel Goldstein. 2003. "Do Defaults Save Lives?" *Science*, 4–5.

Johnson, Eric, John Hershey, Jacqueline Meszaros, and Howard Kunreuther. 1993. "Framing, Probability Distortions, and Insurance Decisions." *Journal of Risk and Uncertainty.*

Johnson, Eric, Steven Bellman, and Gerald Lohse. 2002. "Defaults, Framing and Privacy: Why Opting In-Opting Out." *Marketing Letters.*

Kahneman, Daniel. 2011. *Thinking, Fast and Slow.* Vol. 1, Farrar, Straus and Giroux.

Keller, Punam Anand, Bari Harlam, George Loewenstein, and Kevin G. Volpp. 2011. "Enhanced active choice: A new method to motivate behavior change." *Journal of Consumer Psychology*, 21(4): 376–383.

Lee, David S., and David Card. 2008. "Regression discontinuity inference with specification error." *Journal of Econometrics*, 142(2): 655–674.

Lofgren, Asa, Peter Martinsson, Magnus Hennlock, and Thomas Sterner. 2012. "Are experienced people affected by a pre-set default option – Results from a field experiment." *Journal of Environmental Economics and Management*, 63(1): 66–72.

Madrian, Brigitte, and Dennis Shea. 2001. "The power of suggestion: Inertia in 401(k) participation and savings behavior." *Quarterly Journal of Economics*, 116(4): 1149–1187.

Simonson, Itamar. 1989. "Choice Based on Reasons: The Case of Attraction and Compromise Effects." *Journal of Consumer Research*, 16(2): 158–174.

Thaler, Richard, and Cass Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness.* New Haven:Yale University Press.

Appendix

Our final dataset was constructed by first performing a number of consistency checks and then removing data that appeared to be generated by electronic tests or other types of data errors. The full sample of 170,896,479 was reduced to 13,820,735 observations by performing a number of procedures. We made the following consistency adjustments:

1) The pick-up came after the drop-off time in 0.14% (241,964) of observations. We replaced these pick-up times with their drop-off times, and vice-versa.
2) The drop-off time came after the pick-up time of the subsequent trip in 0.36% (618,570) of observations. We set the drop-off time equal to the pick-up time of the subsequent trip for all of these cases.

The full sample of 170,896,479 was reduced to 163,348,600 by dropping all observations for which:

1) There was a duplicate observation in terms of all original variables (750; 0.0004%).
2) The payment type was "No Charge" (509,194; .30%) or "Dispute" (94,784; .06%).
3) The ride duration was either equal to zero or longer than 3 hours (619,604; 0.36%).
4) The distance was either equal to zero or greater than 100 miles (929,498; 0.55%).
5) The surcharge was greater than $1 (75,295; 0.04%).
6) Corresponding to drivers that drove fewer than 100 rides in 2009 (58,495; 0.03%).
7) Multiple cars were associated with the same driver during the same shift (1,298,412; 0.77%).
8) The driver's shift was longer than 20 hours (3,872,241; 2.31%).
9) The driver's shift was shorter than 30 minutes (89,606; 0.05%).

We then dropped the 5.95% (9,718,999) of observations that were on cars equipped with the third credit card machine vendor. Next, we dropped observations for which either the pick-up location or drop-off location could not be mapped to a census tract in NY, NJ, CT, or PA (2,022,218; 0.13%). To ensure that the regression discontinuity is identified off representative rides, we dropped all rides that had toll amounts applied. This dropped the 4,882,731 (3.22%) rides which were associated with a toll amount greater than zero. We then made the largest sample reduction, removing the 108,620,194 rides paid by cash, as the data did not include tip information for these rides. From this sample of 38,104,458 rides, we further limited to those rides for which the base amount (the sum of the fare, tolls, surcharge, and tax) was equivalent to the "fare". Performing this reduction ensured that rides on either side of the discontinuity were comparable in terms of the time of day, time of year, and the fees faced by the customer. This reduction left 13,929,933 rides that occurred prior to November 1, 2009 and between the hours of 6am to 4pm on Monday through Friday or 6am to 8pm on Saturday and Sunday. Finally, we removed rides that didn't correspond to a multiple of $0.40 (the unit of fare accrual) added to $2.50 (the flat entry fee), leaving 13,820,784 observations in the final sample.