Incentivizing Commuters to Carpool: A Large Field Experiment with Waze

Maxime C. Cohen
McGill University, Montreal, Canada, maxime.cohen@mcgill.ca

Michael-David Fiszer
Google/Waze, Tel Aviv, Israel, mdfz@google.com

Avia Ratzon
Google/Waze, Tel Aviv, Israel, aviar@google.com

Roy Sasson
Google/Waze, Tel Aviv, Israel, rsasson@google.com

Traffic congestion is a serious global issue. A potential solution, which requires zero investment in infrastructure, is to convince solo car users to carpool. In this paper, we leverage the Waze Carpool service and run the largest ever digital field experiment to nudge commuters to carpool. We identify users who can save a significant commute time by carpooling through the use of an high-occupancy vehicle (HOV) lane, users who can still use an HOV lane but with a low time saving, and users who do not have access to a HOV lane on their commute. We then send in-app notifications to examine the tradeoff between mentioning the HOV lane, highlighting the time saving, emphasizing the monetary incentive, and showing a generic message. We find a strong relationship between the affinity to carpool and the potential time saving through an HOV lane. Specifically, we estimate that mentioning the HOV lane increases the click-through rate and conversion rate by 133–185% and 64–141%, respectively relative to sending a generic message.

Key words: Carpooling, Field Experiment, Saving Commute Time via HOV, Time vs. money

1. Introduction

Most governments are devoting considerable efforts to tackle traffic congestion. A study by INRIX conveys that "Americans lost an average of 97 hours a year due to congestion, costing them nearly $87 billion in 2018, an average of $1,348 per driver."[1] Thus, it is not surprising that governments allocate substantial investments to alleviate traffic congestion. Common policies (beyond expensive investments in infrastructures) include congestion pricing, restricted days, and high-occupancy vehicle (HOV) lanes, just to name a few. An alternative solution is to simply encourage solo car users to carpool.

According to data published by the U.S. Census Bureau, the vast majority of Americans go to work by driving alone in their car.[2] Over three quarters (76.3%) choose to commute this way,

[1] http://inrix.com/press-releases/scorecard-2018-us/

[2] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_1YR_B08006&prodType=table

with nearly identical numbers for men and women (this figure has been steady in 2015–2017). This translates to nearly 116 million vehicles transporting exactly one person each. Yet according to Waze data, close to two thirds of drivers have at least one other regular Waze driver with a perfectly matching commute, that is, driving from the same origin to the same destination, within less than 500 meters radium from each.

What can persuade commuters to carpool and help fight traffic congestion? In this paper, we study this question by testing several persuasion factors on a sample of 537,370 U.S. commuters. Specifically, we investigate to what extent highlighting the HOV lane and the resulting commute time reduction is an effective incentive.

1.1. Setting and Research Questions

Waze, the free community-driven GPS navigation software app owned by Google, launched a service to help users find matches and carpool.[3] Initially centered on casual carpool requests from riders to Waze drivers with similar routes, the product rapidly shifted to focus on the commute use-case, with an emphasis on planning a weekly schedule and the social aspects of sharing a ride. Users can register as drivers or riders (or both) and are matched to other users with similar commuting patterns. Subsequently, users can send offers to each other to share a ride. Drivers can thus not only accept incoming requests, but also invite other users with similar routes to join their carpool. The platform takes care of tradeoffs between rider walking and driver detour, handles payments, and proposes optimized pick-up and drop-off locations and pricing which can be modified within certain constraints.

Changing commuters' habits is not easy but if done at scale, can have a considerable impact in reducing congestion and pollution. As of October 2018, Waze has access to more than 110 million drivers globally including more than 30 million in the U.S.[4] The number of Waze drivers is abundant enough to provide the density and liquidity necessary for a significant fraction of drivers to leave their car at home—but how can one efficiently convince users to transform their daily commute? For users who regularly use their car, it seems nearly impossible to convince them to open their car to a stranger or to leave their car at home and hop on to someone else's car.

Interestingly, Waze's data shows that once a driver does share a ride for the first time, retention (i.e., carpooling again) is very high. Cracking the motivations and factors that will lead a driver to make the first step is thus critical.

The Waze Carpool platform is the ideal medium to study how different types of incentives can successfully convince commuters to carpool. For instance, internal econometric models show the

[3] https://www.waze.com/carpool

[4] https://www.timescall.com/2018/10/11/google-owned-waze-aims-to-end-traffic-with-new-carpooling-app/

extent to which drivers prefer matches that are precisely on-route, or involve higher shared mileage and thus, higher cost savings. Motivations can thus be financial, social, or related to reducing commute time, for instance by using the HOV lane.

In this paper, we study the tradeoff between two types of incentives: highlighting monetary compensation and time saving. We identify three types of users: (i) users who can save a significant commute time by carpooling through the use of an high-occupancy vehicle (HOV) lane, (ii) users who can still use the HOV lane but with a low time saving, and (iii) users who do not have access to an HOV lane on their commute. For each user, we leverage the Waze data and algorithms to carefully estimate the potential time saving had this user used the HOV lane for his/her commute. Our experimental population comprises 537,370 users across four U.S. states (our analyses are aggregated over hundreds of thousands of users and are meant to be interpreted as statistical averages). We then apply the following set of interventions:

- For commuters with a high time saving—who could save on average 6–40 minutes if they would carpool and use the HOV lane—we randomly split them into four conditions (we explain in greater detail the way we selected those users in Section 4). Each condition involves sending the user an in-app notification with an invitation to try the carpool service. Specifically, we use the following four framings: (A) mentioning the HOV lane and the potentially (high) time saving, (B) mentioning the HOV lane, (C) using a generic carpool offer, and (D) not sending anything.
- For commuters with a low time saving (i.e., users who could save on average 2–5 minutes if they would carpool and use the HOV lane), we randomly split them into the same four conditions.
- For commuters who do not have access to an HOV lane on their commute, we randomly split them into three conditions: (A) mentioning a monetary incentive (receiving $10 as a welcoming bonus to try the carpool service), (B) using a generic carpool offer, and (C) not sending anything. We further carefully select the commuters in this category to be similar to the commuters in the other two categories (see more details in Section 3).

By comparing the different framings used in our experiment, we aim to understand what are the successful triggers that can persuade commuters to carpool. It also allows us to study the tradeoff between saving commute time and earning compensation.

1.2. Summary of Results

We conduct several ANOVA tests and regression analyses to estimate the impact of our field experiment. Our findings can be summarized as follows.

Mentioning the HOV lane is the most successful framing. We consistently observe that highlighting the fact that commuters can use the HOV lane is effective. It increases the click-through rate and conversion rate by 133–185% and 64–141%, respectively relative to a generic message. This result holds both for users with high and low time savings. This effect is further amplified for users with a longer commute distance.

Mentioning the HOV lane is enough. In addition to the HOV lane, highlighting the potential time saving does not yield an additional marginal impact. One possible implication is that mentioning the HOV lane is enough, and that users can directly translate this information to saving commute time. This result holds both for users with high and low time savings.

Users with a higher time saving show a higher intent. We capture users' carpool intent using two metrics: click-through rate (CTR) and on-boarding rate (OBR). As expected, users with a higher time saving are more receptive to the carpool offer in terms of CTR and OBR. Interestingly, this occurs for both framings A and B (in which the HOV lane is highlighted) but not for framing C (generic offer).

Quantifying the economic impact of various factors: Using our data, we quantify the economic significance of several features. For example, we find that an additional 10 minutes in average time saving boosts the carpool intent (captured by OBR) by 18%. We also observe that a fixed commute schedule (captured by a low standard deviation in the leave time) translates to a higher carpool intent.

Mentioning the monetary incentive is not effective. Highlighting the $10 incentive to non-HOV users does not increase their response (CTR and OBR) to the carpool offer relative to sending a generic message. This finding suggests that offering compensation is not enough to nudge commuters to carpool. In addition, incentivizing users with a high potential benefit from carpooling is more effective than offering compensation to low-intent users, only when explicitly highlighting the benefit of carpooling (in our case, the use of the HOV lane).

Users are more reactive to saving commute time than earning compensation. Our results suggest that the users in our experiment are more receptive (i.e., higher CTR and OBR) to saving time on their commute than receiving monetary compensation. Thus, it seems better to deploy tailored incentives instead of massively subsidizing (or taxing) solo commuters.

2. Literature Review

This paper is related to several research streamlines: field experiments in online platforms, nudging, carpooling as a means of reducing congestion, and the tradeoff between saving time and financial compensation.

Field experiments in online platforms: In recent years, it has become common practice for online platforms to routinely run A/B tests to generate high-quality data and learn users' preferences. Companies such as Microsoft, Amazon, Booking.com, Facebook, and Google, each conduct more than 10,000 online controlled experiments annually, with many tests engaging millions of users (Kohavi and Thomke 2017). For more details on this topic, we refer the reader to Kohavi and Thomke (2017) and to the paper by Kohavi et al. (2013). Several researchers have recently used field experiments to compare different interventions targeted to online platforms' users. For example, in the context of ride-sharing, Cohen et al. (2018) examines the results of a field experiment that sent promotions to users who experienced a poor quality of service. Singh et al. (2017) conduct field experiments to compare charity-linked promotions to discount-based promotions in the context of an online taxi-booking platform. To our knowledge, our paper provides the first digital field experiment aiming to convince commuters to carpool.

Nudging: A traditional lever to incentivize people to do specific actions is the use of economic incentives. The field of behavioral science aims to influence people's behavior by using nudges, that is, altering the environment to favor the desired outcome (see, e.g., Thaler and Sunstein 2009, Halpern 2015). Nudges are typically easy and inexpensive to implement. Governments and public agencies have used behavioral nudging to address several policy problems such as increasing retirement savings, college enrollment, energy conservation, and adult outpatient influenza vaccinations (see Benartzi et al. 2017, and the references therein). To our knowledge, in the context of boosting the number of carpoolers (to reduce traffic congestion), no large nudging strategy was ever deployed. Waze has the unique ability to nudge commuters to alter their driving habits for the better. Waze's ability to nudge commuters does not incur a significant cost (it is nearly free) and can be carefully designed by selecting the appropriate nudging action for the right set of users. In this paper, we focus on users who can save commute time by carpooling through the use of an HOV lane, and test several nudging strategies. Our results can lead to important behavioral insights that may help reduce traffic congestion, and at the same time help commuters make better commuting choices for their own utility.

Carpooling: There is a vast literature on carpooling so that a comprehensive review is beyond the scope of this paper. One of the first empirical papers on this topic is Teal (1987). The author uses data from the 1977–1978 Nationwide Personal Transportation Survey to study the characteristics of carpoolers and to offer explanations of why commuters carpool. Ferguson (1997) uses data from the same survey between 1970 and 1990 to explain the carpooling decline in America. The decline is attributed to several factors such as increasing household vehicle availability, falling fuel costs, and higher average educational attainments among commuters. More recently, Shaheen et al. (2016) examine the motivations and characteristics of casual carpoolers in the San Francisco Bay Area by conducting interviews and surveys (with a total of 519 respondents). As expected, casual carpoolers' motivations include convenience, time savings, and financial savings. The recent work by Ostrovsky and Schwarz (2019) study the complementary nature of carpooling and self-driving cars, focusing on market equilibrium. The authors claim that congestion pricing will play an essential role in inducing the timely adoption of self-driving cars and carpooling.

Closer to our paper, there is an extensive literature on carpooling in the context of congestion pricing and HOV lanes (see, e.g., Giuliano et al. 1990, Yang and Huang 1999, Konishi and Mun 2010). Those studies investigate the extent to which policies such as HOV lanes are effective in terms of increasing carpooling. Giuliano et al. (1990) compare data from the Route 55 HOV in California to a control group of freeway commuters. The authors show that only the carpooling rate for peak period commuters has increased. The authors also conclude that travel time savings must be high to attract new carpoolers, which is one of the motivations in our paper. Yang and Huang (1999) propose a theoretical model for carpooling behavior and optimal congestion pricing in a multilane highway. They show that in the absence of HOV lanes, a uniform toll for all vehicles (independent of their number of occupants) should be charged. With HOV lanes, however, the optimal strategy requires differentiating the toll per vehicle across segregated lanes. Finally, Small et al. (2006) and Konishi and Mun (2010) study the trade-off between HOV and HOT (high occupancy vehicles and toll) lanes. Small et al. (2006) empirically analyze the behavior of motorists traveling on California State Route 91 and show the importance of considering customer heterogeneity. Konishi and Mun (2010) develop a model to examine under which conditions introducing HOV lanes is socially beneficial and whether converting HOV lanes to HOT lanes improves road efficiency.

Another line of related research is the efficacy of incentives to increase carpooling. Vanoutrive et al. (2012) analyze the popularity and determinants of carpooling in Belgium in the context of the workplace and company-induced measures. The authors observe higher levels of carpooling at less accessible locations and in sectors such as construction, manufacturing, and transport. More recently, Neoh et al. (2017) synthesize 22 existing empirical studies (with over 79,000 observations) to create a review of the carpooling literature. Their analysis identifies 24 carpooling factors including the number of employees, partner matching programs, gender, and a fixed work schedule. While there is an extensive literature on incentives and motivations for carpooling, our study is the first large scale digital field experiment (with more than half a million users). The scale of our data allows us to sharpen our current understanding on carpooling and to inform policy making.

Time vs. money: The trade-off between time and money has been extensively studied in various topics related to consumer research (see, e.g., Cross 1993, Okada and Hoch 2004, and the references therein). In the context of carpooling, for example, Beroldo (1999) and Shaheen et al. (2016) find using survey data that a large portion of respondents (both drivers and riders) rank saving time and saving money as the main reasons for carpooling. To our knowledge, there is no clear consensus to whether commuters care more about reducing their commute time or about earning compensation. Our paper aims to provide a first answer to this question. Our results suggest that the users in our dataset are more receptive to saving commute time than to saving money.

Structure of paper. In Section 3, we discuss the data and setting considered in this paper. Section 4 outlines our experimental design. Our econometric results are presented in Section 5. Finally, we report our conclusions and discuss the implications of our findings in Section 6.

3. Data and Setting

As mentioned, this paper is in the context of the Waze carpool service. Having access to more than 110 million users, Waze is seeking to persuade drivers who use its service (called Wazers) to become carpoolers. To this end, several marketing efforts were invested in advertising the new carpool service. We call the Waze drivers who have installed the Waze carpool app, *on-boarded* drivers. This means that these drivers have signed up for the carpool service by completing their home and work addresses and preferred commute times.

3.1. Pre-Experiment Analysis

Our goal is to investigate which factors can entice Wazers to become carpoolers. We first use a large historical dataset to examine several correlations between on-boarding to the carpool service and drivers' attributes. In the context of this paper, we are interested in drivers who have an HOV lane on their daily commute.

We consider a random sample of U.S. Wazers who live in an area with an HOV lane. Specifically, we focus on four U.S. states that have a high HOV occupancy: California (CA), Georgia (GA), Massachusetts (MA), and Washington (WA). The list of HOVs under consideration can be found in Appendix A. We restrict our attention to representative Waze users who are active commuters. To this end, we focus on a random sample of users who completed at least 10 navigations in the last 30 days and at least two navigations with similar origin and destination (e.g., home and work locations or the most common navigation) on weekday morning hours (from 6 AM to 12 PM). This allows us to focus on the times when the HOV lane restriction is relevant. We also restrict our sample to users who have a commute distance between five and 100 kilometers and a duration between five and 100 minutes. For each user, we use the data from all morning weekday navigations in the last 30 days to compute three key variables:

1. **Average leave time:** For each navigation, we record the time at which the user left the origin location (i.e., the time when the user started his/her commute). We then compute the average in a continuous fashion. For example, if the average leave time is 7.75, it means that the user left on average at 7:45 AM.

2. **Standard deviation of leave time:** Similarly, we compute the standard deviation of the leave time (in hours) by using all the morning weekday navigations in the last 30 days.
3. **Average time saving:** For each navigation, we leverage the Waze data and algorithms to compute the estimated travel duration both with and without the HOV lane (assuming that there is an HOV lane on the route). We then compute the difference between the duration with and without the HOV—called the time saving—and take the average over the different navigations for each user. For robustness, we repeat this process for three different days to ensure that the time saving is stable and not driven by unrepresentative events. It is worth mentioning that Waze's data and algorithms offer a unique opportunity to compute the HOV time saving metric at scale.

Depending on the value of the average time saving, we split the users into three categories:

1. **Users with a positive time saving:** These users can potentially save time by using the HOV lane. They currently cannot use it given that they are solo commuters, but if they would carpool by riding with a passenger, they may reduce their commute time.
2. **Users with a negative time saving:** These users cannot save time by using the HOV lane. For example, the fastest route from home to work does not include an HOV lane, so that taking the itinerary with the HOV lane will take a longer time.
3. **Users who do not have access to an HOV lane in their commute:** These users live in a market where there is at least one HOV lane in their neighborhood, but the HOV lane is not located on their daily commute.

For example, in Figure 1, we query the Waze application for a navigation from San Jose, CA to Mountain View, CA at 10 AM.[5] The system suggests three possible routes (see top left of Figure 1). The optimal route has an estimated time of arrival (ETA) of 30 minutes. The third route suggests using the HOV lane (on US-101 N) for an ETA of 21 minutes. Thus, the estimated time saving in this example amounts to nine minutes.

We eliminate users with a time saving lower than -40 and higher than 40 minutes to avoid outliers (these users represent a negligible fraction of our sample). Ultimately, we use a random sample with 806,790 users. The average time saving distribution for these users is shown in Figure 2. We also report the time saving distribution for each market (i.e., state) and each one-hour window leave time (i.e., the average time at which users start their commute) in Figures 3 and 4, respectively. As we can see, the time saving distribution is centered around zero so that users are split into the ones who can save commute time by using the HOV lane and the ones that cannot. Figure 4 shows that as the morning progresses, the variance in time saving decreases.

[5] https://www.waze.com/livemap

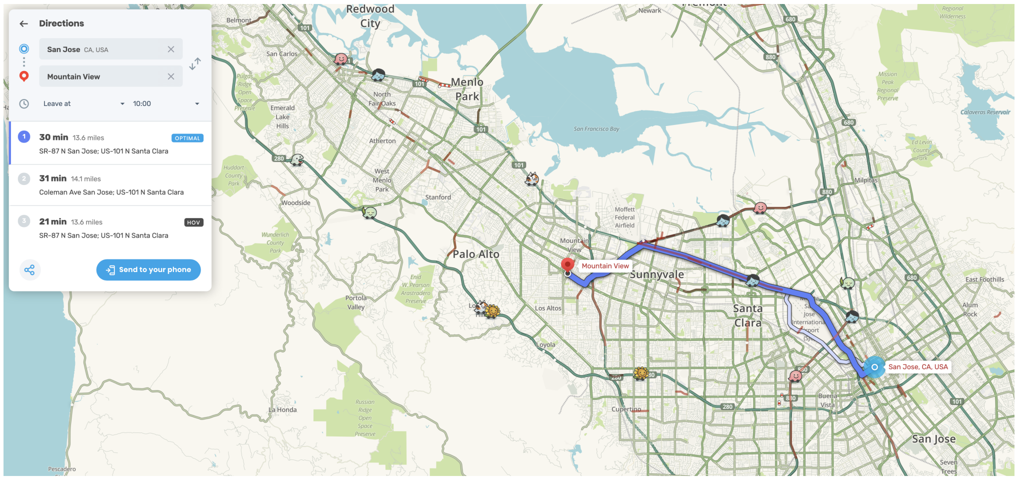


Figure 1 **Example of time saving by using the HOV lane (date accessed: July 23, 2019)**

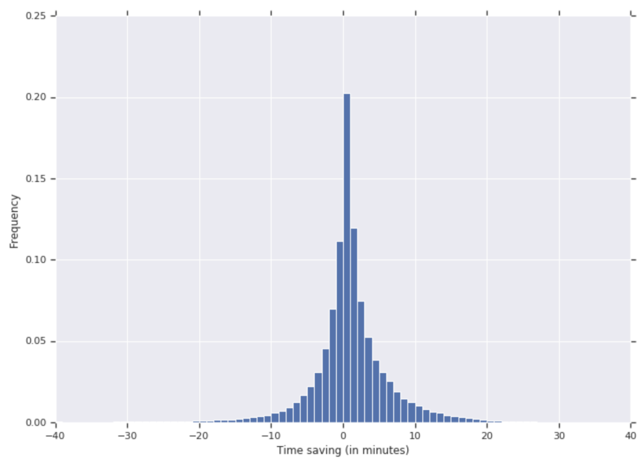


Figure 2 **Time saving distribution**

To further visualize the variation in time saving as a function of the leave time, we plot for each one-hour interval the proportion of users who save on average more than two minutes, between two minutes and minus two minutes, and less than minus two minutes (see Figure 5). We see that during the early morning commute hours (peak times), more than 35% of users with an HOV lane on their route have a meaningful time saving by taking the HOV lane. This percentage then linearly declines with time until 11 AM. We should note that the percentage of users with a positive time saving is underestimated due to the fact that users are likely to optimize their commute based on traffic conditions (e.g., they commute later or earlier than they would have liked, to avoid traffic).

We next investigate to what extent the likelihood of on-boarding to the carpool service correlates with the time saving. We start by plotting the on-boarding rate (OBR), that is, the ratio between on-boarded drivers and the total number of drivers as a function of the average time saving (see

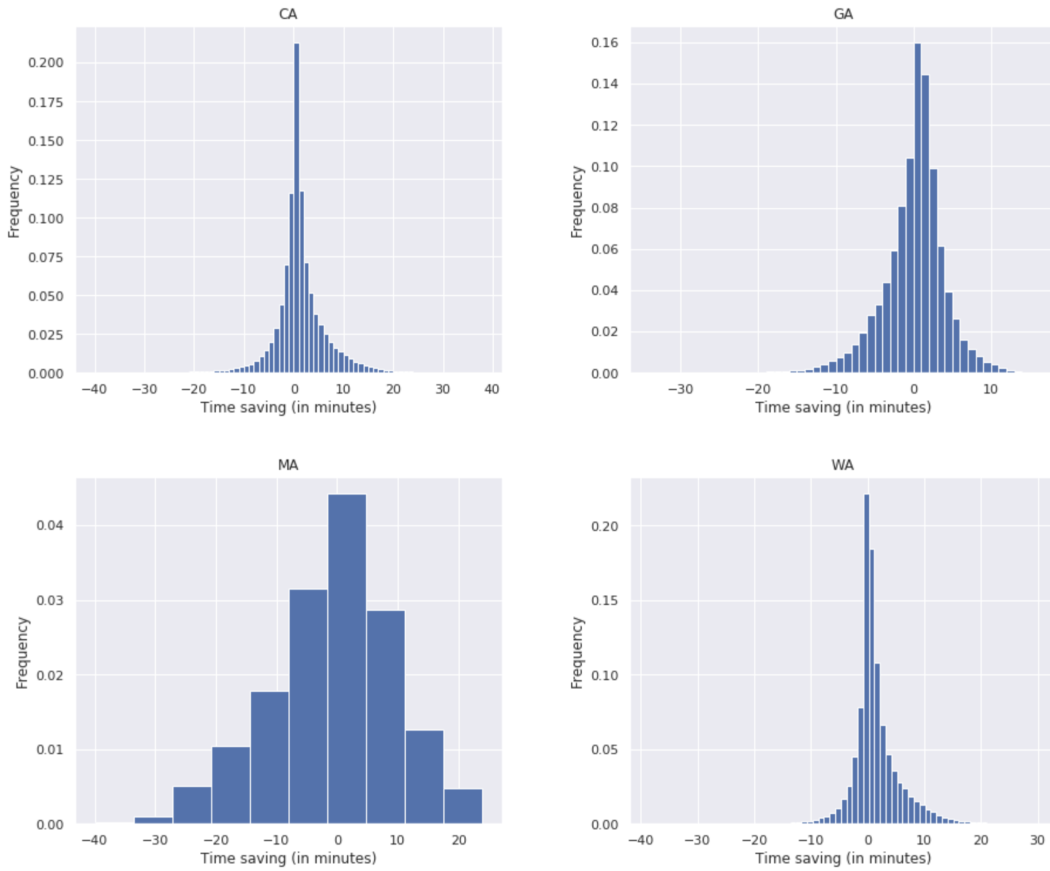


Figure 3 **Time saving distribution for each market (CA, GA, MA, and WA)**

Figure 6).[6] The bucket sizes in Figure 6 was chosen to ensure that we have enough observations in each bucket (the same qualitative pattern holds when using other bucket sizes). We find that the OBR increases with the average time saving. This suggests that users with a higher time saving are more likely to be interested in the carpool service. More specifically, users with a time saving of at least 10 minutes are twice as likely to have on-boarded relative to users with a negative (or close to zero) time saving. We then compare the average time saving for on-boarded and non-on-boarded drivers: t-statistic=-18.8 with a p-value (much) less than 1%.

Finally, to account for additional covariates which might affect the individual decision maker to carpool, we estimate six regression specifications. We consider two types of models:

1. A linear probability ordinary least square model:

$$Y_i = \alpha X_i + \epsilon_i, \tag{1}$$

[6] We normalize all our figures related to on-boarding rates. A conversion factor has been applied so that the highest number is assigned the value of 1 and all other values are adjusted by the same normalizing factor—maintaining the same relationship between different points.

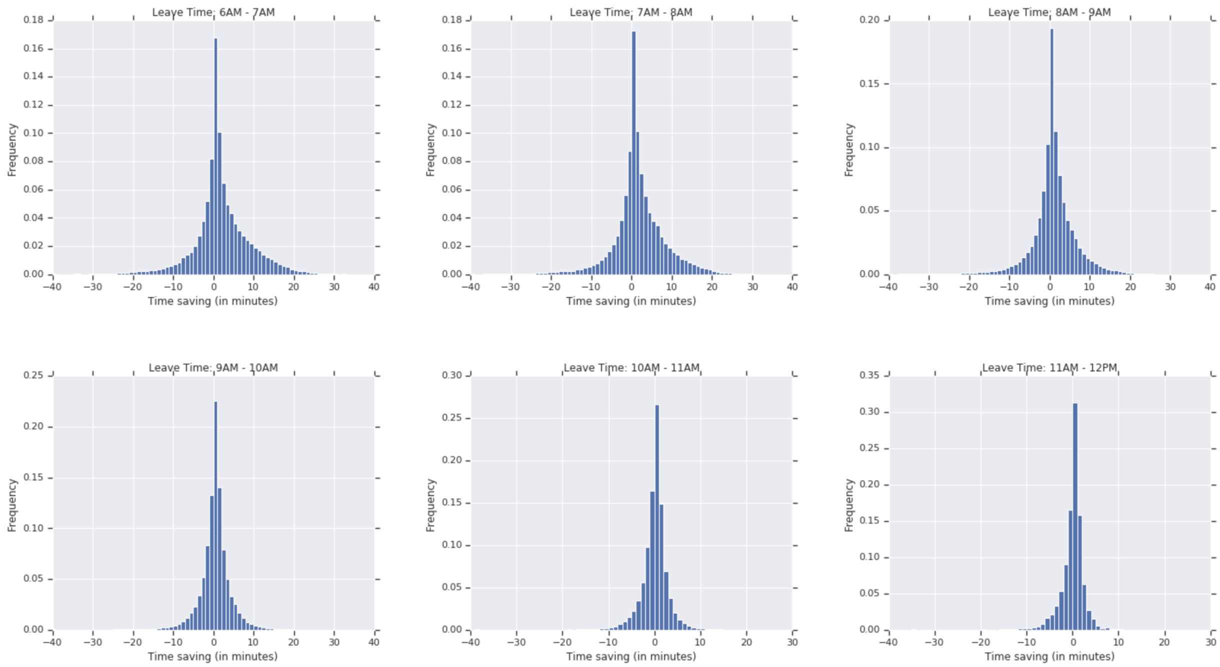


Figure 4 Time saving distribution for different leave times

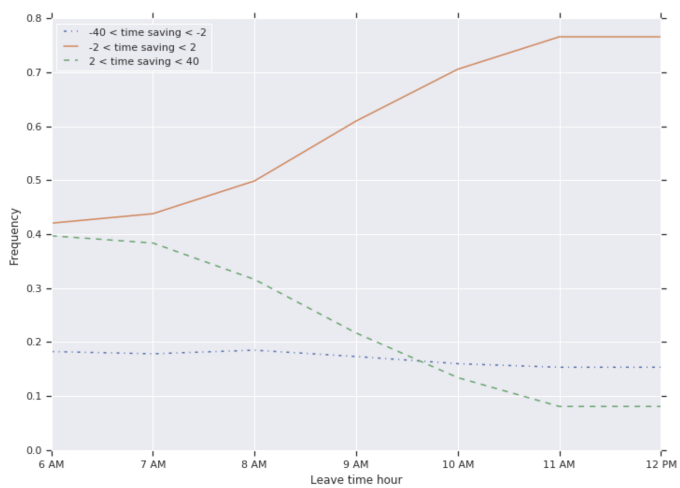


Figure 5 Time saving evolution with time

where i corresponds to the user index and Y_i is an indicator variable that captures whether user i is an on-boarded driver. The independent variables X_i include the device type (Android versus iOS), the market (CA, GA, MA, and WA), the average distance in the last 30 days, the average leave time, the number of days since joined, and the number of sessions in the last 30 days. Finally, ϵ_i^k is a stochastic i.i.d Gaussian term and α is the estimated vector of coefficients associated with the independent variables.

2. A logit probability model:

$$\text{logit}(Y_i) \sim f(X_i; \alpha), \tag{2}$$

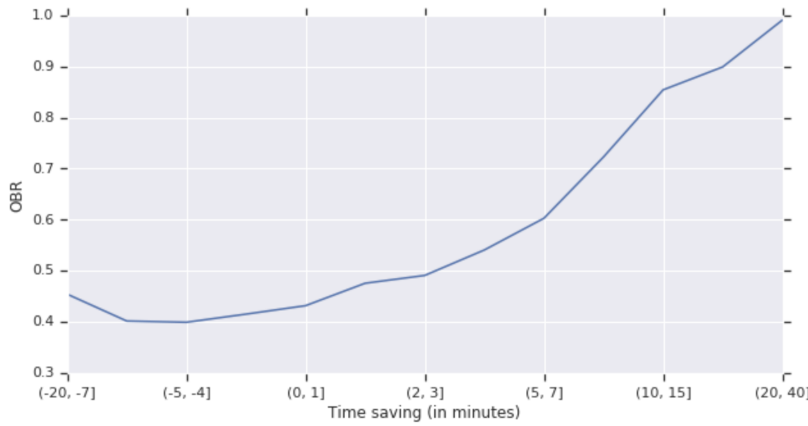


Figure 6 **On-boarding rate (normalized) as a function of time saving**

where the link function $f(\cdot)$ is logit.

The estimated coefficients of Equations (1) and (2) are reported in Table 1.[7] Note that we only consider users with a positive or negative time saving and remove the ones who do not have access to an HOV lane in their commute—so that we remain with 574,559 observations. In the first two columns, we control for the independent variables mentioned above. In the third and fourth columns, we also include the standard deviation of the leave time. In the last two columns, we use the range (i.e., maximum minus minimum) of the leave time instead of the standard deviation (users with a single observation are assigned a standard deviation and a range equal to zero).

The results of Table 1 suggest the following:

- The average distance and number of sessions are positively correlated with OBR. This validates the intuition that commuters who have a longer commute distance (or who are more active - either with Waze's services, or generally as commuters) are more likely to be interested in the carpool service.
- Commuters from CA and WA have a significantly higher OBR relative to users from GA and MA. Using the estimates from the first column of Table 1 together with the average values of the control variables, we find that commuters from CA and WA have an OBR that is 30% higher relative to GA users. This supports the fact that carpooling is a more widespread practice in the west coast where several HOV lanes can be found and commuters typically have a longer commute which cannot easily be completed via public transportation.
- More importantly, the regression estimates confirm that the average time saving is positively correlated with OBR, hence validating our intuition that users who can save more time by

[7] In the regression tables for the on-boarding rate, we have scaled all the estimated parameters (along with their standard errors) by a positive constant to avoid revealing business sensitive information. We use a different constant for OLS and logit models.

Table 1 Regression estimates for pre-experiment analysis (dependent variable: normalized OBR)

	OLS (1)	Logistic (2)	OLS (3)	Logistic (4)	OLS (5)	Logistic (6)
log(avg_distance)	0.008***	0.27***	0.008***	0.288***	0.008***	0.286***
	(0.0008)	(0.034)	(0.0008)	(0.034)	(0.0008)	(0.034)
log(avg_leave_time)	-0.074***	-2.77***	-0.068***	-2.494***	-0.068***	-2.508***
	(0.004)	(0.132)	(0.004)	(0.138)	(0.004)	(0.138)
log(days_since_joined)	0.012***	0.504***	0.012***	0.496***	0.012***	0.498***
	(0.0004)	(0.02)	(0.0004)	(0.02)	(0.0004)	(0.02)
log(sessions_30d)	0.022***	0.874***	0.022***	0.86***	0.022***	0.862***
	(0.0008)	(0.028)	(0.0008)	(0.028)	(0.0008)	(0.028)
avg_time_saving	0.002***	0.038***	0.002***	0.036***	0.002***	0.036***
	(0.00008)	(0.02)	(0.00008)	(0.02)	(0.00008)	(0.02)
log(std_leave_time + 1)			-0.006***	-0.202**		
			(0.002)	(0.084)		
is(std_leave_time_NA)			-0.008***	-0.282***		
			(0.002)	(0.046)		
log(range_leave_time + 1)					-0.004**	-0.11**
					(0.002)	(0.05)
is(range_leave_time_NA)					-0.008***	-0.286***
					(0.002)	(0.05)
GA	-0.03***	-1.328***	-0.03***	-1.33***	-0.03***	-1.33***
	(0.002)	(0.056)	(0.002)	(0.056)	(0.002)	(0.056)
MA	-0.034***	-1.498***	-0.034***	-1.472***	-0.034***	-1.474***
	(0.002)	(0.098)	(0.002)	(0.098)	(0.002)	(0.098)
WA	0.002	0.036	0.002	0.034	0.002	0.034
	(0.002)	(0.056)	(0.002)	(0.056)	(0.002)	(0.056)
Constant	0.034***	-8.448***	0.024***	-8.808***	0.026***	-8.774***
	(0.01)	(0.354)	(0.01)	(0.36)	(0.01)	(0.358)
Observations	574,559	574,559	574,559	574,559	574,559	574,559
R^2	0.007		0.007		0.007	
Log Likelihood		-70,961.090		-70,940.280		-70,940.800
Residual Std. Error	0.164 (df = 574549)		0.164 (df = 574547)		0.164 (df = 574547)	
F Statistic	445.013*** (df = 9;574549)		367.856*** (df = 11;574547)		367.655*** (df = 11;574547)	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

using the HOV lane may be more interested in the carpool service. More precisely, our regression estimates suggest that an additional 10 minutes in average time saving boosts the OBR by 18% (again, using the estimates from the first column of Table 1 together with the average values of the control variables).

- Finally, we find that the standard deviation of the leave time has a negative effect on OBR. This is an interesting finding as one could have posited two opposite hypotheses: (1) A high standard deviation in leave time means more variation in the starting work hour (e.g., unpredictable meeting times) so that such users are less likely to carpool. In the same vein, a higher standard deviation in leave time may also indicate that the user has more flexibility in his/her commute time (so that s/he is not constrained to a specific time and can potentially beat traffic conditions). (2) A high standard deviation in leave time can translate to more flexibility in work hours so that such users are more likely to carpool (e.g., by having the flexibility to pick up a passenger). The results in Table 1 are in favor of the first hypothesis.

All our statistical tests strongly support the fact that a high time saving is positively correlated with carpool intent (captured by the OBR metric). This observation motivates us to further study the relationship between carpool intent and time saving and to examine whether it is causal. The

Waze carpool platform offers a great opportunity to run a randomized controlled experiment while controlling for several important factors, as we discuss next.

3.2. Experiment Data

Following the process outlined in Section 3.1, we consider users across four U.S. states: CA, GA, MA, and WA. We further focus on users who either have a positive time saving or no access to an HOV lane (i.e., we remove the users with a negative time saving). We also remove all users who already on-boarded to the Waze carpool service. Finally, we carefully ensure that the users in our field experiment are not part of another experiment at the same time as our field experiment. We remain with a sample of 537,370 users. We call this set of users our experimental population.

4. Experimental Design

As discussed, our goal is to encourage Waze users via an app notification to try the carpool service, that is, to take a passenger for a future ride. Our experiment allows us to investigate how different users respond to various types of offers or incentives (i.e., different framings).

4.1. Experimental Population

Each user in the experimental population belongs to one of three categories: HOV users with high time saving (called H users), HOV users with low time saving (called L users), and non-HOV users (called N users). We consider the random sample of users mentioned in Section 3.2. N users are simply the ones who do not have access to an HOV lane on their commute. Thus, the remaining users have an HOV lane on their commute and a positive average time saving. We split the remaining users depending on the value of their average time saving: H users (resp. L users) correspond to the top 27 percentile (resp. bottom 73%).[8] The numbers of each user type are reported in Table 2 (the number of impressions will be discussed in Section 4.3).

Table 2 Number of users in our field experiment

User type	Number of users	Number of impressions
H users	84,798	49,504
L users	224,518	131,778
N users	228,054	120,021
Total	537,370	301,303

We highlight that all the interventions used in this field experiment were beneficial to the users since they receive an offer to try the carpool service while highlighting different benefits of carpooling. In addition, our analyses are aggregated over hundreds of thousands of users and are meant to be interpreted as statistical averages.

[8] We have used the cutoffs 27 and 73 as those numbers lead to a clear separation of the average time saving value (the exact time saving threshold is not revealed due to confidentiality but does not affect any of our analyses).

4.2. Balancing Groups

To ensure that our experiment is properly randomized (i.e., no selection bias between the groups for each user type), we carefully split the users from each category to several conditions. Specifically, we balance our sample with respect to the variables in Table 3. We visualize the balancing on these dimensions in Figures 7 and 8. A session is defined as any interaction between the user and the app (Waze). It includes: ETA check, navigation, and driving when the app is open. A navigation occurs when the user completes a drive to a destination searched via the app. Finally, a navigation with an HOV lane captures a navigation in which at least one of the proposed itineraries includes an HOV lane.

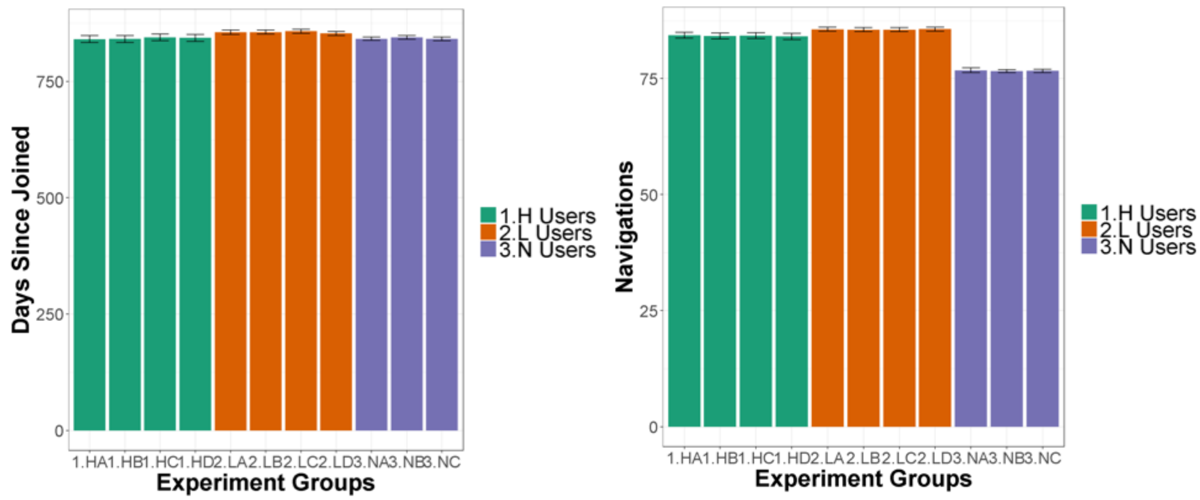
Table 3 Balancing variables in our field experiment

Variable
Number of sessions in the last 30 days
Number of navigations with an HOV lane in the last 30 days
Days since joined Waze
Driven distance in the last 30 days (in km)
Total navigation time in the last 30 days (in minutes)

We also trust the randomization to balance the different conditions for each type of users across other relevant variables. As a sanity check, we examine the balancedness with respect to the variables in Table 3 as well as to the following attributes: Market (i.e., state), average leave time, device (Android vs. iOS), and percentage of users who saw an impression. In Figures 7 and 8, we report the average number of days since joining date, number of navigations in the last 30 days, distance driven in the last 30 days, and (estimated) time saving. The plots for the other variables can be found in Appendix B.

As we can see from all the figures, for each user category, all the relevant variables are well balanced across conditions (there are no statistically significant differences between conditions, for each user type). Note that the different conditions are perfectly balanced for each user type but not necessarily balanced across different user types. This is to be expected given that different categories of users admit inherent differences (e.g., HOV users may systematically have a longer commute time relative to non-HOV users, as confirmed by Figure 8a). Our regression analysis in later sections will account for variations of these variables, beyond the random assignments to treatment groups within each category of users.

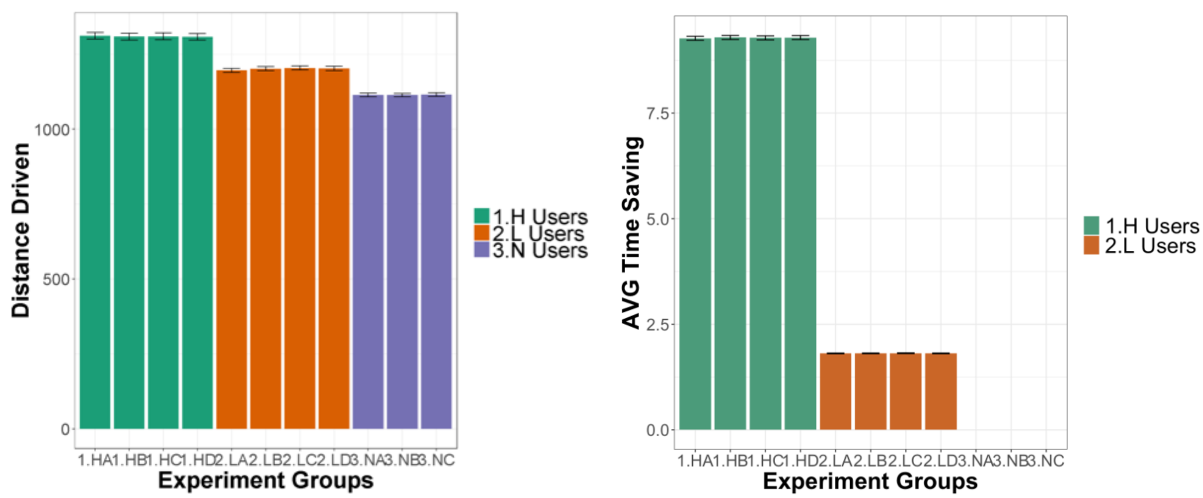
Recall that N users come from similar geographical locations as H and L users. As explained in Section 3, we focus on users who live in an area with an HOV lane across four U.S. states. However, since N users are fundamentally different than H and L users, it is natural to observe some differences in several variables.



(a) Number of days since joining date

(b) Number of navigations in the last 30 days

Figure 7 **Average number of days since joining date and number of navigations in the last 30 days**



(a) Distance driven in the last 30 days (in km)

(b) Time saving by using an HOV lane

Figure 8 **Average distance driven in the last 30 days (in km) and (estimated) time saving by using an HOV lane**

Finally, the average time saving for H users is significantly higher than for L users (see Figure 8b), as expected. The averages are 9.6 minutes and 2.1 minutes for H and L users, respectively (the range of values are $[6, 40]$ and $[2, 5]$, respectively).

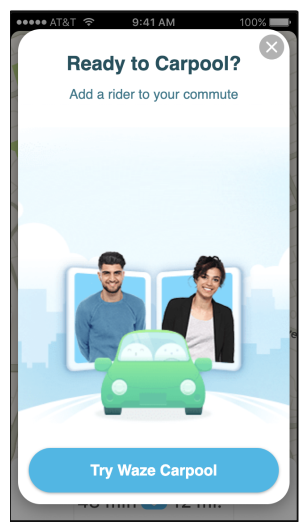
4.3. Treatments

Our field experiment involves several treatments or interventions. Each treatment entails sending an app notification—called an encouragement—to the user to advertise the carpool service. The encouragement is "sent" programmatically to all users at the same time, in our case on June 10, 2019. Users will see the encouragement displayed on the homepage of the Waze app the first two times they open the app (before that the live map appears) during the period in which our experiment is running (June 10 to July 3). As a result, it is possible that users will see the encouragement at different times (we will later explicitly control for such time effects). We split the users into several conditions as follows:

- **Commuters with a high time saving (H users)** are split into four conditions. Specifically, we use the following four framings for the encouragement: (A) mentioning the HOV lane and the potentially (high) time saving, (B) mentioning the HOV lane, (C) using a generic carpool invitation, and (D) not sending anything. The screenshots of the encouragements along with the text used in the three treated conditions are shown in Figure 9.
- **Commuters with a low time saving (L users)** are split into the same four conditions (see Figure 19 in Appendix C for the exact messages).
- **Commuters without access to an HOV lane (N users)** are split into three conditions: (A) mentioning the monetary incentive (receiving $10 as a welcoming bonus to try the carpool app), (B) using a generic carpool invitation, and (C) not sending anything (the screenshots of the encouragements are shown in Figure 20 in Appendix C).[9]

(a) Condition A

(b) Condition B



(c) Condition C

Figure 9 Copy of the message used in our field experiment for H users

Our experiment was live between June 10 and July 3, 2019, that is, a total of 24 days. During this period, when any of the 537,370 users opens the Waze app for the first two times, s/he will see the encouragement inviting the user to try the carpool service. If a user was shown the encouragement, we call it an *impression*. Then, the user can either click on the "Try Waze Carpool" button or on the exit button (see Figure 9). The user can also ignore the encouragement. Several

[9] Note that all users (regardless of our experiment) receive the same $10 monetary incentive as a welcoming bonus. The only manipulation we use is to highlight the incentive in the encouragement versus not mentioning it.

reasons may lead to a driver not seeing the encouragement, including cases when the driver did not open the app during the experiment period, the driver reinstalled (or updated) the app, and the driver changed smartphone. Note that by definition, users in the control condition are not sent an encouragement and thus will not see an impression. Overall, around 78% of our users saw the impression. Fortunately, this number is constant across all conditions and user types, as seen in Figure 10.

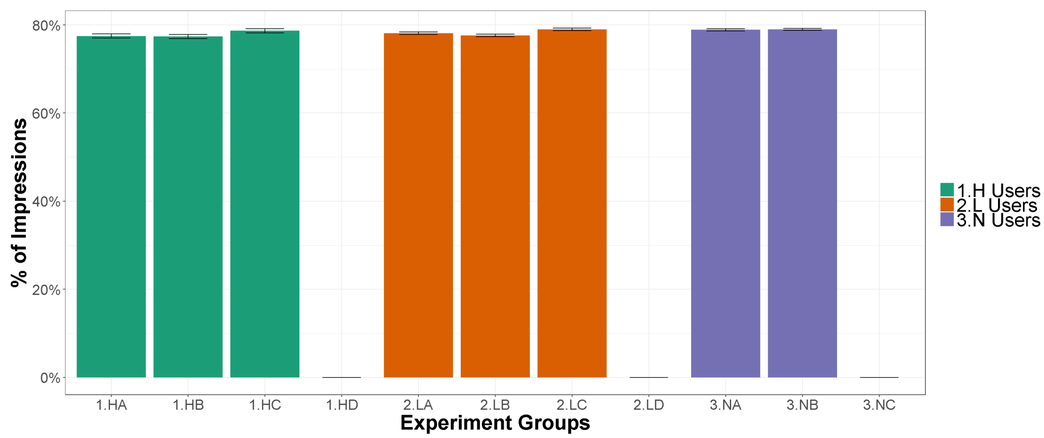


Figure 10 **Average number of impressions across the different groups**

4.4. Performance Metrics

To measure the treatment effect, we consider two performance metrics: the click-through rate (CTR) and the on-boarding rate (OBR). The CTR is defined as the number of users who clicked on the "Try Waze Carpool" button from the encouragement divided by the number of users who were shown an impression. This is a common metric often used in the context of online advertising (see, e.g., Richardson et al. 2007). A higher CTR typically indicates a higher intent/interest for the service. The OBR is the number of users who signed up (or on-boarded) to the Waze carpool service (by creating an account and completing their home and work addresses), and hence captures the conversion rate. We consider three versions of the OBR, depending on the normalization: (1) OBR1 is defined as the number of on-boarded users divided by the number of users who clicked, (2) OBR2 is defined as the number of on-boarded users divided by the number of users who were shown an impression, and (1) Absolute OBR is defined as the absolute number of on-boarded users. The latter metric allows us to include users from the control condition. Indeed, users in the control condition did not receive an encouragement by design, and hence the concepts of impression and click are not relevant for such users. Our four dependent variables are summarized in Table 4.

Table 4 Dependent variables used in this paper

Dependent variable	Description
CTR	Users who clicked divided by users who were shown an impression
OBR1	Users who on-boarded divided by users who were shown an impression
OBR2	Users who on-boarded divided by users who clicked
Absolute OBR	Users who on-boarded

5. Results

In this section, we report the results of our field experiment for each condition, user type, and the four performance metrics described in Table 4. We first show the results of one-way ANOVA tests (see, e.g., Maxwell and Delaney 2004). We then estimate several regression specifications to showcase the robustness of our results when controlling for various factors. Finally, we refine our findings by investigating potential heterogenous treatment effects.

5.1. Basic Results

We start by presenting one-way ANOVA tests on each performance metric, by pooling the observations across all three user types. The results for CTR are reported in Figure 11 ($F(7,301,300) = 200.3$, $p < 0.01$). For each figure, we include the 95% confidence interval corresponding to each average value. We uncover the following findings:

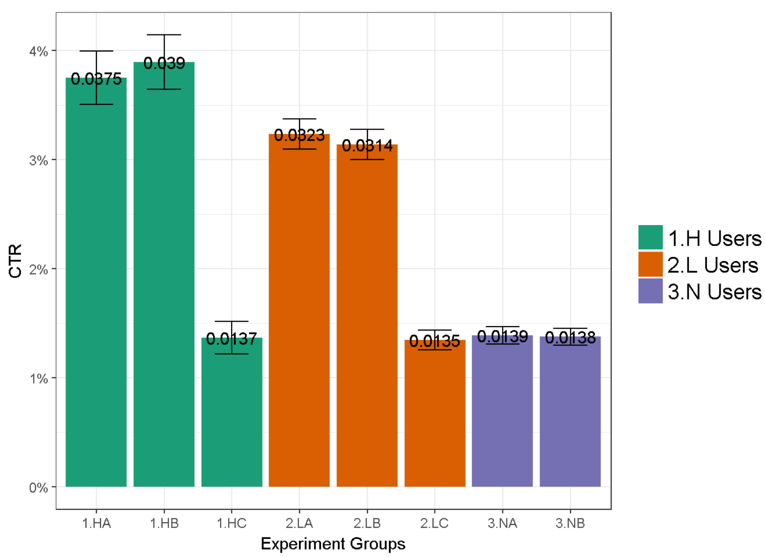


Figure 11 ANOVA test for CTR

- Highlighting the fact that commuters can use the HOV lane significantly boosts the CTR relative to a generic carpool offer. Specifically, conditions A and B for H users increase the CTR by 174% and 185%, respectively relative to a generic message. For L users, these numbers become 139% and 133%. These effects remain statistically significant when performing a separate ANOVA test for each category of users (H and L).

- Mentioning the HOV lane is enough in the sense that highlighting the potential time saving does not yield an additional marginal impact. Specifically, conditions A and B are not statistically different from each other. One possible implication is that mentioning the HOV lane is enough and that users can directly translate this information to saving commute time (note that it is still possible that the framing of the message used in condition B was not well executed). This result holds both for H and L users (using either the pooled sample or separate samples).
- Users with a high time saving (i.e., H users) are more receptive to the carpool offer, and hence have a higher CTR than L users. Interestingly, this result holds for framings A and B (in which the HOV lane is highlighted) but not for framing C (generic message).
- Highlighting the $10 incentive to N users (condition NB) does not increase the CTR relative to sending a generic message. This finding suggests that offering compensation is not enough to nudge commuters to carpool in our experiment. In addition, incentivizing users with a high potential benefit from carpooling is more effective than offering compensation to low-intent users, only when explicitly highlighting the benefit of carpooling (in our case, the use of the HOV lane). Interestingly, conditions HC and LC (i.e., generic offer for HOV users) are not statistically different from NA and NB (i.e., non-HOV users). This reinforces the importance of explicitly highlighting the benefit of carpooling.
- Our results suggest that users are more receptive to saving time on their daily commute than to receiving compensation. Interestingly, highlighting compensation for N users does not impact the CTR (i.e., N users in A and B conditions are not statistically different in terms of CTR). In addition, users who can save time but the HOV was not explicitly mentioned (i.e., HC and LC users) are also not different from N users. However, when the HOV is explicitly mentioned in the encouragement, the CTR increases significantly. As a result, both the type of users and the framing of the intervention play an important role in converting commuters to carpoolers.

We next consider the OBR metric. As discussed in Section 4.4, we consider three variants of the OBR, depending on the normalization. In Figure 12, we use OBR1, that is, the on-boarded rate normalized by the number of impressions. One can see that the results for OBR1 follow the same pattern as the results for CTR, and are statistically significant ($F(7,301,300) = 15.08$, $p < 0.01$). In Figure 13, we consider OBR2 where the on-boarded rate is normalized by the number of clicks. In this case, while all the results follow the same qualitative behavior as the two other metrics, few pairwise comparisons are losing their statistical significance (the pooled sample F-statistics is still significant at the 1% level: $F(7,6,519) = 4.228$, $p < 0.01$). The fact that our results also hold for OBR2 suggest that even after controlling for users who have clicked, the carpool intent of users in conditions A and B remains stronger. This allows us to mitigate the potential concern that users may click on the banner by a lack of understanding or to make it disappear.

Finally, in Figure 14, we consider the Absolute OBR without any normalization. This allows us to include the control condition in which no encouragement was sent (i.e., condition D for H and L users and condition C for N users). The same qualitative results hold and are statistically significant ($F(10,537,365) = 15.42$, $p < 0.01$). Note that some users from condition HD (users who did not receive any treatment) have on-boarded to the carpool service. This follows from the fact that users can onboard organically through the Waze application.

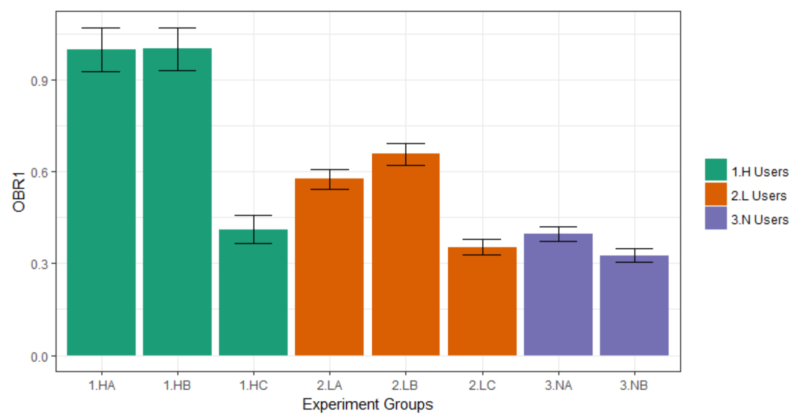


Figure 12 **ANOVA test for OBR1 (normalized)**

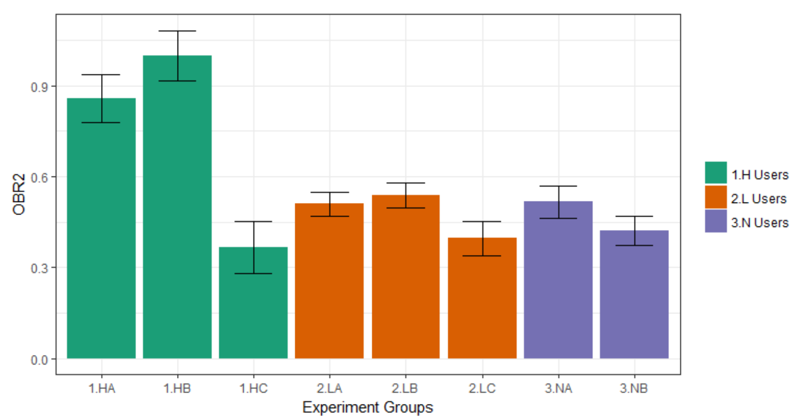


Figure 13 **ANOVA test for OBR2 (normalized)**

The results for the three OBR metrics confirm the insights we found for CTR. Specifically, highlighting the fact that commuters can use the HOV lane significantly boosts the OBR relative to a generic carpool offer. For example, conditions A and B for H users increase OBR1 by 141% (in both cases) relative to a generic message. For L users, these numbers become 64% and 93% for conditions A and B, respectively.

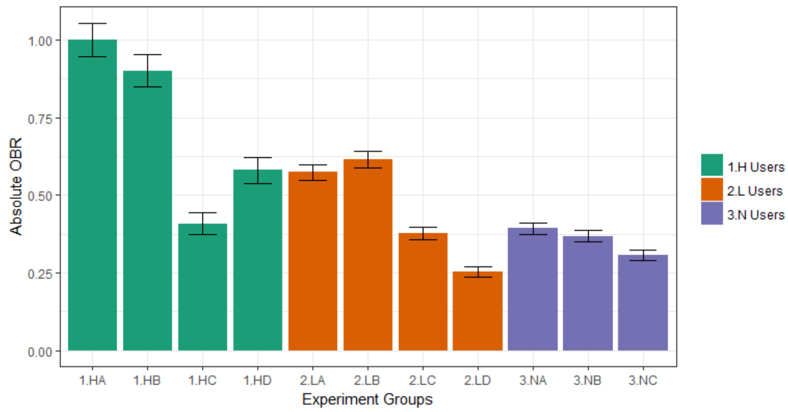


Figure 14 ANOVA test for Absolute OBR (normalized)

We next perform a difference in differences analysis between groups, to account for any potential inherent differences between users who can benefit a lot from the HOV lane (H users) and others (L and N users). We consider six models to compare the following differences: HA - HC vs. LA - LC, HB - HC vs. LB - LC, HA - HC vs. NB - NA, HB - HC vs. NB - NA, LA - LC vs. NB - NA, and LB - LC vs. NB - NA. This analysis retrieves the same numbers as in the ANOVA tests while providing a few refinements. The results for CTR are reported in Table 5 (we obtain similar qualitative results for OBR1 and Absolute OBR, whereas for OBR2, we lose statistical significance for some of the coefficients). Comparing the differences between groups allow us to draw the following insights:

- The H and L coefficients are not statistically significant. This means that the boost in CTR cannot be attributed to the fact that users have a high (or low) time saving from using the HOV lane. It thus conveys that there are no differences in terms of CTR between H and L users.
- The A and B coefficients are positive and statistically significant. This confirms that the treatments used in A and B conditions are more effective than using a generic carpool offer. For example, users in condition A (see first column of Table 5) have a $[1 + (0.019 - 0.013)/0.013]$ $= 146\%$ higher CTR than users in condition C.
- The interaction terms H $\times$ A and H $\times$ B (in the first two columns) are positive and statistically significant. This confirms that the boost in CTR is driven by the intervention (i.e., receiving an encouragement that highlights the HOV lane), and not by the difference in time saving between H and L users. Specifically, H users in condition A have an additional lift in CTR of $(0.005/0.019){=}26\%$ relative to L users in condition A.
- All interactions terms in the last four columns are positive and statistically significant with an economic impact that ranges between 128% and 178%. This suggests that highlighting

the HOV lane (relative to a generic offer) is more successful than explicitly highlighting the monetary incentive (also relative to a generic offer).

- The results for OBR1 are in the same spirit (see Table 11 in Appendix D).

Table 5 Regression estimates for CTR when comparing the differences

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
H	0.0002 (0.001)	0.0002 (0.001)	−0.0002 (0.001)	−0.0002 (0.001)		
L					−0.0004 (0.001)	−0.0004 (0.001)
A	0.019^{***} (0.001)		−0.0001 (0.001)		−0.0001 (0.001)	
H × A	0.005^{**} (0.002)		0.024^{***} (0.002)			
B		0.018^{***} (0.001)		−0.0001 (0.001)		−0.0001 (0.001)
H × B		0.007^{***} (0.002)		0.025^{***} (0.002)		
L × A					0.019^{***} (0.001)	
L × B						0.018^{***} (0.001)
Constant	0.013^{***} (0.001)	0.013^{***} (0.001)	0.014^{***} (0.001)	0.014^{***} (0.001)	0.014^{***} (0.001)	0.014^{***} (0.001)
Observations	121,297	121,011	153,122	153,103	208,217	207,950
R^2	0.005	0.005	0.003	0.004	0.003	0.003
Residual Std. Error	0.151 (df = 121293)	0.151 (df = 121007)	0.127 (df = 153118)	0.127 (df = 153099)	0.131 (df = 208213)	0.131 (df = 207946)
F Statistic	185.191^{***} (df = 3; 121293)	186.392^{***} (df = 3; 121007)	171.222^{***} (df = 3; 153118)	190.873^{***} (df = 3; 153099)	231.254^{***} (df = 3; 208213)	210.239^{***} (df = 3; 207946)

$^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

5.2. Regression Analysis

We next test the robustness of the results from the ANOVA tests. We estimate several regression specifications, allowing us to control for various factors. As in Section 3.1, we estimate two model specifications for each of the four dependent variables:

1. A linear probability ordinary least square model:

$$Y_i^k = \beta^k TC_i^k + \gamma^k X_i^k + \epsilon_i^k, \tag{3}$$

where k corresponds to the group or user type (H, L, or N), i represents the user index, and Y_i^k is one of the four dependent variables for user i from group k. The independent variables are divided in two categories: (1) TC_i^k stands for treatment conditions and includes a categorical variable for each condition (A, B, C, and D) and (2) X_i^k represents the control variables. Specifically, we control for the device type (Android versus iOS), the market (CA, GA, MA, and WA), the average distance in the last 30 days, the average leave time, the number of days since joined, and the total number of sessions in the last 30 days. Finally, ϵ_i^k is a stochastic i.i.d Gaussian term, β^k is the estimated vector of treatment effects for group k, and γ^k is the estimated vector of coefficients associated with the control variables.

2. A logit probability model:

$$\text{logit}(Y_i^k) \sim f(TC_i^k, X_i^k; \beta^k, \gamma^k), \tag{4}$$

where the link function $f(\cdot)$ is logit.

To showcase the robustness of our results, we estimate the specifications in Equations (3) and (4) under various configurations. We first estimate a separate regression for each user type (H, L, and N). We then consider the pooled sample while controlling for the user type (we also estimate the model by pooling H and L users). We estimate each model specification without controls, with partial controls, and with all the controls. Finally, we also estimate a model with impression-time fixed effects (i.e., the day when the user saw the impression) to somewhat account for the context in which the encouragement is observed by the user, such as weekend versus weekday.

Due to space limitations, we only report the regression results for CTR and OBR1 using the pooled sample (see Tables 6 and 7). We estimated several additional regression results (for CTR using H users only, L users only, and both H and L users) and found consistent results. Each regression table includes four columns where we report the estimates for both OLS and logistic models with and without controls.

In Tables 6 and 7, the baseline is set to NB users (i.e., non-HOV users who received an encouragement that highlights the monetary compensation). Recall that control users (conditions HD, LD, and NC) cannot be included as such users do not receive an encouragement. We also use the users from California as our baseline. Below are our main findings:

- Highlighting the monetary compensation to N users does not have a statistically significant impact relative to using a generic carpool message. This suggests that compensation is not a critical driver to nudge commuters to carpool (at least in the way it was framed for the users in our experiment). This result holds across all four performance metrics.
- As before, mentioning the HOV lane in the encouragement has a significant effect in boosting CTR and OBR1 (it also consistently holds for OBR2 and Absolute OBR). This time the impact is measured relative to NB users. For H and L users, the CTR increases by 61% and 46%, respectively (holding the other variables constant). It then suggests that the combination of identifying users who can benefit from the HOV lane with explicitly mentioning the benefit (in our case, the option of using the HOV lane) works better than offering a monetary incentive. On the other hand, HC and LC users are not significantly different from NB users. This means that the nudge is effective, only when explicitly highlighting the benefit.
- The magnitude of the effect between A and B conditions (for H and L users) are very similar. If we switch the baseline to one of these two conditions, we retrieve the finding that conditions A and B are not statistically different from each other.

Table 6 Regression estimates for CTR using the pooled sample

	OLS (1)	OLS (2)	Logistic (3)	Logistic (4)
HA	0.024***	0.021***	1.026***	0.892***
	(0.001)	(0.001)	(0.054)	(0.059)
HB	0.025***	0.022***	1.066***	0.932***
	(0.001)	(0.001)	(0.1053)	(0.059)
HC	-0.0001	-0.003**	-0.008	-0.142*
	(0.001)	(0.001)	(0.075)	(0.1079)
LA	0.019***	0.016***	0.873***	0.754***
	(0.001)	(0.001)	(0.044)	(0.05)
LB	0.018***	0.015***	0.843***	0.726***
	(0.001)	(0.001)	(0.044)	(0.05)
LC	-0.0003	-0.003***	-0.023	-0.14**
	(0.001)	(0.001)	(0.054)	(0.059)
NA	0.0001	0.0001	0.009	0.008
	(0.001)	(0.001)	(0.049)	(0.049)
log(avg_distance)		0.003***		0.134***
		(0.001)		(0.026)
log(avg_leave_time)		0.011***		0.499***
		(0.002)		(0.105)
log(days_since_joined)		-0.002***		-0.074***
		(0.0003)		(0.014)
log(sessions_30d)		0.001		0.03
		(0.0005)		(0.022)
GA		-0.006***		-0.271***
		(0.001)		(0.037)
MA		-0.003***		-0.132***
		(0.001)		(0.045)
WA		-0.008***		-0.7352***
		(0.001)		(0.056)
Constant	0.014***	-0.0002	-4.271***	-4.982***
	(0.001)	(0.006)	(0.035)	(0.276)
Observations	301,303	301,303	301,303	301,303
R^2	0.005	0.006		
Log Likelihood			-30,796.900	-30,641.760
Residual Std. Error	0.145 (df = 301295)	0.145 (df = 301287)		
F Statistic	200.312*** (df = 7;301295)	115.532*** (df = 15;301287)		

$^{*}\ p<0.1$, $^{**}\ p<0.05$, $^{***}\ p<0.01$.

- Interestingly, we uncover a new effect for CTR after including the controls (see second and fourth columns of Table 6). Users in conditions HC and LC (i.e., HOV users who received a generic carpool offer) have a 9% lower CTR relative to NB users. This effect is statistically significant only after adding the controls and for CTR (OBR1 has a negative sign too, albeit not statistically significant). This reinforces our insight that incentivizing high-intent users is effective only when explicitly mentioning the benefit of carpooling. Sending a generic message to higher-intent users is not enough and can potentially be worse than offering compensation to lower-intent users.
- The independent variables: average commute distance and average leave time seem to have a significant economics effect on CTR:

Table 7 Regression estimates for (normalized) OBR1 using the pooled sample

	OLS (1)	OLS (2)	Logistic (3)	Logistic (4)
HA	0.075***	0.05***	2.244***	1.826***
	(0.01)	(0.01)	(0.332)	(0.366)
HB	0.075***	0.05***	2.246***	1.832***
	(0.01)	(0.01)	(0.332)	(0.366)
HC	0.0075	-0.0025	0.464	-0.048
	(0.01)	(0.01)	(0.44)	(0.466)
LA	0.025***	0.025**	1.136***	0.766**
	(0.0075)	(0.0075)	(0.298)	(0.33)
LB	0.025***	0.025***	1.404***	1.038***
	(0.0075)	(0.0075)	(0.29)	(0.322)
LC	0.0025	-0.005	0.16	-0.206
	(0.0075)	(0.0075)	(0.336)	(0.364)
NA	0.0075	0.0075	0.388	0.388
	(0.0075)	(0.0075)	(0.302)	(0.302)
log(avg_distance)		0.005		0.234
		(0.005)		(0.164)
log(avg_leave_time)		-0.0025		-0.076
		(0.025)		(0.676)
log(days_since_joined)		-0.0025		-0.066
		(0.0025)		(0.088)
log(sessions_30d)		0.01***		0.396***
		(0.0025)		(0.142)
GA		-0.01		-0.332
		(0.005)		(0.232)
MA		-0.025**		-0.57**
		(0.0075)		(0.282)
WA		-0.005		-0.216
		(0.01)		(0.334)
Constant	0.025***	0.0075	-13.24***	-14.3***
	(0.005)	(0.05)	(0.224)	(1.774)
Observations	301,303	301,303	301,303	301,303
R^2	0.0004	0.0004		
Log Likelihood			-4,426.213	-4,413.814
Residual Std. Error	0.045 (df = 301295)	0.045 (df = 301287)		
F Statistic	15.079*** (df = 7;301295)	8.731*** (df = 15;301287)		

* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

1. An additional 400 kilometers in the last 30 days (which can be seen as 10 extra kilometers in the one-way commute distance) increases CTR by 7.2% relative to NB users (holding the other predictor variables constant).
2. Similarly, an additional hour in the leave time increases CTR by 10% relative to NB users. This suggests that the schedule of commuters (captured by the average morning leave time) is highly correlated with the carpool intent.

- Users in California are generally more responsive to the carpooling encouragement (in terms of CTR) relative to the three other states. This can be explained by the fact that California commuters are more likely to carpool. Similarly, users with a higher average driven distance in the last 30 days are more receptive to the carpooling encouragement (in terms of CTR).

The above findings are robust across both model specifications and to the inclusion of impression-time fixed effects. We also obtain the same qualitative insights when using different ways of pooling the data, e.g., using H or L users only and combining H and L users (the regression tables are omitted for conciseness). We find the same qualitative results for OBR1 (recall that the numbers in Table 7 are normalized). For OBR2 (omitted for conciseness), the sign of all the coefficients are consistent with the other performance metrics. However, only the effects for H users are statistically significant. Last, the regression results for Absolute OBR provide a different perspective as now the baseline of the regression can be set to users in the control condition. In addition of retrieving all the previous qualitative insights (along with statistical significance), we also find that relative to NC users (i.e., non-HOV users who did not receive any treatment), HA, HB, LA, LB, and HD users have a higher CTR (see Table 10 in Appendix D). This confirms that not mentioning the HOV and incentivizing non-HOV users are not effective interventions.

We conclude that the insights discussed in Section 5.1 continue to hold even after controlling for various factors related to the user and to external attributes. The fact that our results are robust for different dependent variables and under a multitude of model specifications strengthen the validity of our findings.

We next perform one last robustness test. Instead of controlling for the market, we control for the zip-code (home or work). Explicitly controlling for the zip-code can help control for some time-invariant unobserved heterogeneity among different users. For example, home zip-codes can somewhat capture the heterogeneity in income levels and work zip-codes can account for the type of profession. Overall, we have 3,510 unique home zip-codes and 3,029 unique work zip-codes. The results are presented in Table 8. The first four columns are the same models as before (when we control for market fixed effects) while also controlling for the standard deviation of the leave time. The last three columns show the estimates for a fixed effects linear model (FELM) when we control for the zip-code. In Model (5), we control for the home zip-code, in Model (6), for the work zip-code, and in Model (7), for both zip-code. As we can see from Table 8, all our results still hold when controlling for zip-codes. Similarly, when using the sample with only H and L users, we obtain consistent results. Finally, most of the results are consistent for OBR1, OBR2, and Absolute OBR (omitted for conciseness).

5.3. Heterogenous Treatment Effects

In this section, we estimate the models from Equations (3) and (4) for the CTR while adding interaction terms between the treatment conditions (A, B, C, D) and one of the following variables: distance, leave time, average time saving, number of navigations with HOV in the last 30 days, days since joined, and number of sessions in the last 30 days. We next summarize our findings.

Table 8 Regression estimates for CTR using the pooled sample (including zip-code fixed effects)

	OLS (1)	OLS (2)	Logistic (3)	Logistic (4)	FELM (5)	FELM (6)	FELM (7)
HA	0.024***	0.021***	1.026***	0.898***	0.018***	0.021***	0.018***
	(0.001)	(0.001)	(0.054)	(0.059)	(0.001)	(0.001)	(0.001)
HB	0.025***	0.022***	1.066***	0.936***	0.019***	0.022***	0.019***
	(0.001)	(0.001)	(0.053)	(0.059)	(0.001)	(0.001)	(0.001)
HC	-0.0001	-0.003**	-0.008	-0.137*	-0.006***	-0.003**	-0.006***
	(0.001)	(0.001)	(0.075)	(0.079)	(0.001)	(0.001)	(0.001)
LA	0.019***	0.016***	0.873***	0.752***	0.014***	0.016***	0.015***
	(0.001)	(0.001)	(0.044)	(0.05)	(0.001)	(0.001)	(0.001)
LB	0.018***	0.015***	0.843***	0.723***	0.013***	0.015***	0.3013***
	(0.001)	(0.001)	(0.044)	(0.05)	(0.001)	(0.001)	(0.001)
LC	-0.0003	-0.003***	-0.023	-0.144**	-0.004***	-0.003***	-0.004***
	(0.001)	(0.0001)	(0.054)	(0.059)	(0.001)	(0.001)	(0.001)
NA	0.0001	0.0001	0.009	0.008	0.0001	0.0004	0.0001
	(0.001)	(0.001)	(0.049)	(0.049)	(0.001)	(0.001)	(0.001)
log(avg_distance)		0.003***		0.12***	0.003***	0.003***	0.003***
		(0.001)		(0.026)	(0.001)	(0.001)	(0.001)
log(avg_leave_time)		0.005**		0.227**	0.006**	0.004*	0.005**
		(0.002)		(0.11)	(0.002)	(0.002)	(0.002)
log(std_leave_time + 1)		0.00004		0.017	0.0002	-0.0001	0.0002
		(0.001)		(0.072)	(0.001)	(0.001)	(0.001)
is(std_leave_time_NA)		0.005***		0.266***	0.005***	0.005***	0.005***
		(0.001)		(0.039)	(0.001)	(0.001)	(0.001)
log(days_since_joined)		-0.002***		-0.07***	-0.001***	-0.002***	-0.001***
		(0.0003)		(0.014)	(0.0003)	(0.0003)	(0.0003)
log(sessions_30d)		0.001***		0.059***	0.001***	0.001**	0.001**
		(0.0005)		(0.022)	(0.0005)	(0.0005)	(0.0005)
GA		-0.006***		-0.275***			
		(0.001)		(0.037)			
MA		-0.003***		-0.15***			
		(0.001)		(0.045)			
WA		-0.008***		-0.348***			
		(0.001)		(0.056)			
Constant	0.014***	0.006	-4.271***	-4.652***			
	(0.001)	(0.006)	(0.035)	(0.277)			
Observations	301,303	301,303	301,303	301,303	301,303	301,303	301,303
R^2	0.005	0.006			0.017	0.017	0.027
Log Likelihood			−30,796.900	−30,594.600			
Residual Std. Error	0.145 (df = 301295)	0.145 (df = 301285)			0.145 (df = 298439)	0.145 (df = 298708)	0.145 (df = 296218)
F Statistic	200.312*** (df = 7;301295)	107.640*** (df = 17;301285)					

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

- We observe that the treatment effect is amplified for H users with a higher average time saving (the estimated coefficients of the pool regression can be found in Table 9). This finding holds for H users but not for L users. Indeed, all L users have a similar (low) time saving (i.e., no substantial variation across users), whereas the range of time saving for H users is more substantial. More precisely, we find that an additional 20% in time saving increases the CTR by 12.15% and 17.5% for HA and HB users, respectively relative to LC users.
- We find that the treatment effect is also amplified by the driven distance in last 30 days (the regression table is omitted for conciseness). Specifically, we find that an additional 400 kilometers in the last 30 days (which can be seen as 10 extra kilometers in the one-way commute distance) for users in conditions (HA, HB, LA, LB) increases the CTR by (26.8%, 14.5%, 14.5%, 21.7%) relative to NB users.
- We did not find that the magnitude of the treatment effect is moderated by the standard deviation of the leave time, as we did in Table 1 from Section 3.1. Two plausible explanations come to mind: (i) the treatment effect on CTR is not moderated by the flexibility in the leave time but the OBR in general is, (ii) the statistical power of our experiment is not strong

Table 9 Regression estimates for CTR including interaction terms with average time saving

	OLS (1)	Logistic (2)
HA	0.0002 (0.008)	0.498* (0.276)
HB	-0.007 (0.008)	0.338 (0.273)
HC	0.005 (0.008)	0.277 (0.437)
LA	0.02*** (0.002)	0.982*** (0.113)
LB	0.016*** (0.002)	0.867*** (0.114)
log(avg_time_saving + 1)	0.002 (0.002)	0.115 (0.09)
log(avg_distance)	0.005*** (0.001)	0.183*** (0.033)
log(avg_leave_time)	0.015*** (0.003)	0.564*** (0.123)
log(days_since_joined)	-0.002*** (0.0004)	-0.068*** (0.0316)
log(sessions_30d)	0.001* (0.001)	0.049* (0.026)
GA	-0.008*** (0.001)	-0.6312*** (0.045)
MA	-0.006*** (0.002)	-0.22*** (0.081)
WA	-0.008*** (0.001)	-0.342*** (0.057)
HA × log(avg_time_saving + 1)	0.009** (0.004)	0.166 (0.143)
HB × log(avg_time_saving + 1)	0.013*** (0.004)	0.253* (0.141)
HC × log(avg_time_saving + 1)	-0.003 (0.004)	-0.191 (0.207)
LA × log(avg_time_saving + 1)	-0.001 (0.002)	-0.093 (0.107)
LB × log(avg_time_saving + 1)	0.002 (0.002)	-0.0003 (0.108)
Constant	-0.021** (0.009)	-5.716*** (0.15)
Observations	181,282	181,282
R^2	0.005	0.006
Log Likelihood		-21,912.35
Residual Std. Error	0.1161 (df = 181262)	
F Statistic	49.383*** (df = 19;181282)	

* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

enough to identify this heterogenous treatment effect for OBR (the number of on-boarded drivers in each condition is not large enough to convey statistical significance).

The above findings are not surprising and confirm our intuition that users who can benefit more from using an HOV lane or drive longer distances are responding better to the HOV-explicit encouragements tested in our experiment. We also observe that all other variables do not moderate the treatment effect.

6. Conclusions and Implications

To our knowledge, this paper presents the first large digital field experiment to nudge commuters to become carpoolers. We select a sample of 537,370 Waze users across four U.S. states and send them an in-app notification to invite them to try a carpool service. We test different framing alternatives to better understand what are the successful triggers to convert commuters into carpoolers. We identify (i) users who can save a significant commute time by carpooling through the use of an HOV lane, (ii) users who can still use the HOV lane but with a low time saving, and (iii) users who do not have access to an HOV lane in their daily commute. We consider several treatments that correspond to varying the framing of the notification: mentioning the HOV lane, highlighting the time saving, explicitly mentioning the monetary incentive, and sending a generic message.

We find that mentioning the HOV lane is the most successful framing and increases the click-through rate and conversion rate by 133–185% and 64–141%, respectively relative to a generic message. We also find that compensation is not a critical driver to nudge commuters to carpool. This suggests that explicitly mentioning the benefit (in our case, the HOV lane) works better than offering a monetary incentive. We also conclude that incentivizing high-intent users is effective only when explicitly highlighting the benefit of carpooling. Sending a generic message to high-intent users is not enough and can be worse than offering compensation to low-intent users.

The results presented in this paper bear interesting implications on the design of carpooling platforms. First, while drivers seem to intuitively know how much time they might save by taking the HOV lane, it is still valuable to emphasize the link between the desired course of action—carpooling—and the eligibility to use the HOV lane. Second, the experiment in this paper focuses on general intent; when offered an option to pick up a rider along the way, being able to communicate that using the HOV lane outweighs the pick up detour may significantly increase carpool conversion and lead to more efficient prices. Third, the time saving can be predicted and communicated nearly in real time, and thus help drivers plan their carpool commutes.

We conclude this section by highlighting some of the ramifications our results have on policy. Our study ultimately suggests that users are more receptive to saving commute time than to receiving compensation. We also infer that both the type of user and the framing of the intervention play important roles in converting commuters to carpool. Government agencies or large corporations can use the results presented in this research to incentivize their workers to carpool together. Instead of offering financial incentives, they can focus on highlighting the time saving benefits of carpooling (e.g., using the HOV lane or offering a nearby parking spot to reduce the time searching for parking). Convincing even a small portion of the 116 million U.S. solo commuters to carpool can have tremendous environmental and societal impacts. In the same vein, given that there are more than 110 million Waze drivers, if we could convince 1% of them to carpool, this will reduce the number of vehicles on the road by more than 1 million and the CO2 emissions by 3.4 ton.[10] As a result, sharpening our understanding on the key drivers that can nudge commuters to carpool is vital. This paper provides a first step towards addressing this question.

Our focus in this work is on converting commuters to carpool drivers (i.e., taking a stranger in their car). An equally important research question—left for future research—is to investigate what determinants and motivating factors can convince commuters to leave their car at home and become carpool riders.

Acknowledgments

We would like to thank Yarin Aidelman, Rapha Cohen, Baek Jung Kim, Joshua Loftus, Peleg Samuels, and Hal Varian for insightful feedback and discussions, which have helped improve this paper.

References

Benartzi S, Beshears J, Milkman KL, Sunstein CR, Thaler RH, Shankar M, Tucker-Ray W, Congdon WJ, Galing S (2017) Should governments invest more in nudging? *Psychological science* 28(8):1041–1055.

Beroldo S (1999) Casual carpooling 1998 update, rides for Bay Area Commuters.

Cohen MC, Fiszer MD, Kim BJ (2018) Frustration-based promotions: field experiments in ride-sharing, working paper.

Cross GS (1993) *Time and money: The making of consumer culture* (Routledge London).

Ferguson E (1997) The rise and fall of the american carpool: 1970–1990. *Transportation* 24(4):349–376.

Giuliano G, Levine DW, Teal RF (1990) Impact of high occupancy vehicle lanes on carpooling behavior. *Transportation* 17(2):159–177.

Halpern D (2015) *Inside the nudge unit: How small changes can make a big difference* (Random House).

Kohavi R, Deng A, Frasca B, Walker T, Xu Y, Pohlmann N (2013) Online controlled experiments at large scale. *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1168–1176 (ACM).

Kohavi R, Thomke S (2017) The surprising power of online experiments. *Harvard Business Review* 95(5):74–81.

Konishi H, Mun Si (2010) Carpooling and congestion pricing: Hov and hot lanes. *Regional Science and Urban Economics* 40(4):173–186.

Maxwell SE, Delaney HD (2004) *Designing experiments and analyzing data: A model comparison perspective*, volume 1 (Psychology Press).

[10] In this calculation, we assume an average (roundtrip) commute distance of 32 miles and that the average passenger vehicle emits 404 grams of CO2 per mile (these numbers are supported by various sources).

Neoh JG, Chipulu M, Marshall A (2017) What encourages people to carpool? an evaluation of factors with meta-analysis. *Transportation* 44(2):423–447.

Okada EM, Hoch SJ (2004) Spending time versus spending money. *Journal of consumer research* 31(2):313–323.

Ostrovsky M, Schwarz M (2019) Carpooling and the economics of self-driving cars. *Proceedings of the 2019 ACM Conference on Economics and Computation*, 581–582 (ACM).

Richardson M, Dominowska E, Ragno R (2007) Predicting clicks: estimating the click-through rate for new ads. *Proceedings of the 16th international conference on World Wide Web*, 521–530 (ACM).

Shaheen SA, Chan ND, Gaynor T (2016) Casual carpooling in the san francisco bay area: Understanding user characteristics, behaviors, and motivations. *Transport Policy* 51:165–173.

Singh J, Teng N, Netessine S (2017) Philanthropic campaigns and customer behavior: Field experiments on an online taxi booking platform. *Management Science* 65(2):913–932.

Small KA, Winston C, Yan J, Baum-Snow N, Gómez-Ibáñez JA (2006) Differentiated road pricing, express lanes, and carpools: Exploiting heterogeneous preferences in policy design [with comments]. *Brookings-Wharton Papers on Urban Affairs* 53–96.

Teal RF (1987) Carpooling: who, how and why. *Transportation Research Part A: General* 21(3):203–214.

Thaler RH, Sunstein CR (2009) *Nudge: Improving decisions about health, wealth, and happiness* (Penguin).

Vanoutrive T, Van De Vijver E, Van Malderen L, Jourquin B, Thomas I, Verhetsel A, Witlox F (2012) What determines carpooling to workplaces in belgium: location, organisation, or promotion? *Journal of transport geography* 22:77–86.

Yang H, Huang HJ (1999) Carpooling and congestion pricing in a multilane highway with high-occupancy-vehicle lanes. *Transportation Research Part A: Policy and Practice* 33(2):139–155.

Appendix A: List of HOV lanes in our Field Experiment

Here is the list of HOV lanes we consider for selecting the users in our analysis:

- California: SR-57 N, I-880 N, I-805 N, US-101 N, SR-22 E, SR-237 E, and SR-57 S.
- Washington: I-90 W, SR-520 E, Exit 11: I-90 W / Seattle, and SR-520 W.
- Massachusetts: I-93 (North Expressway), and I-93 (Southeast Expressway).
- Georgia: I-75, I-85, and I-20.

Appendix B: Additional Balancedness Checks

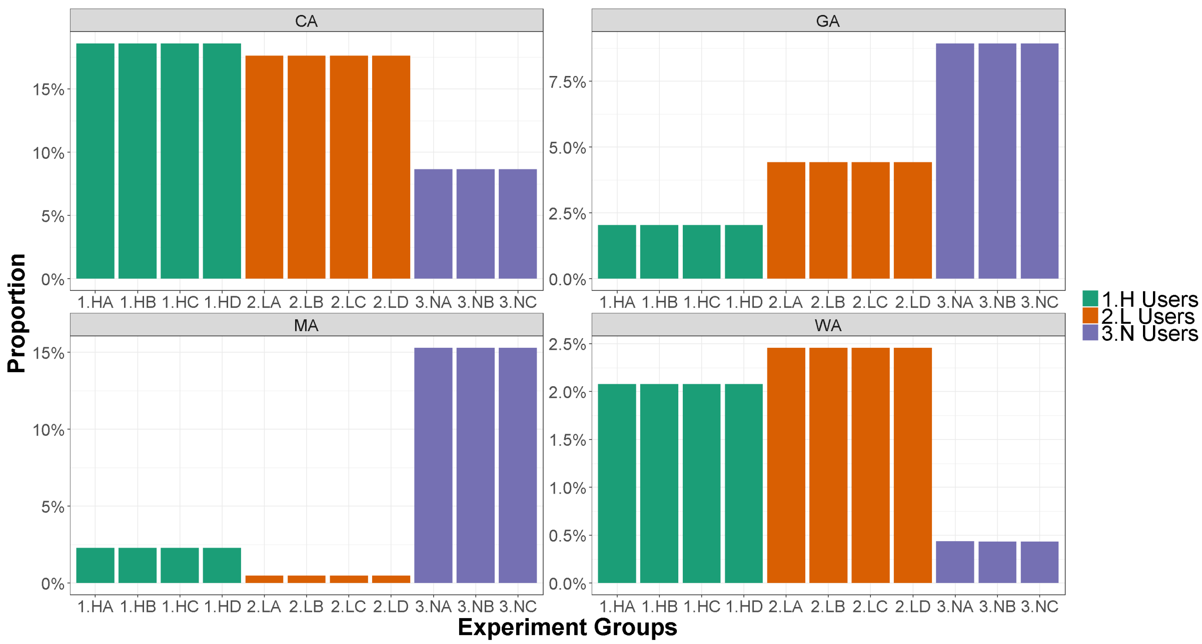
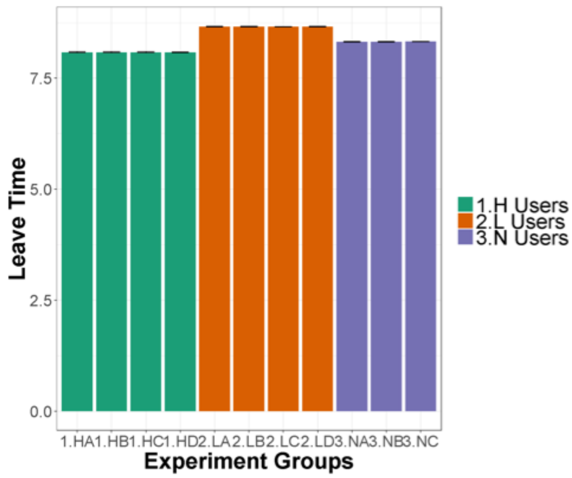
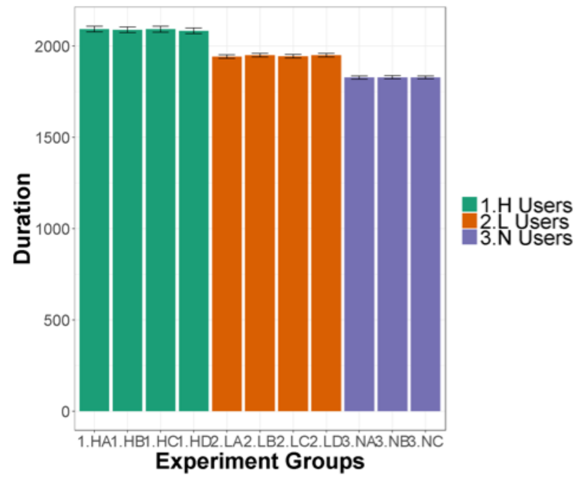


Figure 15 **Proportion in the four markets across the different groups**



(a) Leave time



(b) Duration (in minutes)

Figure 16 **Average leave time and duration (in minutes) in the last 30 days**

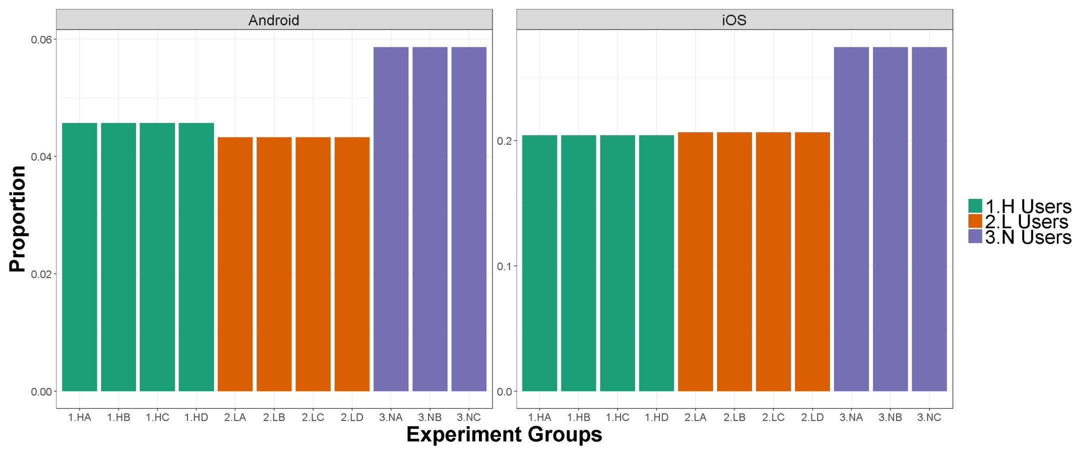


Figure 17 **Proportion of Android versus iOS users across the different groups**

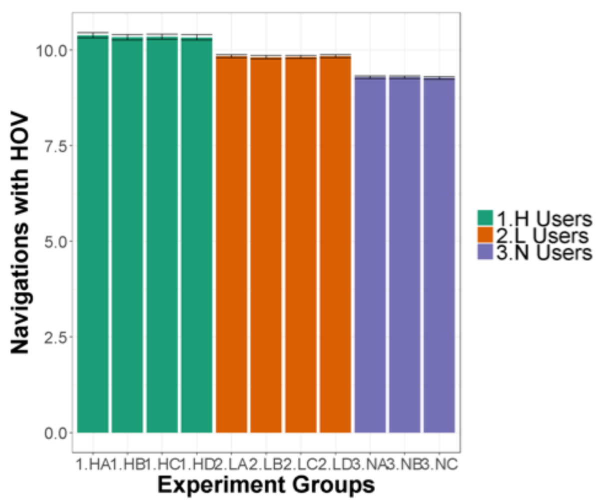
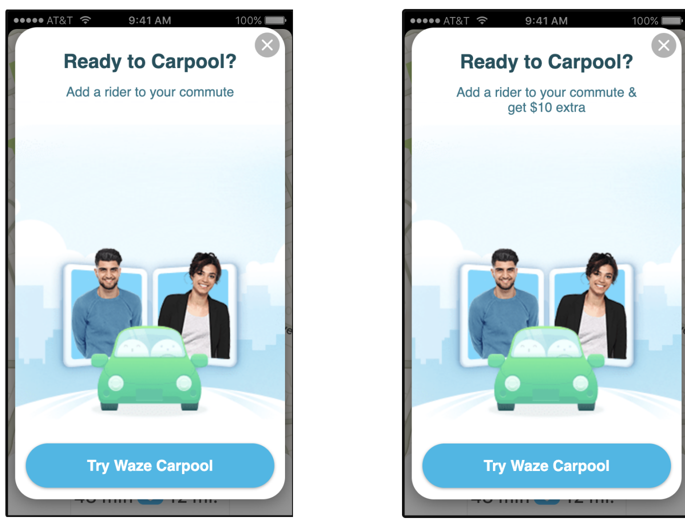


Figure 18 **Average number of navigations with an HOV lane in the last 30 days across the different groups**

Appendix C: Screenshots of the Encouragements

(a) Condition A (b) Condition B (c) Condition C

Figure 19 Copy of the message used in our field experiment for L users



37

(a) Condition A (b) Condition B

Figure 20 Copy of the message used in our field experiment for N users

Appendix D: Additional Regression Results

Table 10 Regression estimates for (normalized) Absolute OBR using the pooled sample

	OLS (1)	OLS (2)	Logistic (3)	Logistic (4)
HA	0.075***	0.05***	2.37***	2.034***
	(0.0075)	(0.0075)	(0.308)	(0.332)
HB	0.05***	0.05***	2.162***	1.828***
	(0.0075)	(0.0075)	(0.316)	(0.34)
HC	0.01	0.0025	0.57	0.236
	(0.0075)	(0.0075)	(0.408)	(0.426)
HD	0.025***	0.0025**	1.28***	0.944**
	(0.0075)	(0.0075)	(0.36)	(0.382)
LA	0.025***	0.025***	1.256***	0.982***
	(0.0075)	(0.0075)	(0.278)	(0.3)
LB	0.025***	0.025***	1.398***	1.126***
	(0.005)	(0.0075)	(0.274)	(0.298)
LC	0.0075	0.001	0.42	0.148
	(0.005)	(0.0075)	(0.308)	(0.328)
LD	-0.005	-0.01*	-0.392	-0.666*
	(0.0075)	(0.0075)	(0.308)	(0.328)
NA	0.0075	0.0075	0.496*	0.496*
	(0.005)	(0.005)	(0.284)	(0.284)
NB	0.005	0.005	0.368	0.368
	(0.005)	(0.005)	(0.288)	(0.288)
log(avg_distance)		0.005*		0.244*
		(0.0025)		(0.132)
log(avg_leave_time)		-0.0025		-0.158
		(0.0125)		(0.554)
log(days_since_joined)		-0.0025		-0.064
		(0.0025)		(0.072)
log(sessions_30d)		0.01***		0.424***
		(0.0025)		(0.114)
GA		-0.0075		-0.292
		(0.005)		(0.192)
MA		-0.0075		-0.302
		(0.005)		(0.218)
WA		0.0025		0.072
		(0.005)		(0.262)
Constant	0.025***	0.0075	-13.498***	-14.666***
	(0.005)	(0.025)	(0.212)	(1.448)
Observations	537,370	537,370	537,370	537,370
R^2	0.0003	0.0004		
Log Likelihood			-6,800.244	-6,781.847
Residual Std. Error	0.042 (df = 537359)	0.042 (df = 537351)		
F Statistic	15.426*** (df = 10;537359)	10.676*** (df = 18;537351)		

$^{*}\ p<0.1$, $^{**}\ p<0.05$, $^{***}\ p<0.01$.

Table 11 Regression estimates for OBR1 when comparing the differences

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
H	0.0002 (0.0004)	0.0002 (0.0004)	0.0001 (0.0004)	0.0001 (0.0004)		
L					−0.0002 (0.0003)	−0.0002 (0.0003)
A	0.001*** (0.0003)		−0.0003 (0.0002)		−0.0003 (0.0002)	
H × A	0.001** (0.001)		0.003*** (0.001)			
B		0.001*** (0.0003)		−0.0003 (0.0002)		−0.0003 (0.0002)
H × B		0.001* (0.001)		0.003*** (0.001)		
L × A					0.001*** (0.0004)	
L × B						0.002*** (0.0004)
Constant	0.001*** (0.0002)	0.001*** (0.0002)	0.002*** (0.0002)	0.002*** (0.0002)	0.002*** (0.0002)	0.002*** (0.0002)
Observations	121,297	121,011	153,122	153,103	208,217	207,950
R^2	0.0003	0.0004	0.0004	0.0004	0.0001	0.0002
Residual Std. Error	0.046 (df = 121293)	0.048 (df = 121007)	0.042 (df = 153118)	0.042 (df = 153099)	0.041 (df = 208213)	0.041 (df = 207946)
F Statistic	13.730*** (df = 3; 121293)	14.307*** (df = 3; 121007)	18.928*** (df = 3; 153118)	18.974*** (df = 3; 153099)	5.951*** (df = 3; 208213)	10.399*** (df = 3; 207946)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.