Putting Teams into the Gig Economy: A Field Experiment at a Ride-sharing Platform

By Wei Ai, Yan Chen, Qiaozhu Mei, Jieping Ye, and Lingyu Zhang *

The gig economy provides workers with the benefits of autonomy and flexibility, but it does so at the expense of work identity and co-worker bonds. Among the many reasons why gig workers leave their platforms, an unexplored aspect is the organization identity. In this study, we develop a team formation and inter-team contest field experiment at a ride-sharing platform. We assign drivers to teams either randomly or based on homophily in age, hometown location, or productivity. Having these teams compete for cash prizes, we find that: (1) compared to those in the control condition, treated drivers work longer hours and earn 12% higher revenue during the contest, with a larger effect (19%) for teams comprised of drivers who are more communicative and responsive; and (2) drivers in responsive teams continue to work longer hours and earn higher revenue during the two weeks after the contest ends. Together, our results show that platform designers can leverage team identity and team contests to increase revenue and worker engagement in a gig economy.

Keywords: team, contest, identity, gig economy, ride-sharing, field experiment

JEL Classification Numbers: C72, C93, D91

As trends in work sourcing move us toward a gig economy, this economy is widely considered to be the future face of work, despite questions about its sustainability (Ravenelle, 2019). While workers in traditional sectors derive their identities from their work and share their experiences with co-workers, those whose livelihood relies on the gig economy often find that "these are jobs that don't lead to anything," citing a lack of work identity and bonds with co-workers as well as an inability to move upward based on strong performance (*The New Yorker*, May 15, 2017). Ride-sharing platforms, such as Uber, have experienced high attrition rates (Scheiber, 2017). Among the many reasons why gig workers leave their platforms, an unexplored aspect is the organization identity.

To analyze these and other concerns associated with the gig economy, we apply social identity theory (Tajfel et al., 1971; Tajfel and Turner, 1979; Akerlof and Kranton, 2000) to a large gig platform, DiDi Chuxing (DiDi henceforth), where

* Ai: College of Information Studies, University of Maryland, College Park, MD 20742, USA, email: aiwei@umd.edu; Chen and Mei: School of Information, University of Michigan, Ann Arbor, MI 48109, USA, email: yanchen@umich.edu, qmei@umich.edu; Ye: Beike, Beijing, China, email: jieping@gmail.com; Zhang: AI Labs, DiDi Chuxing, Beijing 100094, China, email: zhanglingyu@didiglobal.com

individual drivers offer ride sharing in China. Specifically, we design a field experiment to study team formation and inter-team competition within DiDi. In our experiment, we examine how the creation of a team identity impacts driver revenue. Furthermore, since DiDi is a flat organization with no group structure, we are also able to investigate how different team formations impact team member communication and productivity.

Our research applies insights from the social psychology and economics literature on identity (Brewer, 1999; Akerlof and Kranton, 2000, 2010). This research shows that, when people feel a stronger sense of common identity with a group, they exert higher effort and make more contributions to public goods to reach a more efficient outcome in the laboratory, using either induced (Eckel and Grossman, 2005; Charness, Rigotti and Rustichini, 2007; Chen and Chen, 2011) or natural identities (Goette et al., 2012; Chen et al., 2014). Moving from the lab to the field, results have been mixed. In an early public goods field experiment, Erev, Bornstein and Galili (1993) use team competition in fruit harvesting and find that team competition increases productivity. Likewise, Ai et al. (2016) report the results of a large-scale field experiment designed to test the hypothesis that team membership can increase pro-social lending for an online microfinance community, Kiva.org. They find that team recommendations increase the likelihood that a lender joins a team, and joining a team increases lending substantially following the intervention. In a charitable giving context, Kessler and Milkman (2018) show that individuals are more likely to donate when a facet of their identity associated with a norm of generosity is primed in field experiments run by the American Red Cross. This study demonstrates how identity primes can be implemented in practice to encourage public good provision. In comparison, Gee, Schreck and Singh (2020) randomly assign potential donors into teams and vary the social distance within these teams. They find no evidence that reduced social distance increases giving. In a recent survey of the theory and experiments in identity economics, Charness and Chen (2020) suggest that identity-based teams in the field might be a useful behavioral mechanism to increase pro-social behavior. Applying this framework to our setting, we anticipate that a driver who has a strong sense of team identity will work harder to help his team get ahead compared to drivers who do not belong to any team (Brewer and Silver, 2000).

In examining how different team formations may have different effects on communication and coordination, we use an algorithm that maximizes either similarity or diversity within a team. We conjecture that similarity might facilitate team member communication and coordination, leading to intra-team bonding and team stability (Ruef, Aldrich and Carter, 2003; Yuan and Gay, 2006; Kim and Aldrich, 2002). Indeed, empirical network science studies provide evidence for homophily, or the tendency of people to associate with others whom they perceive as similar to themselves in some way (McPherson, Smith-Lovin and Cook, 2001; Girvan and Newman, 2002). By contrast, we conjecture that diversity might bolster team performance, due to different perspectives in problem-solving and better complementarity among team members (Krishnan, Miller and Judge, 1997; Page, 2007).

In addition to examining different team formation strategies, we draw on insights from contest theory (Konrad, 2009; Vojnović, 2016) and experiments (Sheremeta, 2018) to explore how to structure team contests, which have been shown to be among the most effective ways to strengthen team identity in the laboratory (Eckel and Grossman, 2005).

Lastly, our work contributes to the rapidly growing literature on the gig economy, and ride-sharing in particular, which has uncovered important insights related to labor market outcomes (Hall and Krueger, 2018), the value of flexible work (Chen et al., 2019), consumer surplus (Cohen et al., 2016), and decentralized dynamic matching efficiency (Liu, Wan and Yang, 2018). Our findings contribute to this stream of research by showing that a team-based approach can significantly increase driver revenue, their bonds with co-workers, and their team identity.

I. Experiment Design

To test the effectiveness of team formation and inter-team competition on productivity, we design a multistage natural field experiment using the ride-sharing platform DiDi. Founded in 2012, DiDi is the dominant ride-sharing company in China. The platform employs more than 31 million drivers globally and offers app-based transportation options for 550 million users across Asia, Latin America and Australia, making it the largest ride-sharing platform in the world.[1] In China, DiDi drivers comprise of workers laid off from their traditional jobs, veterans, migrant workers from rural areas, and commuters who offer rides during their daily commute. On the DiDi platform, drivers receive 81% of the revenue they generate and give the remaining 19% to the platform.

We collaborate with the DiDi AI Labs to conduct our field experiment to address one of its persistent problems, i.e., low engagement among its workforce, which is also prevalent among other gig economy platforms (Ravenelle, 2019). Our experiment consists of three stages: recruitment, team formation, and team contest. In what follows, we present our design choices in each stage.

A. Recruitment and Power Analysis

We conduct our experiment in the southern city of Dongguan in the summer of 2017, which has 480,000 registered DiDi drivers. We select our pool of drivers based on their productivity in a two-week period (July 18 - 31, 2017) prior to the announcement of the contest, using the following criteria to filter the drivers. First, the driver has finished one or more trips on at least five weekdays and two weekend days during the two-week period. Second, the driver finishes five or more trips on average on the days she works during the two-week period. This filtering

[1]Statistics are from the company Website: https://www.didiglobal.com/about-didi/about-us, retrieved on March 5, 2020.

process yields a total of 28,394 eligible drivers. From this pool, we randomly select 24,000 drivers to receive a text message invitation. The remaining 4,394 drivers comprise our *no-contact* group.

To determine the number of drivers needed in our experiment, we use the observational data from 9,000 randomly selected drivers in Beijing in January and February 2017.[2] We find that drivers in that random sample complete on average 11.7 orders per day ($\sigma = 4.7$). Since we expect an effect size of 10% ($\delta = 1.17$), with $\alpha = 0.05$ and $\beta = 0.10$ (90% power), this requires us to have 340 drivers per experimental condition, assuming equal variance across experimental conditions (List, Sadoff and Wagner, 2011). As each team has seven drivers, the number of drivers per treatment should be a multiple of seven. This leads us to selecting a sample size of 350 drivers per experimental condition.

We elicit participation interest by sending drivers text messages. In each text message, we ask if the driver would like to register for a team contest, in which they might find new friends and earn 1,000 CNY or more together as a team if they win.[3] Additionally, we ask if a driver is interested in becoming a team captain, who will earn an additional 100 CNY upon fulfilling a captain's duties. Our announcement received 2,343 positive responses, 531 of which were interested in being a team captain. These text messages are included in the Appendix A.

For our experiment, we randomize our positive responses into the following groups: (1) 1,750 drivers are randomized into our treatment group. These drivers are subsequently partitioned into teams of seven (250 teams in total). (2) 350 drivers are randomized into the control group. These drivers are not placed in a team. During the contest period, they continue to earn piece rate. (3) The remaining 243 drivers serve as backups in case drivers in the treatments drop out before the start of the contest. Indeed, in our experiment, 15 drivers were reported by their captains as not responsive or no longer available for the contest. We mark these 15 drivers as "dropouts" and replace them with similar drivers from the backup group.[4] Figure 1 presents the experimental procedure.

B. Team Formation

During the team formation stage, we first randomly assign the 1,750 treatment drivers into five treatments, with the constraint that there are sufficient number of drivers in each treatment who are willing to serve as a team captain. We then group drivers in each treatment into 50 teams. We choose five team formation

[2] Using this particular data set for our power calculation is based on the availability of processed data during the experiment design phase.

[3] Contest rules are announced after the team formation process and before the start of the contest. See the Appendix A, for recruiting materials. Around the time of our experiment, the exchange rate was 1 USD $\approx$ 6.7 CNY.

[4] Similarity is based on pre-contest productivity using the number of trips in a 14-day window while ensuring additional treatment-specific criteria. For example, if the dropout driver is assigned to a team based on hometown similarity, we require that the new driver is from the same province. If the dropout driver is assigned to a team based on age similarity, we require that the substitute driver is in the same age group.

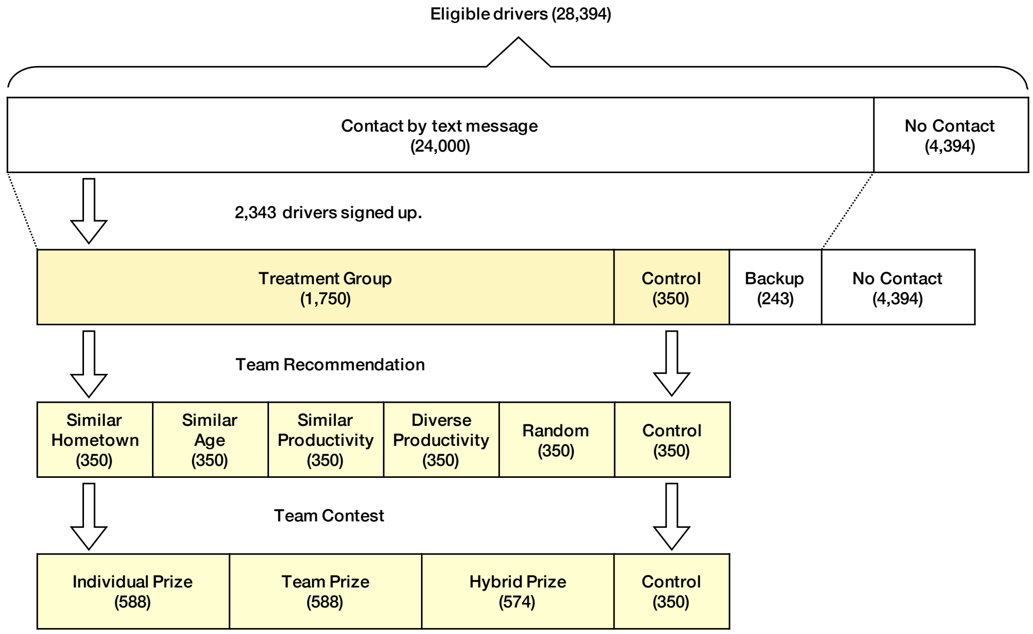

FIGURE 1. EXPERIMENTAL PROCEDURE

algorithms based on homophily or diversity considerations.

Our first condition, *hometown similarity*, is based on previous findings that location similarity is the most effective characteristic in getting a microfinance community member to join a specific lending team (Ai et al., 2016). In our study, we use hometown similarity, a form of location similarity, assigning drivers from the same (or a nearby) province to the same team. Prior studies indicate that hometown location is a salient identity among migrant workers in China (Zhang and Xie, 2013). We partition the drivers in to teams to ensure the seven drivers in the same team are all from the same (or a nearby) province.

Our second condition, *age similarity*, is based on prior research illustrating that age cohorts, such as the generation "born in the 1980s," are meaningful identity groups in China as each cohort has different socialization experiences (Harmel and Yeh, 2015). Therefore, we form our age similarity teams to reflect an age span of five years, e.g., 1980-1984, 1985-1989, with the exception of the 1960-1969 cohort, which spans a decade due to the relatively lower number of drivers born in the 1960s in our sample. Within each age cohort, we randomize the drivers into teams of seven. We randomize several times until we reach a partitioning that ensure at least one driver in each team has volunteered to be a team captain.

Third, we include *productivity similarity* as one of our algorithms as it is the preferred team formation algorithm by the platform. For this condition, we partition the 350 drivers into 10 buckets (35 each) based on their number of trips completed in the two weeks prior to the announcement of the team contest. Within each bucket, we randomize the drivers into teams of seven. We randomize several times until we reach a partitioning that ensure at least one driver in each team has volunteered to be a team captain.

Finally, we draw on recent scholarly research supporting the advantages of diversity (Page, 2007) and use two strategies to create diverse teams. To achieve *productivity diversity* in our teams, we partition the 350 drivers into seven buckets based on their number of trips completed in the two weeks prior to the announcement of the team contest in our experiment. We then randomly select one driver from each bucket to form a team. We randomize several times until we reach a partitioning that ensure at least one driver in each team has volunteered to be a team captain. Therefore, each team consists of drivers from all seven buckets.

Our final strategy, *random formation*, reflects the diversity achieved from a random grouping of drivers. For this condition, we randomly partition the drivers in to teams of seven and we repeat the randomization until at least one driver in each team has volunteered to be a team captain.

In sum, our team formation stage yields a total of 1,750 treatment drivers formed into 250 teams, with seven drivers in each team and 50 teams in each treatment. As we have 531 volunteers for 250 captain positions, we randomly select one volunteer to be the captain whenever a team has more than one. This randomization enables us to estimate the effect of being a captain on productivity and other metrics.

C. Within-team Communication

Within each team, we identify a team captain who is notified of this position, given the phone number of each team member, and asked to complete a pre-contest survey. The survey requires captains to communicate with each driver in the team to get the last three digits of their license plate number as well as several key pieces of demographic information (see Appendix B, for survey questions). Meanwhile, team members are given the captain's phone number and told that the captain might call them. Each team captain earns 100 CNY, which is public information announced in the initial text message.

The team communication task is designed to nudge the captains to communicate and collaborate with their team members, as several laboratory experiments demonstrate that a collaborative problem solving task involving group communications is a reliable way to strengthen group identity (Chen and Li, 2009; Chen and Chen, 2011; Chen, Chen and Riyanto, 2020) and to improve within-group coordination in contests (Cason, Sheremeta and Zhang, 2012). As the DiDi platform does not contain any team communication tools, we expect that most teams communicate by phone or WeChat,[5] an expectation which is verified by our post-

[5] WeChat is the dominant social media and communication app in China, which allows group communication and group calls.

experiment interviews. Captains who submit the survey through an online form are given 50 CNY as a bonus regardless of the correctness of their answers.

If a captain submits the survey, we mark the team as *responsive*, as completing the survey requires the captain to contact team members who need to respond to the captain. In comparison, a *non-responsive* team is defined as such if the captain does not submit the survey. This can happen because (1) the captain never contacts team members; or (2) the captain tries to contact them but team members are unresponsive; or (3) the captain contacts team members, team members respond, but the captain eventually fails to submit the survey. In the first two cases, either the captain or team members are non-responsive. In the last case, the team members communicate with the captain but the captain forgets to submit. While our design cannot disentangle these scenarios, communication within "non-responsive teams," such as in case (3), will only bias our heterogeneous treatment effects downwards (Tables 3 and 4). In our sample, 60.8% of our captains submit their survey. Conditional on submitting the survey, 81.1% of the answers are correct, indicating that communication did happen among responsive teams

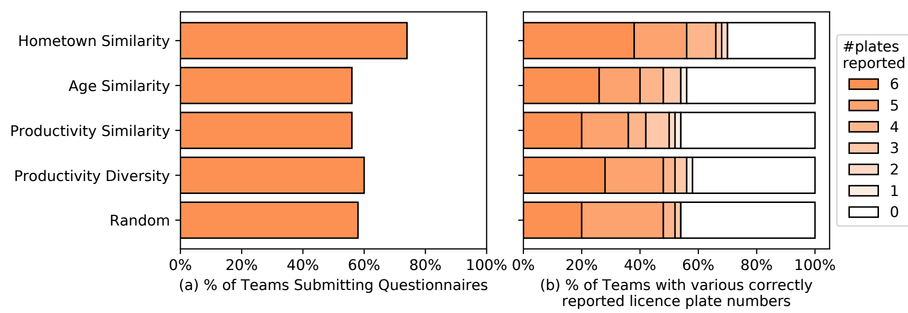

Figure 2. Team Responsiveness in Different Treatments. Team Responsiveness is coded based on the pre-contest survey. Panel (a) codes the responsiveness binarily, with a team deemed responsive if the captain submits the survey on team member characteristics. Panel (b) codes responsiveness based on the number of correctly-reported license plate numbers.

Figure 2 reports the proportion of responsive teams using both a binary variable to indicate whether a captain submits the survey (left panel) and the number of correctly-reported license plate numbers (right panel), where darker shades indicate more correct answers. From the left panel, we see that teams composed of members from the same hometown have the highest proportion of responsiveness. This result is consistent with prior research that shows that location similarity is a strong predictor of whether a member of an online community joins a team (Ai et al., 2016). The corresponding regression analysis is reported in Table F1 in the Appendix F. The results in Table F1 again show that teams based on hometown similarity show the highest level of responsiveness. Quantitatively, these teams are 19% more likely to be responsive than age-similar or productivity-similar teams after controlling for demographics. Along the intensive margin, however, we do not find any significant difference among the teams whose captains have submitted the survey. One possible reason for this finding may be that the captains decide to submit their surveys only if they have sufficient information. Indeed, more than 75% of the captains who submit the survey correctly report at least five out of six license plate numbers. While the team formation algorithms are not revealed to any captains or drivers, we expect that people from the same hometown can infer this information from their distinct accents.

D. Contest stage

Our contest rules are based on team contests of inherently individualistic sports, such as tennis and chess, where team outcomes are determined by multiple pairwise battles. Specifically, we set up a contest where drivers from two rival teams form pairwise matches to engage in distinct component battles. Within each team, we use an algorithm to automatically rank drivers by their number of trips completed in the two weeks prior to the announcement of the contest, and pair the most productive driver in team A with the most productive one in team B; and so on. In this contest, a team wins if and only if its drivers win a majority of their battles. This team contest format is often deployed for table tennis and badminton as well, which are popular sports in China that our drivers are familiar with. Fu, Lu and Pan (2015) provide a theoretical analysis of this type of team contests under complete information and sequential moves. In our ride-sharing context, since we conduct a field experiment, our drivers also earn piece rate in addition to any prize money. Furthermore, each driver finds out about the outcome at the end of the contest day, making it an incomplete information, simultaneous move team contest. These features differ from the settings in Fu, Lu and Pan (2015).

To determine our pairwise team matching, we sort the 250 teams decreasingly by the sum of team members' individual number of trips in the two weeks prior to the announcement of the contest. From the groupings of most to least productive teams, two adjacent teams are paired for each contest, independent of their team formation strategy. This matching process ensures that each pair of teams in each contest is as similar as possible, preserving the symmetry assumption from the theoretical model. We randomly assign each team-pair into one of three prize allocation conditions with equal probability.

Table 1 illustrates the prize structure in each of the individual, team, and hybrid prize allocation conditions. Under the <u>individual prize</u> condition, the driver who wins the contest receives a 30 CNY prize, regardless of team performance. Under the <u>team prize</u> condition, each driver in a team that wins a majority (4 or more) of its contests receives a 30 CNY prize. Under the <u>hybrid prize</u> condition, drivers receive both individual and team prizes of 15 CNY each. The prizes are set such

Table 1—Prize Structure. This table indicates the prize that drivers get if they win the individual contests (individual wins), if their team wins a majority of the contests (team wins), or both. The prize is calculated for each contest based on the number of trips a matched pair of drivers make on that day.

Prize Structure	Individual Wins	Team Wins
Individual-Prize Treatment	30	-
Team-Prize Treatment	-	30
Hybrid-Prize Treatment	15	15

that the expected reward per driver remains the same across treatments, which is 15 CNY under the symmetry assumption. The allocation rules are explained in the newsletter released to the drivers prior to the contest (see Appendix C, for the full-text translation of the newsletters). Finally, each contest consists of seven component battles, where the drivers compete on the number of trips they finish in one day of competition.

Table 2—Summary Statistics and Randomization Check

	No Contact	Control	Hometown Similarity	Age Similarity	Productivity Similarity	Productivity Diversity	Random	P-value
Daily Rev.	223.8	269.0	264.4	266.6	255.9	265.9	260.9	0.858
(CNY)	(109.7)	(116.0)	(117.2)	(111.8)	(107.8)	(116.1)	(105.2)	
Local	0.307	0.24	0.231	0.254	0.300	0.254	0.231	1.000
	(0.461)	(0.428)	(0.422)	(0.436)	(0.459)	(0.436)	(0.422)	
Age	36.866	35.16	35.483	34.786	35.057	34.609	35.506	0.690
	(7.775)	(7.637)	(7.273)	(7.277)	(7.620)	(7.201)	(7.462)	
DiDi Age	0.806	0.869	0.868	0.879	0.868	0.850	0.849	0.916
	(0.560)	(0.559)	(0.587)	(0.582)	(0.565)	(0.595)	(0.557)	
Male	0.976	0.983	0.983	0.963	0.991	0.977	0.971	1.000
	(0.152)	(0.130)	(0.130)	(0.189)	(0.092)	(0.150)	(0.167)	
# Drivers	4,397	350	350	350	350	350	350	

Note: Standard deviations appear in parentheses. P-values test the equality between the control and the pooled treatment groups. No contact is not part of the control or treatment conditions.

Table 2 provides the summary statistics by experimental condition and reports randomization checks. Our drivers are on average 35 years old, having been driving for DiDi for 10 months, predominantly male (98%) and a quarter of them are local (from Dongguan). Tests of equality between the treatment and control conditions indicate that our randomization works ($p > 0.10$). Note that the no-contact condition is not part of the experiment.

The contest was implemented between August 13 - 21, 2017, with one day off between every two contest days. Before each contest day, we reset the contest and repeat it five times with the same pairing of teams. The contest results are calculated at the end of each contest day and communicated to each driver on the following day. In addition to the contest days, we also obtain data two weeks prior to and four weeks after the contest.

II. Hypotheses

To motivate our hypotheses, we set up a simple theoretical framework of team contests, with and without team identity, and characterize the solutions under each treatment in the Appendix E. Based on our theoretical analysis, we formulate the following hypotheses.

Hypothesis 1 (Contest Effect). Drivers in an team contest will exert greater effort than those in the control condition.

Hypothesis 1 is based on Observation 1 in Appendix E. Without team identity, the contest prizes provide the additional monetary incentives for drivers to increase their effort.

We now consider the effects of team identity. Eckel and Grossman (2005) demonstrate that inter-team competition is among the strongest methods that induce team identity in the laboratory. According to Tajfel and Turner (1979), an important part of the social identification process is social comparison. Once we put drivers into teams and they have identified with that team, they then tend to compare their team with the rival team, and maintain their self-esteem by comparing favorably with the rival team. Based on this theory, we use a simple reduced-form method to incorporate team identity into the contest framework, which enable us to derive Observations 3 and 4.

From many laboratory experiments, we learn that ingroup communication increases the strength of group identity, which leads to more contributions in a public goods game (Eckel and Grossman, 2005), and greater effort in coordination games (Chen and Chen, 2011). Observation 3 in Appendix E predicts that an increase in the strength of team identity leads to greater effort, which forms the basis for the following hypothesis.

Hypothesis 2 (Communication). Drivers in responsive teams will exert greater effort than those in non-responsive teams.

Based on prior field experiments on team competition (Ai et al., 2016), we expect that teams based on homophily will perform better. In our experiment, teams based on hometown similarity comprise drivers from the same (or a nearby) province, whereas those based on age similarity share similar socialization experiences. Both are meaningful and salient identities. Again, based on Observation 3 in Appendix E, we expect members of these teams to exert greater effort than those from teams with weaker identities, such as randomly formed teams.

Hypothesis 3 (Team Composition). Teams based on homophily are more productive than randomly formed teams.

Lastly, Observation 2 in Appendix E indicates that, if drivers only care about monetary prizes, the individual prize treatment should induce greater effort than either of the other two prize conditions. However, when team identity is incorporated into the contest framework, the effort ranking might be different. Specifically, as both the individual and, to a lesser extent, hybrid prize rules prime the importance of the individual, whereas the team prize rule primes the importance of the team, we expect that drivers will have a stronger team identity under the team prize rule. Observation 4 postulates that, under the team prize rule, a sufficiently strong team identity can lead to greater effort compared to that under either of the other rules. Based on Observation 4, we formulate the following hypothesis.

Hypothesis 4 (Prize Structure). When team identity is sufficiently strong, drivers under the team prize rule will exert greater effort compared to those under either the individual or hybrid prize rule.

III. Results

In this section, we present the results from our field experiment. We first examine the effect of our contest on driver working hours and revenue. We then examine the impact of team formation on team performance, followed by the effects of the prize allocation conditions. Finally, we end with a discussion of the effects of leadership experience.

We first investigate the average treatment effect, i.e., the effects of team contest on the number of hours worked, number of completed trips, and driver revenue. Figure 3 presents our results for driver daily revenue before, during, and after the contest period by experimental condition.[6] The top panel presents the comparison across three conditions: drivers who were never contacted (*no contact*, light dashed line), those who expressed interest but were not assigned to a team (*control*, black dashed line), and those who expressed interest and were assigned to a team (*treatment*, solid green line). The bottom panel further breaks the treated drivers into those in responsive (solid orange line) versus non-responsive (blue dashed line) teams.

We refer to the five days of our team contest as *contest days* and the 14 days prior to (post) the contest as the *pre-* (*post-*) *contest* periods. Finally, to investigate whether our effects last more than two weeks, we create the *4-week post-contest* period. Our choice of windows ensures that we always compare the same day of the week pre-contest, contest, and post-contest. During our experiment, we record daily data on each driver including the number of hours worked,

[6]The figure looks almost identical if we replace daily revenue with the number of hours worked or the number of completed trips.

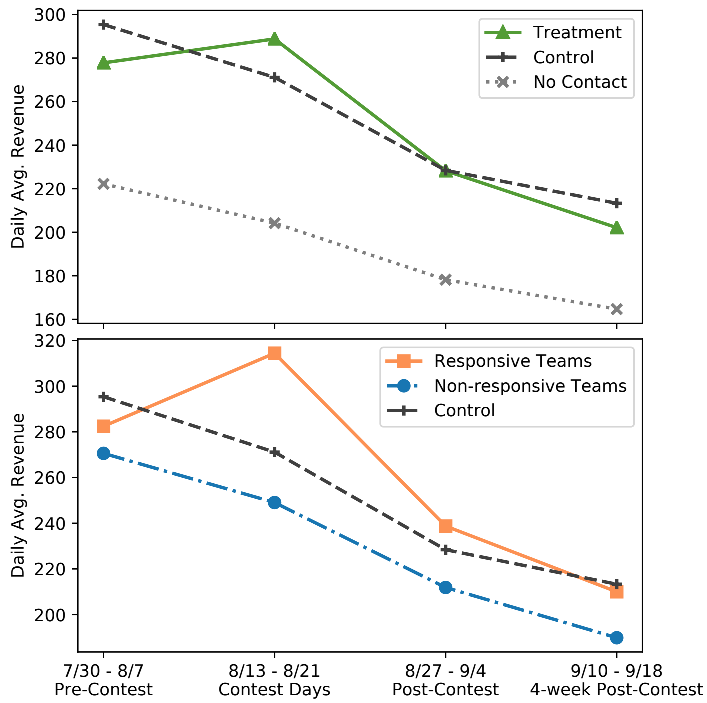

Figure 3. Driver Daily Revenue Before, During, and After the Contest. Contest Days refer to August 13, 15, 17, 19, and 21, the dates on which the contests were conducted. We shift the dates by -14, +14, and +28 days to obtain the Pre-Contest, Post-Contest, and 4-week Post-Contest periods. Note that driver revenue is calculated only on the 5 days in each period accordingly. In the upper panel, drivers in the No Contact group are those who meet our criteria but are not randomly selected to receive an invitation to participate in our experiment. Drivers in the Treatment groups are those who sign up for the experiment and are assigned to a team. Drivers in the Control group are those who sign up for the experiment but are not assigned to a team nor do they participate in the contest. In the lower panel, we break drivers in the Treatment group into Responsive versus Non-responsive teams based on whether the team captain submits the pre-contest survey.

the number of completed trips, and the revenue generated. On the DiDi platform, drivers receive 81% of the revenue they generate and give the remaining 19% to the platform.[7]

Returning to Figure 3, we see from the upper panel that those who sign up to join a team, regardless of whether they are assigned to a treatment or control condition, generate higher revenue than those who are never contacted (grey dashed line). Figure 3 also shows that both the control group and the no-contact group exhibit a similar decreasing trend over the eight-week time period of our experiment, a pattern similar to the platform's typical attrition rate.[8] Our results in the bottom panel of Figure 3 show that drivers assigned to a responsive team demonstrate a much larger increase in revenue during the contest period compared to those assigned to a non-responsive team.

Table 3—Average and Heterogeneous Treatment Effects on Working Hours. Difference-in-differences linear panel regressions. We compare each of the three time periods with the pre-contest period.

	Dependent variable: Δ of Daily Working Hours					
	Average Treatment Effects			Heterogeneous Treatment Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Treated	0.772^{***}	0.379^{*}	0.134			
	(0.192)	(0.197)	(0.221)			
	[0.001]	[0.125]	[0.7]			
Responsive				1.205^{***}	0.484^{**}	0.188
				(0.207)	(0.205)	(0.230)
				[0.001]	[0.057]	[0.623]
Non-responsive				0.0996	0.217	0.0509
				(0.207)	(0.226)	(0.248)
				[0.71]	[0.606]	[0.838]
Constant	-0.521^{***}	-1.579^{***}	-1.225^{***}	-0.521^{***}	-1.579^{***}	-1.225^{***}
	(0.162)	(0.180)	(0.203)	(0.162)	(0.180)	(0.203)
# Drivers	2,100	2,100	2,100	2,100	2,100	2,100
Observations (#Drivers×#Days)	10,500	10,500	10,500	10,500	10,500	10,500
H_0: Responsive = Non-responsive				$p < 0.001$	$p = 0.1148$	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

[7]In the subsequent analysis, we focus on revenue comparisons, with the understanding that alternative comparisons using driver income or platform profit would reach similar conclusions.

[8]High attrition rate also appears to be a problem with Uber drivers (Scheiber, 2017).

Table 4—Average and Heterogeneous Treatment Effects on Daily Revenue. Difference-in-differences linear regressions. We compare each of the three time periods with the pre-contest period.

	Dependent variable: Δ of Daily Revenue (CNY)					
	Average Treatment Effects			Heterogeneous Treatment Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Treated	35.24^{***}	17.36^{*}	6.369			
	(9.319)	(9.679)	(10.09)			
	[0.001]	[0.166]	[0.68]			
Responsive				56.21^{***}	23.25^{**}	9.607
				(9.999)	(10.12)	(10.55)
				[0.001]	[0.066]	[0.654]
Non-responsive				2.706	8.237	1.348
				(10.14)	(10.88)	(11.23)
				[0.889]	[0.675]	[0.905]
Constant	-24.24^{***}	-66.96^{***}	-82.06^{***}	-24.24^{***}	-66.96^{***}	-82.06^{***}
	(7.892)	(8.844)	(9.192)	(7.892)	(8.844)	(9.193)
# Drivers	2,100	2,100	2,100	2,100	2,100	2,100
Observations (#Drivers×#Days)	10,500	10,500	10,500	10,500	10,500	10,500
H_0: Responsive = Non-responsive				$p < 0.001$	$p = 0.063$	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

To quantify the average and heterogeneous treatment effects on outcome, Y, respectively, we construct the following difference-in-differences models:

$$\Delta \mathrm{Y}_{i,t} = \beta_0 + \beta_1 * \text{Treated} + \epsilon_{i,t}, \tag{1}$$

$$\Delta \mathrm{Y}_{i,t} = \beta_0 + \beta_1 * \text{Responsive} + \beta_2 * \text{Non-responsive} + \epsilon_{i,t}, \tag{2}$$

where $\Delta \mathrm{Y}_{i,t}$ represents the outcome change of the t-th day in the current period compared to the t-th day in the pre-contest period. We report the results of these models in Tables 3 (number of hours worked) and 4 (daily revenue) in the main text, and Table F3 (daily number of trips completed) in Appendix F. Each table includes both the average (specifications 1-3, eq. 1) and heterogeneous treatment effects (4-6, eq. 2).

Table 3 reports the treatment effect on the number of hours worked. Pooling drivers across all treatment and control conditions, we find that treated drivers work longer hours (additional 0.77 hours, or 46 minutes, per day for all treated drivers). In Table 4, we find that the daily revenue for treated drivers increases by 35 CNY (12%) during the contest period compared to those in the control condition.[9] Furthermore, this effect persists during the two-week post-contest period, albeit with half of the effect size and marginally significant.

In these and subsequent analyses, we report the false discovery rate adjusted q-values in square brackets to correct for multiple hypothesis testing (Benjamini and Hochberg, 1995) using the Stata code provided by Anderson (2008). We follow the convention of using 5% (respectively 10%) cutoff for p-values (respectively q-values) to claim statistical significance (Efron, 2010). We summarize our first result below.

Result 1 (Average Treatment Effect: Team Contest). During the contest period, drivers in teams work 46 minutes longer per day and generate 12% higher revenue compared to those in the control condition.

By Result 1, we reject the null in favor of Hypothesis 1 that drivers in a team contest will exert greater effort than those in the control condition. We find that the effect size is both statistically and economically significant.

Separating our results by team responsiveness (specifications 4-6) in Table 3, we find that the increased hours worked for those in responsive teams is 1.2 hours ($p < 0.001$), whereas the treatment effect for non-responsive teams is not significantly different from zero. Similarly, in Table 4, we find that the increased daily revenue for those in a responsive team is 56 CNY (19%, $p < 0.001$), whereas the treatment effect for non-responsive teams is again not significantly different from zero. In both tables, the treatment effect for those in responsive teams persists during the two-week post-contest period (hours worked: 0.484, $p < 0.05$; revenue: 23 CNY, $p < 0.05$). Comparing the two types of teams, we find that drivers in responsive teams work significantly longer and generate significantly higher revenue than those in non-responsive teams ($p < 0.001$ for both outcome variables).

Result 2 (Heterogeneous Treatment Effect: Communication). During the contest period, drivers in *responsive* teams work 1.2 hours longer per day and generate 19% higher revenue compared to those in the control condition. Furthermore, they continue to work longer hours and earn higher revenue during the two weeks after the contest ends.

By Result 2, we reject the null in favor of Hypothesis 2 that drivers in responsive teams will exert greater effort than those in non-responsive teams. In our experiment, treated drivers' increase in revenue primarily comes from longer hours worked rather than faster driving or location preferences. In an empirical study, Cook et al. (2018) find that the gender earnings gap among Uber drivers can be explained by three factors: driving experience, location preference and speed. Our study confirms the effect of driving experience, as the variable, DiDi

[9] In this and subsequent comparisons, we use the average daily revenue of drivers in the control group in the five days corresponding to contest days during the two weeks before the contest (295.37 CNY) as the baseline.

Age, has a statistically significant and economically substantial effect on revenue (Table F2 in Appendix F). While we do not observe location preferences in our data, speed is often out of the control of our drivers due to traffic congestion.

Lastly, the average and heterogeneous treatment effects are robust to demographic controls (Table F2 in Appendix F). Adding demographic controls reveals a significant association between work experience (DiDi Age) and revenue (17 CNY, $p < 0.01$) both during and two weeks after the contest.

As the treatment effect could be driven by a combination of team identity and competition, we provide some evidence that it is not entirely driven by competition for monetary prizes. First, treated drivers continue to generate higher revenue in the two weeks post contest, absent of any monetary prize or formal competition. Furthermore, this spillover effect is larger for responsive teams, consistent with evidence from laboratory experiments that in-group communication strengthens team identity (Chen and Li, 2009). Lastly, our post-contest survey (Appendix D) indicates that over 88% of the drivers either like or very much like the team contest (Question 2), citing team belonging (66%), making friends (70%), a sense of honor from winning (61%), and monetary incentives (68%) as the top benefits. When asked whether they prefer a temporary or a long-lasting team (Question 12), 79.2% of the drivers choose "a long-lasting team, so team members can keep in touch after the contest." The long-lasting bonds among team members are further corroborated in a post-experiment interview with 14 drivers conducted by DiDi staff and the first author, where drivers mention finding friends from their hometown as one of the top benefits of the contest, and that teammates continue to socialize after the contest.

In our experiment, we are also interested in whether different ways of forming teams have different effects on our results. Table 5 presents our results using team formation strategy as the independent variables in specifications (1) - (3). For robustness check, we use alternative measures of team diversity as the independent variables in specifications (4) - (6). More specifically, we measure driver diversity based on the standard deviation in driver age, productivity, and DiDi age within a team; we measure hometown diversity using the average distance (in kilometers) between any two drivers within the same team. Our results in Table 5 show that, irrespective of our independent variables, team formation has no significant effect on driver revenue during the contest. Interestingly though, we find that teams based on age similarity exhibit significantly higher revenue during the two-week period after the contest, earning 33 CNY more on average compared with drivers in randomly-formed teams (specification 2, $p < 0.05$, $q = 0.122$). This observation is confirmed by the negative correlation between standard deviation of age and team productivity (specification 5, $p < 0.05$). We summarize our analysis below.

Result 3 (Team Formation). Teams based on age similarity are more productive than randomly formed teams two weeks post contest.

By Result 3, we reject the null in favor of Hypothesis 3 that teams based on homophily are more productive than those from randomly formed teams. In this case, only teams based on age similarity exhibit a significant effect.

Next, we examine the effects of prize structure on driver revenue. Table 6 presents pairwise comparisons between the three prize structures under each team formation algorithm. The outcome variable being compared is the difference in driver revenue, which is the same as the Δ of Daily Revenue (CNY) in Tables 4 and 5. Based on the multiplicity-adjusted p-values (List, Shaikh and Xu, 2019), variations in the prize structure have a significant effect on changes in revenue only for teams formed based on hometown similarity. Specifically, team prize leads to a greater increase in daily revenue than either the hybrid (68.4 CNY, multiplicity-adjusted $p = 0.0003$) or individual prizes (41.8 CNY, multiplicity-adjusted $p = 0.062$).

Result 4 (Prize Structure). For teams based on hometown similarity, team prize leads to a 23% (14%) increase in daily revenue than the hybrid (individual) prize.

By Result 4, we reject the null in favor of Hypothesis 4. This result is largely consistent with findings from laboratory experiments that between-group competition can enhance cooperation and coordination within a group (Sheremeta, 2018), although we are not aware of any laboratory implementation of the exact payoff structure as that in our field experiment.

Finally, we are interested in the productivity of those who volunteer to be captains in our study. Our results in Table F4 in Appendix F, show that those who had completed more trips in the two weeks prior to the announcement of the contest, as well as those who had been working for DiDi for a longer period of time (i.e., their DiDi age) were significantly more likely to volunteer to lead. In 141 teams with two or more drivers who express interests in being a captain, only one in each team is randomly appointed as the team captain. OLS regressions show that among our base of 298 volunteers, those who are randomly chosen to be captains are more active than those who are randomized out, earning 34 CNY more per day on average ($p < 0.10$) during contest days (Table F5 in the Appendix F), although this result is only marginally significant.

Related to the treatment effects on captains, we address the question of whether team members benefit from being in a team as much as the captains who are both more productive and more experienced before teams are formed. To answer this question, we re-evaluate the treatment effects on team members by excluding the captains from the analysis (Table F6 in Appendix F). We find that both the average and the heterogeneous treatment effects remain as significant and large as those with captains included. This indicates that the contest benefits less productive drivers as much as the more productive ones.

IV. Discussion

Our study uses a natural field experiment at a ride-sharing platform in China to understand how team formation and other factors impact team responsiveness, driver working hours and productivity. Applying social identity theory to the ride-sharing context, we use different team formation strategies to place drivers in teams and compare our treatment and control groups on their revenue earned during and after a contest. Our results show that, compared to those in the control condition, treated drivers work 0.8 longer hours, complete 2.4 more trips, and earn 12% higher revenue per day during the contest, with a much larger effect (1.2 longer hours, 3.5 more trips, 19% more revenue) for responsive teams who communicate more with each other prior to the contest. Furthermore, we find that drivers in responsive teams as well as those in teams comprised of drivers with similar age continue to work longer hours and generate higher revenue during the two weeks after the contest, absent of any cash prize or formal competition.

We conclude with a few observations on the impact of our experiment on the actual organization of the platform and the gig economy in general. Encouraged by the results of our experiment, DiDi shipped two of our team-formation algorithms (hometown similarity and age similarity) into production within their platform. In 2018 alone, DiDi conducted 1,548 team contests across 180 cities in China, involving over two million drivers. These contests, typically one-week long, helped the platform to meet high demands from tourists during national holidays, and increased both driver income as well as their retention (Ye et al., 2019). While our experiment examines the effect of team formation on one platform, our results indicate that team identity shows great promise as a design tool that can be leveraged to increase worker productivity and engagement in the modern gig economy. Future research could use our study as a foundation for exploring the full potential of social identity theory, examining the impact of longer contests and more persistent teams.

TABLE 5—SIMILARITY AND DIVERSITY ON DRIVER REVENUE: DID REGRESSIONS ON TREATED DRIVERS. THE DEPENDENT VARIABLE IS THE DIFFERENCE IN DRIVER REVENUE (COMPARED WITH THE PRE-CONTEST TIME WINDOW). FOR (1-3), THE OMITTED CATEGORY IS THE RANDOM TREATMENT.

	Dependent variable: Δ Daily Revenue (CNY)					
	By Treatment Group			By Diversity Metrics		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Age Similarity	0.933	33.19^{**}	9.806			
	(16.91)	(12.70)	(11.05)			
	[0.957]	[0.122]	[0.474]			
Hometown Similarity	5.838	20.70	17.12			
	(18.35)	(13.16)	(13.62)			
	[0.820]	[0.468]	[0.474]			
Productivity Similarity	-14.65	21.47^{*}	13.85			
	(17.15)	(12.04)	(12.67)			
	[0.474]	[0.463]	[0.474]			
Productivity Diversity	-17.50	17.50	11.33			
	(15.62)	(12.25)	(13.09)			
	[0.474]	[0.468]	[0.474]			
Age Std. Dev.				-0.417	-3.357^{**}	-0.123
				(1.647)	(1.346)	(1.279)
Avg. Hometown Distance				0.0297	-0.00706	-0.0196
				(0.0242)	(0.0227)	(0.0203)
Productivity Std. Dev.				0.0953	-0.0347	-0.00401
				(0.122)	(0.0882)	(0.0961)
DiDi Age Std. Dev.				-0.0646	-0.0370	-0.0852
				(0.0914)	(0.0852)	(0.0799)
Constant	16.07	-68.17^{***}	-86.12^{***}	4.701	-15.89	-48.15^{**}
	(13.69)	(9.377)	(8.566)	(29.68)	(21.04)	(22.52)
# Drivers	1,750	1,750	1,750	1,750	1,750	1,750
Observations	8,750	8,750	8,750	8,750	8,750	8,750

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

Table 6—Effects of Prize Structure on Daily Revenue. For each team formation algorithm, we conduct pairwise comparisons between the three prize structures. The outcome variable is the difference in driver revenue, which is the same as the Δ Daily Revenue (CNY) in Tables 4 and 5. Based on multiplicity-adjusted p-values (List, Shaikh and Xu, 2019), the prize structure has a significant effect on revenue only under the Hometown Similarity algorithm.

Recommendation Algorithm	Prize A	Prize B	Δ outcome (A - B)	Bootstrapped p-value (Unadjusted)	Multiplicity-Adjusted p-value
Age Similarity	Team	Hybrid	-26.73	0.075	0.510
	Team	Individual	-28.44	0.072	0.550
	Hybrid	Individual	-1.70	0.925	0.995
Hometown Similarity	Team	Hybrid	+68.42	0.0003	0.0003
	Team	Individual	+41.79	0.005	0.062
	Hybrid	Individual	-26.63	0.073	0.537
Productivity Diversity	Team	Hybrid	+15.43	0.358	0.897
	Team	Individual	-1.08	0.952	0.952
	Hybrid	Individual	-16.52	0.363	0.88
Productivity Similarity	Team	Hybrid	-6.24	0.686	0.989
	Team	Individual	-19.79	0.167	0.763
	Hybrid	Individual	-13.55	0.347	0.925
Random	Team	Hybrid	+23.20	0.150	0.746
	Team	Individual	+17.61	0.273	0.896
	Hybrid	Individual	-5.59	0.762	0.99

Appendix A. Text Messages

1) Contest announcement text message, sent to all eligible drivers:

   DiDi driver team contest is about to start, [an opportunity] to get to know new friends and win more than 1,000 yuan in bonus. Click here to register and get a position in the participation quota!

2) Team formation text messages

   a) Message sent to team captains:

      You have been selected as a DiDi team contest team captain. Please make sure to click this link to fill out information about your team members by 10:00 PM on August 12th. If you do not submit, we will deem your team formation as unsuccessful, and you cannot receive the team captain bonus. Contact information for your team members [is as follows]: {Name1, cell phone number}, {Name2, cell phone number}, {Name3, cell phone number}, {Name4, cell phone number}, {Name5, cell phone number}, {Name6, cell phone number}.

   b) Message sent to team members:

      Congratulations! You have successfully registered for the DiDi team contest. Please cooperate with your team captain to provide information [such as] the last digits of your license plate number to complete the team formation [information] submission. If your captain has not contacted you yet, please take the initiative to contact and cooperate with your team captain. [The] team captain's contact information is: {Name, cell phone number}. A successful team formation facilitates winning the bonus. If you encounter any problems, please click this link to let us know. (Please ignore this message if you have already completed the process.)

3) Text message sent to drivers in the control group:

   Thank you for registering for the DiDi driver team contest! Many people registered for this popular activity. According to the activity rules, you were not selected to participate in this activity based on the lottery [selection]. Do not be discouraged. You will have priority to participate in our next activity.

4) Announcement of contest rules:

   DiDi driver team contest will formally start on Sunday! Click this link to check the activity rules. Contest is on for odd-numbered days. Teams compete and win cash bonus at ease!

5) Announcement of daily updates:

> The first day results of driver team contest is ready to be revealed! Did you beat your opponent? Did you team win? Click the newsletter to find out.

Appendix B. Pre-Contest Survey

Captains, please contact each team member, confirm and fill out the [requested] information. We suggest that team captains contact team members by phone, and then set up a WeChat group to facilitate information acquisition from and communications with the team members. Move on to victory!

Team formation confirmation deadline: August 18th.

Please fill out the information completely and correctly. Captains who successfully confirm the team formation will get an extra bonus of 50 yuan after the contest.

1) As a captain, please contact every team member and obtain the last 3 digits of their license plate numbers. [six blank lines, one for each team member]
2) Together with your team members, please come up with a powerful name for your team: _____
3) To get to know each other better and to increase the likelihood of winning, please confirm with your team members: Where is the farthest hometown from Dongguan among your team members?
4) To get to know each other better and to increase the likelihood of winning, please confirm with your team members: What is the maximum age in your team?
5) If you encounter any problem in the team formation process, please explain here. If you encounter issues with any of your team members, please explain and leave the team member's phone number here.

Appendix C. Newsletter Implementing Treatments

The newsletter for each treatment is identical except for the section, "Prize Allocation Rules." Therefore, we first present the newsletter for the Individual Prize treatment here in full.

Contest Rules:

- Two teams form a match.
- Rules for winning:

- At the individual contest level, the driver who completes more trips during the contest day wins the individual contest. If the number of trips tie, the driver who generates more revenue wins.
- At the team contest level, the team whose members win a majority of their individual contests during the contest day wins the team contest.

- Competitor matching:
  - Individual matching: a team member is ranked within a team based on the number of completed trips in the previous game day, and compete with the member of the same rank in the rival team.
  - Team matching: The system pairs teams with comparable productivity based on the total number of completed trips in the two weeks prior to the contest. The pairing persists [througout the contest].
- Contest days: August 13, 15, 17, 19, 21; 5 days in total.

PRIZE ALLOCATION RULES:

[Individual Prize treatment]
The team member who wins the individual contest receives 30 CNY, with a maximum total of 150 CNY for the five contest days.

ADDITIONAL RULES:

1) Carpool trips and pre-scheduled trips also count towards the contest.
2) Prize amounts, include the team captain's reward, will be posted within five business days.
3) If the system detects cheating behavior, the respective team member's completed trips will be treated as zero.
4) Before the start of the contest, if a team member can no longer participate due to unexpected circumstances, the system will replace this team member with another driver of similar characteristics after confirmation from both the captain and the respective team member.
5) After the contest begins, team members can no longer be replaced. For team members who can no longer participate due to an unexpected circumstance, the captain should provide timely feedback and communications. Upon confirmation, the respective team member's number of completed trips in future contest days will be set as zero, and any [of the team member's eligible] prize will be divided equally among the remaining teammates.

This is the end of the newsletter for the Individual Prize treatment.
The newsletter for each treatment is identical except for the section, "Prize Allocation Rules." Therefore, we only present this section of the newsletter for the Team Prize and Hybrid Prize treatments.

Prize Allocation Rules:

[Team Prize treatment]
The team that wins a team contest receives 210 CNY, with a maximum total of 1050 CNY for the five contest days (with the prize divided equally among the team members).

Prize Allocation Rules:

[Hybrid Prize treatment]
The team member who wins an individual contest receives 15 CNY, whereas the team that wins the team contest receives 105 CNY (with the prize divided equally among the team members). The individual and team prizes are calculated separately.

Appendix D. Post-Contest Survey and Responses

The survey response rate is 25.3% (443 out of 1,750). Table F7 presents factors which affect the likelihood of responding to the post-contest survey. The number and percent of drivers choosing a certain choice are indicated in the brackets.

1) Did you participate in the team contest in Dongguan from August 13th to August 21st?
    a) Yes. (99.3%)
    b) I am not sure. (0.7%)
2) To what extent do you like this team contest? Please rate on a scale from 1 (dislike extremely) to 5 (like very much).
    (1) Dislike extremely (4, 0.9%)
    (2) Dislike (6, 1.4%)
    (3) Neutral (42, 9.5%)
    (4) Like (60, 13.5%)
    (5) Like very much (331, 74.7%)
3) Why do you like this team contest? Please select all that apply. (Limited to the 508 drivers who choose 4 or 5 in Question 2.
    a) I have a sense of team belonging. (252, 56.9%)
    b) The contest is interesting and thrilling. (176, 39.7%)
    c) I get to make more friends. (260, 58.7%)
    d) Winning the contest gives me a sense of honor. (233, 52.6%)
    e) I get the monetary bonus. (255, 57.6%)

    f) Other reasons. Please specify. (17, 3.8%)

4) Why do you dislike this team contest? Please select all that apply. (Limited to the 69 drivers who choose 1, 2, or 3 in Question 2.)
    a) The team members are not collaborative or united enough. (23, 44.2%)
    b) The team is not active enough to justify its existence. (33, 63.5%)
    c) The captain did not have good leadership or management skills. (27, 51.9%)
    d) The contest rules are too complicated for me to understand. (4, 7.7%)
    e) The contest rules are unfair. (8, 15.4%)
    f) The monetary bonus is not large enough to attract me. (26, 50.0%)
    g) Other reasons. Please specify. (8, 15.4%)
5) As a team member/captain , what did you get from this team contest? Select all that apply.
    a) I made more friends. (308, 69.5%)
    b) I improved my leadership skills. (only for captains, 63, 70.0% among captains)
    c) I improved my communication skills. (208, 47.0%)
    d) I improved my collaboration skills with other drivers. (256, 57.8%)
    e) I became more experienced and skillful about taking the DiDi orders. (218, 49.2%)
    f) I got consolation from my teammates when I was down. (144, 32.5%)
    g) Other reasons. Please specify. (24, 5.4%)
6) Which of the rules in this contest do you like? Please select all that apply.
    a) There was one day off between two contest days. (209, 47.2%)
    b) Scores were announced immediately after each contest day. (232, 52.4%)
    c) There were both driver-level and team-level competitions. (297, 67.0%)
    d) The team could discuss and decide the lineup together. (66, 31.6% among the 209 applicable participants)
    e) The lineup changed between contest days. (144, 32.5%)
    f) Other reasons. Please specify. (5, 1.1%)
    g) None. (19, 4.3%)
7) How did your team get along in this contest? Please select all that apply.

a) Although each team member was different, we got along well. (228, 51.5%)
b) Our team shared common characteristics and had common conversation topics. (196, 44.2%)
c) Everyone contributed for our team's honor during the contest. (301, 67.9%)
d) Inactive team members influenced others' enthusiasm [for the contest]. (149, 33.6%)

8) What would you choose if you could participate in the contest again?
a) I will choose to be a team member. (273, 61.6%)
b) I will choose to be a team captain. (123, 27.8%)
c) I have not made up my mind. (47, 10.6%)

9) Why did you prefer NOT to be a team captain? (Applicable only to drivers who choose team member in Question 8.)
a) I do not want to initiate communications with strangers. (7, 4.1%)
b) I do not know how to lead a team. (73, 42.9%)
c) The extra bonus for team captains is not enough. (18, 10.6%)
d) Team captains take up too much extra work. (40, 23.5%)
e) I am inexperienced with team management and I need more practice. (99, 58.2%)
f) Other reasons. Please specify. (11, 6.5%)

10) What do you think a team captain should do?
a) Lead by example. (314, 70.9%)
b) Be positive and energetic. (287, 64.8%)
c) Help teammates to be more active. (315, 71.1%)
d) Help the team win the contest. (283, 63.9%)
e) Provide feedback and suggestions to the DiDi platform on behalf of the team. (251, 56.7%)
f) Other reasons. Please specify. (10, 2.3%)

11) What is your preferred way to join a team?
a) I prefer to join the team WeChat group and communicate with other teammates online. (55, 12.4%)
b) I prefer to call others and ask to join their team. (245, 55.3%)

c) I prefer to wait for others to call me and invite me to join their team. (140, 31.6%)
d) Other reasons. Please specify. (3, 0.7%)

12) Which of the following teams would you prefer?
    a) Temporary teams, so I can join a different team in each contest. (92, 20.8%)
    b) A long-lasting team, so team members can keep in touch after the contest. (351, 79.2%)

13) Which of the following team structure would you prefer?
    a) I do not care if there is a team captain or not, as long as all team members can work together. (320, 72.2%)
    b) I prefer to have a team captain and team members, with each member taking on a different role. (121, 27.3%)
    c) Other reasons. Please specify. (2, 0.5%)

14) Which of the following prize structure do you think is the best for team contests (given that the total amount of the prize is fixed)?
    a) Team prize. The prize should be allocated equally or proportional to the contributions of the team members. (168, 37.9%)
    b) Get rid of the team prize. The prize should emphasize individual contributions. (251, 56.7%)
    c) In addition to the team prize, there should be some prize to reward individuals who contribute a lot. (24, 5.4%)

15) Do you have other suggestions for team activities?

Appendix E. A Theoretical Framework

In this section, we set up a theoretical framework to motivate our hypotheses. The reader can find similar characterizations of symmetric pure strategy Bayesian Nash equilibrium strategy in simultaneous contests under incomplete information in textbooks, such as Vojnović (2016) Chapter 2. Using our field experiment context, we sketch a simple version here for completeness.

For player i, let $x_i \in [0, \bar{x}]$ denote his effort, which can be approximated by the number of hours he drives each day. The platform imposes a maximum of ten hours of driving per day, which justifies the upper bound, $\bar{x}$. Let $c_i \in (0, \bar{c}]$ be his marginal cost of effort, and $w > 0$ be the piece rate for all players. In the control condition where a driver earns piece rate, a risk neutral driver will maximize his expected income, $wx_i - c_i x_i$, and will max out his disposable hours, $x_i \equiv x_0(c_i) > 0$ if $w > c_i$, and $x_i = 0$ otherwise. We omit the case with risk aversion and other plausible assumptions on preferences, as they have been dealt with in the theoretical (Fibich, Gavious and Sela, 2006) and experimental literature (Noussair and Silver, 2006). We now characterizes players' equilibrium effort function under each of the three contest rules.

<u>Case 1. Individual Prize.</u> Under the individual prize contest rule, a player wins a cash prize, V, if he completes more trips than his match, regardless of whether his team wins or not. For simplicity, we assume that a player working longer hours will complete more trips, which is justified in our data, e.g. by comparing results in Table 3 with those in Table F3. If team identity is not taken into consideration, this reduces to a modified two-player all-pay contest with an extra component of piece rate, wx_i. Player i's objective function is:

$$U_i = \max_{x_i} P(x_i \geq x_j)V + wx_i - c_i x_i. \tag{E1}$$

Again, $x_i \equiv x_0(c_i)$ if $w > c_i$. We now examine the more interesting case of $c_i \in (w, \bar{c}]$. Define $v_i = V/(c_i - w)$, Equation (E1) can be written as:

$$\max_{x_i} P(x_i \geq x_j)v_i - x_i.$$

In a symmetric Bayesian Nash equilibrium, each player employs the same strictly decreasing effort function $x(\cdot) : (w, \bar{c}] \to [0, \bar{x}]$. Let players' marginal costs be independent and identically distributed according to a prior distribution, $F(c)$, with a corresponding density function $f(c)$. The probability that player i wins is $P(x_i \geq x_j) = P(c_j > c_i) = 1 - F(c_i)$. Using the effort function, $x(\cdot)$, the driver's objective function can be written in the following standard form:

$$\max_{x_i} \left[1 - F(x^{-1}(x_i))\right] v_i - x_i.$$

The first-order condition with respect to x_i is:

$$-f(x^{-1}(x_i))\frac{dx^{-1}(x_i)}{dx_i}v_i = 1,$$
$$-f(c_i)\left(\frac{dx_i}{dc_i}\right)^{-1}v_i = 1,$$

which leads to the following differential equation characterizing the solution, $x_I(c_i)$:

$$x_I'(c_i) = -f(c_i)v_i. \tag{E2}$$

If we know the functional form of $F(\cdot)$, we can plug the corresponding density function f into Equation (E2) and solve for the equilibrium effort function, $x_I(\cdot)$, which we demonstrate in the following numerical example.

In this example, we assume $w = 1$, $\bar{c} = 2$, and c_i is uniformly distributed over $(1, 2]$. Equation (E2) becomes:

$$x_I'(c_i) = -f(c_i)v_i = -\frac{V}{c_i - w}.$$

The solution is characterized as $x_I(c_i) = -V\ln(c_i - w) + C$. Since $U_i(c_i = 2) = (w - c_i)x_i = -x_i \geq 0$, we have $x_i(2) = C = 0$. We then obtain the following close-form solution for the effort function:

$$x_I(c_i) = V[-\ln(c_i - w)]. \tag{E3}$$

<u>Case 2. Team Prize.</u> Under the team prize rule, a team wins if and only if a majority of its drivers win their component battles. We denote the probability that i is pivotal as P_i. In this case, player i's utility function is:

$$\max_{x_i}\left[P_iP(x_i \geq x_j) + (1 - P_i)P(i\text{'s team wins})\right]v_i - x_i. \tag{E4}$$

Player i is pivotal when there are exactly three other players winning and three other players losing in his team. As the *ex ante* probability of any player winning is $1/2$, the probability that i is pivotal is $P_i = \binom{6}{3}(\frac{1}{2})^6 = \frac{5}{16}$, and the probability that i's team wins is $\frac{1}{2}$. Therefore, the objective function (E4) can be simplified as:

$$\max_{x_i}\left\{\frac{5}{16}\left[1 - F(x^{-1}(x_i))\right] + \frac{11}{16} \times \frac{1}{2}\right\}v_i - x_i.$$

The corresponding first-order condition is:

$$-\frac{5}{16}f(x^{-1}(x_i))\frac{dx^{-1}(x_i)}{dx_i}v_i = 1,$$

which leads to the following characterization of the solution, $x_T(c_i)$:

$$x'_T(c_i) = -\frac{5}{16} f(c_i) v_i.$$

Using the same numerical example, we obtain the following closed-form solution:

$$x_T(c_i) = \frac{5}{16} V[-\ln(c_i - w)].$$

<u>Case 3. Hybrid Prize.</u> Under the hybrid prize rule, a player wins a cash prize $V/2$, if he exerts more effort than his match, regardless of whether his team wins or not; he gets an additional $V/2$ if his team wins. In this case, player i's objective function is a convex combination of those in the first two cases:
(E5)

$$\max_{x_i} \left[1 - F(x^{-1}(x_i))\right] \frac{v_i}{2} + \left[P_i P(x_i \geq x_j) + (1 - P_i) P(i\text{'s team wins})\right] \frac{v_i}{2} - x_i.$$

Taking the first-order condition and simplifying, we obtain the following characterization of the solution, $x_H(c_i)$:

$$x'_H(c_i) = -\frac{21}{32} f(c_i) v_i.$$

Using the same numerical example, we obtain a corresponding closed-form solution:

$$x_H(c_i) = \frac{21}{32} V[-\ln(c_i - w)].$$

We now summarize the above characterizations, using $t \in \{I, T, H\}$ to represent the individual, team and hybrid prize contest rules, respectively. Based on Equations (E1), (E4) and (E5), we make the following observation.

<u>Observation 1.</u> Under any of the three contest rules, a player exerts greater effort than a corresponding player under the control condition, i.e., $x_t(c_i) \geq x_0(c_i)$, for $t \in \{I, T, H\}$.

It is also straightforward to see that, under any of the three contest rules, a player with a higher marginal cost exerts less effort, i.e., $\partial x_t / \partial c_i < 0$. Furthermore, a larger prize induces higher effort, i.e., $\partial x_t / \partial V > 0$, for $t \in \{I, T, H\}$. We now rank the equilibrium effort under the three rules.

<u>Observation 2.</u> Without any team identity, players under the individual prize rule exert greater effort than a corresponding player under the hybrid prize rule, who in turn, exert greater effort than that under the team prize rule, i.e., $x_I(c_i) \geq x_H(c_i) \geq x_T(c_i)$.

<u>Team Identity.</u> Eckel and Grossman (2005) demonstrate that inter-team contests are among the strongest methods that induce team identity in the laboratory.

According to Tajfel and Turner (1979), an important part of the social identification process is social comparison. Once we put drivers into teams and they have identified with that team, they then tend to compare their team with the rival team, and maintain their self-esteem by comparing favorably with the rival team. Based on this theory, we use a simple reduced-form method to incorporate team identity into the contest framework. Specifically, we use $\alpha_t \geq 1$, $t \in \{I, T, H\}$, to denote the strength of a player's team identity under contest rule t, which translates into how much weight he puts on his team winning the contest. Therefore, the objective functions under each of the three rules are modified as follows.

- Under the individual prize rule, even though a player wins the prize if and only if he exerts higher effort than his rival, his winning contributes to his team's performance. Therefore, he might care more about winning his individual battle. In this case, the objective function (E1) becomes

$$\max_{x_i} \alpha_I P(x_i \geq x_j)V + wx_i - c_i x_i,$$

with the following characterization of the solution, $x_I(c_i, \alpha_I)$:

$$x_I'(c_i, \alpha_I) = -\alpha_I f(c_i) v_i.$$

The same numerical example yields the following solution:

$$x_I(c_i, \alpha_I) = \alpha_I V[-\ln(c_i - w)] \geq x_I(c_i).$$

- Under the team prize rule, a player wins the prize if and only if his team wins a majority of the component battles. A player with a stronger team identity cares more about his team winning. The objective function (E4) becomes:

$$\max_{x_i} \alpha_T \left[P_i P(x_i \geq x_j) + (1 - P_i)P(i\text{'s team wins})\right] v_i - x_i,$$

which leads to the following characterization of the solution, $x_T(c_i, \alpha_T)$:

$$x_T'(c_i, \alpha_T) = -\frac{5}{16}\alpha_T f(c_i) v_i.$$

The same numerical example yields the following solution:

$$x_T(c_i, \alpha_T) = \frac{5}{16}\alpha_T V[-\ln(c_i - w)] \geq x_T(c_i).$$

- Under the hybrid prize rule, with team identity, the objective function (E5) becomes:

$$\max_{x_i} \alpha_H \left\{ \left[1 - F(x^{-1}(x_i))\right] + \left[P_i P(x_i \geq x_j) + (1 - P_i) P(i\text{'s team wins})\right] \right\} \frac{v_i}{2} - x_i.$$

The solution, $x_H(c_i, \alpha_H)$, is characterized by:

$$x'_H(c_i, \alpha_H) = -\frac{21}{32} \alpha_H f(c_i) v_i.$$

The same numerical example yields the following closed-form solution:

$$x_H(c_i, \alpha_H) = \frac{21}{32} \alpha_H V[-\ln(c_i - w)] \geq x_H(c_i).$$

We now summarize our analysis of team identity.

Observation 3. Under any of the three contest rules, an increase in the strength of team identity leads to a higher effort, i.e., $\partial x_t / \partial \alpha_t > 0$, for $t \in \{I, T, H\}$.

Lastly, we discuss on the relative strengths of team identity under the three rules. As both the individual and, to a lesser extent, hybrid prize rules prime the importance of the individual, whereas the team prize rule primes the importance of the team, we expect that players will have a stronger team identity under the team prize rule, i.e., $\alpha_T > \alpha_I$, and $\alpha_T > \alpha_H$. In the numerical example, if $\alpha_T > \frac{16}{5} \alpha_I$, we have $x_T > x_I$. Similarly, when $\alpha_T > \frac{21}{10} \alpha_H$, we have $x_T > x_H$. Therefore, compared to the case without any team identity (Observation 2), the effort ranking with team identity might be different. In particular,

Observation 4. Under the team prize rule, a sufficiently strong team identity can lead to a higher effort compared to that under either of the other rules, i.e., when α_T is sufficiently high, we can have $x_T > x_I$ and $x_T > x_H$.

The observations in this section form the basis for our hypotheses in Section II.

Appendix F. Additional Analyses

This section presents additional data analyses and robustness checks.

In Table F1, specifications (1) and (2) use a Probit regression to examine the treatment effect along the extensive margin, with the likelihood of submitting the survey as the dependent variable. By contrast, specifications (3) and (4) use an OLS regression to examine the treatment effect along the intensive margin, with the number of license plates reported correctly as the dependent variable. The results in Table F1 show that teams based on hometown similarity show the highest level of responsiveness. Quantitatively, these teams are more likely to be responsive than age-similar teams, productivity-similar teams, or randomly-composed teams after controlling for demographics (0.191, $p < 0.05$, $q = 0.141$). Along the intensive margin, however, we do not find any significant difference among the teams whose captains have submitted the survey. One possible reason for this finding may be that the captains decide to submit their surveys only if they have sufficient information. Indeed, more than 75% of the captains who submit the survey get at least five out of six license plate numbers correctly.

Table F1—Treatment Effects on Team Responsiveness. The extensive margin measures whether the team captain submits the questionnaire, reporting the average marginal effects of Probit estimates. The intensive margin measures the number of license plates reported correctly. The omitted group is Productivity Similarity.

	Extensive margin, Probit $Y = P(\text{Responsive})$		Intensive margin, OLS $Y = \#\text{Correct Plate Numbers}$	
	(1)	(2)	(3)	(4)
Age Similarity	0	-0.00323	0.429	0.327
	(0.0952)	(0.0964)	(0.404)	(0.412)
	[1]	[1]		
Hometown Similarity	0.186*	0.191**	0.437	0.403
	(0.0967)	(0.0967)	(0.378)	(0.382)
	[0.141]	[0.141]		
Productivity Diversity	0.0387	0.0304	0.431	0.358
	(0.0954)	(0.0956)	(0.397)	(0.402)
	[1]	[1]		
Random	0.0193	0.00728	0.326	0.265
	(0.0953)	(0.0957)	(0.400)	(0.405)
	[1]	[1]		
Avg. Daily Revenue		0.0337		0.0290
(100 CNY)		(0.0260)		(0.105)
Age		0.00303		-0.0106
		(0.00427)		(0.0175)
DiDi Age (Year)		-0.0119		-0.0240
		(0.0523)		(0.209)
Local		-0.0176		-0.469
		(0.0750)		(0.308)
Male		-0.000324		-0.785
		(0.225)		(0.893)
Constant			4.536***	5.769***
			(0.285)	(1.156)
Observations (# teams)	250	250	152	152
H_0: Hometown Similarity	p=0.0542	p=0.0455		
= Age Similarity	[0.141]	[0.141]		
H_0: Hometown Similarity	p=0.0862	p=0.0586		
= Random	[0.173]	[0.141]		

Note: Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

Table F2—Average and Heterogeneous Treatment Effects on Daily Revenue with Demographic Controls. Difference-in-differences linear panel regressions. We compare each of the three time periods with the pre-contest period.

	Dependent variable: Δ of Daily Revenue (CNY)					
	Average Treatment Effects			Heterogeneous Treatment Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Treated	35.39^{***}	17.62^{*}	6.308			
	(9.306)	(9.637)	(10.04)			
	[0.001]	[0.154]	[0.682]			
Responsive				56.30^{***}	23.37^{**}	9.518
				(9.971)	(10.07)	(10.49)
				[0.001]	[0.063]	[0.634]
Non-responsive				2.900	8.700	1.320
				(10.10)	(10.83)	(11.15)
				[0.871]	[0.634]	[0.906]
Age	0.741	0.801	0.829^{*}	0.636	0.772	0.813^{*}
	(0.518)	(0.529)	(0.465)	(0.505)	(0.521)	(0.465)
DiDi Age (year)	17.16^{***}	18.19^{***}	6.455	17.30^{***}	18.23^{***}	6.476
	(6.606)	(6.709)	(6.252)	(6.539)	(6.711)	(6.243)
Local	14.23^{*}	3.998	24.99^{***}	15.09^{**}	4.234	25.12^{***}
	(7.621)	(7.833)	(7.657)	(7.418)	(7.838)	(7.680)
Male	35.42	27.18	34.82	34.92	27.04	34.74
	(29.03)	(29.82)	(26.08)	(29.44)	(30.10)	(26.09)
Constant	-103.4^{***}	-138.6^{***}	-157.0^{***}	-99.57^{***}	-137.5^{***}	-156.4^{***}
	(36.56)	(36.71)	(30.96)	(36.71)	(36.82)	(31.09)
# Drivers	2,100	2,100	2,100	2,100	2,100	2,100
Observations (#Drivers×#Days)	10,500	10,500	10,500	10,500	10,500	10,500
H_0: Responsive = Non-responsive				$p < 0.001$	$p = 0.0673$	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

Table F3—Average and Heterogeneous Treatment Effects on Daily Number of Trips. Difference-in-differences linear panel regressions. We compare each of the three time periods with the pre-contest period.

	Dependent variable: Δ of Daily Trips					
	Average Treatment Effects			Heterogeneous Treatment Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Treated	2.392***	1.219**	0.462			
	(0.513)	(0.542)	(0.559)			
	[0.001]	[0.057]	[0.461]			
Responsive				3.493***	1.494***	0.574
				(0.555)	(0.567)	(0.584)
				[0.001]	[0.026]	[0.419]
Non-responsive				0.684	0.791	0.289
				(0.560)	(0.617)	(0.627)
				[0.334]	[0.334]	[0.646]
Constant	-2.032***	-4.408***	-5.082***	-2.032***	-4.408***	-5.082***
	(0.434)	(0.493)	(0.513)	(0.434)	(0.493)	(0.513)
# Drivers	2,100	2,100	2,100	2,100	2,100	2,100
Observations (#Drivers×#Days)	10,500	10,500	10,500	10,500	10,500	10,500
H_0: Responsive = Non-responsive				$p < 0.001$	$p = 0.1343$	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

Table F4—Team Captain Volunteers. Probit estimates. Reported results are average marginal effects. We include all drivers who signed up for the competition.

	Dependent Variable Volunteering to be Captain
Pre-Contest Avg. Daily Revenue (100 CNY)	0.0253***
	(0.00760)
Male	0.0377
	(0.0600)
Local	-0.0298
	(0.0201)
Age	-0.00124
	(0.00117)
DiDi Age (years)	0.0297**
	(0.0150)
# Drivers	2,343

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table F5—Effect of Being Appointed as a Captain: Difference-in-differences linear regressions. Dependent variable: Difference of driver revenue (compared with the pre-contest time window). Subjects are drivers who volunteer to be captains and are assigned to teams with multiple volunteers. Note that only one volunteer in each team is randomly selected to be the captain. Kolmogorov–Smirnov tests find no significant difference in prior revenue, age, DiDi age, or gender between the the selected captains and other volunteers ($p > 0.1$).

	Dependent Variable: Δ of Daily Revenue (CNY)		
	(1)	(2)	(3)
Time Period:	Contest	2-week Post Contest	4-week Post Contest
Assigned Captain	34.181*	23.647	-5.278
	(19.534)	(19.673)	(18.624)
Constant	-17.910*	-57.146***	-65.077***
	(12.589)	(13.759)	(14.707)
# Volunteers	298	298	298
Obs.	1,490	1,490	1,490

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions.

Table F6—Average and Heterogeneous Treatment Effects on Daily Revenue Excluding Team Captains. Difference-in-differences linear regressions. We compare each of the three time periods with the pre-contest period.

	Dependent variable: Δ of Daily Revenue (CNY)					
	Average Treatment Effects			Heterogeneous Treatment Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Period	Contest	2-week Post Contest	4-week Post Contest	Contest	2-week Post Contest	4-week Post Contest
Treated	35.23^{***}	16.74^{*}	5.997			
	(9.337)	(9.699)	(10.16)			
	[0.001]	[0.192]	[0.668]			
Responsive				54.36^{***}	22.16^{**}	9.311
				(9.880)	(10.18)	(10.65)
				[0.001]	[0.091]	[0.668]
Non-responsive				5.564	8.334	0.857
				(10.40)	(11.04)	(11.52)
				[0.668]	[0.668]	[0.941]
Constant	-24.24^{***}	-66.96^{***}	-82.06^{***}	-24.24^{***}	-66.96^{***}	-82.06^{***}
	(7.892)	(8.844)	(9.192)	(7.893)	(8.844)	(9.193)
# Drivers	1,850	1,850	1,850	1,850	1,850	1,850
Observations (#Drivers×#Days)	9,250	9,250	9,250	9,250	9,250	9,250

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the contest (individual) level for treatment (control) conditions. False discovery rate adjusted q-values are in square brackets to correct for multiple hypothesis testing.

Table F7—Post-contest Survey Response: Probit. Coefficients are average marginal effects. We include all drivers who participated in the team contest.

	Dependent Variable Responding to Post-contest Survey
Is Captain	0.0983***
	(0.0276)
Responsive Team	0.0565***
	(0.0219)
# Individual Wins	0.0348***
	(0.00764)
# Team Win	0.0131*
	(0.00699)
Average Daily Revenue (100 CNY)	-0.0349***
	(0.00946)
Male	0.0333
	(0.0696)
Local	0.0188
	(0.0231)
Age	0.00431***
	(0.00138)
DiDi Age	-0.00598
	(0.0178)
# Drivers	1,750

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

*

REFERENCES

Ai, Wei, Roy Chen, Yan Chen, Qiaozhu Mei, and Webb Phillips. 2016. "Recommending teams promotes prosocial lending in online microfinance." *Proceedings of the National Academy of Sciences*, 113(52): 14944–14948.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Akerlof, George A., and Rachel E. Kranton. 2010. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton, New Jersey:Princeton University Press.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B (Methodological)*, 57(1): 289–300.

Brewer, Marilynn B. 1999. "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?" *Journal of Social Issues*, 55(3): 429–444.

Brewer, Marilynn B., and Michael D. Silver. 2000. "Group Distinctiveness, Social Identification, and Collective Mobilization." In *Self, Identity and Social Movements.* , ed. S. Stryker, T. J. Owens and R. W. White. Minneapolis, MN:University of Minnesota Press.

Cason, Timothy N., Roman M. Sheremeta, and Jingjing Zhang. 2012. "Communication and efficiency in competitive coordination games." *Games and Economic Behavior*, 76(1): 26 – 43.

Charness, Gary, and Yan Chen. 2020. "Social Identity, Group Behavior and Teams." *Annual Review of Economics.*

Charness, Gary, Luca Rigotti, and Aldo Rustichini. 2007. "Individual behavior and group membership." *The American Economic Review*, 97: 1340 – 1352.

Chen, M. Keith, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen. 2019. "The Value of Flexible Work: Evidence from Uber Drivers." *Journal of Political Economy*, 127(6): 2735–2794.

Chen, Roy, and Yan Chen. 2011. "The Potential of Social Identity for Equilibrium Selection." *The American Economic Review*, 101(6): 2562–2589.

Chen, Roy, Yan Chen, and Yohanes E. Riyanto. 2020. "Best practices in replication: A case study of common information in coordination games." *Experimental Economics.*

Chen, Yan, and Sherry Xin Li. 2009. "Group Identity and Social Preferences." *The American Economic Review*, 99(1): 431–457.

Chen, Yan, Sherry Xin Li, Tracy Xiao Liu, and Margaret Shih. 2014. "Which hat to wear? Impact of natural identities on coordination and cooperation." *Games and Economic Behavior*, 84(0): 58 – 86.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2016. "Using Big Data to Estimate Consumer Surplus: The Case of Uber." National Bureau of Economic Research Working Paper 22627.

Cook, Cody, Rebecca Diamond, Jonathan Hall, John A List, and Paul Oyer. 2018. "The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare Drivers." National Bureau of Economic Research Working Paper 24732.

Eckel, Catherine C., and Philip J. Grossman. 2005. "Managing Diversity by Creating Team Identity." *Journal of Economic Behavior & Organization*, 58(3): 371–392.

Efron, Bradley. 2010. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction.* Stanford University Online Manuscript.

Erev, Ido, Gary Bornstein, and Rachely Galili. 1993. "Constructive Intergroup Competition as a Solution to the Free Rider Problem: A Field Experiment." *Journal of Experimental Social Psychology*, 29(6): 463–478.

Fibich, Gadi, Arieh Gavious, and Aner Sela. 2006. "All-pay Auctions with Risk-averse Players." *International Journal of Game Theory*, 34(4): 583–599.

Fu, Qiang, Jingfeng Lu, and Yue Pan. 2015. "Team contests with multiple pairwise battles." *American Economic Review*, 105(7): 2120–2140.

Gee, Laura K., Michael J. Schreck, and Ankriti Singh. 2020. "From lab to field: Social distance and charitable giving in teams." *Economics Letters*, 109128.

Girvan, M, and M E J Newman. 2002. "Community structure in social and biological networks." *Proceedings of the National Academy of Sciences of the United States of America*, 99(12): 7821–6.

Goette, Lorenz, David Huffman, Stephan Meier, and Matthias Sutter. 2012. "Competition Between Organizational Groups: Its Impact on Altruistic and Antisocial Motivations." *Management Science*, 58(5): 948–960.

Hall, Jonathan V., and Alan B. Krueger. 2018. "An Analysis of the Labor Market for Uber's Driver-Partners in the United States." *ILR Review*, 71(3): 705–732.

Harmel, Robert, and Yao-Yuan Yeh. 2015. "China's Age Cohorts: Differences in Political Attitudes and Behavior." *Social Science Quarterly*, 96(1): 214–234.

Kessler, Judd B., and Katherine L. Milkman. 2018. "Identity in Charitable Giving." *Management Science*, 64(2): 845–859.

Kim, Phillip H, and Howard E Aldrich. 2002. "Teams that work together, stay together: Resiliency of entrepreneurial teams." Vol. 2006.

Konrad, Kai A. 2009. *Strategy and Dynamics in Contests.* New York, New York:Oxford University Press.

Krishnan, Hema A, Alex Miller, and William Q Judge. 1997. "Diversification and top management team complementarity: Is performance improved by merging similar or dissimilar teams?" *Strategic Management Journal*, 18(5): 361–374.

List, John A., Azeem M. Shaikh, and Yang Xu. 2019. "Multiple hypothesis testing in experimental economics." *Experimental Economics.*

List, John A., Sally Sadoff, and Mathis Wagner. 2011. "So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design." *Experimental Economics*, 14: 439.

Liu, Tracy Xiao, Zhixi Wan, and Chenyu Yang. 2018. "The Efficiency of A Dynamic Decentralized Two-sided Matching Market." University of Rochester Working Paper.

McPherson, M, L Smith-Lovin, and JM Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual Review of Sociology*, 27(2001): 415–444.

Noussair, Charles, and Jonathon Silver. 2006. "Behavior in All-Pay Auctions Under Incomplete Information." *Games and Economic Behavior*, 55(1): 189–206.

Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies.* Princeton, New Jersey.:Princeton University Press.

Ravenelle, Alexandrea J. 2019. *Hustle and Gig: Struggling and Surviving in the Sharing Economy.* . First edition ed., Oakland, California:University of California Press.

Ruef, Martin, Howard E Aldrich, and Nancy M Carter. 2003. "The structure of founding teams: Homophily, strong ties, and isolation among US entrepreneurs." *American Sociological Review*, 195–222.

Scheiber, Noam. 2017. "How Uber Uses Psychological Tricks to Push Its Drivers' Buttons." *New York Times.*

Sheremeta, Roman M. 2018. "BEHAVIOR IN GROUP CONTESTS: A REVIEW OF EXPERIMENTAL RESEARCH." *Journal of Economic Surveys*, 32(3): 683–704.

Tajfel, Henri, and John Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations.* , ed. Stephen Worchel and William Austin, 94–109. Monterey, CA:Brooks/Cole.

Tajfel, Henri, Michael Billig, R. Bundy, and Claude L. Flament. 1971. "Social Categorization and Inter-Group Behavior." *European Journal of Social Psychology*, 1: 149–177.

Vojnović, Milan. 2016. *Contest Theory: Incentive Mechanisms and Ranking Methods.* Cambridge, United Kingdom:Cambridge University Press.

Ye, Teng, Wei Ai, Yan Chen, Qiaozhu Mei, and Jieping Ye. 2019. "Status competition for ride sharing: A large-scale field experiment at DiDi." University of Michigan Working Paper.

Yuan, Y Connie, and Geri Gay. 2006. "Homophily of network ties and bonding and bridging social capital in computer-mediated distributed teams." *Journal of Computer-Mediated Communication*, 11(4): 1062–1084.

Zhang, Chunni, and Yu Xie. 2013. "Place of Origin and Labour Market Outcomes Among Migrant Workers in Urban China." *Urban Studies*, 50(14): 3011–3026.