Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023

Human Technology Interaction

Dec 11th, 12:00 AM

Value of Autonomous Last-mile Delivery: Evidence from Alibaba

Lixia Wu
Cainiao Smart Logistics Network Ltd., wallace.wulx@cainiao.com

Peng Li
Cainiao Smart Logistics Network Ltd., lipeng.lipeng@cainiao.com

Brian Rongqing han
University of Illinois Urbana Champaign, brianhan@illinois.edu

Meng Li
University of Houston, mli@central.uh.edu

Yanan Zhang
Shanghai Jiaotong University, sjtu_erica1318@sjtu.edu.cn

Value of Autonomous Last-mile Delivery: Evidence from Alibaba

Completed Research Paper

Lixia Wu
Cainiao Smart Logistics Network Ltd.
969 Wenyi West Road, Zhejiang, China
wallace.wulx@cainiao.com

Peng Li
Cainiao Smart Logistics Network Ltd.
969 Wenyi West Road, Zhejiang, China
lipeng.lipeng@cainiao.com

Brian Rongqing Han
University of Illinois at Urbana-Champaign
1206 S 6th St, Champaign, Illinois, USA
brianhan@illinois.edu

Meng Li
University of Houston
4250 Martin Luther King Blvd, Texas, USA
mli@bauer.uh.edu

Yanan Zhang
Shanghai Jiaotong University
1954 Huashan Road, Shanghai, China
sjtu_erica1318@sjtu.edu.cn

Abstract

This paper provides the first empirical evidence of consumer responses to autonomous last-mile delivery using Alibaba's recent implementation in Chinese university campuses as a case study. The study leverages customer-level data from three universities over three years, employing a difference-in-differences (DID) approach combined with dynamic matching to estimate the impact of autonomous delivery adoption on order quantities. The results reveal a significant increase in the number of orders following autonomous delivery adoption with a 21% growth. The efficiency and flexibility of autonomous vehicles reduce consumers' travel costs, driving long-term usage and increased sales. However, the value of autonomous delivery diminishes when a fee is charged. The study contributes to our understanding of the value of autonomous last-mile delivery and its potential advantages over traditional courier delivery.

Keywords: Empirical operations, last-mile delivery, autonomous vehicles, difference-in-differences

Introduction

Autonomous vehicles have received increasing interest, particularly in last-mile delivery. Kroger uses autonomous vehicles to deliver groceries directly to consumer's door (Morris 2023). Robots operated by Starship Technologies have completed more than two million deliveries and are dropping off dinner on four-dozen college campuses in the United States (Vartabedian 2022). Besides the cost-saving aspect of using robots instead of humans, autonomous vehicles are expected to streamline routine operations more consistently and efficiently. In turn, the enhanced quality of last-mile delivery services may further improve customer satisfaction.

Despite the growing investment into autonomous last-mile delivery (using "self-driving vehicles" to deliver packages at the last mile), the value of its implementation is unknown. Specifically, can autonomous last-mile delivery increase sales? Several studies have already shown that last-mile delivery is an important element of logistics, which can significantly increase or decrease e-commerce sales (Cui et al. 2020; Cui et al. 2023; Han et al. 2022; Luo et al. 2020). One of the main reasons is that last-mile delivery closely connects with customers at the final stage of delivery, and any operational improvement or failure can be instantly felt by customers. Therefore, the quality of last-mile delivery is a crucial part of overall logistics service quality. However, these studies focused only on human-courier delivery (using "vehicles driven by human drivers" to deliver packages at the last mile). At the same time, previous studies on autonomous vehicles mostly focus on the operational aspect such as vehicle routing (Cao and Qi 2023; Carlsson and Song 2018; Reed et al. 2022), contributing to a better understanding of its "potential" impact in real life. It makes an empirical investigation of its real-life impact on consumers particularly necessary and urgent because we need to validate whether autonomous delivery indeed achieves efficient delivery and perhaps offers more flexibility. More importantly, the enhanced customer experience can be further translated into more orders or sales, which is still a question.

This paper seeks to provide the first empirical evidence on consumer responses to autonomous last-mile delivery based on Alibaba's recent implementation of autonomous vehicles for package delivery on university campuses in China. By 2023, Alibaba's logistics arm Cainiao Network has established more than 120,000 last-mile stations as pickup points across mainland China including university campuses. Packages (including purchases from Alibaba's online platforms and other retailers) are directed to the last-mile station so students and staff on campus can self-pick up their packages. In 2021, Alibaba started deploying autonomous vehicles to university campuses for package delivery. Instead of self-pickup, consumers can order autonomous last-mile delivery to a specified location (e.g., dorm building) in a time window (e.g., one hour time window). Using customer-level data (of more than 130,000 individuals over three years) from three universities (that have deployed autonomous vehicles) in a major metropolitan city, our goal is to estimate the impact of autonomous last-mile delivery adoption on order quantities, compared with self-pickup, and understand why.

Following previous literature (Bapna et al. 2018; Xu et al. 2017), we use a difference-in-differences (DID) approach coupled with dynamic matching to obtain an as clean as possible estimate of the impact of adoption. We find that compared with self-pickup, consumers order significantly more after adopting autonomous last-mile delivery. The number of orders increased by 21% ($p < 0.01$). We conduct various robustness checks on our identification strategy and find consistent results.

One unique aspect of our dataset is that before the deployment of autonomous vehicles, the last-mile station on each campus also hired a human courier to deliver packages for a few months. It creates an opportunity for us to estimate impact of human couriers on last-mile delivery. We estimate that compared with self-pickup, human courier delivery increases consumer orders by about 9% ($p < 0.01$). Although not a direct comparison, the analysis suggests that the value of autonomous last-mile delivery does not come from the "delivery effect" itself. Autonomous vehicles, compared to human couriers, may drive additional value.

To understand the underlying mechanism, we further examine the operational impact of autonomous vehicle by investigating whether they can deliver packages efficiently and offers additional flexibility. First, we find a 10% decrease in lead time, i.e., from packages' arrival at the last-mile station to package receiving time. It suggests that consumers can receive packages sooner, which also indicates a faster turnover at the last-mile stations compared with self-pickup. Second, we estimate the impact of autonomous last-mile delivery on the number of orders by each hour of the day and find that the effect comes from both the self-pickup peak hours (such as noon or after 5:00 pm) and non-peak (such as 2:00 pm in the afternoon or before 9:00 pm). It shows that the nonstop working hours of the autonomous vehicle bring significant flexibility for customers.

Because of their efficiency and flexibility, autonomous vehicles provide high-quality delivery services, effectively decreasing consumers' travel costs (of self-pickup), which drives long-term usage and increased sales. For example, we find a significantly larger effect for those who are further away from the last-mile station. However, the value of autonomous last-mile delivery disappears if it is not free. Since delivery is not free for packages from retailers other than Alibaba's online platforms (such as Taobao.com and Tmall.com), we find that autonomous last-mile delivery has no effect on these packages. Therefore, autonomous last-mile delivery brings great value only when it is free.

Before proceeding, we summarize our key contributions as follows.

- We provide the first empirical evidence of the positive impact of autonomous last-mile delivery on consumer orders. Our results also suggest that its value is potentially larger than traditional courier delivery.
- We show that autonomous delivery offers efficient and flexible last-mile delivery, which effectively reduces consumers' travel costs. In other words, the high-quality delivery services (offered by autonomous vehicles) drive the increased number of orders. However, such an effect disappears if we charge consumers for the autonomous delivery service.

Overall, we believe our results can provide important insights into last-mile delivery operations in a general sense beyond campuses, such as metropolitan areas in Europe or China. Our study showcases that if autonomous vehicles are capable of package delivery in occasionally crowded streets, we can expect positive value generated from such operational efficiency uniquely brought by machines. In other words, if autonomous vehicles can deliver packages more consistently and in larger volumes every day, customers are willing to adopt them and eventually order more frequently.

Literature Review

Our study contributes to several streams of literature. The first stream of literature aims to understand the relationship between retailing and logistics. (Cohen and Lee 2020; Fan et al. 2018; Lee and Whang 2001; Terwiesch et al. 2005). Earlier research has investigated the impact of supply chain glitches on firms' operational performance (Hendricks and Singhal 2003; Hendricks and Singhal 2005). In today's omnichannel world, studies have also shown that how customers receive products can significantly benefit the overall revenue of retailers (Gallino and Moreno 2019) through innovative business practices, such as buy-online-pickup-in-store (Gallino and Moreno 2014) or ship-to-store (Gallino et al. 2017). More recently, researchers have established a link between delivery speed and online sales (Cui et al. 2020; Cui et al. 2023; Han et al. 2022; Luo et al. 2020). We contribute to the literature by providing the first empirical evidence of the impact of autonomous delivery on online sales. Our findings connect to the literature by showing the linkage between high-quality delivery services and sales. Furthermore, we add to the literature by showing that autonomous logistics services can offer more efficiency and additional flexibility compared to traditional courier delivery.

The second stream of literature examines human responses to different applications of artificial intelligence (AI) driven technology, which can be further characterized into firm-side and consumer-side studies. Previous literature has investigated human-AI collaboration (Dietvorst et al. 2018) in various component in firm operations (Caro and de Tejada Cuenca 2023; Cui et al. 2022; Kesavan and Kushwaha 2020; Sun et al. 2022), fintech (Fu et al. 2021; Ge et al. 2021; Lu and Zhang 2023), marketing (Luo et al. 2021), healthcare (Jussupow et al. 2021), media (Claussen et al. 2019), or more generally firm performance (Brynjolfsson and Mitchell 2017; Luo et al. 2021). Others have focused more on consumer responses to AI services (Luo et al. 2019; Yalcin et al. 2022) or co-creation (Zhang et al. 2023). These studies all have focused on the informational role of AI. Much fewer studies have examined the application in which AI makes decisions and takes action. With autonomous driving, AI fully replaces humans to carry out routine tasks. Recent work studies the operational impact of autonomous driving on traffic (Zhang et al. 2020). We contribute to the literature by first looking into how consumers respond to AI applied in a physical business process.

Finally, our work also relates to the more general IS literature on the economic impact of technology adoption. For example, previous studies on the mobile economy have found that customers' adoption of mobile apps can lead to more purchases and usage (Sun et al. 2019; Xu et al. 2017; Xu et al. 2023), social engagement (Jung et al. 2019), consumption of news (Xu et al. 2014), and digital services (Liu et al. 2016). Our study contributes by focusing on the economic effect of autonomous delivery in offline operations.

Data

We focus on a major metropolitan city where three university campus stations have implemented autonomous vehicles for package delivery by Cainiao Smart Logistics Network (the logistics subsidiary company of Alibaba Group). The raw data includes package-level delivery information from these campus last-mile stations from September 2019 to February 2023. Before 2021, customers could only receive packages through self-pickup at the campus stations. Starting in May 2021, these stations adopted human couriers for last-mile delivery, which were later replaced by autonomous vehicles. Notably, self-pickup is consistently available to consumers. Figure 1 provides an overview of the research setting, while Figure 2 describes the process of placing an autonomous delivery order.

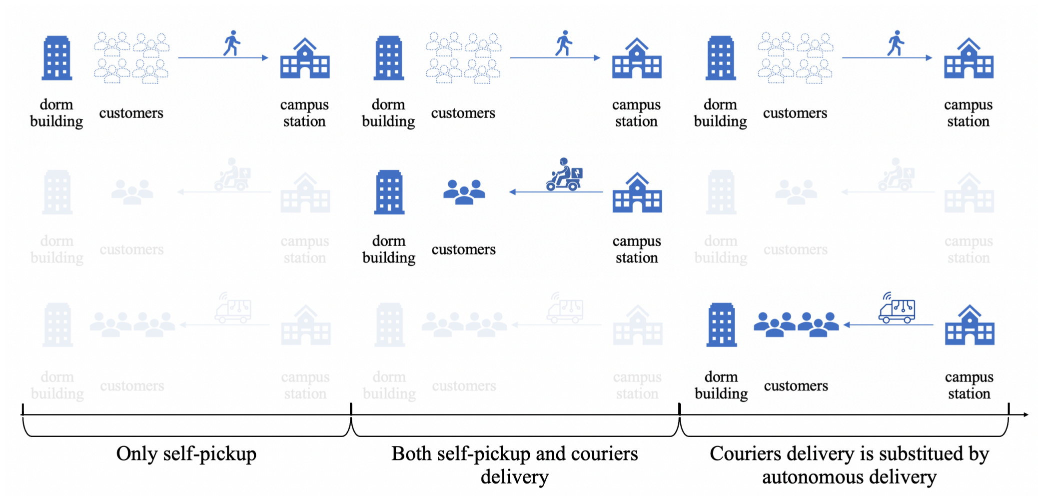

Figure 1. Research Setting

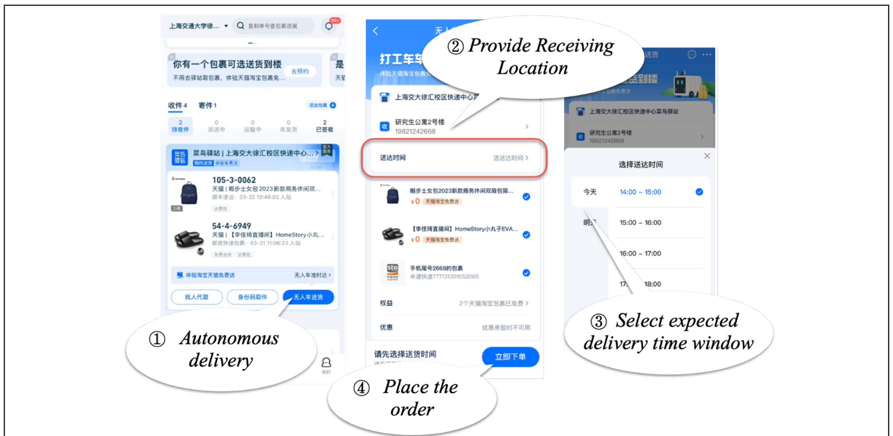

Figure 2. Ordering Autonomous Delivery

Our data includes the timestamp of package arrival (at the stations), timestamp of delivery, user ID, delivery channel (autonomous vehicles, human couriers, and self-pickup), the distance between the receiving address and campus station, and whether the purchase was made through Alibaba's platforms (e.g., Taobao.com and Tmall.com). After eliminating 1.15% of missing values, we have a total of 2,733,135 package-level records from 138,594 consumers, comprising 2,440,954 records (89.3%) for self-pickup, 285,573 records (10.4%) for autonomous delivery, and 6,608 records (0.24%) for human couriers. Table 1 summarize the number of packages by delivery method across the three stations.

Our data include 138,594 customers throughout the study period. As illustrated in Table 2, a majority of them chose self-pickup to obtain packages, while a portion of them adopted autonomous vehicles, and a minimal number used human couriers. We exclude those who have used human couriers. Those who only used self-pickup are considered "non-adopters". Those who, at some point, adopted autonomous delivery are considered "adopters". We summarize the number of consumers by adoption time in Table 3. We then construct a user-month panel data of order quantity for these users, excluding winter and summer breaks. Figure 3 shows the average percentage of packages delivered by autonomous vehicles by adopters in different semesters across the three stations. We can see a growing trend after their first adoption.

	Number of packages from		
	human couriers delivery	autonomous delivery	self-pickup
Station 1	4,218	107,377	1,299,207
Station 2	1,143	168,362	609,812
Station 3	1,247	9,834	531,935

Table 1. Number of packages from different signing channels

	Number of consumers used			
	autonomous delivery (adopter)	human couriers delivery	both	only self-pickup (non-adopter)
Station 1	12,678	238	948	38,525
Station 2	23,108	90	460	26,955
Station 3	4,122	376	326	30,768

Table 2. Number of Consumers Using Different Signing Channels

	Groups by adoption semester	First adoption month	No. of consumers
Station 1	Fall 2021 adopter	2021-10-01	4,866
	Spring 2022 adopter	2022-02-01	2,812
	Fall 2022 adopter	2022-09-01	4,958
	Non adopter	–	38,525
Station 2	Spring 2022 adopter	2022-02-01	13,522
	Fall 2022 adopter	2022-09-01	9,582
	Non adopter	–	26,955
Station 3	Spring 2022 adopter	2022-03-01	2,283
	Fall 2022 adopter	2022-09-01	1,833
	Non adopter	–	30,398

Table 3. Consumers by Groups

Before proceeding, Figure 4 presents a set of model-free evidence on the impact of autonomous delivery adoption on the number of orders. We plot the average monthly orders by group and find an increasing trend after the month in which the autonomous vehicles are deployed. It signals a significant effect of autonomous vehicles. Although the graphical trends may offer general insights into the phenomenon, the conclusions drawn are subject to endogeneity concerns that can arise from unobserved systematic biases that are inherent in autonomous vehicle adoption decisions. We employ several econometric techniques to account for these concerns, which we will describe in detail in section 4.

Main Result

First of all, we estimate the following generalized difference-in-differences model to quantify the effect of adopting autonomous vehicles on the outcome measures of interest,

$$Y_{it} = \alpha + \beta \times PostAdoption_{it} + UserFE_i + TimeFE_t + \varepsilon_{it}, \tag{1}$$

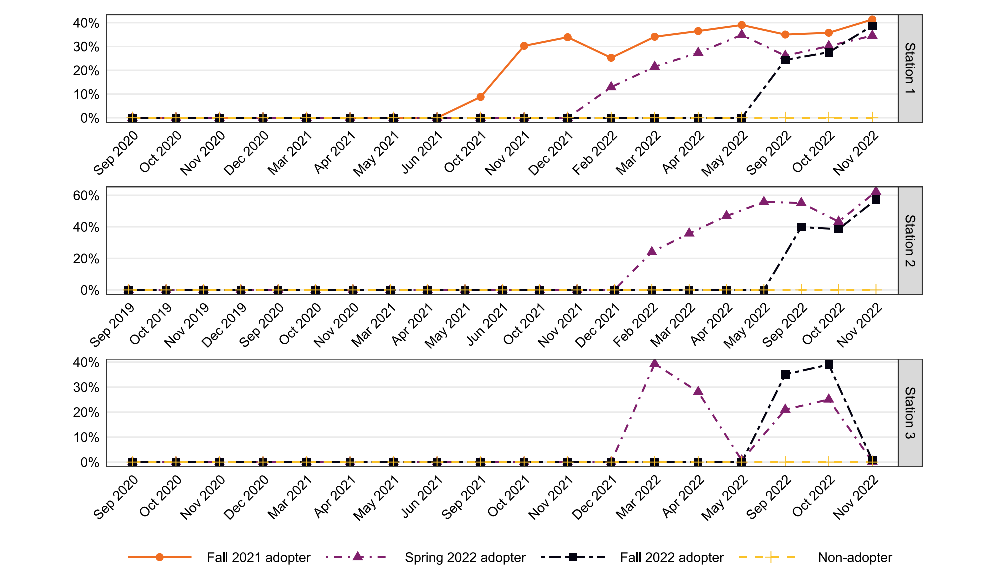

Figure 3. Proportion of Autonomous Delivery Packages by Groups

where the outcome variable Y_{it} is number of orders for user i in month t. And outcome variable is log-transformed—namely, the *Log*(*Number of orders* + *1*) —to produce an elasticity interpretation with respect to whether the consumer has adopted the autonomous vehicles. For certain adopter, the month when she adopted autonomous vehicles is defined as period 0. Then $PostAdoption_{it}$ is a binary variable that indicates the post adoption period (i.e. period 0, 1, 2 , ...) for each adopter. For instance, for a user who first adopted the autonomous vehicles in October 2021, the user's $PostAdoption_{it}$ indicator will be a 1 for that month and subsequent months and a 0 for periods before October 2021. For a user who never adopted the autonomous vehicles, the indicator will be a 0. Because our panel data have repeated observations every period for each user, $UserFE_i$ could control the user-fixed effects to account for any unobservable and time-invariant user characteristics. Moreover, $TimeFE_t$ is the monthly fixed effect and ε_{it} are clustered by users to account for potential correlation over time. We are interested in the coefficient β measuring the effect of adopting autonomous vehicles.

Moreover, we note that endogeneity concerns can arise from selection bias (i.e., certain groups are more likely to adopt than others), leading to the unfair comparison between adopters and non-adopters. To reduce potential differences across the adopters and non-adopters and alleviate the endogeneity, we further use DID approach coupled with matching according to Equation (2),

$$Y_{it} = \alpha + \gamma \times After_{it} + \beta \times (Adopter_i \times After_{it}) + UserFE_i + TimeFE_t + \varepsilon_{it}, \tag{2}$$

where $Adopter_i$ is denoted with 1 if a user has ever adopted the autonomous vehicles within the study period and 0 if otherwise. This binary term controls for the time-invariant group differences that may be present between adopters and non-adopters. For certain adopter, the month when she adopted autonomous vehicles is defined as period 0. Then $After_{it}$ is a binary variable that indicates the post adoption period (i.e. period 0, 1, 2 , ...) for each adopter and the matched counterpart. For instance, for a user who first adopted the autonomous vehicles in October 2021, the user's $After_{it}$ indicator will be a 1 for that month and subsequent months and a 0 for periods before October 2021. Matched users for this focal adopter will have similar values for this binary variable. Similar to Equation (1), coefficient β measure the effects of adopting autonomous vehicles. The coefficient β for the interaction term, $Adopter_i \times After_{it}$, represents the difference-in-difference estimator that captures how the number of orders of adopters changes after adoption in contrast to that of non-adopters.

Samples are matched using two main types of propensity score matching (PSM) procedures—namely, static

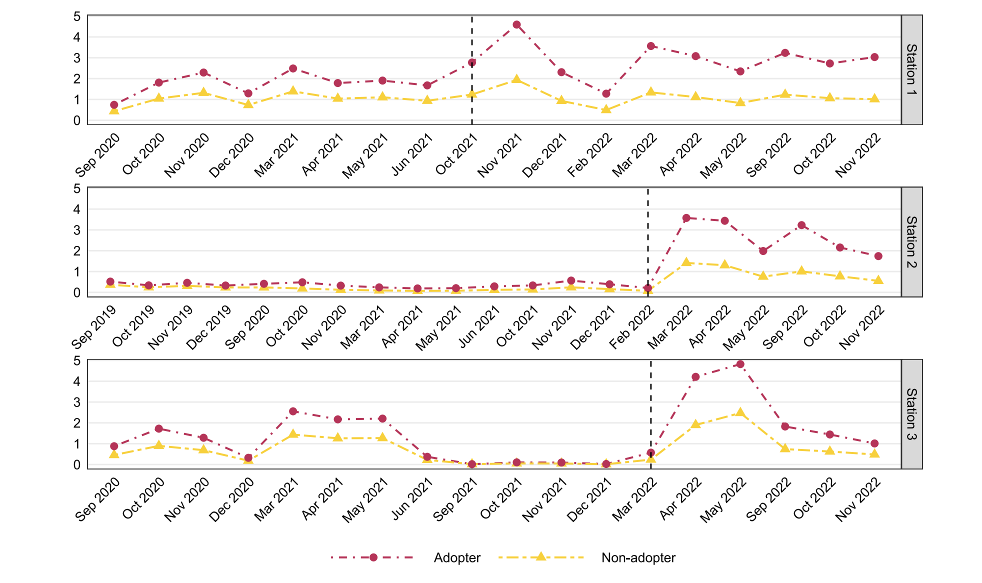

Figure 4. Number of Packages Per User by Groups

matching and dynamic matching. Under static matching, propensity scores of autonomous delivery adopters and non-adopters in months after the introduction of autonomous vehicles are tabulated. Specifically, we use eight months of package records prior to the introduction of autonomous vehicles to match users and these adopters are matched with non-adopters who resemble them most closely in terms of their overall propensity scores. In our case, static matching has two drawbacks. On the one hand, all users' propensity scores depended on the identical pre-period (i.e. eight months before the introduction of autonomous vehicles by campus stations). Considering that the adopting date is different among adopters (see Table 3), it is possible that users who adopted the autonomous vehicles in the latter months are matched with imperfect non-adopter to some extent. Under these circumstances, a more valuable pre-period (i.e. eight months just before adoption month) will be ignored so that endogeneity issues are not effectively solved. On the other hand, static matching requires that the user has picked up at least one package during the eight-month pre-period. Consequently, quite a few users are discarded before the static matching process. So we further use the dynamic matching (Xu et al. 2017), allowing adopters to be matched by adoption month. And the matched pool of non-adopters changes according to the adopter's groups by adoption month. For instance, for adopters who first adopted the autonomous vehicles in a certain month (e.g., December 2021), we only consider the non-adopters who have ever picked up the package in that month, which provides more flexibility and choice. The endogeneity issue could be solved to a large extent with dynamic matching compared with static matching, and we are more interested in results from dynamic matching, which we will show in the main result table.

When estimating the propensity scores, our covariates include the monthly number of packages and the monthly number of packages from Alibaba's platforms, which could represent the users' consumption habit. Given that purchase-related attributes of users change with time, the matching procedure is likely to derive better matches, because it accounts for time-varying factors in the matching process. In addition, average package delivery time and median package delivery time, which could represent users' sign-off behavior and whether relatively far distance from the campus stations, are also considered. The same set of covariates is utilized across all matching procedures. Our baseline matching utilizes one-to-one matching with replacement to derive the closest matched non-adopter, under a caliper size 0.20 times the standard deviation of the propensity scores. We use the logistic link to estimate the distance measure. Further, we utilize various matching algorithms for dynamic matching to assess the robustness of the results with respect to different matched samples. To evaluate the success of propensity score matching, we compare the overall distributions of propensity scores of the matched and unmatched samples under dynamic matching. Figure 5 shows that matched non-adopters have a propensity score distribution more similar to the adopters than to the unmatched adopters.

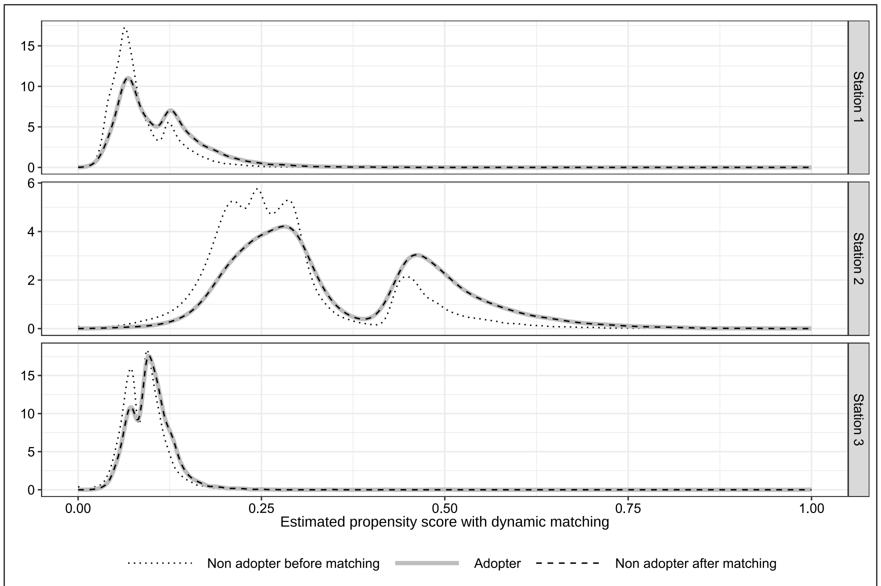

Figure 5. Distribution of Propensity Scores for Adopters and Non-adopters (Both Matched and Unmatched)

Table 4 presents our main estimation result obtained from Equation (1) and Equation (2). Specifically, Column (1) shows the result of the generalized difference-in-differences model from Equation (1) on the unmatched sample data. Autonomous delivery increased the number of orders by 53% ($p < 0.01$), we note that the result is significant but implausible in Column (1), which could be involved in the endogeneity concerns potentially. Column (2) shows the result of the difference-in-differences model from Equation (2) coupled with static matching, in which the endogeneity issues are mitigated to some extent. In this situation, autonomous delivery increased the number of orders by 39% ($p < 0.01$). The result is significant and the magnitude of the effect is not dramatically great. Whereas, we stress that Column (2) is still not entirely convincing due to the existence of different adoption months and the resulting small sample size just mentioned before. Column (3) shows the result of the difference-in-differences model from Equation (2) coupled with dynamic matching, which could obtain an as clean as possible estimate of the impact of adoption. Autonomous delivery actually increased the number of orders by 21% ($p < 0.01$). In short, we find that compared with self-pickup, consumers order significantly more after adopting autonomous last-mile delivery.

Importantly, to allow for a correct interpretation of the difference-in-differences estimator, the critical assumption is that there must be parallel trends between the adopters and the matched non-adopters. We estimate an event study specification on matched data in Equation (2-1):

$$\begin{aligned} Y_{it} = \alpha &+ \delta_{-6} \times Before\_6Months_{it} + \cdots \\ &+ \delta_{-2} \times Before\_2Months_{it} + \delta_{+0} \times After\_0Months_{it} + \cdots \\ &+ \delta_{+4} \times After\_4Months_{it} + \beta_{-6} \times (Adopter_i \times Before\_6Months_{it}) + \cdots \\ &+ \beta_{-2} \times (Adopter_i \times Before2\_Months_{it}) + \beta_{+0} \times (Adopter_i \times After\_0Months_{it}) + \cdots \\ &+ \beta_{+4} \times (Adopter_i \times After\_4Months_{it}) + UserFE_i + TimeFE_t + \varepsilon_{it}, \end{aligned} \tag{2-1}$$

Here, the right-hand side allow for 5 leads ($\beta_{-2}, \ldots, \beta_{-6}$) or anticipatory effects and 4 lags ($\beta_{+1}, \beta_{+2}, \ldots, \beta_{+4}$) or post-treatment effects. For certain adopter, the month when she adopted autonomous vehicles is defined as period 0. $Before\_jMonths_{it}$, $j = (2, 3, \ldots, 5)$ equals one for each adopter in the jth month before adoption, matched users for this focal adopter will have similar values for this binary variable. $After\_jMonths_{it}$,

	Number of orders		
	(I)	(II)	(III)
$PostAdoption_{it}$	0.530^{***}		
	(0.004)		
$After_{it}$		0.029^{***}	0.155^{***}
		(0.007)	(0.006)
$Adopter_i \times After_{it}$		0.392^{***}	0.214^{***}
		(0.008)	(0.007)
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	2,593,389	739,506	996,846
R^2	0.420	0.387	0.406
Adjusted R^2	0.388	0.355	0.374
Residual Std. Error	0.521	0.680	0.661

Robust standard errors clustered by each user are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

Table 4. Main Estimation Results

$j = (1, \ldots, 3)$ equals one for each adopter in the jth month after adoption, matched users for this focal adopter will have similar values for this binary variable. But at the end-point, $Before\_6Months_{it}$ equals one for all months that are 6 or more months before adoption and $After\_4Months_{it}$ equals one for all months that are 4 or more months after adoption. We also include user-fixed effects $UserFE_i$ which absorbs both observable and unobservable time-invariant differences, and month-fixed effects $TimeFE_t$, which absorbs time-varying common influence to all users. The standard errors are clustered at the user level. And we omit period -1 to describe the evolution of the difference between adopters and matched non-adopters with respect to period -1. Consistent with model-free evidence in Figure 4, Figure 6 shows the result that our dynamic matching sample exactly satisfies the parallel trend assumption. Period $(\leq -6, -5, \ldots, -2)$ shows no significant difference between adopters and non-adopters compared with period -1. Yet, the period $(0, \ldots, 3, \geq 4)$ is significantly different, which represents heterogeneous effects over time relative to autonomous delivery adoption. In the meantime, it suggests that consumers continue to use autonomous delivery after the initial try-out for its novelty.

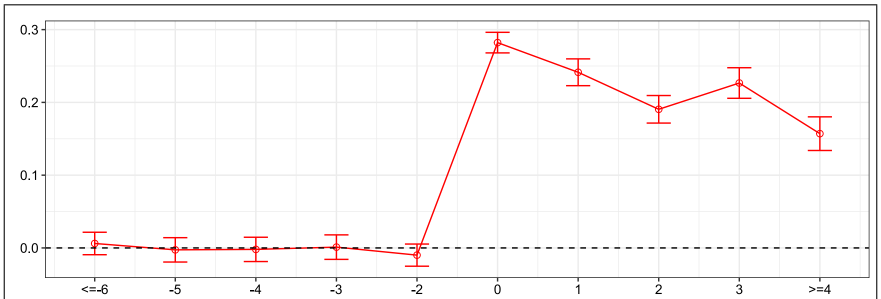

Figure 6. Coefficients of the Monthly Difference-in-Difference Estimates Before and After Adoption

Robustness Check

Alternatives of Main Results

In our main model, we consider the user sample where the influence from human couriers has been ruled out, whereas the issue of unfair comparison uniquely arising from campus stations where the subsequent adopters who adopted the autonomous vehicles are compared with non-adopters who have graduated at that time and would not place orders in the future may generate the biased estimates. To further verify the credibility of the main results, we exclude the users who have left during the study period. Eventually, there exist two types of users in the sample: 1) users who were present throughout the study period, 2) users who were new for the school year when the autonomous vehicles have been introduced, e.g., the new students for the school year 2021-22, in the meanwhile the autonomous vehicles have been introduced in October 2021. We summarize the number of alternatives consumers by adoption time in Table 5. And we further repeat the procedure in the main result to analyze the alternative sample. The result of this analysis is shown in Table 6. The coefficient did not differ much from the coefficient in the main result and the number of orders increased by 21 % ($p < 0.01$), which shows consistency with our main result. We can conclude that our main result is trustworthy and reliable.

	Groups by adoption semester	First adoption month	No. of customers
Station 1	Fall 2021	2021-10-01	2,068
	Spring 2022	2022-02-01	1,339
	Fall 2022	2022-09-01	1,727
	Non adopter	–	9,151
Station 2	Spring 2022	2022-02-01	8,713
	Fall 2022	2022-09-01	2,922
	Non adopter	–	6,032
Station 3	Spring 2022	2022-03-01	1,451
	Fall 2022	2022-09-01	943
	Non adopter	–	9,596

Table 5. Alternatives Consumers by Groups

	Number of orders		
	(I)	(II)	(III)
$PostAdoption_{it}$	0.471***		
	(0.005)		
$After_{it}$		0.161***	0.293***
		(0.008)	(0.007)
$Adopter_i \times After_{it}$		0.299***	0.210***
		(0.008)	(0.007)
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	843,957	536,412	673,230
R^2	0.465	0.473	0.486
Adjusted R^2	0.436	0.445	0.459
Residual Std. Error	0.628	0.641	0.632

Robust standard errors clustered by each user are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

Table 6. Alternatives Main Results

Identical Pre- and Post-period for All Individuals

We estimated an another model with an identical pre- and post-period intervention/cutoff date for all users based on the autonomous vehicles introduction in campus station. In other words, we compared adopters and non-adopters along their outcomes before and after the autonomous vehicles' introduction. Furthermore, we employ static and dynamic matching to alleviate the endogeneity concern, and the coefficient (24%, $p < 0.01$) from Table 7 is similar to the main result, which suggests that effects are robust and consistent with previous findings.

$$Y_{it} = \alpha + \beta \times (Adopter_i \times AfterIntroduction_t) + UserFE_i + TimeFE_t + \varepsilon_{it}, \tag{3}$$

	Number of orders		
	(I)	(II)	(III)
$Adopter_i \times AfterIntroduction_t$	0.382***	0.372***	0.243***
	(0.003)	(0.006)	(0.005)
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	2,593,389	739,506	996,846
R^2	0.412	0.386	0.403
Adjusted R^2	0.380	0.354	0.371
Residual Std. Error	0.525	0.680	0.663
Robust standard errors clustered by each user are in parentheses. *p<0.1; **p<0.05; ***p<0.01.			

Table 7. Identical Pre- and Post-period for All Individuals

Variety of Matching Methods

Our main analysis relies on the commonly used one-to-one nearest neighbor matching with replacement algorithm and requiring common support. Additionally, we employ a variety of matching algorithms for dynamic matching including one-to-one matching without replacement, nearest three neighbors, or a larger caliper size of 0.05 and a tighter one of 0.01 times the standard deviation of the propensity scores to verify the robustness of the results. A range of estimates from 21% to 22% ($p < 0.01$) are robust toward more stringent requirements in Table 8.

	Number of orders								
	0.2 caliper size			0.05 caliper size			0.01 caliper size		
	With Replacement	Nearest Three Neighbors	Without Replacement	With Replacement	Nearest Three Neighbors	Without Replacement	With Replacement	Nearest Three Neighbors	Without Replacement
$After_{it}$	0.155***	0.136***	0.143***	0.156***	0.138***	0.139***	0.157***	0.138***	0.137***
	(0.006)	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)
$Adopter_i \times After_{it}$	0.214***	0.217***	0.213***	0.215***	0.217***	0.212***	0.214***	0.216***	0.213***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)
User Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	996,846	1,991,589	978,324	993,348	1,980,705	957,948	977,448	1,932,774	937,188
R^2	0.406	0.401	0.408	0.405	0.400	0.406	0.405	0.400	0.406
Adjusted R^2	0.374	0.369	0.376	0.373	0.368	0.375	0.373	0.368	0.374
Residual Std. Error	0.661	0.654	0.658	0.660	0.653	0.658	0.657	0.648	0.655
Robust standard errors clustered by each user are in parentheses. *p<0.1; **p<0.05; ***p<0.01.									

Table 8. Variety of Matching Methods

Look Ahead PSM

We also employed a look-ahead matching technique as a robustness check to further compare the coefficients derived under the static and dynamic matching process. Under the look-ahead matching procedure, the selection of users includes an additional restriction requiring the matching candidates to be non-adopters at the time of matching but to be adopters at a future period. We picked the adopted user from the alternative user sample to make the comparison between the early adopters and later adopters. To execute this matching procedure, we create a new setup where the overall study period is divided into two parts by the cutoff date. Specifically, users who adopted autonomous vehicles after this date will become the matching candidates for the users who adopted autonomous vehicles before the cutoff date, and we only consider the first part of study period. In a similar vein, Table 9 shows the effect is 23% ($p < 0.01$), which indicates that the estimates under this method are substantively consistent.

	Number of orders		
	(I)	(II)	(III)
$PostAdoption_{it}$	0.736***		
	(0.012)		
$After_{it}$		0.190***	0.590***
		(0.014)	(0.014)
$Adopter_i \times After_{it}$		0.433***	0.226***
		(0.014)	(0.015)
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	286,418	205,184	233,874
R^2	0.529	0.508	0.505
Adjusted R^2	0.495	0.474	0.469
Residual Std. Error	0.563	0.598	0.614

Robust standard errors clustered by each user are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

Table 9. Look Ahead PSM

Placebo Test

We now conduct a placebo test by a random implementation to assess whether the main results arise spuriously. Specifically, we randomly apply the adoption to the user-month pairs and run model upon this "pseudo" adoption. Simultaneously, this analysis is replicated 1,000 times. Since the randomly assigned adoption are fake, a significant "effect" at the 5% level should be found at most 5% of the time (50 times). From the 1000 runs, we find that the fraction of simulations in which the null hypothesis is rejected is 4.7% of the time (47 times). These results suggest that our results are unlikely to be driven by random chance, and thus help reduce concerns regarding identification of the effects described.

Mechanisms

Although the previous sets of analyses and additional checks consistently support our finding that autonomous delivery increases the number of orders effectively, they do not inform us of how this relationship came about. To get a nuanced and deeper understanding of how autonomous vehicle adoption drives this result, we consider further tests to unveil the mechanisms of how autonomous delivery induced positive improvement of the orders and changes in consumers behaviors.

Operational Impact

There are several plausible reasons suggesting that the adoption of autonomous vehicles could lead to a positive impact on the number of orders. For users need to pick up packages, autonomous vehicles could to be an effective logistics tool in last-mile delivery.

On the one hand, it provides ease of access and convenience of time and geographical space by allowing users to obtain packages with no effort merely place orders on the phone. Consequently, it takes a shorter time for users to obtain desired packages, which means the lead time is decreased. We conduct a descriptive analysis by aggregating and averaging the package delivery time per user in each semester (e.g. Fall 2021). Figure 7 shows that lead time (i.e., from packages' arrival at the last-mile station to package receiving time) decreased by 10% (the decrease of 2.2 hours is relative to the initial lead time of 16.6 hours for adopters after matching), which confirmed the effectiveness and efficiency of the autonomous vehicles as a logistics tool. On the other hand, time flexibility offered by autonomous vehicles is highly possible to be the driving factor that affects the orders. Before the introduction of autonomous vehicles, users need to make the pick-up decision at their convenience, usually, pick-up occurred during the noon break (e.g. 12:00-13:00). When it comes to the autonomous vehicles, users could place orders whenever they preferred during the operational time leading

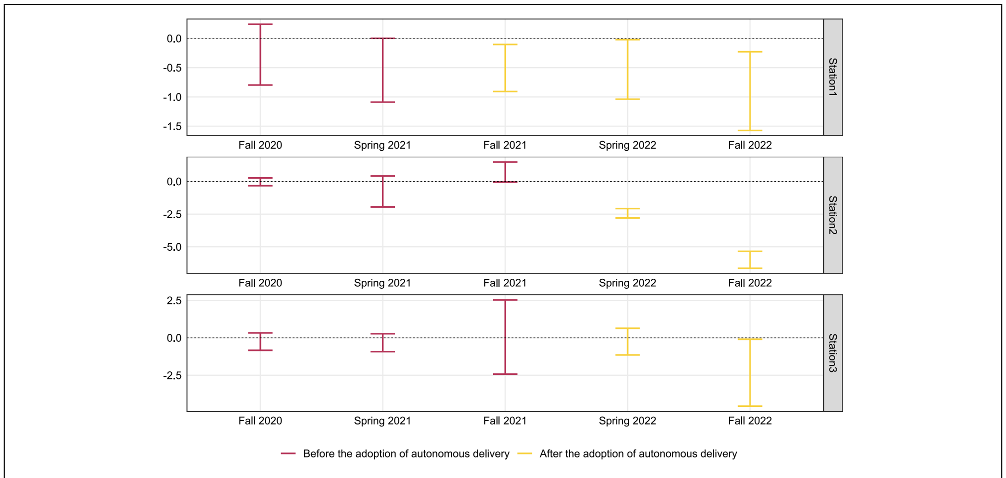

Figure 7. Lead Time Difference Between Adopters and Non-adopters

to dramatic changes of pick-up time distribution. To explicitly examine how this time flexibility affects consumers' pick-up decisions, we interact hourly dummies with the difference-in-difference variable and rely on the coefficient of this three-way interaction to capture the effects that autonomous vehicle adoption has on the number of orders hourly. It is noted that hourly fixed effects are added to this specification to account for hour-specific trends. We show the coefficients of the above analysis in Figure 8. Almost all hours of the day besides the hours in the early morning experience a significant increase in the number of orders at each hour after the adoption of autonomous vehicles.

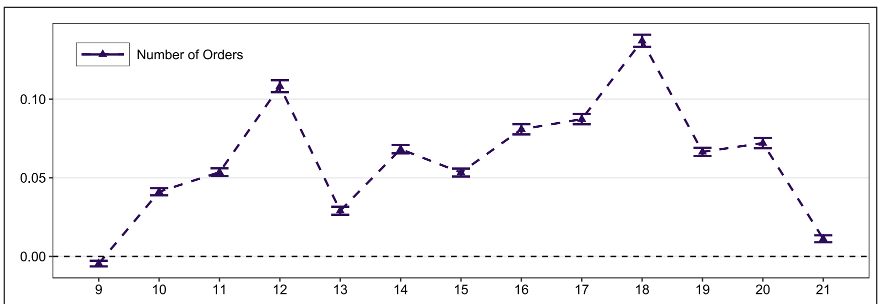

Figure 8. Coefficients of the Hourly Difference-in-Difference Estimates After Adoption

Consumer Responses

As mentioned above, autonomous vehicles lifted the space constraints for users. It is reasonable that the impact of autonomous delivery varies by how long it takes for the user to the campus station (i.e. travel distance). For our analyses, we bin the distance measures according to the 30% percentile into a binary variable $FarDistance_i$ that indicates the travel distance of the user is relatively far and interact with the difference-in-difference variable. Table 10 report that the users who have relatively far distance experience a much more increase in the number of order. This effect (4%) is statistically significantly different from the effects for the closer user.

Actually, autonomous delivery is charged by piece rate, consumers need to pay for 2.5 RMB for each package, nevertheless, if the package is from Alibaba's platforms, then it will be free of charge. Therefore, it is reasonable to suspect the increase in the number of orders is mainly from Alibaba's platforms. To conduct this analysis, we estimate a similar model where the outcome variable is the number of Tao's packages (package from Alibaba's platforms) or Non-Tao's packages. In Table 10, we see that users increase the number of orders from Tao by 21%($p < 0.01$), whereas we do not find evidence of effects when it comes to Non-Tao.

	Dependent variable:		
	(I) Number of orders	(II) Number of Tao orders	(III) Number of Non-tao orders
$After_{it}$	0.099***	0.155***	0.001***
	(0.007)	(0.006)	(0.0002)
$Adopter_i \times After_{it}$	0.206***	0.214***	−0.00001
	(0.009)	(0.007)	(0.0002)
$FarDistance_i \times After_{it}$	0.181***		
	(0.010)		
$Adopter_i \times FarDistance_i \times After_{it}$	0.038***		
	(0.015)		
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	996,846	996,846	996,846
R^2	0.408	0.406	0.065
Adjusted R^2	0.376	0.374	0.015
Residual Std. Error	0.660	0.661	0.023

Robust standard errors clustered by each user are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

Table 10. Consumer Responses

Human Couriers

In addition to examining the effects of autonomous last-mile delivery, we are interested in understanding how these effects compare to more traditional solutions, namely human couriers delivery. One unique aspect of our dataset is that before the deployment of autonomous vehicles, the last-mile station on each campus also hired a human courier to deliver packages for a few months. It creates an opportunity for us to estimate impact of human courier on last-mile delivery. To this end, we estimate a version of Equation (1) and Equation (2) coupled with static and dynamic matching where the interest is whether a adoption of couriers delivery. In this analysis, we consider study period which is prior the introduction of autonomous vehicles in order to avoid the confounding from the adoption of the autonomous delivery.

	Groups by adoption semester	Adoption month	No. of consumers
Station 1	Spring 2021	202105 – 202106	768
	Fall 2021	202110	386
	Non-adopter	–	51,235
Station 2	Spring 2021	202106	41
	Fall 2021	202110–202112	493
	Non-adopter	–	50,079
Station 3	Spring 2021	202105 – 202106	583
	Fall 2021	202109–202112	117
	Non-adopter	–	34,892

Table 11. Human Couriers Delivery Consumers by Groups

Table 11 summarize the number of consumers by human couriers adoption time. As can be seen, the sample size of couriers delivery adopters is extremely small with respect to the autonomous vehicles, and the duration of couriers delivery is just few months which is from May 2021 to December 2021. In spite of this, it still creates an opportunity for us to estimate impact of human courier on last-mile delivery and we underline that the utilization intensity of human couriers delivery during the duration indicated that the human

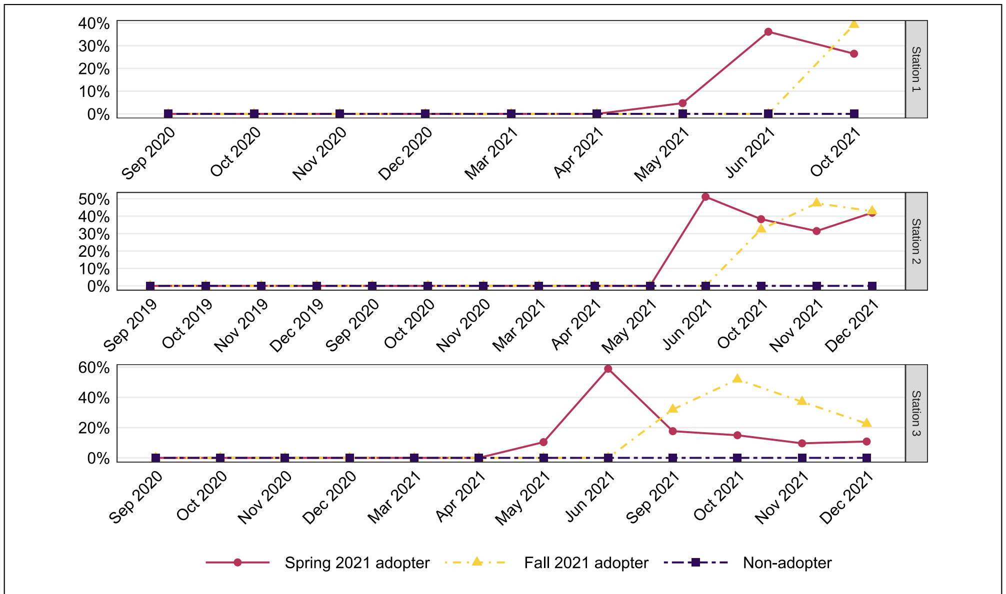

Figure 9. Proportion of Human Couriers Delivery Packages by Groups

couriers delivery adopters are comparable with non-adopters. Figure 9 shows that the surge in proportion of human couriers delivery packages occurred respectively in the early stage of adopters, then the trends went down slowly and flattened out over the next few months.

In Table 12, we examine whether human couriers delivery yields similar effects on the number of orders. We find that the value of last-mile delivery is 9% ($p < 0.01$) which is statistically significant but a smaller magnitude with respect to the coefficient in main result(21%, $p < 0.01$). Although not a direct comparison, the analysis suggests that the value of autonomous last-mile delivery does not come from the "delivery effect" itself. Autonomous vehicles, compared to human couriers, may drive additional value. Also we check the parallel trend assumption in a version of Equation (2-1) where the interest is whether a adoption of couriers delivery. It shows that our human couriers estimation results is convincing.

	Number of orders		
	(I)	(II)	(III)
$PostAdoption_{it}$	-0.057^{***}		
	(0.017)		
$After_{it}$		0.101^{***}	0.234^{***}
		(0.017)	(0.016)
$Adopter_i \times After_{it}$		0.237^{***}	0.092^{***}
		(0.017)	(0.017)
User Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	695,004	43,862	44,626
R^2	0.503	0.552	0.529
Adjusted R^2	0.456	0.515	0.491
Residual Std. Error	0.568	0.689	0.697

Robust standard errors clustered by each user are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

Table 12. Estimation Results (for Human Couriers Delivery)

References

Bapna, R., Ramaprasad, J., and Umyarov, A. 2018. "Monetizing freemium communities: Does paying for premium increase social engagement?," *MIS Quarterly* (42), pp. 719–735.

Brynjolfsson, E. and Mitchell, T. 2017. "What can machine learning do? Workforce implications," *Science* (358:6370), pp. 1530–1534.

Cao, J. and Qi, W. 2023. "Stall economy: The value of mobility in retail on wheels," *Operations Research* (71:2), pp. 708–726.

Carlsson, J. G. and Song, S. 2018. "Coordinated logistics with a truck and a drone," *Management Science* (64:9), pp. 4052–4069.

Caro, F. and de Tejada Cuenca, A. S. 2023. "Believing in analytics: Managers' adherence to price recommendations from a DSS," *Manufacturing & Service Operations Management* (25:2), pp. 524–542.

Claussen, J., Peukert, C., and Sen, A. 2019. "The editor vs. the algorithm: Targeting, data and externalities in online news," *Data and Externalities in Online News* (June 5, 2019).

Cohen, M. A. and Lee, H. L. 2020. "Designing the right global supply chain network," *Manufacturing & Service Operations Management* (22:1), pp. 15–24.

Cui, R., Li, M., and Li, Q. 2020. "Value of high-quality logistics: Evidence from a clash between SF Express and Alibaba," *Management Science* (66:9), pp. 3879–3902.

Cui, R., Li, M., and Zhang, S. 2022. "AI and Procurement," *Manufacturing & Service Operations Management* (24:2), pp. 691–706.

Cui, R., Lu, Z., Sun, T., and Golden, J. M. 2023. "Sooner or later? Promising delivery speed in online retail," *Manufacturing & Service Operations Management* (0:0).

Dietvorst, B. J., Simmons, J. P., and Massey, C. 2018. "Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them," *Management science* (64:3), pp. 1155–1170.

Fan, J., Tang, L., Zhu, W., and Zou, B. 2018. "The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce," *Journal of International Economics* (114), pp. 203–220.

Fu, R., Huang, Y., and Singh, P. V. 2021. "Crowds, lending, machine, and bias," *Information Systems Research* (32:1), pp. 72–92.

Gallino, S. and Moreno, A. 2014. "Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information," *Management Science* (60:6), pp. 1434–1451.

Gallino, S. and Moreno, A. 2019. *Operations in an omnichannel world,* Springer.

Gallino, S., Moreno, A., and Stamatopoulos, I. 2017. "Channel integration, sales dispersion, and inventory management," *Management Science* (63:9), pp. 2813–2831.

Ge, R., Zheng, Z., Tian, X., and Liao, L. 2021. "Human–robot interaction: When investors adjust the usage of robo-advisors in peer-to-peer lending," *Information Systems Research* (32:3), pp. 774–785.

Han, B. R., Sun, T., Chu, L. Y., and Wu, L. 2022. "COVID-19 and E-commerce operations: evidence from Alibaba," *Manufacturing & Service Operations Management* (24:3), pp. 1388–1405.

Hendricks, K. B. and Singhal, V. R. 2003. "The effect of supply chain glitches on shareholder wealth," *Journal of operations Management* (21:5), pp. 501–522.

Hendricks, K. B. and Singhal, V. R. 2005. "Association between supply chain glitches and operating performance," *Management science* (51:5), pp. 695–711.

Jung, J., Bapna, R., Ramaprasad, J., and Umyarov, A. 2019. "Love unshackled: Identifying the effect of mobile app adoption in online dating," *Mis Quarterly* (43), pp. 47–72.

Jussupow, E., Spohrer, K., Heinzl, A., and Gawlitza, J. 2021. "Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence," *Information Systems Research* (32:3), pp. 713–735.

Kesavan, S. and Kushwaha, T. 2020. "Field experiment on the profit implications of merchants' discretionary power to override data-driven decision-making tools," *Management Science* (66:11), pp. 5182–5190.

Lee, H. and Whang, S. 2001. "Winning the Last Mile of E-Commerce," *MIT Sloan Management Review* (42) 2001.

Liu, J., Abhishek, V., and Li, B. 2016. "The impact of mobile adoption on customer omni-channel banking behavior," in: *Proceedings of* 37^{th} *International Conference on Information Systems,* Dublin, Ireland.

Lu, T. and Zhang, Y. 2023. "$1 + 1 > 2$? Information, Humans, and Machines," Working Paper.
Luo, J., Rong, Y., and Zheng, H. 2020. "Impacts of logistics information on sales: Evidence from Alibaba," *Naval Research Logistics (NRL)* (67:8), pp. 646–669.
Luo, X., Qin, M. S., Fang, Z., and Qu, Z. 2021. "Artificial intelligence coaches for sales agents: Caveats and solutions," *Journal of Marketing* (85:2), pp. 14–32.
Luo, X., Tong, S., Fang, Z., and Qu, Z. 2019. "Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases," *Marketing Science* (38:6), pp. 937–947.
Morris, C. 2023. "Your groceries will get stocked by robots and delivered by self-driving trucks (if you shop with Kroger)," (available at https://fortune.com/2023/03/15/your-groceries-will-get-stocked-by-robots-and-delivered-by-self-driving-trucks-if-you-shop-with-kroger/; accessed Mar. 15, 2023).
Reed, S., Campbell, A. M., and Thomas, B. W. 2022. "The value of autonomous vehicles for last-mile deliveries in urban environments," *Management Science* (68:1), pp. 280–299.
Sun, J., Zhang, D. J., Hu, H., and Van Mieghem, J. A. 2022. "Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations," *Management Science* (68:2), pp. 846–865.
Sun, T., Shi, L., Viswanathan, S., and Zheleva, E. 2019. "Motivating effective mobile app adoptions: Evidence from a large-scale randomized field experiment," *Information Systems Research* (30:2), pp. 523–539.
Terwiesch, C., Ren, Z. J., Ho, T. H., and Cohen, M. A. 2005. "An empirical analysis of forecast sharing in the semiconductor equipment supply chain," *Management science* (51:2), pp. 208–220.
Vartabedian, M. 2022. "Starship, a Robot Food-Delivery Startup, Nets $100 Million," (available at https://www.wsj.com/articles/starship-a-robot-food-delivery-startup-nets-100-million-11646184462; accessed Mar. 2022).
Xu, J., Forman, C., Kim, J. B., and Van Ittersum, K. 2014. "News media channels: Complements or substitutes? Evidence from mobile phone usage," *Journal of Marketing* (78:4), pp. 97–112.
Xu, K., Chan, J., Ghose, A., and Han, S. P. 2017. "Battle of the channels: The impact of tablets on digital commerce," *Management Science* (63:5), pp. 1469–1492.
Xu, Y., Ghose, A., and Xiao, B. 2023. "Mobile Payment Adoption: An Empirical Investigation of Alipay," *Information Systems Research* (0:0).
Yalcin, G., Lim, S., Puntoni, S., and van Osselaer, S. M. 2022. "Thumbs up or down: Consumer reactions to decisions by algorithms versus humans," *Journal of Marketing Research* (59:4), pp. 696–717.
Zhang, M., Sun, T., Luo, L., and Golden, J. M. 2023. "Consumer and AI Co-creation: When and Why Nudging Human Participation Improves AI Creation." Working Paper.
Zhang, Y., Zheng, J., and Tan, Y. 2020. "From automobile to autonomous: Does self-driving improve traffic conditions?," Working Paper.