The impact of tradable rush hour permits on peak demand: evidence from an on-campus field experiment

Abstract:

Tradable permits have received growing attention as a new travel demand management intervention to manage rush-hour travel behavior and related negative social, economic, and environmental impacts. This study provides the first real-life evidence of tradable permits' ability to manage actual scheduling decisions in a congested morning peak. By conducting a 2-week field experiment with 91 students in Beijing, we investigate the effectiveness of the tradable permit scheme in terms of reducing "rush-hour" breakfasts, as well as the trading behavior of participants. The results of nested logit models show that the tradable permit scheme significantly reduces rush-hour breakfasts by about 20%. These results are robust to controlling for other factors, such as individual, commuting, attitudes, game- and market-related characteristics. Our results further suggest that participants are not perfectly rational when responding to the tradable permit scheme. This study informs policymakers regarding the design and implementation of a tradable rush hour permit scheme.

Keywords: Tradable permits, Field experiment, Behavioral response, Nested logit model

1 Introduction

Rush-hour travel behavior is one of the main concerns for transportation economics. Individuals' rush-hour behavior will lead crowd gathered in limited time and space, and then cause congestion and related negative impacts. For example, empirical evidence shows that traffic congestion can be significantly responsible towards more CO2 emissions (Bharadwaj et al., 2017). Researchers pay lots of attention to road traffic congestion given its severe social, economic, and environmental effects. However, congestion in a relatively small space, for example, campus canteen, also has negative impacts, such as stampede and other security risks (Tang et al., 2019a). In China, the rush-hour crowding often occurs in school canteens, although little attention has been paid to. In some schools, students have to wait in queue for about 45 minutes before they can have their lunch[1]. Some students even reported the canteen congestion issue to the government for solutions, since they need to wait for 7 minutes to get the food while the break time is only 15 minutes[2].

Actually, the formation of canteen congestion is in many ways similar to that in road traffic. The students have the role of drivers, and the food windows compare to lanes, reflecting overall capacity. When the food demand exceeds the canteen capacity, congestion will occur in the form of queues. Some studies described and explored the formation mechanism of the pedestrian flow in a canteen setting using theoretical models and simulations (e.g., Ravner, 2014; Tang et al., 2019a; Tang et al., 2019b). Yet the exploration on possible behavior interventions to solve this rush-hour travel behavior is limited. However, policy interventions proposed for managing traffic congestion can also be used in the canteen context, because congestion is determined

[1] http://news.sohu.com/20060919/n245410655.shtml, accessed at 14/03/2023.
[2] https://www.pds.gov.cn/contents/1027/19229.html, accessed at 14/03/2023.

by the rush-hour behavior of individuals and one effective way to change that is to encourage rescheduling. Interventions that can effectively reschedule students' rush-hour canteen behavior, may therefore also give valuable policy insights for managing dynamic peak road traffic.

To encourage peak-avoidance behavior, transportation researchers have proposed an efficient and effective policy solution: a congestion charge (early, seminal contributions include Pigou, 1920; and Vickrey, 1969). However, only a few cities have implemented congestion charging, which is likely due to the public's perception of it as an additional tax (Lindsey and Santos, 2020; Green et al., 2020). Tradable permits have received growing attention in the transportation literature as a policy alternative to congestion charge, with initial work by Verhoef et al. (1997) and many more recent contributions (see, e.g., Akamatsu and Wada, 2017; de Palma et al., 2018; Yang and Wang, 2011). An important advantage of a tradable permit scheme over congestion charge is that the former can easily be rendered revenue neutral, which ensures that there is no net financial flow from road users to the government or vice versa; this would likely increase public support for the policy.

Tradable permits schemes are, both in the practice of policy making and in textbook discussions, typically focused on firms. Important examples include the European Union Emissions Trading Scheme (EU-ETS) and the national Emissions Trading Scheme (ETS) in China. Fleming (1996) proposed that a personal tradable permits scheme can be a supplementary for the firm-level tradable permits scheme. However, the applicability for households or individuals remains an important empirical question. Most previous literature on tradable mobility permits (or credits) takes a theoretical or simulation approach; relatively few empirical studies have been conducted. Prior empirical work (such as Brands et al., 2020 and Tian et al., 2019) has used a lab setting and therefore "virtual" or experimental-game behavior to study tradable permits. In contrast, our paper contributes to the literature by being the first to investigate actual behavioral responses in a tradable peak permit scheme. We conducted a 2-week tradable permit experiment among 91 first-year students at Beijing Jiaotong University (BJTU) during their summer semester in July 2019. Students live on campus, and typically have breakfast in the canteen before their first lecture, which causes the canteen to be crowded during the morning "rush hour."

Participants were randomly assigned to one of two groups: They either started as the incentivized group in the first week and had no incentive in the second week or vice versa. Within each group, half of the participants received a relatively high monetary starting budget in the web application and a low number of permits, and the other half received a relatively low starting budget and a high number of permits. When incentivized, getting breakfast between 7:20 and 8:00 a.m. cost one tradable permit, but no permit was needed outside this time window. Participants were incentivized to trade smartly and avoid rush hour, since they would receive the remainder of their monetary budget at the end of the experiment. The market for permits followed the design proposed and described by Brands et al. (2020), in which permits can be bought and sold from a bank at a single price in a web application. This study applies that design—which proved successful in a lab setting with virtual mobility choices and preferences that were defined by the researcher by specifying payoffs—to a field application with real behavior and real preferences. In our setting, preferences for timing govern participants' actual behavior, and permits are introduced to affect that behavior.

We investigate the effectiveness of the tradable permit scheme in terms of reducing rush-hour breakfasts and the trading behavior of participating students. Our results indicate that the tradable permit scheme reduces the number of rush-hour breakfasts significantly, as intended, and that students mostly respond by rescheduling their breakfast—i.e., having breakfast before or after the peak. To our knowledge, this study is the first to provide real-life evidence on the effectiveness of tradable permits to manage rush-hour behavior. It corroborates the previously demonstrated effectiveness and behavioral insights on tradable mobility permits from theoretical work and lab experiments. Furthermore, our results support the notion that tradable permits could indeed be an effective measure for policymakers seeking to manage rush-hour travel behavior.

The remainder of the paper is structured as follows. Section 2 provides an overview of the relevant literature and Section 3 describes the experimental set-up. The data collected are discussed in Section 4, and in Section 5 we present the estimated econometric models and results. Section 6 concludes.

2 Literature

In recent years, starting with Verhoef et al. (1997), the use of tradable permits (also referred to as tradable credits) to address transportation externalities has received increasing attention from researchers and policymakers, due to their potential to combine effectiveness with social and political feasibility. Multiple studies have analyzed the efficiency of various categories of tradable permit schemes using theoretical models (see, e.g., Fan and Jiang, 2013; Grant-Muller and Xu, 2014). Yang and Wang (2011) have investigated the effect of a link-specific tradable credit scheme on equilibrium traffic flow in a setting with homogeneous travelers. Miralinaghi and Peeta (2016) use a multiperiod equilibrium modeling framework with the same assumption of homogeneous travelers and propose a multi-period link-specific credit scheme.

In pursuit of a more realistic evaluation of tradable permits, other recent studies have expanded such modeling frameworks by including heterogeneity in terms of the value of time (VOT). Wang et al. (2012) divided road users into different classes with different VOTs. They expand their own link-based tradable permit scheme, introduced by Yang and Wang (2011), by changing the uniform credit distribution to a user-class-based credit distribution. Xiao et al. (2013) propose a time-varying credit charge at the bottleneck and separately examine the equilibrium conditions and welfare effects of an optimal tradable credit scheme with identical and nonidentical commuters. Nonidentical commuters are represented by their VOT, which is a continuous function of income. Tian et al. (2013) further extend this work by solving a competitive two-mode bottleneck problem that incorporates both departure time choices and model split. Akamatsu and Wada (2017) use an equilibrium model to explore the properties of a tradable permits system in a general network equilibrium. They include a comparison in terms of the efficiency of tradable permits and a congestion charge, for both the case of perfect information and imperfect information. Their results suggest that tradable permits and a congestion charge can be made equivalent in the case of perfect information, but tradable permits can offer advantages if information imperfections exist.

Including heterogeneity in terms of VOT relaxes the strict homogeneity assumption and renders models more realistic. However, these models still abstract away from various behavioral biases that have been identified in the behavioral economics and cognitive psychology literature and may have an important effect on actual behavior in this context (Dogterom et al., 2017). Some theoretical work already includes certain aspects of individuals' behavior in the modeling, which provides new insights into the effects of tradable permits on travel behavior. Bao et al. (2014) use a predetermined amount of credits for each origin-destination (O-D) pair as the reference point for each user. If the amount of credits charged for a specific route is higher than this reference point, the user faces a loss; otherwise, they face a gain. User equilibrium and market equilibrium conditions have subsequently been examined while considering loss-aversion effects. Bao et al. (2016) model three groups of users with different VOTs. In line with the theory of mental accounting, different classes of users were modeled to frame or label the credit charge differently. The authors found that when they embedded travelers' framing or labeling of the use of credit, travel demand and credit prices were relatively high compared with conventional models that do not include this framing. Some authors also use behavioral insights in traffic assignment models with a tradable credit scheme. For instance, Han et al. (2020) incorporate cumulative prospect theory in their traffic assignment in a bimodal stochastic transportation network.

Although theoretical studies have already discussed several kinds of possible behavioral biases in the context of tradable permits, there is still a lack of empirical observations of travelers' behavioral patterns. Furthermore, prior empirical studies on tradable permit schemes predominantly rely on stated preference techniques. For example, Harwatt et al. (2011) interviewed 60 employees from the UK about personal carbon trading and provide data on respondents' stated change in travel distance and travel mode. Their results indicate that the behavioral response to such a scheme may be greater than when using increases in fuel prices. Dogterom et al. (2018 a, b) use an online stated adaptation experiment to evaluate public response to kilometer-based tradable driving permits. A total of 308 frequent drivers from the Netherlands joined this experiment and recorded their daily activities and travel patterns for a week. These studies provide empirical insights into the effectiveness of tradable permit schemes in road transport. However, given the limitations of stated preference, a static permit scheme is most commonly used. Such a static scheme does not include a dynamic permit market, and hence trading behavior has not been considered in these studies. However, the market interaction inherent to a tradable permit scheme could be its most important difference from a congestion charge or a license restriction scheme.

The emerging field of experimental economics offers an alternative way to analyze travel behavior along the lines of stated choice (Dixit et al., 2017). Besides allowing for the direct observation of human behavior, such experiments also enable the inclusion of market interactions (Smith, 1962). Some recent studies have used laboratory experiments with human subjects to explore tradable permit schemes. Aziz et al. (2015) conducted an online experimental game to study travelers' routes and departure time choices when subject to a personal travel carbon quota. Participants were recruited among graduate students from Purdue University and were divided into three income groups. Each group had a different number of work trips, shopping trips, and leisure trips per week, and the VOT corresponding to different travel purposes varied as well. Participating students were asked to choose the route and departure time for each trip for 5 weeks. At the end of each week, they could trade carbon allowances in a binary auction market. The results show that different income groups have different sensitivities to the carbon cost increase for different travel purposes. Low- and middle-income users are highly sensitive to the increase in carbon costs of non-work travel, and high-income people are less sensitive to the increase in the carbon cost of work travel. Tian et al. (2019) designed an online interactive experiment that allowed participants to interact extensively with each other and with intelligent virtual agents in the credit trading and route choice stages. The study uses a route-based tradable mobility credits scheme with an auction market. The results suggest that the collected data on responses to tradable mobility credits contain behavioral effects such as loss aversion, an immediacy effect, and a learning effect.

Compared with the lab experiments, which provide a virtual travel context to participants, field experiments can provide a more natural and familiar context. By studying behavior in real choice situations, rather than in virtual settings, field experiments can be expected to provide more representative and reliable insights into the workings of tradable permits in real contexts. In particular, participants trade off real determinants of actual utility, such as those related to scheduling preferences, against the incentives offered by a tradable permit scheme. Also, by directly manipulating the context and randomly grouping samples, field experiments can ensure comparability between treatment and control groups, which enables researchers to examine the pure treatment effect (Dixit et al., 2017; Bruhn and McKenzie, 2009). An important next step in research on tradable mobility permit schemes is collecting and analyzing revealed preference data on people subject to a tradable permits scheme, which we do in this paper.

3 Field experiment

3.1 Market design and application tests

In this study, we use the market design for tradable permits introduced by Brands et al. (2020), which operated well in their lab-in-the-field experiment with tradable parking permits. The market design uses a 'bank' that can be accessed via a web application that enables participants to buy and sell permits at the prevailing permit price anytime and anywhere with their smartphone. A simple algorithm is used to set the permit price. As shown in Eq. (1), the price-setting algorithm is a function of a prespecified target quantity Q for the specified time interval, during which the permits can be used (e.g., working week); the prevailing price at the time of the transaction; and a parameter that determines the size of the price change (δ).

$$P_t = \begin{cases} P_{t-1} + \delta & if\ Z_t > Q - U_t \\ P_{t-1} & if\ Z_t = Q - U_t \\ P_{t-1} - \delta & if\ Z_t < Q - U_t \end{cases} \qquad \text{Eq (1)}$$

The price dynamics further depend on the relationship between the number of permits in users' possession (Z_t) and the remainder of the target quantity, which equals Q minus the number of permits used up to that moment (U_t). When the number of permits in possession is more (less) than the remainder of the target quantity, the permit price increases (decreases) by the step size δ.

The advantage of the design is that it allows transaction costs, in terms of the time and effort participants need to invest, low compared with markets in which trading partners must be found by the participants themselves. It does so by having an easily accessible location at which trades can be conducted against a single price, while simultaneously limiting the possibility of unwanted speculation and manipulation. The latter is accomplished by introducing a small transaction fee, requiring that permits be traded one at a time, limiting the influence single individuals can have on the price, and limiting the maximum number of permits that users can own at each moment.

Drawbacks of the design are that budget neutrality is not guaranteed and the use of permits does not necessarily exactly equal the prespecified target quantity Q. However, we use the same design mainly because it is simple for users to understand— like it will be in real applications, which is important in a field experiment. A perfect budget neutrality market, such as auctions, will take longer time for travelers to find a proper seller or buyer, which is obviously unsuitable in this rush-hour context that in our case students are hurry to have lectures. A pilot study of personal carbon trading in Lahti city also uses such price dynamic algorithm rather than a perfect budget neutrality market (Kuokkanen et al., 2020).

Brands et al.'s (2020) market and application design have been adopted and adjusted for our experiment. We modify the values of several parameters of the algorithm and application screens to render them suitable for the context of tradable breakfast permits at the BJTU campus in Beijing, China. Additionally, two rounds of tests of the application were performed to ensure that it functioned well before the formal field experiment started (details can be found in Appendix C).

3.2 Recruitment

Recruitment was conducted during the third week of June 2019 on the BJTU campus. First-year students in the School of Economics and Management were invited to participate in the tradable permit experiment during their two-week summer semester. The experiment was advertised as a free breakfast event in which students were invited to participate in an application-based tradable permits market, enjoy a free breakfast, and have the opportunity to earn real money. The advertisement included a brief overview of the experiment and rules, along with a picture of the breakfast that would be provided. WeChat was used for recruitment. The invitation for the experiment was sent to several WeChat groups and shared among students. Of the 1,920 that viewed the invitation to participate, 117 students filled out the registration questionnaire. We called each of those students to ensure that they had reported correct contact information. In the end, 100 students were chosen to participate in the experiment.

After the recruitment, three teach-ins were organized to explain the experiment's rules and the application (details can be found in Appendix D). The pre-experiment survey was also conducted during the introduction period. In this survey, participants were asked about their usual departure time (from their dormitory) and breakfast time (specifically, when they entered the canteen), the perceived rush hour for breakfast in the canteen, average costs for a normal breakfast, and a stated preference (SP) question about their willingness to pay to have breakfast during rush hour (details can be found in Appendix E).

3.3 Field experiment and post survey

The field experiment was conducted at BJTU between June 1 and June 12, 2019, which is during the summer semester for first-year students. Students have morning lectures at 8:00 a.m. on some days and have different schedules, depending on their specific courses and program. During the experiment, we set up a temporary breakfast station which is at around 200 meters from the classrooms, and only allow take-away, to avoid disturbing non-participating other students or staff. The free breakfast was only provided at this breakfast station, while it was unnecessary for participants to pick up their free breakfast every day. When acting under the permit regime, participants were required to pay one permit to pick up their breakfast during rush hour (7:20 to 8:00 a.m.) on weekdays[3]. No permits were required to pick up their breakfast before or after rush hour, or not show up on that day.

The complete experiment consisted of two periods of five weekdays each. At the beginning of each period, each treated participant received a virtual monetary budget and a number of breakfast permits. The permit market was opened the Sunday before each period. Participants could freely trade breakfast permits on the mobile website and could use their personal budget to buy additional permits or increase it by selling some

[3] See Appendix E for the identification of the rush hour.

of their permits. With each purchase or sale, the market price increased or decreased by ¥0.01 and a transaction fee of ¥0.10 was charged.

Moreover, to avoid undesirable speculation, the maximum number of permits a participant could possess at any given time was controlled to the maximum number the participant can still meaningfully use at that time (Brands et al., 2020). In our case, the number of permits used for each rush trip is only one. So, this number was capped at the number of remaining morning peaks (i.e., 1 × the number of remaining morning peaks). For example, on Monday morning before the peak started, there were five remaining morning peaks in that week, so the maximum allowed number of permits is five. Similarly, on Wednesday morning before the peak started, the maximum allowed number of permits was three, which consisted of one for Wednesday peak and two for the remaining weekday peaks that week. The maximum thus decreased during the week. Permits were automatically sold by the application if users would otherwise end up with more permits in their possession than the maximum.

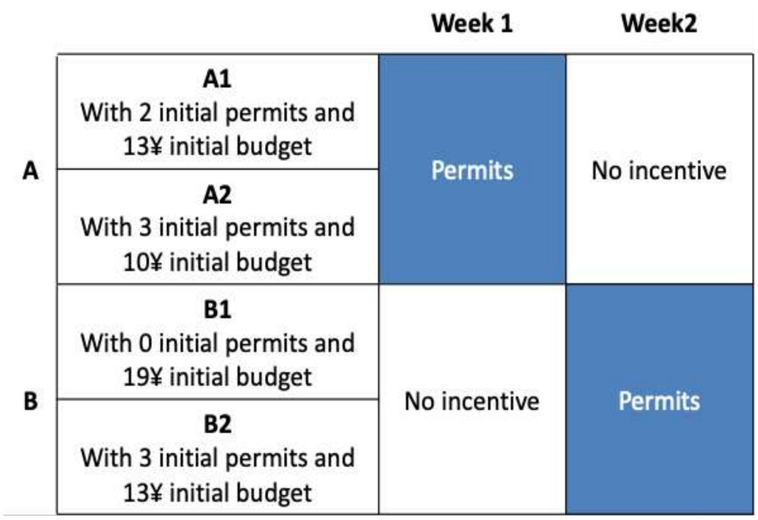

Figure 1: Endowments and schedule of groups

Students were randomly assigned to one of four groups—A1, A2, B1, or B2. Participants in groups A and B received a starting budget and a number of permits at the beginning of period 1 or period 2, respectively. In addition, groups A1 and group B1 received fewer permits than A2 and B2. However, the initial value of their endowments (number of permits * initial permit price + initial monetary budget) was the same, which mean that participants with more permits received less money to start with. Following the information collected from the pre-survey (see Appendix E), the initial permit price for the first week was set at ¥3, with students in group A1 receiving 2 permits and A2 receiving 3 permits. The initial permit price for the second week was set to ¥2, which was based on the price dynamics of the first week. Furthermore, the average number of permits per participant was 1.5 after considering the students' morning lecture schedule. The initial number of permits for students in group B1 was 0 and for those in B2 was 3. The starting budgets, number of initial permits, and trading week for each group are shown in Figure 1.

The no-incentive periods function as a reference and allow us to test the effectiveness of tradable permits to reduce the amount of peak scheduling of participants—i.e., in this specific application, having breakfast during the morning rush hour. In addition, the different allocations between groups A1 and A2 and groups B1 and B2 allow us to examine the impact of personal permit endowments on behavior. In addition, this design is compatible with induced value theory (Smith, 1976, 1982) to ensure internal and external validity. To incentivize participants to avoid rush hour and trade smartly, they were informed that the remainder of their personal budgets at the end of the experiment (after the two-week experiment and the final survey) would be transferred in real money to their personal account. At any moment during the experiment, participants could open the web application and see the prevailing price, the remainder of their personal budget, an overview of all their transactions, and the number of permits in their possession.

During the experiment, three research assistants worked at the breakfast station for each group (i.e., A or B). Each group has a separate desk and does not interfere with the other. Students could pick up their breakfast at the appropriate desk. There was another assistant standing meters before the breakfast station, to guide ways. When students arrive at the appointed desk, one assistant helps them to use the permit if they arrive during the rush hour, one records their arrival time manually, and the other one gives them breakfast. Hence, students in each group were asked to stand in a line and picked up their breakfasts one by one.

Because 9 students dropped out during the first week, the final sample size is 91 students: 25 students in A1; 22 in A2; 23 in B1; and 21 in B2. One week after the field experiment, a post-experiment survey was conducted to collect participants' feedback and attitudes toward tradable permits.

4 Data

4.1 Descriptive statistics

Demographics and descriptive statistics for the 91 participants are shown in Table 2. Of the 91 participants, 73 are female. This is in line with the student population of the School of Economics and Management of BJTU, in which female students account for a large proportion of the student population. Age varies little across participants, with the majority aged 19, and most have a monthly income between ¥1,000 and ¥2,000. Only 15 students disagree that they must have breakfast every day. When they have an early lecture, starting at 8:00 a.m., more than 75% depart from their dormitory during the defined peak, between 7:20 a.m. and 8:00 a.m. On other days, most will depart after the peak. Less than 30% think they can easily depart earlier than usual, and more than half state that it is easy to depart later. During each summer semester, first-year students who live on the east campus (which is relatively far from the breakfast station) are asked to move to the main campus (where the breakfast station is located). Thirty-four participating students changed dormitories and moved to the main campus during the weekend between the first and second weeks of our experiment.

Table 2: Descriptive statistics

Variable	Categories	Respondents
Gender	Male	18
	Female	73
Age	18	6
	19	66
	20	16
	21	3
Gross monthly income	$\leqslant 1000$	21
	$1000 < \& \leqslant 2000$	51
	> 2000	19
Degree of breakfast dependence	Disagree strongly	4
	Disagree	11
	Normal	24
	Agree	21
	Agree strongly	31
Usual departure time with an 8:00 am lecture	Pre	21
	Peak	69
	Post	1
Usual departure time without an 8:00 am lecture	Pre	7
	Peak	13
	Post	71
Ease of departing earlier than usual	Very hard	9
	Hard	26
	Normal	33
	Easy	16
	Very easy	7
Ease of departing later than usual	Very hard	4
	Hard	4
	Normal	32
	Easy	33
	Very easy	18
Change of dormitory during experiment	Yes	34
	No	57

Subjects were randomly assigned to four groups based on their student ID. The results of a balance check are reported in Table 3 and show that our random assignment of subjects does create balanced groups in terms of demographics, breakfast, travel habits, and dormitory.

Table 3: Sample Balance

	A1 (n=25)	A2 (n=22)	B1 (n=23)	B2 (n=21)	*p*
Gender	0.76 (0.44)	0.82 (0.39)	0.83 (0.39)	0.81 (0.40)	0.130
Age	19.28 (0.74)	19.32 (0.65)	18.96 (0.47)	19.14 (0.36)	1.825
Gross monthly income	1,658.00 (648.99)	1,688.64 (609.83)	1,708.70 (805.61)	1,819.05 (541.64)	0.247
Degree of breakfast dependence	3.96 (1.21)	3.32 (1.36)	3.83 (1.19)	3.67 (0.91)	1.259
Usual departure time with an 8:00 am lecture	446.88 (24.52)	451.14 (18.94)	451.57 (16.31)	455.95 (12.91)	0.879
Usual departure time without an 8:00 am lecture	518.24 (75.17)	520.82 (71.91)	530.30 (82.51)	523.33 (70.43)	0.112
Ease of departing earlier than usual	2.96 (1.17)	2.95 (1.17)	2.61 (1.03)	2.86 (0.91)	0.535
Ease of departing later than usual	3.40 (1.08)	3.55 (1.22)	3.74 (0.96)	3.86 (0.57)	0.951
Change of dormitory during experiment	0.72 (0.46)	0.73 (0.46)	0.83 (0.39)	0.81 (0.40)	0.378

Note: Departure time is measured by the number of elapsed minutes since midnight, e.g., 430 means 7:10 a.m.

4.2 Time choices and transactions

Figure 2 presents an overview of the distribution of breakfast pickups (referred to as "trips") over time for each group. The upper panel of Figure 2 shows the time choices of group A for both weeks and the lower panel shows the choices of group B (incentive group in week 1 and 2, respectively). Since the total number of trips varies between groups and between weeks (group A: 199 in week 1 and 184 in week 2; group B: 184 in week 1 and 171 in week 2), we use the percentage of trips as the vertical axis. The peak is marked by two gray dashed lines. The graphs show a clear difference between incentive and control weeks. A decrease in peak time trips and an increase in pre-peak and post-peak trips can be observed when participants were asked to use tradable permits. For group A, 55% picked up their breakfast during the peak in the incentive week and therefore used permits. The number was 73% in the control week, in which no permits were needed. For group B, the weekly percentage of peak trips follows a similar pattern, decreasing from 73% for the control week to 52% during the incentive week. Within each of the two weeks, the behavior of the incentive group is therefore also different from that of the control group. When comparing the behavior of both groups within the same week, the spikes are lower for the incentive group. This suggests that the incentive results in spreading out pickups over time. For example, in the first week, the maximum percentage is less than 16% in group A, but more than 20% in group B. Students seem to be incentivized to avoid the peak by departing earlier or later.

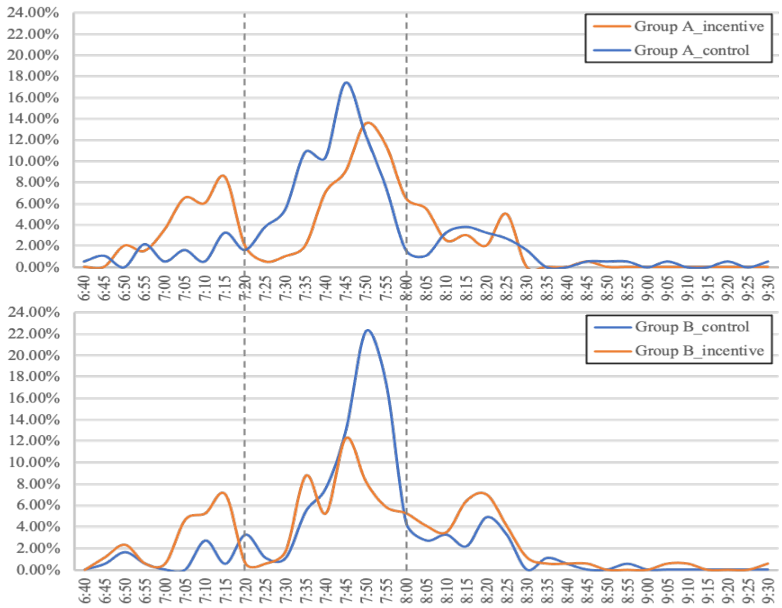

Figure 2: Distribution of time to pick up breakfast[4]

Note. The number of students every 5 minutes have been added up. The first interval is from 6:40 a.m. to 6:45 a.m. (not including 6:45 a.m.), and so on.

Figure 3 displays the cumulative number of buys and sells by participants and the price development over days for both experimental weeks. The left y-axis shows the number of permits (blue and orange bars) and the right y-axis shows the price of permits in Yuan (gray line). The cumulative buys and sells are shown in blue and orange bars separately, and the permit price is shown by the gray lines. The initial permit price of the first week is ¥3, which was based on the stated preference survey and assumed that

[4] Figure F1 shows similar results but uses the number of students at each minute as the vertical axis.

all students have five morning lectures per week. However, the reality is that students have different schedules and thus the number of their actual peak breakfasts is also different from what they stated on the survey. Hence, we did not find permit price fluctuations around the initial price, which was our expected equilibrium price based on the survey results. Whereas in week one the total number of permits initially allocated was larger than the total number of actual peak trips, in week two it was the other way around. Therefore, in week one the cumulative buys are less than the cumulative sells throughout, causing the permit price to decrease over time. In week two, the price increased gradually. Unlike Brands et al.'s (2020) study, in which the equilibrium price could be calculated beforehand—since the payoffs were determined by the researchers and hence known in advance—the price in our experiment was uncertain because preferences were not known exactly. We did not seem to reach an equilibrium price in this experiment. In other words, the initial permit allocation was so generous in the first week that the equilibrium price would be below both the initial value of 3 and the terminal value of ¥2.1. In contrast, it was so strict in the second week that the equilibrium price would be above both the initial value of 2.1 and the terminal value of ¥3.1. However, the permit price does reflect the intuitive relationship between market demand and supply.

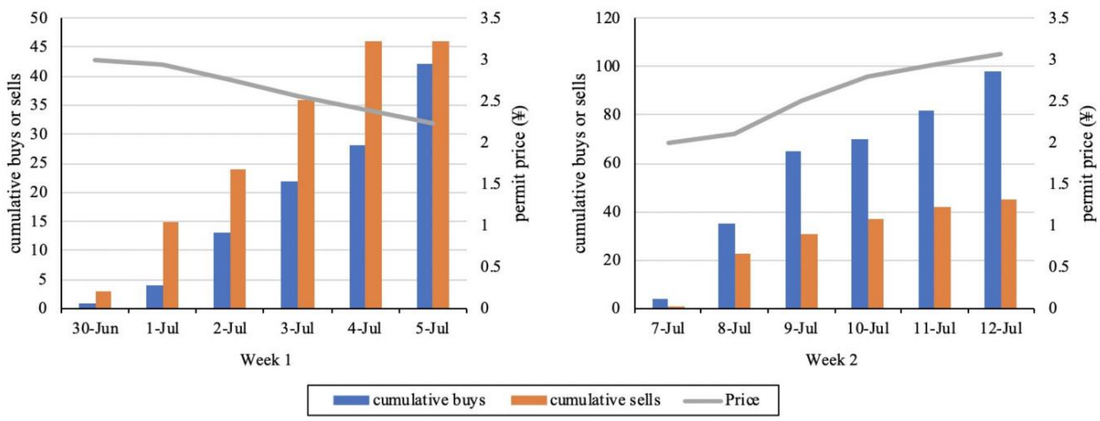

Figure 3: Cumulative transactions and price development

4.3 Feedback from participants

All participants were asked to answer questions about their attitudes toward the experiment and the tradable breakfast permits scheme. Most participants were positive about the experiment. In general, a large majority agreed that the breakfast (75%) and the web-based service (99%) provided during the experiment were good (giving a score of 4 or 5 on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). The mobile website worked well on their phones (66%). Almost all students had read the experiment rules (99%), watched the introductory video (92%), and read all notices in the WeChat groups (85%). In addition, they found the game (92%) and the rules (91%) to be clear and simple.

Table 4: Feedback on the tradable breakfast scheme

	Strongly disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly agree (5)	Average score
It was clear that I must use permits if I want to have breakfast during rush hour.	0%	0%	3.30%	32.97%	63.74%	4.6
It was clear that I could trade peak breakfast permits.	0%	0%	3.30%	34.07%	62.64%	4.59
I can easily determine when to pick my breakfast.	0%	5.49%	24.18%	34.07%	36.26%	4.01
I can easily determine whether it was best to buy or sell a permit.	2.20%	10.99%	37.36%	29.67%	19.78%	3.54
Participating in the game cost me little time or effort.	0%	10.99%	25.27%	48.35%	15.38%	3.68
The implementation of tradable breakfast permits would reduce congestion.	0%	5.49%	28.57%	46.15%	19.78%	3.8
I would be better off if tradable breakfast permits were implemented.	0%	7.69%	24.18%	46.15%	21.98%	3.82
All students would be better off if tradable breakfast permits were implemented.	2.20%	17.58%	46.15%	18.68%	15.38%	3.27
I view tradable breakfast permits as fair.	0%	1.10%	19.78%	56.04%	23.08%	4.01

As shown in Table 4, most students agree that the usage and trading rules of tradable breakfast permits were clear. The average score for these two statements is 4.6 and 4.59, respectively, on the 5-point scale described above. Most could easily determine when to pick up their breakfast (average score: 4.01). It was also relatively easy to decide whether it was best to buy or sell a permit, although the average score for this is somewhat lower (3.54). Overall, participating cost students little time or effort. These results are in line with the findings of Brands et al. (2020). However, unlike in their study, willingness to pay for a peak-time breakfast was not predetermined by the experimental design but rather varied among students and over time. Hence, the rationality of each student's decisions could not be observed directly from their behavior. Therefore, several questions were used to test the rationality of permit usage and transactions, as shown in Figures 4a, 4b, and 4c. Sixty-three percent scored more than 3 points on the statement that they could consciously decide to use permits, 53% agreed that they bought permits at a lower price and sold at a higher price, and more than 60% stated that they would wait to sell until the price was rising and wait to buy until the price was falling.

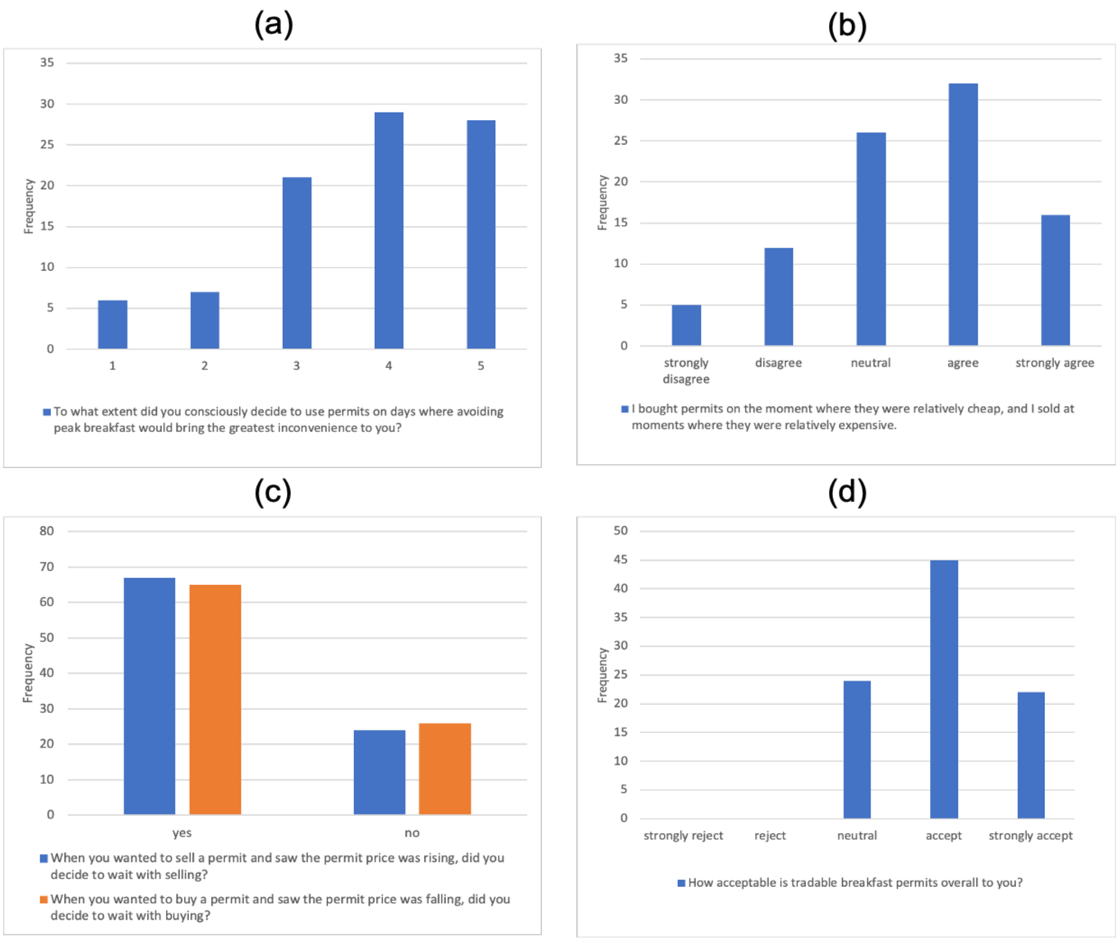

Figure 4: Rationality and acceptability

Finally, participants were also asked to respond to statements about general attitudes toward a tradable breakfast permits policy. Sixty-six percent agreed that the tradable permits scheme could reduce congestion, and 68% stated that they would be better off if a tradable permits scheme were implemented. Most were not sure about whether it would be beneficial for all students, and 31% stated that all students would be better off. Seventy-nine percent viewed the tradable permit scheme as fair. Also, as shown in Figure 4d, none of the participants rejected this kind of tradable permits scheme, and more than 70% would accept it.

5 Results

5.1 Effectiveness of the tradable permits scheme

5.1.1 Nested logit

We model the choices of students as discrete choices with four alternatives; pre-peak, peak, post-peak and no-show. The most basic utility functions for these alternatives, with alternative specific constants (ASC) to represent scheduling preferences and the incentive to avoid the peak in the form of tradable breakfast permits, are:

$$U_{pre} = ASC_{pre} + \beta_{incentive} Incentive + \varepsilon_{pre}$$

$$U_{peak} = ASC_{peak} + \varepsilon_{peak}$$

$$U_{post} = ASC_{post} + \beta_{incentive} Incentive + \varepsilon_{post}$$

$$U_{ns} = \beta_{incentive} Incentive + \varepsilon_{ns}$$

The no-show alternative is used as the reference category, which makes the ASC of the other three alternatives capture the average intrinsic preference for the specific alternative, relative to not showing up. The incentive is included as a variable that may influence the utility of the non-peak alternatives. The results of this most basic model can be found in Table B1. All included results have been produced using Pandas Biogeme (Bierlaire, 2020) to analyze the data we collected during the experiment.

We then allow the model for more flexibility, including variables from which we expect a priori that they influence the choices of participants. For example, whether it was rain on a specific day ($Rain$), whether students do not have a class start at 8:00 a.m. on a specific day ($Late\_Class$), and whether a student's dormitory is further away (D_{dorm}). Intuitively, we could expect there to be a nested structure in the choices students faced, with showing up in one nest of three alternatives (pre-peak, peak, post-peak) and not showing up in the other nest. When estimating nested logit models with different nesting structures, this is indeed the nesting structure that performs best. Using this nesting structure also improves on the MNL models (see Table B2), which is why the remaining results all use this nesting structure. The utility functions of the alternatives of the estimated model are:

$$U_{pre} = ASC_{pre} + \beta_{incentive} Incentive + \beta_{rain} Rain + \varepsilon_{pre}$$

$$U_{peak} = ASC_{peak} + \beta_{rain} Rain + \varepsilon_{peak}$$

$$U_{post} = ASC_{post} + \beta_{incentive} Incentive + \beta_{dorm\_post} D_{dorm} + \beta_{late\_post} Late\_Class + \varepsilon_{post}$$

$$U_{ns} = \beta_{incentive} Incentive + \beta_{late\_ns} Late\_Class + \varepsilon_{ns}$$

Results of the model above are presented in Table 5 and show that the incentive does indeed render non-peak alternatives (i.e., pre, post, and no-show) more attractive. All other estimated parameters have the expected sign. Having a class that starts late renders the no-show and post-peak alternatives relatively more attractive, as would be expected. Heavy rain makes it less likely that an alternative will be chosen (because rain only occurred during the pre-peak and peak in our experiment, it is included in the utility functions of these two alternatives only). Some students switched dormitories between the experimental weeks, from one that was farther from where breakfast was provided to one that was closer. Living farther away makes it more likely that the student will choose to pick up breakfast after the peak.

Table 5: Nested Logit model

	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
Incentive	0.525	0.161	3.26	0.001	0.164	3.21	0.001
dormitory_post	0.468	0.157	2.98	0.003	0.171	2.74	0.006
Lateclass_no_show	1.510	0.229	6.59	0.000	0.237	6.36	0.000
Lateclass_post	1.220	0.383	3.2	0.001	0.422	2.9	0.004
Rain	-1.150	0.359	-3.22	0.001	0.410	-2.81	0.005
Alternative Specific Constants (ASC)							
Peak	2.140	0.150	14.3	0.000	0.152	14.1	0.000
Post	0.703	0.378	1.86	0.063	0.412	1.71	0.088
Pre	1.300	0.224	5.82	0.000	0.234	5.58	0.000
MU_Show	2.180	0.668	3.26	0.001	0.719	3.03	0.003
Log Likelihood			-996.012				

AIC	2,010.025
BIC	2,053.346
Rho-square-bar	0.203

To better explain students' behavior, other control variables were added to the model in a stepwise manner. We included individual characteristics, game-related characteristics, and characteristics related to attitudes toward tradable permits.

Table 6: Nested model with other control variables

	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
Incentive	0.403	0.182	2.22	0.027	0.211	1.91	0.056
dormitory_post	0.340	0.165	2.07	0.039	0.199	1.71	0.087
Lateclass_no_show	1.810	0.246	7.37	0.000	0.262	6.93	0.000
Lateclass_post	0.809	0.381	2.12	0.034	0.466	1.74	0.082
Rain	-0.783	0.382	-2.05	0.040	0.486	-1.61	0.107
Individual and commuting characteristics							
Income_no_show	0.389	0.105	3.72	0.000	0.107	3.63	0.000
Brf_dep_no_show	-0.677	0.094	-7.2	0.000	0.094	-7.22	0.000
Brfscore_no_show	-0.492	0.131	-3.76	0.000	0.136	-3.62	0.000
Depearly_pre	0.159	0.080	1.99	0.046	0.094	1.69	0.091
Depl_depecl_pre	-0.163	0.076	-2.13	0.033	0.091	-1.78	0.075
Depl_no_show	-0.258	0.104	-2.47	0.014	0.101	-2.54	0.011
Deplcl_lclass_no_show	-1.360	0.509	-2.68	0.007	0.481	-2.84	0.005
Game-related characteristics							
Earnmoney_no_show	-0.661	0.171	-3.85	0.000	0.170	-3.9	0.000
Reexplan_peak	-0.694	0.382	-1.82	0.069	0.367	-1.89	0.058
Renotice_no_show	-1.090	0.258	-4.24	0.000	0.264	-4.14	0.000
Website_no_show	0.227	0.104	2.18	0.030	0.111	2.05	0.0401
Tradable permit attitudes-related characteristics							
Accep_no_show	0.304	0.172	1.77	0.077	0.170	1.8	0.073
Allbetter_no_show	-0.347	0.124	-2.8	0.005	0.124	-2.79	0.005
Ibetter_no_show	0.522	0.153	3.4	0.001	0.148	3.53	0.000
Alternative Specific Constants (ASC)							
Peak	-3.130	1.090	-2.87	0.004	1.090	-2.88	0.004
Post	-4.800	1.120	-4.27	0.000	1.160	-4.13	0.000
Pre	-4.520	1.080	-4.17	0.000	1.100	-4.12	0.000
MU_Show	3.470	1.620	2.14	0.032	1.960	1.77	0.077
Log Likelihood			-856.8451				
AIC			1,759.69				
BIC			1,870.399				
Rho-square-bar			0.303				

5.1.2 Individual and commuting characteristics

The model improves in terms of AIC/BIC and $\overline{\rho}^2$ if we include characteristics such as income, whether a student stated being able to depart earlier or later than usual, and their dependence on breakfast. Importantly, the estimated coefficient for the incentive is not considerably affected by including these other control variables. The description of each variable used in Table 6 can be found in the Appendix A.

During the experiment, an incentive was in place to motivate participants to show up. To receive the earned rewards after the experiment ended, students were required to show up at least once during the incentive week. Furthermore, a penalty of ¥1 was imposed on not showing up on 2 or more days during the incentive week, which was deducted from their final budget. Students with high income may care less about this penalty and about the provided breakfasts' being free, which may explain why our results show that students who have higher monthly income are less likely to show up during the experiment. Students who stated that they need to have breakfast are less likely to choose the no-show alternative—i.e., they are more likely to show up. Students who rate the provided breakfast highly are also more likely to show up.

Flexibility in terms of how easily a student can depart early or late from their dormitory in the morning is also likely to affect their choices. However, realized choices can also affect the perception of one's own flexibility and result in simultaneous causality. Therefore, the coefficients on whether a student could easily depart early or late cannot be interpreted causally. They do, however, capture a pattern in the data, which shows that those who state that it is easy to depart earlier than usual choose the pre-peak alternative relatively often. On the other hand, students who regularly depart during the peak when they have an early class choose the pre-peak alternative less often if they can easily depart later than usual. A regular departure time when a student has a later class has significant effects on their choices. The interaction between a late departure time and a late class shows that students who usually depart at the peak when they have a late class are more likely to show up on a day with late classes.

5.1.3 Game-related characteristics

Our inclusion of questions about understanding that real money can be earned, having read the instructions, having read notices on WeChat, and whether the website worked well also does not considerably affect estimates for the incentive. Students for whom it was clear they could earn real money from the experiment were less likely to choose the no-show alternative. Students who read the instructions before the experiment were more likely to avoid the peak. Those who read the daily notices in the WeChat group had a higher tendency to show up during the experiment. However, the website's functioning well has an unexpected sign: Those who believed the application worked well were less likely to show up during the experiment. This may be because students who do not show up may have used the application less frequently and are therefore less likely to have experienced problems with it. Since the online application was constructed in the Netherlands and hosted on a server there, sometimes the connection was not stable for Chinese users. Some students could not open the website and failed to click the "Use Permit" button at the right time. Students who did not show up, of course, would not experience such a problem.

5.1.4 Tradable permit attitudes-related characteristics

Because we considered attitudes toward tradable permits as possible determinants of behavioral responses, we also included the scheme's acceptability to participants and their views on the benefits of tradable permit schemes. The estimated coefficient for the incentive is also robust to including these variables. The data show that students who have a high level of acceptability for tradable breakfast permits are more likely to choose the no-show alternative. This is, however, only significant at the 10% level and cannot be interpreted causally, since having an attractive outside option may influence their attitude toward tradable permits instead of their choices' being influenced by their attitude. Similar caution is warranted for interpreting estimates of the statements "All students would be better off if a tradable breakfast permits scheme were implemented" and "I would be better off ...", which have opposite signs. Students who believed that all students would be better off chose the no-show and pre-peak alternatives less often, whereas those who believed that they themselves would be better off chose the no-show or pre-peak alternatives more often. Causality can run either way here or in both directions.

After including all of these other variables that help to further explain the students' behavior, we see that the incentive was effective in motivating students to avoid the morning rush hour.

5.1.5 Permit market-related characteristics

We also tested the effect of variables related to permit trading on peak behavior. To focus on students' behavior when using tradable permits, we only use data from incentive weeks (Group A in week 1 and Group B in week 2) in the following part. Since we only use incentive-week data, the variable *incentive* is no longer included in these models.

The model in which we include students' trading activity is presented in Table 7 Model 1. The results show that students who are active on the application are also less likely to choose the no-show alternative. Students who have more purchases on a specific day are more likely to show up during the peak (or alternatively, those who are more likely to show up during the peak tend to buy additional permits). In contrast, students who have more sales on a specific day will be more likely to choose one of the other alternatives. The higher the daily average permit price, the more likely it is that students choose the no-show alternative. The number of initial permits and initial monetary budget do not significantly affect students' behavior.

Table 7: Nested models using data from incentive weeks

	Model 1		Model 2		Model 3		Model 4	
	Value	Rob. S.D.	Value	Rob. S.D.	Value	Rob. S.D.	Value	Rob. S.D.
dormitory_post	0.473*	0.266	0.476*	0.249	0.388	0.25	0.415	0.259
Lateclass_no_show	1.76***	0.461	1.59***	0.45	1.48***	0.441	1.64***	0.444
Lateclass_post	2.47***	0.768	2.11***	0.745	1.96***	0.721	2.27***	0.714
Rain	-2.17***	0.59	-1.87***	0.622	-1.79***	0.664	-2.12***	0.684
Dactive_no_show	-0.106**	0.044	-0.107**	0.044	-0.106**	0.044	-0.105**	0.044
Dbuy_peak	1.28***	0.463	1.27***	0.465	1.08**	0.437	1.24***	0.454
Dsell_no_show	1.35*	0.703	1.3*	0.668	1.16*	0.656	1.27*	0.668
Dsell_post	1.37**	0.647	1.31**	0.594	1.18**	0.583	1.3**	0.601
Dsell_pre	1.51**	0.704	1.42**	0.645	1.27**	0.625	1.42**	0.646
Davprice_no_show	1.13**	0.534	1.12**	0.529	1.15**	0.52	1.19**	0.521
Inipermit_peak			0.203**	0.088				
Loss_rfini_peak					-0.74**	0.291		
Loss_rffu_peak							-0.36	0.253
Day2_cumuse_peak					1.24***	0.462	1.21**	0.473
Day3_cumuse_peak					0.413*	0.244	0.446	0.275
Day4_cumuse_peak					0.753***	0.282	0.733***	0.267
Day5_cumuse_peak					0.472***	0.176	0.421***	0.158
Alternative Specific Constants (ASC)								
Peak	3.84***	1.44	3.45**	1.48	3.64**	1.42	3.72***	1.42
Post	2.36	1.61	2.58	1.61	2.74*	1.55	2.64*	1.54
Pre	3.56**	1.47	3.65**	1.46	3.73***	1.42	3.75***	1.42
MU_Show	1.08***	0.318	1.26***	0.415	1.3***	0.443	1.14***	0.333
Log Likelihood	-488.2878		-485.5271		-475.1219		-476.7355	
AIC	1,004.576		1001.054		988.2437		991.471	
BIC	1,062.26		1062.859		1066.529		1069.757	
Rho-square-bar	0.204		0.206		0.217		0.214	

5.2 Behavior biases

5.2.1 Test of rationality

In the absence of transaction costs, permits and money are completely interchangeable in a standard economic model with rational agents. This would imply that receiving relatively many permits and little budget or relatively few permits and more budget should not affect participants' choices. In our experiment, participants who have more (fewer) permits initially also receive less (more) money, such that the monetary value of the total endowment given the starting price (the initial budget plus the number of permits multiplied by the starting price) is the same. The idea from standard theory would be that this initial allocation should not affect participants' behavior, since they could have the same distribution of permits and budget by simply trading at the beginning. However, as shown in Table 7 Model 2, the initial number of permits does affect participants' choices: The number of initial permits has a significant positive effect on the probability of choosing to pick up breakfast during the peak. One possible explanation for this is the divergence between the realized permit prices and the initial price, which could mean that the observed difference in choices is a result of an income effect. Since the week-average permit price is either higher or lower than the initial price, the total monetary value of endowments may differ across participants. Participants with a higher value endowment could use more permits and be more likely to choose the peak alternative. However, the difference between the week-average permit price and the initial price is not very large, which makes it unlikely that this small difference results in participants' valuing their permits differently. Furthermore, including the total monetary value of the endowment as an explanatory variable does not result in significant estimates, and therefore does not seem to considerably influence participants' choices.

Another possible explanation could be that respondents have an inequivalent valuation between permits and money. Since the initial number of permits has a significant positive effect, students with more permits would be more likely to show up at the peak. This implies that travelers (in this case, students) spend the permits they received initially more easily than their money, which suggests that participants value permits less than their market price. This result is in line with the assumption made by Bao et al. (2016). Although both the permits and initial budget can be regarded as a windfall for participants in our experiment, permits are valued less than their identical amount of out-of-pocket money.

5.2.2 Tests of reference dependence

An important concept in behavioral economics is reference dependence. According to Kahneman and Tversky (1979), because of limitations on decision-makers' ability to cognitively solve difficult problems, their preferences are not determined by states of wealth but by changes relative to a reference point. The relative gain or loss situation will then affect the decision-maker's utility and choices. Reference-dependent behavior has received attention in the transportation literature (see, e.g., Mabit et al., 2015; Borger and Fosgerau, 2008; Li and Hensher, 2015; 2008; 2013). Its influence on individuals' responses to transport policies has also been discussed before, also in the context of tradable mobility permits (see, e.g., Bao et al., 2014; Dogterom et al., 2017; Tian et al., 2014; 2017; 2019). In this section, we examine the reference dependency of participants' choices.

We start with a simple model that uses a static exogenous reference point, which is participants' initial permit budget (the number of permits they are endowed with at the start), and assume that participants make decisions only based on past experience. In this simple model, if the total number of permits consumed is fewer than the initial permit budget, participants face a gain relative to their reference point. In contrast, if the total number of permits consumed is more than the initial permit budget, participants face a loss. Then, the gain or loss situation for participant *i* during their decision process on each day can be defined as the difference between their initial number of permits and their cumulative permit usage so far, which is shown as Eq. (5.1).

$$d_{in} = K_i - \sum_{j=1}^{n-1} k_{ij} \qquad (5.1)$$

$$Loss_{ini} = \begin{cases} 0, & if\, d_{in} \geq 0 \\ 1, & if\, d_{in} < 0 \end{cases} \qquad (5.2)$$

where K_i denotes the initial number of permits of student i and equals 0, 2, or 3 in our case. k_{ij} equals 1 if student i shows up during the peak on day j. If their cumulative permit usage is less than the initial number of permits, students face a gain. If their cumulative permit usage is more than the initial number of permits, they face a loss.

The dummy variable $Loss_{ini}$ (as shown in Eq. (5.2)) has been included to capture possible reference dependence. However, given that participants made time choices for a week rather than one day, the reason a participant who is in a gain/loss situation is less/more likely to choose the peak on that day could be that they did not use many permits before and will continue in this behavior, rather than adjusting to the gain/loss situation they face on that day. It is important to separate the gain/loss effect on daily decisions from the persistence of their behavior. Hence, the interactions of day-of-week and the number of permits used so far have been included in the model to capture participants' specific preferences for permit use. The results are shown in Table 7 Model 3. When controlling for participants' persistence of behavior, the daily gain/loss situation compared with the initial permits budget still affects decisions. Students who face a loss situation are significantly less likely to show up during the peak.

However, participants may be less myopic than we assumed above. They may consider the entire experimental week and take future consumption of permits into account when making current decisions. Therefore, we also test the effect of a dynamic endogenous reference point, which is defined as Eq. (5.3). We use the expected monetary value of the expected future permit consumption of each student as a reference point. The expected monetary value of the expected future permit consumption φ_{in} of student i on day n is defined as

$$\varphi_{in} = \frac{\sum_{j=1}^{n-1} k_{ij}}{n-1} * [N_r - (n-1)] * p_{n-1} \tag{5.3}$$

where k_{ij} denotes whether student i uses a permit on day j, and equals 1 if student i picks up their breakfast during the peak on day j. N_r denotes the total days in one tradable permit period and equals 5 in our case. p_{n-1} denotes the average permit price over the previous day. The gain or loss is calculated as

$$g_{in} = p_{n-1} * \left(K - \sum_{j=1}^{n-1} k_{ij} \right) - \varphi_{in} \tag{5.4}$$

$$Loss_{fu} = \begin{cases} 0, & if\, g_{in} \geq 0 \\ 1, & if\, g_{in} < 0 \end{cases} \tag{5.5}$$

where g_{in} can be explained as the difference between the monetary value of participants' current number of permits owned and expected future permit consumption; this captures the fact that participants not only look back but also look ahead. Given their experience on past days, they can predict their future permit usage, which is used as the reference point on a given day. If their current endowment on that day could cover their future usage, they will face a gain; otherwise, they will suffer a loss.

The definition of this reference point is in line with Tian et al. (2019), who conducted a lab experiment on tradable permits. A small difference is that in their study, they use $\sum_{j=1}^{n} k_{ij}$ instead of $\sum_{j=1}^{n-1} k_{ij}$ in calculating both the reference points (Eq. (5.3)) and the gain/loss (Eq. (5.4)). However, in our experiment, the effect of the gain/loss situation should influence participants' decisions before they make daily time choices, and thus we use $\sum_{j=1}^{n-1} k_{ij}$ in the equations to calculate the situation participants face before their daily decision. Table 7 Model 4 shows that including these definitions of losses and gains does not improve the model: $\overline{\rho}$ does not change substantially, while the AIC and BIC increase and decrease, respectively.

Given these results, we can infer that students' behavioral responses to tradable permits are not as complicated as we supposed. Unlike the results in Tian et al. (2019), participants in our study simply used their initial permit endowment as the reference point, rather than as a dynamic reference point, which requires considering both past and future behavior and therefore entails more complex calculations. Although participants do show behavioral biases, such as reference dependence, the model without this consideration can well explain their behavior. The increase of $\overline{\rho}^2$ from Table 7 Model 3 to Model 4 is slight, which implies that participants in our study are nearly rational when using permits.

6 Conclusion and discussion

It is widely recognized that both toll-based and quantity-based transport demand management measures have pros and cons. Modern technology provides a chance for transport policymakers to combine the advantages of both a congestion-charge system and a quantity-based policy, such as license plate restrictions, in a tradable permits scheme. Quantity control characteristics can then be combined with the freedom of trading in a market. Such schemes have been applied in the environmental sector for many years. Unlike most applications of tradable permits scheme in the environmental sector, a tradable mobility permits scheme focuses on affecting individuals' behavior, rather than that of firms. This implies that also the market for permits will be populated by individuals rather than firms, and it remains to be seen to what extent the trading and use of permits will then comply with textbook expectations of rational utility maximizing behaviour, or instead will be more random, e.g., due to a lack of understanding of the system. Academic interest in the use of tradable permit schemes to manage transport issues has been growing over recent years, especially for road traffic congestion. Many studies have examined the efficiency of tradable permit schemes in diverse hypothetical contexts using various theoretical approaches.

Nevertheless, more empirical evidence is needed to understand the performance of tradable permit schemes in reality.

This study contributes to the literature by providing the first real-life evidence of tradable permits to manage rush-hour behavior by applying a system that was tested in a lab environment to a real application, with actual scheduling decisions during the morning peak. We conducted a 2-week field experiment in July 2019 among a group of students from Beijing Jiaotong University in order to test the effectiveness of a tradable rush-hour permits scheme in natural circumstances in which the participants normally experience congestion during rush hour. Specifically, participants were directed to use one permit if they picked up their breakfast during the predefined rush hour.

Our results indicate that the proposed tradable permits scheme effectively manages rush-hour mobility choices. A noticeable drop in the number of peak trips was observed in each incentive group each week. Compared with the control weeks, about 20% of peak trips were avoided (by departing earlier, later, or not showing up) when using the tradable permits scheme. One limitation is that we didn’t estimate the real waiting time reduction, which given the small sample size relative to total student numbers will be negligible. Furthermore, more than 70% of participants believe that the tradable permits scheme is acceptable, and nearly the same number of participants believe that they would be better off under such a scheme. Participants also had a positive attitude about the effectiveness and equity of the tradable permits scheme.

We also investigated participants’ heterogeneous responses and several kinds of behavioral biases that may occur when using tradable permits. First, participants with different residential locations, schedules, flexibility to change their departure time, and regular departure times had different responses to the tradable permits scheme. When designing a tradable rush-hour permits scheme for road traffic, a targeted permits allocation plan that differs among locations and employers can be considered. Future studies can be conducted to test whether specific ways of allocating permits across participants could further improve effectiveness. Weather also affects participants' departure time choices, and therefore affects their behavioral response to the tradable permits scheme. An implication for policymakers is that equilibrium permit prices can be expected to vary over seasons, and within seasons over days, whether or not quantities are dynamically optimized to reflect changing societal scarcity conditions. Such price variation itself is in fact an efficient property of permit prices, and itself no reason to worry.

Second, we observed some form of endowment effect or mental accounting and reference dependence in this experiment; for instance, participants reported different perceived values for permits and their equivalent market price. We share Bao et al.'s (2014) concern that a windfall label might render permits less effective than an identical road-pricing scheme, since it may induce more travel demand. One solution to deal with this would of course be to issue fewer permits. However, when considering the labour market meanwhile, a smaller demand reduction than under road pricing in fact may be desirable under pre-existing labour taxes. Further research can examine the possible welfare effects of these biases and test whether they occur on the road among road users and also over a longer period of time.

The web-based permit market works well in this field experiment. According to the app data recorded during the experiment, the trend in price dynamics reflects market demand. According to the survey results, most participants believed that they fully understood the rules of the tradable permit scheme and could easily and rationally trade in the permit market. The test of market-related variables using only incentive-week data further showed that the number of participants' purchases and sales has a rational relationship with their revealed behavior. Unlike an auction market in which participants could form the permit price by themselves from the start, the permit price in our study, which uses a bank, starts from an exogenous initial price. If the initial price deviates substantially from the final equilibrium price, the permit price may have a monotonous fluctuation in the short term, as we found in this experiment. However, such variations in the short term are to be expected. For instance, policymakers can foresee variations when peak demand varies with weather conditions. The policy would be efficient when the equilibrium price range is stable in the long run (Brands et al., 2020), while varying with changing scarcity conditions. As what we find in the study, although the price always decreases or increases during the respective incentive weeks, the marginal change in the price per unit of time decreases over time. We can expect stability if exogenous demand and supply factors are sufficiently stable; certainly when demands are downward sloping while user cost is upward sloping, as it would be under congestion.

Although the respondents in our research consisted of a small group of students and the behavioral response and attitudes from car users may be quite different, these results are encouraging. It is encouraging to see that tradable permits can indeed impact participants' response to congestion, change their behavior, and more effectively manage rush-hour travel demand. The results are promising for the application of tradable mobility permits in a real traffic context, and provide grounds for further experimentation by researchers and policymakers. Future studies can expand our findings by using worker subjects with heterogeneous characteristic (such as income ranges, car ownership, etc.), focusing on other strategic behavior (such as mode choice, route choice), exploring more and other types of permit applications, and having a larger sample size to estimate the real congestion reduction by tradable permits.

Acknowledgments

This study was supported by the Chinese National Natural Science Foundation (72071017), Capital Think Tank Project (2022ZKKT011) and Chinese Scholarship Council (CSC) Scholarship (202007090026).

References

Akamatsu, T. and K. Wada (2017). Tradable network permits: A new scheme for the most efficient use of network capacity. *Transportation Research Part C: Emerging Technologies 79*, 178–195.

Aziz, H. A., S. V. Ukkusuri, and J. Romero (2015). Understanding short-term travel behavior under personal mobility credit allowance scheme using experimental economics. *Transportation Research Part D: Transport and Environment 36*, 121–137.

Bao, Y., Z. Gao, M. Xu, and H. Yang (2014). Tradable credit scheme for mobility management considering travelers' loss aversion. *Transportation Research Part E: Logistics and Transportation Review 68*, 138–154.

Bao, Y., Z. Gao, and M. Xu (2016). Traffic assignment under tradable credit scheme: An investigation considering travelers' framing and labeling of credits. *IEEE Intelligent Transportation Systems Magazine 8*(2), 74–89.

Bharadwaj, S., S. Ballare, Rohit, and M. K. Chandel (2017). Impact of congestion on greenhouse gas emissions for road transport in mumbai metropolitan region. *Transportation Research Procedia 25*, 3542-3555.

Bierlaire, M. (2020). A short introduction to pandasbiogeme. Technical report TRANSP-OR 200605, Transport and Mobility Laboratory, ENAC, EPFL.

Borger, B. D. and M. Fosgerau (2008). The trade-off between money and travel time: A test of the theory of reference-dependent preferences. *Journal of Urban Economics 64*(1), 101 – 115.

Brands, D. K., E. T. Verhoef, J. Knockaert, and P. R. Koster (2020). Tradable permits to manage urban mobility: market design and experimental implementation. *Transportation Research Part A: Policy and Practice 137*, 34–46.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics 1*(4), 200 – 232.

de Palma, A., S. Proost, R. Seshadri, and M. Ben-Akiva (2018). Congestion tolling-dollars versus tokens: A comparative analysis. *Transportation Research Part B: Methodological 108*, 261–280.

Dixit, V. V., A. Ortmann, E. E. Rutstr¨om, and S. V. Ukkusuri (2017). Experimental economics and choice in transportation: Incentives and context. *Transportation Research Part C: Emerging Technologies 77*, 161–84.

Dogterom, N., D. Ettema, and M. Dijst (2017). Tradable credits for managing car travel: a review of empirical research and relevant behavioural approaches. *Transport Reviews 37*(3), 322–343.

Dogterom, N., D. Ettema, and M. Dijst (2018a). Activity-travel adaptations in response to a tradable driving credits scheme. *Transport Policy 72*, 79–88.

Dogterom, N., D. Ettema, and M. Dijst (2018b). Behavioural effects of a tradable driving credit scheme: Results of an online stated adaptation experiment in the netherlands. *Transportation Research Part A: Policy and Practice 107*, 52–64.

Fan, W. and X. Jiang (2013). Tradable mobility permits in roadway capacity allocation: Review and appraisal. *Transport Policy 30*, 132–142.

Grant-Muller, S. and M. Xu (2014). The role of tradable credit schemes in road traffic congestion management. *Transport Reviews 34*(2), 128–149.

Green, C. P., J. S. Heywood, and M. N. Paniagua (2020). Did the London congestion charge reduce pollution? *Regional Science and Urban Economics 84*, 103573.

Han, F., X.-m. Zhao, and L. Cheng (2020). Traffic assignment problem under tradable credit scheme in a bi-modal stochastic transportation network: A cumulative prospect theory approach. *Journal of Central South University 27*(1), 180–197.

Harwatt, H., M. Tight, A. L. Bristow, and Guhnemann (2011). Personal carbon trading and fuel price increases in the transport sector: an exploratory study of public response in the uk. *European Transport Trasporti Europei 47*, 47–70.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263 – 292.

Li, Z. and D. A. Hensher (2013). Behavioural implications of preferences, risk attitudes and beliefs in modelling risky travel choice with travel time variability. *Transportation 40*(3), 505 – 523.

Lindsey, R. and G. Santos (2020). Addressing transportation and environmental externalities with economics: Are policy makers listening? *Research in Transportation Economics 82*, 100872. Special Issue in honour of Finn Jørgensen.

Mabit, S. L., E. Cherchi, A. F. Jensen, and J. Jordal-Jørgensen (2015). The effect of attitudes on reference-dependent preferences: Estimation and validation for the case of alternative-fuel vehicles. *Transportation Research Part A: Policy and Practice 82*, 17 – 28.

Miralinaghi, M. and S. Peeta (2016). Multi-period equilibrium modeling planning framework for tradable credit schemes. *Transportation Research Part E: Logistics and Transportation Review 93*, 177–198.

Pigou, A.C., 1920. Wealth and Welfare. Macmillan, London.

Ravner, L. (2014). Equilibrium arrival times to a queue with order penalties. *European Journal of Operational Research 239*, 456–468.

Smith, V. (1976). Experimental economics: induced value theory. *American Economic Review 66*(2), 274 – 279.

Smith, V. (1982). Microeconomic systems as an experimental science. *American Economic Review 72*(5), 923 – 955.

Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of Political Economy 70*(2), 111–137.

Tang, T.-Q., B.-T. Zhang, and C.-Z. Xie (2019a). Modeling and simulation of pedestrian flow in university canteen. *Simulation Modelling Practice and Theory 95*, 96-111.

Tang, T.-Q., B.-T. Zhang, J. Zhang, and T. Wang (2019b). Statistical analysis and modeling of pedestrian flow in university canteen during peak period. *Physica A: Statistical Mechanics and its Applications 512*, 29-40.

Tian, L.-J., H. Yang, and H.-J. Huang (2013). Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review 54*, 1–13.

Tian, Y., Y.-C. Chiu, and J. Sun (2019). Understanding behavioral effects of tradable mobility credit scheme: An experimental economics approach. *Transport Policy 81*, 1–11.

Vickrey, W. S. (1969). Congestion theory and transport investment. *American Economic Review*, 251–260

Verhoef, E., P. Nijkamp, and P. Rietveld (1997). Tradeable permits: their potential in the regulation of road transport externalities. *Environment and Planning B: Planning and Design 24*(4), 527–548.

Wang, X., H. Yang, D. Zhu, and C. Li (2012). Tradable travel credits for congestion management with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review 48*(2), 426–437.

Xiao, F., Z. S. Qian, and H. M. Zhang (2013). Managing bottleneck congestion with tradable credits. *Transportation Research Part B: Methodological 56*, 1–14.

Yang, H. and X. Wang (2011). Managing network mobility with tradable credits. *Transportation Research Part B: Methodological 45*(3), 580–594.

Appendix A

Table A1: Variable description

Variable	Description	
Basic MNL model		
ASC_Peak	constant used in "peak" utility function	
ASC_Post	constant used in "post-peak" utility function	
ASC_Pre	constant used in "pre-peak" utility function	
variable_no_show	the variable has been added in the "no-show" utility function	
variable_peak	the variable has been added in the "peak" utility function	
variable_pre	the variable has been added in the "pre-peak" utility function	
variable_post	the variable has been added in the "post-peak" utility function	
incentive	equals 1 if students need to use permits on a specific day	
Additional variables in extend MNL model		
dormitory	equals 1 if students live in a dormitory which is further away from where the breakfast was provided	
lateclass	equals 1 if students do not have a class start at 8:00 a.m. on a specific day	
rain	equals 1 when it was rain on a specific day	
Individual and commuting variables		
income	students' monthly income	
brf_dep	You must have breakfast on each day	strongly disagree =1, strongly agree =5
brfscore	scores for breakfast which we provided	1 (bad), 5 (good)
depearly	can depart earlier than usual	very uneasy =1, very easy =5
depl	can depart later than usual	very uneasy =1, very easy =5
depecl	regular departure time when have 8:00a.m. class	in peak =1, others =0
deplcl	regular departure time when do not have 8:00a.m. class	in peak =1, others =0
depl_depecl	an interaction of can depart later than usual / regular departure time when have 8:00a.m. class	
deplcl_lclass	an interaction of regular departure time when do not have 8:00a.m. class / late class	
Game-related variables		
earnmoney	It was clear to me that I could earn real money with the game.	strongly disagree =1, strongly agree =5
reexplan	Have you read the explanation of the rules	yes=1, no=0
renotice	Have you read the notices about the rules in the Wechat group	have read all notices =1, others =0
website	The mobile website (or "app") worked well on my phone	strongly disagree =1, strongly agree =5
Attitudes-related variables		
accep	How acceptable is tradable breakfast permits overall to you?	strongly reject = 1, strongly accept =5
allbetter	All students would be better off if a tradable breakfast permit were implemented.	strongly disagree =1, strongly agree =5

ibetter	I would be better off if a tradable breakfast permit were implemented.	strongly disagree =1, strongly agree =5
Market-related variables		
dactive	number of daily activities on the website	
davprice	day-average permit price	
dbuy	number of daily purchases	
dsell	number of daily sells	
Additional variables in Test of rationality		
inipermit	number of initial permits	
Additional variables in Test of reference dependence		
day 2	equals 1 if it is on Tuesday	
day 3	equals 1 if it is on Wednesday	
day 4	equals 1 if it is on Thursday	
day 5	equals 1 if it is on Friday	
cumuse	number of permits used so far	
day x_cumuse	an interaction of day x and comuse	
loss_rfini	loss_ini	if dn < 0 = 1, others =0
loss_rffu	loss_fu	if gn < 0 = 1, others =0

Appendix B: Multinomial logit

Results of the basic multinomial logit (MNL) model are presented in Table B1 and show that the incentive does indeed make the non-peak alternatives more attractive. Furthermore, students have a preference for picking up their breakfast during the peak instead of not showing up, as can be seen from the positive and significant ASC for peak, while a pre-peak pick-up is relatively less attractive.

Table B1: Basic MNL model

	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
Incentive	0.784	0.136	5.76	8.31e-09	0.136	5.76	8.31e-09
Alternative Specific Constants (ASC)							
Peak	1.23	0.112	10.9	0	0.117	10.5	0
Post	0.163	0.103	1.59	0.113	0.103	1.59	0.113
Pre	-0.361	0.118	-3.06	0.00222	0.118	-3.06	0.00222
Log Likelihood			-1150.245				
AIC			2,308.489				
BIC			2,327.743				
Rho-square-bar			0.085				

The results of the extended MNL model are shown in Table B2.

Table B2: Extended MNL model

	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
Incentive	0.954	0.15	6.34	0.000	0.147	6.51	0.000
dormitory_post	0.762	0.178	4.29	0.000	0.174	4.39	0.000
Lateclass_no_show	2.04	0.2	10.2	0.000	0.212	9.61	0.000
Lateclass_post	2.57	0.199	12.9	0.000	0.2	12.9	0.000
Rain	-2	0.282	-7.11	0.000	0.324	-6.19	0.000
Alternative Specific Constants (ASC)							
Peak	2.29	0.164	14	0.000	0.176	13	0.000
Post	-0.482	0.192	-2.5	0.012	0.196	-2.45	0.014
Pre	0.599	0.159	3.77	0.000	0.164	3.66	0.000
Log Likelihood			-1,001.441				
AIC			2,018.882				
BIC			2,057.389				
Rho-square-bar			0.2				

Appendix C: App tests

Before the formal field experiment started, two rounds of tests of the application were performed to ensure that it functioned well (the first test ran from May 26 to May 31 and the second from June 9 to June 13, 2019). Each test lasted 6 days, from Sunday to Friday. All participants were asked to send a “good morning” message in the experiment’s WeChat group between 7:00 and 9:00 a.m. (WeChat is a popular chat application in China). A permit was needed to send the message on weekdays between 8:00 and 9:00 a.m. (the specified rush hour). The permit market could be accessed and used starting on Sunday. Participants were also asked to send a screenshot of “permit use” in the group chat for monitoring. From Monday to Friday, rush hour messages in the WeChat group were rewarded with ¥6, ¥5, ¥4, ¥3, and ¥2 on each of the respective days. The reward for the off-peak message was set at ¥1. Participants were randomly split into two similar-size groups. Individuals in one group received 3 initial permits and ¥10 as their initial budget (with a total equivalent value of $3 \ast 1.5 + 10 = 14.5$). Individuals in the other group received 5 initial permits and ¥7 as their initial budget ($5 \ast 1.5 + 7 = 14.5$). Given the reward design, we expected each person to use 4 permits and that the equilibrium permit price would fall between ¥1 and ¥2. Transaction costs were set at ¥0.1. Students’ final budget was the sum of the trading budget in the application plus what they had earned by sending messages. The student with the highest budget received their budget in real money in order to encourage participants to maximize their final budget. All participants confirmed they clearly understood these rules before the tests started.

In the first round, 14 students joined the test. Student feedback was used to improve the application’s adaptability to different phone models and to reformulate the explanation of tradable permits to render it easier to understand. The information gathered confirmed that a value of ¥0.01 for the size of the price change in the price-setting algorithm (δ) and a transaction fee of ¥0.1 worked well in the Chinese context. The permit price moved within our expected range. After updating the application and to test the market with more users, another 10 students joined the test in the second round. Considering that each user may have a different valuation for their 'peak trip' in the field experiment, we changed the order of peak rewards for each student randomly. The researchers involved were also included in the second round and undertook some actions (e.g., buying and selling repeatedly) to test the robustness of the market. In general, the application worked well in the Chinese context, the permit price moved within the expected range, no undesirable speculation was observed, and most participants acted rationally.

Appendix D: Teach-ins

Instructions and a short instructional video were sent to participants. Because some final tests for the spring semester were scheduled during the teach-in week, about 60 students did not attend the teach-ins. Therefore, the rules and introductory video were uploaded in a WeChat group that was generated for the experiment, and every participant was free to ask the remaining questions online. In addition, each student was asked to answer two calculation questions to make sure they understood the trading rules. The first question was, “What is the maximum number of breakfast permits you can have on Wednesday?” This question was used to test their understanding of the maximum number of day-specific permits a participant can own at any given moment—which is, following Brands et al. (2020), equal to the number of remaining morning peaks in that week—to avoid undesired speculation. The second question was, “If you currently have 3 permits and ¥20, how many permits and monetary budget will you have after selling one permit at ¥2.35?”. This simple question is used to test their understanding of the transaction fee. Students who did not give the right answer have been asked to read the rules again. Finally, all students were asked to sign a terms and conditions form that included the rules and how the collected data would be used. They were also asked to commit to finishing the whole experiment, including the pre- and post-surveys.

Appendix E: Pre-survey results

As shown in Figure E1, departures incur from 6:25 a.m. when students have morning lectures. Given that the university canteen provides breakfast after 6:30, and around 15 minutes is needed to deliver the breakfast to our temporary breakfast station, the time window within which students could pick up breakfast was set to be between 6:50 and 8:30 a.m.. Breakfast pick-ups after 8:00 a.m. were also allowed because of the consideration of students who get up late or don't have morning lectures on some days.

Figure E1 also shows that a sharp increase in departures occurs from 7:20 a.m. Besides, given our experience and previous feedbacks from students, a lot of students enter the classroom just at time. Hence, the rush hour, during which trading participants would need a permit to pick up breakfast, was set to be between 7:20 and 8:00 a.m.

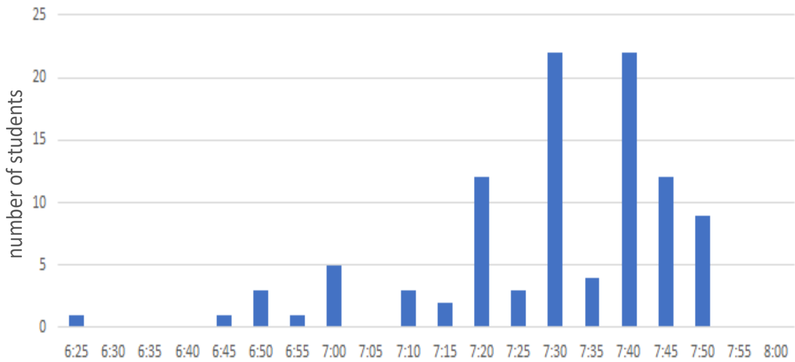

Figure E1: If you have lectures starting at 8:00 a.m., at what time do you usually depart from your dormitory?

We also asked the average cost of participants' normal breakfast in the canteen, which was around ¥4. This information was used to select a breakfast such that it amounted to a value of about ¥5 per day.

At the end of the pre-survey, we designed a simple stated preference (SP) question to reveal students' willingness to pay to have breakfast during rush hour. Respondents were given a description of the general setting in which they would be provided with a free breakfast and told that they would need to use one permit to have breakfast during rush hour but would not need to use a permit before or after rush hour. Before the experiment, they would receive a limited number of permits for free and a sufficiently large monetary budget they could use to buy or sell their permits. The remaining budget would be transferred to each participant after the experiment. We then showed them some possible permit prices—¥0, ¥1, ¥3, and ¥5—and asked them to (1) imagine a week with five 8:00 a.m. lectures and (2) state how many times they would choose to have breakfast during rush hour per week. The SP results, which are reported in Table 1, suggest that about 25% of peak breakfasts would be eliminated with a ¥3 permit price. This was used as input for the starting values of the actual experiment. We used ¥3 as the initial permit price and set the average number of initial permits at 2.5 per person per week.

Table 1: Number of breakfasts during rush hour for different permit prices

Permit price	0 Yuan	1 Yuan	3 Yuan	5 Yuan
Average number of peak breakfasts (per week)	3.6	3.4	2.7	1.6

Appendix F: Distribution of time to pick up breakfast

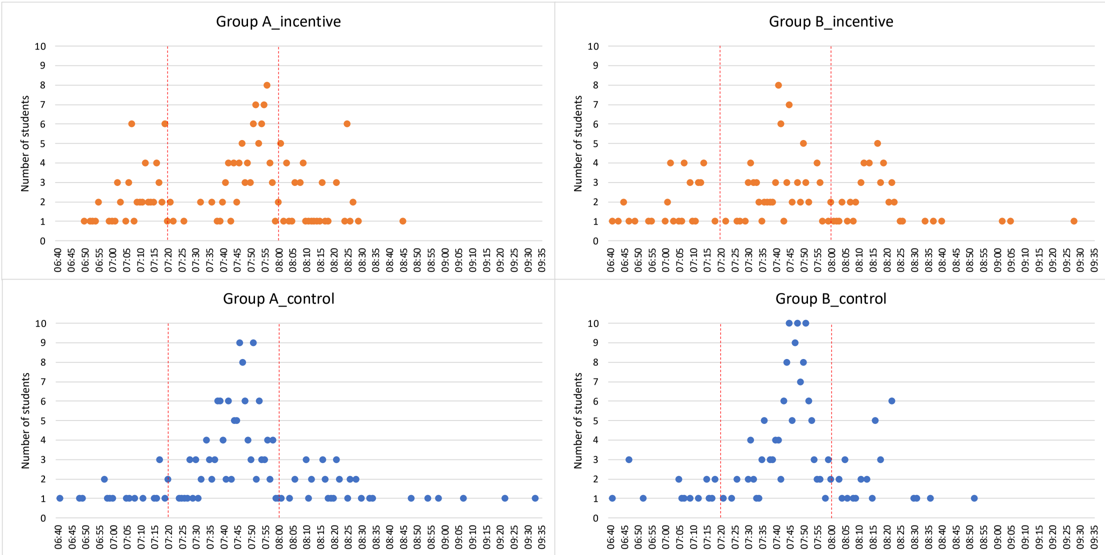

Figure F1: Distribution of time to pick up breakfast (number of students at each minute)