Born to wait?
A study on allocation rules in booking systems*

Lingbo Huang† Tracy Xiao Liu‡ Jun Zhang§

March 24, 2023

Abstract

Many goods and services are allocated through various booking systems. Queue-based booking systems are often thought to allocate goods more efficiently than random allocation because the time spent queuing signals an agent's valuation. This paper demonstrates that the opportunity cost of queuing time can be a significant efficiency loss in queue-based systems. To quantify different sources of efficiency loss, we first develop an experimental framework where agents participate in both a booking system and a production activity. Using a queue-based booking system, our lab experiments confirm that the efficiency loss due to the opportunity cost of queuing time dominates other sources of efficiency loss. However, a lottery-based booking system almost eliminates this efficiency loss. We further develop a novel dual-track booking system that allows participants to choose their preferred booking track, and find that most prefer the lottery track to the queue track.

Keywords: market design; booking system; queue; lottery; opportunity cost of time

JEL Classification: C92, D47

*We thank Bo Chen, Philipp Heller, Dorothea Kübler, Sen Geng and participants from several seminars and conferences for helpful comments. Huang is supported by NSFC (Grant 72203099). Liu gratefully acknowledges financial support by NSFC (72222005) and Tsinghua University (No. 2022Z04W01032). Zhang is supported by NSFC (Grants 72122009, 72033004) and the Wu Jiapei Foundation of the China Information Economics Society (Grant E21103567). The paper title borrows from a New York Times article with the headline "*Born to Wait*" on February 24, 2013 (https://www.nytimes.com/2013/02/24/nyregion/for-new-york-city-parents-a-waiting-list-for-nearly-everything.html).

†Economics Experimental Lab (NEXL), Nanjing Audit University. Email: lingbo.huang@outlook.com.

‡Department of Economics, School of Economics and Management, Tsinghua University. Email: liuxiao@sem.tsinghua.edu.cn.

§Institute for Social and Economic Research, Nanjing Audit University. Email: zhangjun404@gmail.com.

1 Introduction

A first-come, first-served basis is commonly used to manage the distribution of scarce goods or services, and its performance as an allocation rule is widely studied across academic disciplines (e.g., computer science, operations research, and economics). The first-come, first-served process may lead people to spend enormous time and energy on the unproductive activity of queuing. According to an article in the New York Times, "Americans spend roughly 37 billion hours each year waiting in line."[1] Economically speaking, a 2014 survey of two thousand US adults reports that: "Businesses lose some $130 billion in employee productivity every year ($900 per employee) due to the time they waste dealing with service inefficiencies during the work day. 40% of employed adults reported spending at least one hour waiting in line or on a telephone queue trying to resolve a service issue that they could have been spent working."[2,3]

The primary focus of this paper is on the allocation of scarce goods or services via booking systems that either do not use prices or have limited ability to adjust prices to match supply and demand (e.g., bookings for public services, tickets to events, or spots in courses). The use of a queuing system is often justified on the grounds that the time spent in a queue signals an agent's valuation of goods, leading to greater *allocative efficiency* than random allocation. By contrast, the lottery system, another widely-used allocation rule,[4] is criticized for hampering allocative efficiency, even though it guarantees fairness in principle. Our study examines a potentially negative aspect of queuing systems, the time cost incurred by queue participants.[5] Given that people often multi-task, the time they spend in a booking queue

[1] "Why Waiting is Torture," *New York Times*, August 18th, 2012.

[2] See https://www.huffpost.com/entry/waiting-in-line-is-bad-bu_b_12523316; last accessed on May 26, 2022.

[3] Relatedly, the literature on congestion pricing in transport has emphasized that time lost due to traffic congestion is one of the largest externalities associated with automobile use (see Naor (1969); Parry, Walls and Harrington (2007); Heller et al. (2019), among others).

[4] For example, lotteries are used in public school choice contexts to break priority ties among equally eligible students (Abdulkadiroğlu and Sönmez, 2003). Many US colleges and universities use lotteries to allocate on-campus housing placements (Chen and Sönmez, 2002). Major cities in China use lotteries to allocate vehicle licenses (Li, 2018). The Broadway theater system uses lotteries to allocate heavily-discounted tickets (see https://lottery.broadwaydirect.com/; last accessed on May 23, 2022). On a local level, a craft beer brewer in Vermont uses lotteries to sell annual festival tickets (see https://blog.freshtix.com/the-benefits-of-setting-up-a-lottery-for-ticket-buyers/; last accessed on May 26, 2022).

[5] The economics of rationing and queuing have been studied by Tobin (1952); Nichols, Smolensky and Tideman (1971); Barzel (1974); Holt and Sherman (1982); and Suen (1989), among others. As far as we

could have been spent on another production task. Therefore, the forgone opportunity related to working on another production task could cause substantial *productive efficiency* losses. More fundamentally, market designers face a classic trade-off between fairness and efficiency when deciding on allocation rules.[6] To strike a balance between these two goals, market designers must evaluate potential efficiency losses within both the experience of the booking system and the parallel production task.

Learning about individuals' opportunity costs of time is critical to quantifying the aforementioned two types of efficiency, that is, allocative efficiency and productive efficiency. First, to quantify the productive efficiency, we must compute individuals' forgone payoffs from the production task due to the time they spend on booking systems. Second, individuals' opportunity costs of time also affect the time they spend in a booking queue, which complicates the determination of allocative efficiency. As Holt and Sherman (1982) have theoretically shown, if individuals' opportunity costs of time are heterogeneous, the queue rule does not necessarily produce a more efficient allocation of goods than random allocation.[7] Finally, studies on the psychology of queuing find that the act of waiting in a queue can lead to feelings of stress, boredom, and the uneasy sensation that one's life is slipping away, thus leading to a reduction in on-the-job productivity (Larson, 1987; Leroy, 2009).[8] This motivates us also to examine the consequences of different allocation rules for individuals' productive efficiency through their on-the-job productivity, which we term *behavioral efficiency*. Using field data to quantify these types of efficiency is challenging given the difficulty in obtaining individual-level data on opportunity costs of time and on-the-job productivity.

know, this strand of work has been developed in isolation from the market design literature. Taylor, Tsui and Zhu (2003) have theoretically examined when the productive efficiency loss due to queuing outweighs any efficiency gain in the allocation of goods.

[6]Discrete allocation problems have long been studied in the market design literature (see Abdulkadiroğlu and Sönmez (1998), among others). Algorithms such as random serial dictatorship frequently use lotteries to break ties; the performance of these algorithms has been experimentally evaluated (Chen and Sönmez, 2002; Guillen and Kesten, 2012; Hugh-Jones, Kurino and Vanberg, 2014). In these studies, the trade-off between efficiency and fairness typically does not involve efficiency concerns related to the opportunity cost of time.

[7]Theoretically, a participant's queuing time is a function of her valuation of goods as well as her opportunity cost of time. If a participant with a high valuation has a high opportunity cost of time, she may actually spend *less* time queuing compared to a participant with a low valuation and a low opportunity cost of time.

[8]Though not captured in our experiment, such negative feelings could also spill over to subsequent tasks. A small but growing literature in experimental economics has shown that individuals' performance on one task is affected by their choices in other tasks, driven by both cognitive load and spillover effects (see Bednar et al. (2012), among others).

In our study, we introduce an experimental framework to empirically quantify and compare different types of efficiency loss across different booking systems using a queue, lottery, or hybrid allocation process. Specifically, we design a real-effort experiment in which each participant faces two parallel tasks: an appointment booking task and a real-effort production task. In the booking task, each participant must book one appointment slot for which her valuation is private and randomly generated. This booking task consists of two stages, each lasting four minutes. Stage 1 produces the initial allocation of slots, while stage 2 mimics real-life situations in which participants who fail to book a slot in stage 1 may visit the booking system to search for any remaining or canceled slots. In both stages, participants are considered for an available slot either by booking on a first-come, first-served basis or entering a lottery, depending on the treatment conditon. We obtain our pool of available slots for stage 2 in two ways: letting any unassigned slots in stage 1 be available at the beginning of stage 2 and canceling one of the slots allocated in stage 1 at a random moment in stage 2. Note that stage 2 booking system is available only for those who have yet to obtain a slot. In addition to the booking task, participants also work on a real-effort production task represented by counting the number of white dots in a series of dark-shaded squares. Participants can freely switch between the two tasks at any time but cannot work on both tasks simultaneously. Therefore, they essentially face a time allocation problem between the two tasks.

Using a between-subjects design, we first compare two *solo-track* systems that use either the queue rule or the lottery rule exclusively in both stages. We also vary the degree of market competitiveness to test for the robustness of our results. In our theoretical model, we distinguish between three sources of efficiency loss: inefficient allocation of slots (allocative efficiency loss), the opportunity cost of time spent on the booking task (strategic efficiency loss), and changes in on-the-job productivity (behavioral efficiency loss). Consistent with our theoretical predictions, our experimental results show that queue participants spend substantial amounts of time on the booking task in both stages while lottery participants spend only a few seconds submitting their lottery entry and the remainder of their time on the production task. We also find that the strategic efficiency loss under the queue rule outweighs the other two sources by a large margin, leading to a much higher overall efficiency loss under the queue rule than the lottery rule. We further observe that allocative efficiency is actually not higher under the queue rule, either. The reason is that most participants exhibit bimodal behavior under the queue rule: they spend either a few seconds or almost all of their time on the booking task, largely irrespective of their private valuations. Hence, in our experiments the lottery rule outperforms the queue rule in almost every aspect of efficiency, let alone the guaranteed fairness in the allocation of slots.

While our experimental results strongly support the transition from the queue rule to the lottery rule, a shift to a pure lottery may not be practical for several reasons. First, participants may have concerns about the transparency of lottery draws, especially those for highly-competitive goods or services. Second, participants may be concerned about the inability of a lottery to distinguish participants with greater need for the good or service. A possible resolution of these concerns is a combination of the two allocation rules into one system in which participants could gain experience with both rules and then freely decide which rule to use. To this end, we design a novel *dual-track* allocation system. In this hybrid system, slots are provided in two tracks, each implementing one of the two allocation rules, and each participant can freely choose which track she wants to enter (but she cannot choose both). In our experiment, slots are split evenly between the two tracks in stage 1. Stage 2 implements either the queue rule or the lottery rule, depending on the treatment condition.

Under a dual-track system, our theoretical analysis predicts that participants will be more likely to choose the lottery track over the queue track, even though the chance of obtaining a slot is theoretically the same in either track. Our experimental results support this prediction. We further find that participant behavior under each track is similar to that in the corresponding solo-track system. Consequently, those who choose the lottery track earn a higher payoff than those who choose the queue track, offsetting their lower probability of obtaining a slot. Comparing the sources of efficiency loss under the dual- and solo-track systems, we find that the efficiency losses due to opportunity costs of time under the queue system remain substantial. The total loss is lower given the lower number of participants choosing the queue track in our dual-track setting. Finally, we find that the dual-track system helps to reduce allocative efficiency loss by channeling some participants with high valuations to compete for slots in the queue track, consistent with our theoretical prediction.

There are many studies on various queue systems, including two types of queues that differ from the one we study. The first type is a queuing system where a facility continuously provides services to people who arrive over time. For instance, at airports, passengers are checked in based on the order of their arrival. In this situation, an important reason for people to arrive earlier is to be served earlier. Numerous studies in economics and management science have been devoted to this type of queue (e.g., Naor, 1969; Platz and Østerdal, 2017; Che and Tercieux, 2021). In our queuing system, slots on booking systems are released at a pre-defined time, so an earlier arrival does not result in an earlier assignment. The second type is a waiting list, where people who may stochastically arrive enter their names into a list for goods that arrive over time (e.g., public housing, daycare spots, organ transplants), but do not spend time physically queuing. There is no opportunity cost of time as studied in our paper, but there may be other types of waiting costs. Some papers have studied the trade-off between quick matching to cut down waiting costs and slow matching to generate higher match surplus (e.g., Akbarpour, Li and Gharan (2020); Baccara, Lee and Yariv (2020); Schummer (2021); Leshno (2022)).[9] Other papers take agents' waiting times as endogenous choices and design mechanisms to encourage truthful reports (e.g., Schummer and Abizada (2017); Dimakopoulos and Heller (2019)). In these papers, goods are often heterogeneous, which is different from the homogeneous booking slots we study.

Our paper is positioned within the broad experimental literature on matching markets (see Roth (2021) and Hakimov and Kübler (2021) for recent surveys). However, we differ from this literature in our introduction of a new experimental framework for quantitatively evaluating various forms of efficiency loss that arise during the matching process. A related study that also compares different versions of first-come, first-served and lottery rules is that of Hakimov et al. (2021). However, they focus on the behavior of scalpers, or those who resell tickets at above-value cost. In online booking systems, scalpers use speed to occupy all first-come, first-served slots in the initial allocation process. To solve this problem, Hakimov et al. (2021) propose a lottery-based batch system that periodically collects applications and then draws lotteries to allocate the slots within a given batch. This system eliminates the importance of speed and deters scalpers from entering the market, a prediction confirmed in their lab experiment. Their batch system is similar to the lottery rule in our study.[10] Our paper complements their study by demonstrating another advantage of lottery-based booking systems, that is, eliminating the productive efficiency loss that comes from the opportunity cost of time spent in a queue.

The rest of the paper is organized as follows. Section 2 describes our experimental design. Section 3 presents our theoretical analysis and predictions. Section 4 reports our experimen-

[9]Interestingly, Leshno (2022) finds that a buffer queue mechanism with a randomized queuing policy can reduce misallocation and improve welfare over a first-come, first-served queuing policy.

[10]Our analysis is applicable to a broader scope of allocation problems than those covered by Hakimov et al. (2021). Since we do not restrict our framework to online booking systems, our analysis is applicable to allocation problems where the identities of participants are predetermined (e.g., allocation of courses or on-campus housing in colleges and universities), meaning that scalpers are unable to take on the identity of a regular participant.

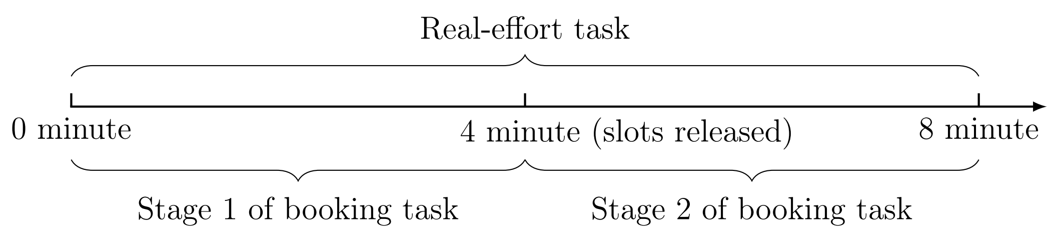
tal results. Section 5 provides concluding remarks.

2 Experimental Design

2.1 Basic design

To compare sources of efficiency loss across queue and lottery allocation systems, we implement a real-effort experiment in a dual-tasking environment. Depending on treatments, participants are randomly matched into groups of five or seven for each round of the experiment. Each session has eight rounds and each round lasts eight minutes. After each round, we randomly rematch subjects to mimic a one-shot setting. In each round, a participant works on two tasks displayed on two different screens. One is an appointment booking task and the other is a novel real-effort counting-dots production task. At the beginning of each round, to mimic real-life situations in which individuals can often initiate which task to work on first, each participant chooses which task will display first on the screen. During the round, a participant can freely switch between the two tasks at any time and as many times as she wishes. Requiring a participant to choose only one task at a time imposes a time allocation trade-off between spending time on the booking versus production task. The timeline of our experiment in each round is shown in Figure 1.

Figure 1: Timeline of the dual-tasking environment



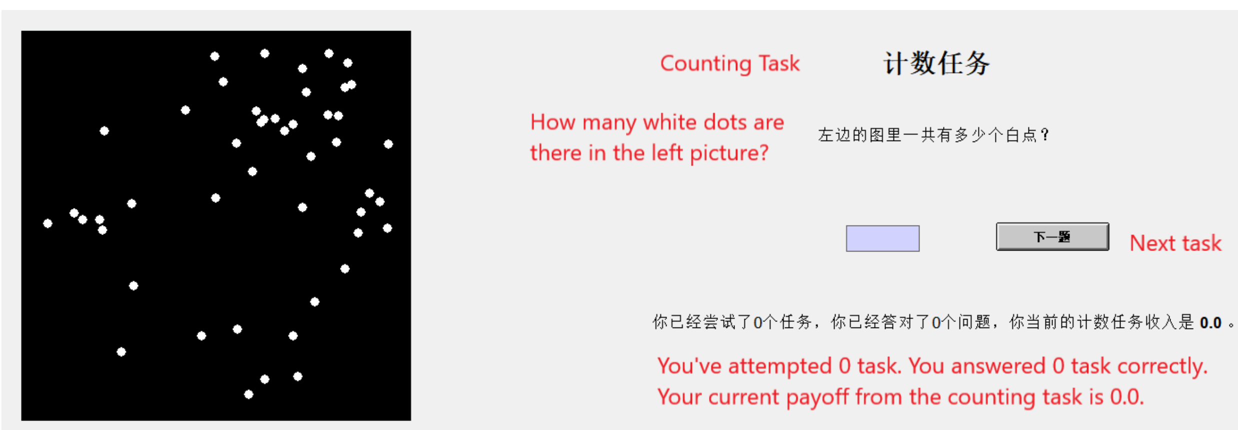
In the booking task, depending on the treatment condition, there are two or three total slots available for each group of five or seven participants. Each participant can acquire at most one slot in each round. At the beginning of each round, each participant is privately informed about her valuation for a slot, which is drawn independently from the uniform distribution over integers between 400 and 600 Experimental Currency Units (ECUs).

In each round, the booking task consists of two stages, each lasting four minutes. In stage 1 of the booking task, participants compete for slots by queuing or entering a lottery, depending on their treatment condition. All slots are released and allocated at the end of stage

1. Any unassigned slots in stage 1 become available for stage 2 of the booking task at the beginning. Moreover, if at least one slot is allocated in stage 1, one slot will be randomly selected and canceled at a random moment in stage 2, when it then becomes immediately available. The participant whose slot is canceled will still obtain a payoff equivalent to her valuation and cannot book a slot again. This setup is designed to mimic real-life situations in which people with allocated appointments may cancel these appointments, making these appointments available to others. To simplify our decision-making environment, we determine a cancellation, as opposed to a participant choosing to cancel, as the cancellation source is not essential for our research purposes.

In the production task, we ask participants to count the number of white dots in a series of dark-shaded squares (see Figure 2). In each square, both the total number (between 35 and 54, inclusive) and the positions of dots are randomly generated. Each correct answer is rewarded by 35 ECUs.[11] This task is designed to mimic the type of task that may compete with a queuing activity for a participant's time. To build task familiarity, participants are asked to work on the production task for five minutes at the beginning of each session without any reward.[12]

Figure 2: Screenshot of the counting-dots task



Performance on the production task may be negatively impacted by both strategic and

[11]The piece rate is chosen so that the expected payoffs from working only on the production task and from obtaining a slot in the booking task are largely comparable. This payoff selection is intended to highlight the trade-off in time allocation between the two tasks.

[12]The average per-minute productivity for our full sample (n = 344) in the five-minute trial round is about 1.12 correctly-answered squares (s.d. = 0.59), compared to 1.72 (s.d. = 0.58) in the payment rounds.

behavioral efficiency loss. Strategic efficiency loss occurs when the booking task takes time from the production task. Behavioral efficiency loss occurs when participants likely need to start again after being distracted by the booking task.

One reason for designing the two-stage process is to observe these two types of productive efficiency loss. In both stages, we can directly observe the time participants spend on the booking system instead of the production task. This allows us to compute the strategic efficiency loss for individual participants in monetary terms. We expect behavioral efficiency loss likely to occur in stage 2 given the distraction of visiting and revisiting the booking system during this time. So to compute the behavioral efficiency loss, we compare a participant's actual per-minute productivity in stage 2 with her average productivity and then compute its monetary value.

In the next two subsections, we first present the solo-track system that uses either one of the two allocation rules and then present the dual-track system that allows for endogenous choices between the two allocation rules.

2.2 The solo-track booking system

In the solo-track booking system, slots are assigned using either one of the two allocation rules: the queue rule or the lottery rule.

The *queue rule* models the first-come, first-served booking systems widely used in real-life situations.[13]

- **Stage 1**: A queue is used to determine who will obtain a slot at the end of stage 1. During this stage, a participant can choose to enter and remain on the booking system at any time to reserve a position in the queue. Those who enter the booking system earlier reserve an earlier position in the allocation queue. However, if a participant switches to the production task and then back to the booking system, this participant must go to the back of the queue.[14] When the fourth minute of the round is reached, if

[13]While it is possible that online booking participants may work on several activities simultaneously, booking behavior in highly competitive scenarios likely precludes multi-tasking. In these scenarios, participants typically concentrate solely on the booking task, refreshing web pages to reveal the most recent information about available slots when the website design does not allow auto-refreshing or when the network is heavily congested. Thus, slot seekers spend time or energy "glued to their device" and waiting before slots are released. It is reasonable to expect that those who spend more time or energy on waiting are more likely to win a slot.

[14]In stage 1, participants are not aware of either their or others' positions in the queue. This feature ensures

the queue length is greater than the number of slots, slots will be assigned according to participant placement in the queue. If the queue length is not greater than the number of slots, every participant in the queue obtains a slot, and any unassigned slots become available in stage 2.

- **Stage 2**: Only participants who have not obtained a slot in stage 1 can participate in the booking task in stage 2. In this stage, available slots are those unassigned after stage 1 as well as a slot created by the random cancellation of an assigned slot. The unassigned slots are available in the booking system from the start of stage 2, whereas the canceled slot becomes available at the moment of cancellation. Participants can enter the booking system at any time and observe the number of currently available slots. If a slot is available, a participant can book it immediately. However, if the system shows no available slot, participants cannot be certain whether this is because all slots have already been assigned or because the cancellation has not yet happened.

The *lottery rule* collects participants' applications to the lottery during each stage and assigns slots randomly to applicants at the end of each respective stage.

- **Stage 1**: Each participant can enter the booking system at any time and apply for entry into the lottery by pressing a button on the screen. All applications are collected into a virtual urn. When the fourth minute is reached, applications are randomly drawn from the urn one by one until all available slots have been allocated. If the number of applications is smaller than the number of available slots, unassigned slots are transferred to the slot pool for stage 2.

- **Stage 2**: Participants who do not receive a slot in stage 1 can apply using the same application process in stage 2. Available slots are comprised of unassigned carryovers from stage 1 as well as a random cancellation of a stage 1 assigned slot. Participants can enter the booking system at any time and observe the number of currently available slots. In stage 2, participants can apply at any point by pressing the application button, even if the system shows no currently available slots. All applications are collected in a virtual urn. At the end of stage 2, applications are randomly drawn from the urn

consistency with the sealed bid nature of our theoretical model. While people do observe this information in physical queues, they may not know the number of available slots for those in the queue or even their queue position in large markets. It is important for our experiment to capture some level of uncertainty and tension related to one's success in obtaining a slot, which we believe is a central feature of first-come, first-served allocation rules.

one by one to fill any available slots.

2.3 The dual-track booking system

While we expect that participants in our experiment will achieve higher productive efficiency under the lottery versus queue rule, it is possible that they may still choose the queue rule if given a chance to choose. First, people may dislike the uncertainty inherent in the lottery rule and may choose the queue rule to feel a sense of control over the process, even at the cost of productive efficiency (Bartling, Fehr and Herz, 2014; Owens, Grossman and Fackler, 2014). Second, participants may have concerns about the transparency of a lottery. While this concern is unlikely to matter in our lab environment, it may have traction outside of a lab setting if manipulation or corruption are suspected, especially in high-stakes situations. Third, completely shutting down the possibility of using the queue rule will not necessarily benefit everyone in the market from a Pareto improvement perspective. For example, a lottery rule may hurt those who put a high valuation on a slot. Finally, an abrupt transition from one rule to another may be perceived as a violation of the moral principle of free-will decision-making.

While addressing all these practical concerns is challenging and out of the scope of our paper, to provide some resolution to the question of rule preference from the market design perspective, we design a novel *dual-track* booking system and implement it in the lab. The basic idea behind the dual-track system is that people can freely choose between the two allocation rules. This system can balance productive efficiency with participant preferences and free choice.[15] Further, our dual-track system allows people to learn both rules, building familiarity with lottery systems for later potential implementation of these systems. Under the dual-track environment, our main research interest is to observe whether participants are more likely to choose the lottery rule after having gained experience with both rules. Moreover, we are interested in understanding how such endogenous choices affect the balance between fairness and efficiency.

[15] Similar dual-track or hybrid systems are observed in the real world. One example is the assignment of vehicle licenses in some major cities of China (Huang and Wen, 2019), such as Guangzhou, Shenzhen, Tianjin, and Hangzhou. There has been a long-standing debate in China about the optimal rule for allocating vehicle licenses. Beijing uses lotteries exclusively, while Shanghai uses only auctions. While auctions are often criticized for generating unreasonably high prices and causing an additional heavy burden on relatively poorer citizens, they are lauded for supposedly superior allocative efficiency and for revenue generation that then goes to improving public transit systems. In our study, since participants "bid" by spending time under the queue rule, there is no public benefit generated from their bid.

The dual-track system differs from the solo-track system in that participants can choose either a lottery or queue track at the beginning of stage 1. Each track has the same number of available slots. After all participants in a group make their track decisions, they are placed into subgroups under their given track, where they are informed of the number of participants in their subgroup. Stage 1 then proceeds in the two separate subgroups with one implementing the queue rule and the other the lottery rule. Note that if one track is not chosen by any participant, the slots in that track transfer over to stage 2.

In stage 2, we merge the two subgroups and implement only one allocation rule. To keep the cancellation procedure comparable to that of the solo-track systems, only one of the assigned slots in stage 1 is randomly selected to be canceled at a random moment. We do not apply a dual-track system in stage 2 as it is likely that the canceled slot will be the only available slot.[16]

The dual-track system combines the queue and lottery rules in stage 1 and implements one of these rules in stage 2, depending on the treatment condition. The system is comprised of the following components.

- **Track decision**: At the beginning of each round, each participant chooses to enter either the track with the queue rule or the track with the lottery rule. This decision is binding for that round.
- **Stage 1**: After all participants have made their track decisions, one track assigns slots according to the queue rule while the other track uses the lottery rule. Any unassigned slots from the two tracks are transferred to stage 2.
- **Stage 2**: Participants who have not received slots in stage 1 are allowed to book a slot in stage 2. Depending on the treatment condition, stage 2 uses either the queue rule or the lottery rule.

2.4 Treatments

In our experiment, we implement a 2×2 design in our main treatments by varying the allocation rule (Queue vs. Lottery) and whether we allow for a solo-track or dual-track system

[16] Suppose we keep our two-track system in stage 2. If the canceled slot comes from one track in which each member has obtained a slot in stage 1, then this slot will be wasted if we do not allow members from the other track to book it. Because we want to use the same cancellation procedure across all treatments, keeping the stage 2 rule comparable across treatments helps avoid such logistic subtleties.

in stage 1. Within the solo-track treatments, we also vary the level of market competitiveness. Under low market competitiveness, each group consists of five participants vying for three slots. We refer to the treatment adopting the queue (lottery) rule and low competitiveness as *Queue5* (*Lottery5*). Under high market competitiveness, each group consists of seven participants contending for two slots. We refer to the treatment adopting the queue (lottery) rule and high competitiveness as *Queue7* (*Lottery7*). The comparisons between Queue5 and Lottery5 and between Queue7 and Lottery7 allow us to study whether the lottery rule improves productive efficiency compared to the queue rule. Further, implementing different levels of market competitiveness allows us to study the robustness of our results.

In the dual-track treatments, we examine only the high market competitiveness environment because it provides greater scope for us to observe if there is any tendency among participants to prefer one track over the other. Specifically, each track in stage 1 under high market competitiveness has exactly one available slot. In the *Dual-Queue* (*Dual-Lottery*) treatment, we implement the queue (lottery) rule in stage 2. Observing how participants make track decisions in stage 1 allows us to examine whether participants generally prefer the lottery rule over the queue rule. We also compare behavior under each track in the dual system with behavior in the corresponding solo system. This allows us to see whether the dual-track system finds the sweet spot between improving overall productive efficiency (compared to the solo-track queue treatments) and respecting distinct individual preferences (participants with high valuations select the queue track). Finally, we examine whether the stage 2 rule affects participants' track decisions and behavior in stage 1. Table 1 summarizes the main features of our experimental design.

Table 1: Experimental Design

Treatments	**Allocation rule**	**Competitiveness**	**# of participants**	**# of matching groups**
Solo-track:				
Queue5	Queue	Low	60	6
Lottery5	Lottery	Low	60	6
Queue7	Queue	High	56	4
Lottery7	Lottery	High	56	4
Dual-track:				
Dual-Queue	Dual (stage 1) Queue (stage 2)	High	56	4
Dual-Lottery	Dual (stage 1) Lottery (stage 2)	High	56	4

2.5 Procedure

The experiment was conducted at the Nanjing Audit University Economics Experimental Lab with a total of 344 university students, using the software z-Tree (Fischbacher, 2007). Each session consists of two independent matching groups of 10 participants (for Queue5 and Lottery5) or 14 participants (for Queue7, Lottery7, Dual-Queue, and Dual-Lottery). Within each matching group, participants are randomly re-matched in each round according to the group size stipulated by the treatment. After every round, all participants receive feedback about whether they were allocated a slot, whether an allocated slot of theirs was cancelled, and what their respective booking and production payoffs are. In the dual-track treatments, they also learn the number of participants and the average payoff in each track. At the end of a session, one round is privately and randomly chosen for each participant and the participant receives her payoff from that round.

During the experiment, as participants arrived, they were randomly seated at a partitioned computer terminal. The experimental instructions were given to participants in printed form and were also read aloud by the experimenter. Participants then completed a comprehension quiz before proceeding. At the end of the experiment, they completed a questionnaire concerning their demographics and a number of psychological measures. For every 10 ECUs, participants earned 1 RMB. A typical session lasted about 2 hours with average earnings of 80.7 RMB, including a show-up fee of 15 RMB.

3 Theoretical Framework and Hypotheses

This section presents the theoretical framework that guides our interpretation of our experimental data. In Section 3.1, we make several simplifying assumptions to allow us to analyze participant time allocations across treatments. In Section 3.2, we discuss our allocation rule evaluation criteria with respect to fairness and efficiency. Section 3.3 presents our set of testable hypotheses.

In our experiment, participants face a trade-off between spending time on the booking task versus spending time on the production task. Since application timing does not impact slot allocation likelihood in the lottery treatment, we focus on participants' strategies in stage 1 of the queue treatments and the dual-track treatments. Under a lottery rule, we expect that participants will minimize the time spent on the booking system: in stage 1, they should visit the booking system only once and remain for a few seconds (which shall be regarded as zero time in theory) to submit their application. Therefore, all slots will be assigned at the end of stage 1, with the canceled slot being the only available slot in stage 2. In stage 2 of the lottery treatment, non-assigned participants again spend a minimum amount of time submitting an application, while assigned applicants spend all their time on the production task. We do not model participants' choice of when to visit the booking system in either stage. Intuitively, they should choose a timing that minimally affects their productivity in the production task; for example, they might switch to the booking system when they want to take a break in the production task.

By contrast, in the queue treatments, participants need to compete for slots. Stage 1 of the booking task can be modeled as an auction in which participants bid by choosing when to arrive at the queue. Winners are those who obtain slots. But it differs from standard auctions in textbooks (see Krishna (2009)) in that participants may have different opportunity costs of time, measured by their productivity in the production task. In addition, their stage 1 strategies are affected by their expectations of outcomes in stage 2. In equilibrium, all participants will bid in stage 1 and therefore only a canceled slot will be available in stage 2 at a random moment. Within this setting, the optimal stage 2 search strategy is unclear. Some may stay on the booking system until the canceled slot appears, while others may choose to visit the booking system periodically. In our framework, we assume that stage 1 participants believe that every participant who still needs a slot in stage 2 has an equal chance of winning. Given this expectation, participants' strategies in stage 1 are essentially independent of their strategies in stage 2. Thus, in Section 3.1, we focus on stage 1 of the queue treatments and, similarly, the dual-track treatments.

3.1 Equilibrium strategies in stage 1 of the booking task

We model our experimental environment as a setup in which there are m slots and n participants, with $n > m \geq 1$. Each participant i demands one slot and values each slot at $v_i \in \mathbf{R}_+$. Each v_i is independently drawn from a commonly known uniform distribution on an interval $[\underline{v}, \overline{v}] \subset \mathbf{R}_+$ and every i knows her valuation v_i.[17] Each i also has a productivity denoted by $w_i \in \mathbf{R}_+$ in the production task. That is, i will obtain a payoff of w_i by spending one unit of time on the production task. We further define $y_i = v_i/w_i$, which represents i's time valuation of slots, or the valuation of slots measured in time units from i's perspective. In our experiments, because v_i is randomly drawn, v_i is independent of w_i for every i. Each

[17] We omit the detail in our experiment that valuations are discrete integers.

stage of the booking task lasts for $T > 0$ units of time.

We make the following two assumptions to simplify our analysis.

Assumption 1. *All participants are risk-neutral and believe that every y_i is independently drawn from an interval $[\underline{y}, \overline{y}] \subset \mathbf{R}_+$ according to a continuously differentiable distribution function F.*

Assumption 2. *In stage 1 of the queue treatments, participants' strategies do not depend on their strategies in stage 2, and participants believe that, if they do not win a slot in stage 1, they have an equal chance of winning any available slot in stage 2.*

Applying the above two assumptions, we can analyze participants' strategies in stage 1 of the queue treatments given their expectation of winning a remaining slot in stage 2. Following Holt and Sherman (1982), we model stage 1 of the queue treatments as an all-pay auction in which participants bid their amount of waiting time in the queue. In the symmetric Bayesian Nash equilibrium, participants' bids will be monotone in their valuations of slots relative to their productivity, i.e., *time valuations*. In real life, participants may be able to wait for an arbitrarily long time if they wish. But in our experiment, their waiting time is capped by the duration of stage 1, T. Whether participants are constrained by this cap depends on the time valuation distribution, F. If the participant with the highest time valuation $\overline{y}$ does not bid more than T in the absence of a cap, then the cap is not a binding constraint. The equilibrium will be characterized by an increasing function $t(y)$, which determines a participant's waiting time in the queue if her time valuation is y. Otherwise, there exists a threshold, $y^\diamond \in [\underline{y}, \overline{y})$, such that all participants with time valuations weakly above $y^\diamond$ will pool by bidding T, while the others will bid less than T. If $y^\diamond \in (\underline{y}, \overline{y})$, like Gavious, Moldovanu and Sela (2002), we will see that $t(y^\diamond) < T$, meaning that the equilibrium bidding function is not continuous at the threshold.[18]Another special case is when $y^\diamond = \underline{y}$, which happens when all participants' time valuations are sufficiently high such that all of them pool at T.

We next let h denote the density function of the m-th order statistics among $n-1$ independent draws from the time valuation distribution, F. We derive participants' equilibrium strategies in stage 1 of the queue treatments in Proposition 1. The proof is in Appendix A.

Proposition 1. *Under Assumption 1 and Assumption 2, there are three cases in the symmetric Bayesian Nash equilibrium in stage 1 of the queue treatments:*

[18]This phenomenon of jump bidding at the threshold has been analyzed by Gavious, Moldovanu and Sela (2002) in the single-object all-pay auction with bid cap.

- *Case 1:* $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds \leq T$. *The cap* T *is not binding. Every participant* i *spends* $t(y_i)$ *units of time in the queue, where* $t(y_i) = \frac{n-m-1}{n-m}\int_{\underline{y}}^{y_i} h(s)sds$*;*
- *Case 2:* $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds > T$ *and* $\frac{m(n-m-1)}{n(n-m)}\underline{y} < T$. *The cap* T *is binding. There exists a threshold* $y^\diamond \in (\underline{y}, \overline{y})$ *such that every participant* i *with* $y_i \geq y^\diamond$ *spends* T *units of time in the queue, while every participant* i *with* $y_i < y^\diamond$ *spends* $t(y_i) = \frac{n-m-1}{n-m}\int_{\underline{y}}^{y_i} h(s)sds$ *units of time in the queue;*
- *Case 3:* $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds > T$ *and* $\frac{m(n-m-1)}{n(n-m)}\underline{y} \geq T$. *Every participant spends* T *units of time in the queue.*

We now analyze the dual-track treatments. Following our experimental design, we assume $k = m/2$ slots (here m is an even number) in each track of stage 1. We make the following assumption that allows us to focus on participants' strategies in stage 1.

Assumption 3. *In stage 1 of the dual-track treatments, participants' strategies do not depend on their strategies in stage 2, and participants believe that, if they do not win a slot in stage 1, they have an equal chance of winning any available slot in stage 2.*[19]

Unlike in the solo-track treatments, the number of participants in each track of the dual-track treatments is endogenous. If fewer than k participants choose any track in stage 1, the unassigned slots in that track transfer to stage 2, meaning it is possible to have more than one available slot in stage 2.

In stage 1, participants choose their stage 1 track; those who choose the queue track also choose their bidding strategy. In the symmetric Bayesian Nash equilibrium, there exists a threshold, $y^* \in (\underline{y}, \overline{y})$, such that a participant i will choose the queue track if and only if $y_i \geq y^*$.[20] A participant with the threshold time valuation y^* will be indifferent between the lottery track and the queue track. We prove that $F(y^*) > 1/2$, meaning that we expect more participants to choose the lottery track rather than the queue track. So participants in equilibrium will know of the time savings in the lottery track. For those choosing the queue track, there further exists an equilibrium bidding strategy like that in Proposition 1. We characterize the equilibrium strategy in Proposition 2 and prove it in Appendix A.

Proposition 2. *Under Assumption 1 and Assumption 3, in the symmetric Bayesian Nash equilibrium in stage 1 of the dual-track treatments, there exists a threshold,* $y^* \in (\underline{y}, \overline{y})$*, such*

[19]Note the second part of the assumption is naturally true if the lottery rule is used in stage 2.

[20]It is clearly not an equilibrium for all participants to always choose either the lottery track or the queue track, so $y^* \in (\underline{y}, \overline{y})$.

that participant i will choose the queue track if and only if $y_i \geq y^*$. Moreover, $F(y^*) > 1/2$, meaning that more participants are expected to choose the lottery track rather than the queue track.

For participants choosing the queue track, as in Proposition 1, there are three cases in equilibrium: (1) the cap T is not binding, and every participant i spends $t(y_i)$ units of time in the queue, where $t(y_i)$ is a strictly increasing function derived in Appendix A; (2) the cap T is binding and there exists another threshold $y^\diamond \in (y^*, \overline{y})$ such that every i with $y_i \in [y^\diamond, \overline{y}]$ spends T units of time in the queue while every i with $y_i \in [y^*, y^\diamond)$ spends $t(y_i)$ units of time in the queue; (3) the cap T is binding and all participants spend T units of time in the queue.

3.2 Evaluation criteria

This subsection discusses our allocation rule evaluation criteria with respect to fairness and efficiency. To measure efficiency, we examine both slot allocation efficiency and participant productive efficiency. We also examine the fairness of the allocation of slots, that is, whether all participants have an equal chance of winning a slot. This definition of fairness is natural in the allocation of indivisible goods without transfers and is widely used in the market design literature.

To represent our evaluation criteria formally, given participants' valuations and productivity, we let $v_{(\ell)}$ denote the ℓ-th highest valuation among n participants. An allocation of slots is a function $\mu : \{1, 2, \ldots, m\} \rightarrow \{0, 1, 2, \ldots, n\}$ such that, for every $\ell \in \{1, 2, \ldots, m\}$, $\mu(\ell)$ is the participant who obtains the ℓ-th slot. If $\mu(\ell) = 0$, it means that the ℓ-th slot is unassigned. Every participant can receive, at most, one slot.

Now suppose that a rule determines an allocation of slots, μ, and every participant i spends $t_i \in \mathbf{R}_+$ units of time on the booking system under this rule. We identify two types of potential efficiency loss in the allocation process. The first is *allocative efficiency loss*. In the most efficient allocation, slots should be allocated to those who value those slots the most. Taking the most efficient allocation as the benchmark, we define allocative efficiency loss of μ as follows:

$$\text{Allocative efficiency loss} = \sum_{\ell=1}^{m} v_{(\ell)} - \sum_{\ell=1}^{m} v_{\mu(\ell)}.$$ [21]

The second type is productive efficiency loss, which can occur through strategic and behavioral efficiency loss. *Strategic efficiency loss* is measured as the opportunity cost of time

[21] If a slot is unassigned, we let $v_0 = 0$.

spent on the booking system:

$$\text{Strategic efficiency loss} = \textstyle\sum_{i=1}^{n} t_i w_i.$$

Although Section 3.1 assumes that participants have constant productivity, we expect that participants' productivity may change over time in our experiments, and in particular, there may be productivity loss due to mental distraction in stage 2 of the queue treatments (see discussion in Section 2). We refer to such loss as *behavioral efficiency loss*. Formally, in stage 2, let w_i^{actual} denote the actual productivity of participant i in the production task, and let t_i^2 denote the amount of time that i spends on the booking task. Then,

$$\text{Behavioral efficiency loss} = \textstyle\sum_{i=1}^{n}(T - t_i^2)(w_i - w_i^{actual}).$$

We next compare the fairness of the lottery and queue rules in our solo-track setting. The lottery rule is fair by definition since all participants have an equal chance of winning a slot in both stage 1 and stage 2 as long as they apply. By contrast, the queue rule may not be fair since participants with higher time valuations may be led to spend more time in the queue and thus enjoy a higher probability of winning a slot. In our experiments, participants' valuations of slots are randomly generated and therefore independent of their productivity. Thus, we expect that participants' time valuations are positively correlated with their valuations. Since participants with higher valuations are more likely to win a slot, fairness is not warranted under the queue rule.

In terms of efficiency, the queue rule is expected to achieve a higher level of allocative efficiency than the lottery rule in our experiments.[22] However, the queue rule may yield greater productive efficiency loss than the lottery rule. Examining the components of productive efficiency, the queue rule requires participants to spend valuable time queuing for slots in stage 1 and competing by speed (and luck) to obtain available slots in stage 2, leading to a loss in strategic efficiency. By contrast, the lottery rule requires only a few seconds on the booking system to press an application button. In addition, behavioral efficiency loss is expected to be greater under the queue rule where participants may feel distracted from their production task during the stage 2 allocation process.

Table 2 summarizes the comparisons between the two solo-track rules in terms of fairness

[22]It is worth noting that, in general environments beyond our experiment, slots need not be allocated more efficiently under the queue rule than under the lottery rule. A participant with a high valuation of slots and a high opportunity cost of time may spend the same amount of time in the queue as a participant with a low valuation and a low opportunity cost. See Taylor, Tsui and Zhu (2003) for more analyses.

and efficiency. Note that the lottery rule may dominate the queue rule in both fairness and efficiency if the productive efficiency loss under the queue rule dominates the allocative efficiency loss under the lottery rule.[23]

Table 2: Fairness and Efficiency under Queue and Lottery

		Queue	Lottery
Fairness of the allocation of slots in the booking system		No	Yes
Allocative efficiency loss in the booking system		Low	High
Productive efficiency loss in the production task	**Strategic**	High	Low
	Behavioral	High	Low

Next, we discuss our evaluation criteria for the dual-track system. As discussed in Section 2, we introduce a dual-track system as both a structure that addresses practical concerns and a useful compromise between the queue and lottery rules. While the lottery rule yields better productive efficiency and greater fairness, the queue rule gives participants who value a slot more than productive efficiency the opportunity to express their preference and potentially increase their welfare.[24] At the same time, as shown in Proposition 2, since we expect more than half of participants to select the lottery track, fairness is still warranted for the majority.

It is worth noting that we do not view our version of the dual-track system as the ideal one, but only as an illustration of how a dual-track system might work. In practice, policymakers can adjust the distribution of slots in the two tracks to achieve their desired balance.

3.3 Hypotheses

In the lottery treatment, we expect participants in both stages will spend only a few seconds on the booking system to press the application button. By contrast, in the queue treatment,

[23]There is another conceptual reason why the lottery rule may dominate the queue rule in efficiency. In our study, we define allocative efficiency through a utilitarian welfare function and call an allocation efficient if slots are allocated to those who value them the most. However, because there are no transfers between participants, the utilitarianism criterion may not apply in some applications. Consider two patients who compete for an appointment slot for the same doctor. First, their valuation of the appointment may be hard to measure because it is subjective and patients may be unable to express true valuations due to budget constraints. Second, even though their valuations are well-measured, it is hard to justify that serving one of them is more efficient than serving the other. This explains why Pareto efficiency is widely used in market design. When the allocative efficiency of slots is not a suitable criterion to compare the two rules, the efficient allocation of time will become the primary criterion and the lottery rule will dominate the queue rule.

[24]This welfare improvement is not guaranteed since they still face a risk of losing the competition for a slot.

we expect participants will spend considerable time queuing in the booking system, as indicated by Proposition 1. Although we do not model participants' strategies in stage 2, even searching only once will entail a greater time cost than pressing an application button in the lottery treatment. This leads to our first hypothesis.

Hypothesis 1. *Participants spend more time on the booking system in both stages of the queue treatment than in the lottery treatment.*

Proposition 1 also implies that participants with higher time valuations will spend more time on the booking system to increase their chance of obtaining a slot in stage 1 of the queue treatment. In contrast, there is no expected correlation between a participant's valuation and time spent in stage 1 of the lottery treatment. In stage 2 of the queue treatment, we assume all participants have an equal chance of winning a slot (Assumption 2). This, along with the stage 1 behavior, leads to the following hypothesis:

Hypothesis 2. *In the queue treatment, participants with higher time valuations will have a greater chance of winning a slot.*

In our analysis of the potential productive efficiency losses between the two solo-track rules, Hypothesis 1 tests our prediction that the strategic efficiency loss is higher in the queue treatment than in the lottery treatment. Our discussion in Section 3.2 predicts that the behavioral efficiency loss is also higher in the queue treatment. In our experimental design, participants' time and monetary valuations of slots are positively correlated. Thus, we predict that the allocative efficiency loss is higher in the lottery treatment than in the queue treatment. We formally state the following hypothesis:

Hypothesis 3. *(a) Productive efficiency loss (both strategic and behavioral) is higher in the queue treatment than in the lottery treatment. (b) Allocative efficiency loss is higher in the lottery treatment than in the queue treatment.*

In the dual-track system, Proposition 2 immediately implies the following hypothesis:

Hypothesis 4. *In stage 1 of the dual-track treatment, more participants choose the lottery track rather than the queue track.*

Finally, according to Proposition 2 , participants with time valuations higher than a given threshold will select the queue track. This leads to our final hypothesis.

Hypothesis 5. *In stage 1 of the dual-track treatment, participants with higher time valuations are more likely to choose the queue track than the lottery track.*

4 Experimental Results

We first present the results from the four solo-track treatments that implement either the queue rule or the lottery rule, and then quantify and compare the different sources of efficiency loss across these treatments. We then present the results from the two dual-track treatments. We organize the presentation of results by the same order of hypotheses derived in the previous section.

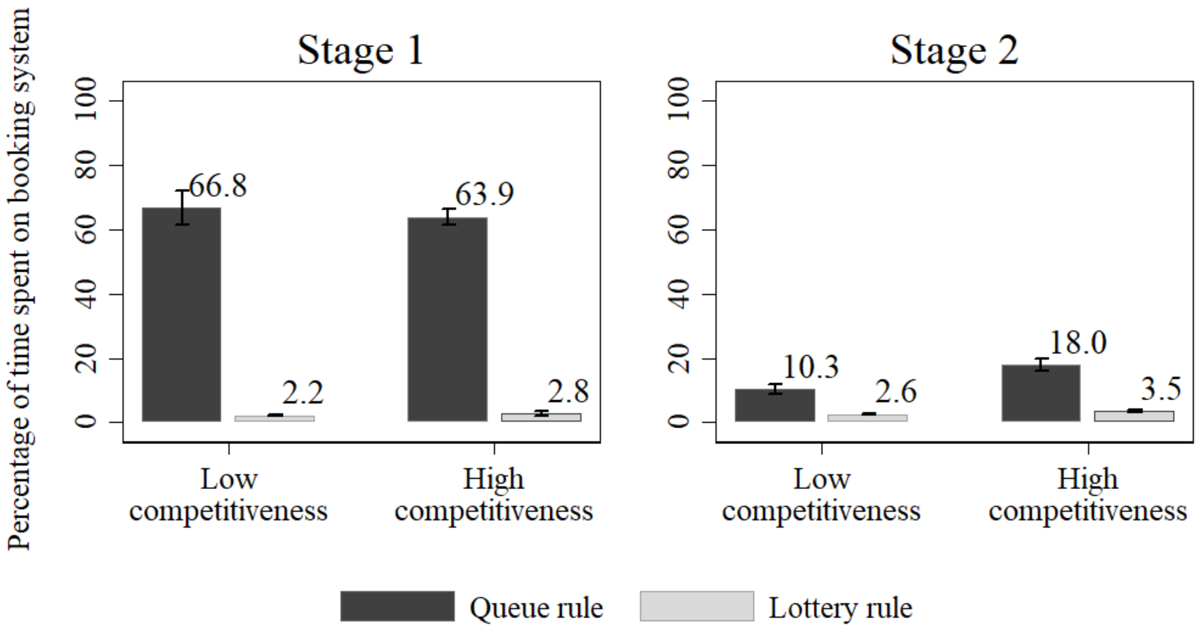
4.1 Solo-track treatments

To test for Hypothesis 1, we first examine how participants allocate their time between the booking task and the production task. Figure 3 displays the percentage of time spent on the booking task in both stages. In stage 1, regardless of the level of market competitiveness, we see that participants in the queue treatment spend almost two-thirds of their time on the booking task. By contrast, participants in the lottery treatment spend only a few seconds on the booking task. This substantial behavioral difference translates into a much lower production task output in the queue versus lottery treatment, as shown in Figure B1 of Appendix B. When market competition is low, the average output in the production task is 2.6 in stage 1 under the queue rule, compared to 6.1 under the lottery rule ($p = 0.004$, Wilcoxon rank-sum test).[25] The gap is even greater when market competition is high, with an average output of 2.2 under the queue rule and 6.8 under the lottery rule ($p = 0.021$). Thus, we reject the null hypothesis in favor of Hypothesis 1. The evidence here strongly suggests that, compared to the lottery rule, the queue rule leads to a substantial (strategic) productive efficiency loss in terms of the opportunity cost of time that would otherwise have been spent on the production task.

Turning to our stage 2 results, we see from Figure 3 (right panel) that, while queue participants still spend significantly more time on the booking task than do lottery participants (10.3% versus 2.6%, $p = 0.004$ in the low competitiveness environment; 18.0% versus 3.5%, $p = 0.021$ in the high competitiveness environment, Wilcoxon rank-sum test), the gap is much smaller than that observed in stage 1. This likely reflects the lower number of participants who still need a slot in this stage. It may also reflect the design feature that participants can freely switch between the booking and production tasks during a round. Moreover, the greater number of participants to slots in the high competitiveness environment means the average time spent on the booking system is significantly higher compared

[25]Unless otherwise stated, we treat each matching group as a unit of observation in all reported statistics.

Figure 3: Percentage of time spent on the booking system in the solo-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level. Low competitiveness is defined as a market with 5 participants and 3 slots (Queue5 and Lottery5), while high competitiveness is defined as a market with 7 participants and 2 slots (Queue7 and Lottery7).

to the low competitiveness environment (18.0% versus 10.3%, $p = 0.019$). As shown in Figure B1, this behavioral pattern is also largely manifested in the stage 2 production task output in the production task. While in the low competitiveness environment the output is not significantly different under the two allocation rules (6.9 versus 6.6, $p = 0.423$), in the high competitiveness environment it is significantly lower under the queue rule than under the lottery rule (5.4 versus 7.0, $p = 0.021$). Finally, we verify that the overall time allocation pattern in stage 2 is driven by those seeking slots in this stage (see Figures B2 and B3 in Appendix B).[26]

[26] Figure B2 shows the percentage of time spent on the booking task separately for those who have (have not) obtained a slot in stage 1. It is clear that the overall pattern in stage 2 is driven mainly by participants seeking slots. Regardless of the level of market competitiveness, these participants spend significantly more time on the booking task under the queue rule than the lottery rule (19.5% versus 3.7%, $p = 0.004$ in the low competitiveness environment; 22.9% versus 3.9%, $p = 0.021$ in the high competitiveness environment, Wilcoxon rank-sum test). By contrast, participants who obtain a slot in stage 1 spend only slightly more time on the booking task under the queue rule than the lottery rule (4.1% versus 1.8%, $p = 0.007$ in the low competitiveness environment; 4.9% versus 2.5%, $p = 0.083$ in the high competitiveness environment).

Result 1. *Regardless of the degree of market competitiveness, participants spend significantly more time on the booking system in both stages of the queue treatment than in the lottery treatment.*

To test for Hypothesis 2, in Table 3 we report random effects probit regression where the dependent variable is whether a subject obtains a slot or not and the independent variable is her time valuation, separately for when a slot is obtained in either one of the stages and when it is obtained in stage 1. Before presenting the results, we explain how the time valuation is calculated for each individual. To do so, we first measure individual productivity per unit of time considering productivity in stage 1 as well as in stage 2 if a participant has obtained a slot in stage 1, since we expect productivity under these two cases is unlikely to be influenced by the booking system.[27] Specifically, in stage 1 we observe participants either work on the production task continually or stay on the booking system, with few instances of distractive switching. Likewise, in stage 2, when some participants have already obtained a slot, their productivity is unlikely to be influenced by the booking system. Thus, to measure individual productivity, we first take the average of each individual's productivity across stage 1 and stage 2 (if the individual has obtained a slot in stage 1) weighted by the respective time spent on the production task. We then take the average of each individual's productivity across all eight paying rounds to obtain our measure of individual (time-invariant) productivity. Finally, we calculate the time valuation for each individual in each round as the ratio of the monetary valuation and individual productivity. Figure B4 in Appendix B shows the distribution of the computed time valuation across all treatments, showing a relatively smooth and left-skewed distribution.

From Table 3, we observe little evidence for a positive correlation between individuals' time

While this may reflect curiosity about the allocation process for the remaining slots, the screen layout under the lottery rule is similar in both stages and no information about allocations is revealed until the end of each stage. Figure B3 plots the corresponding stage 2 output in the production task, showing that queue participants seeking slots in stage 2 produce less than those who have already obtained a slot, though the difference is not statistically significant in the low competitiveness environment (6.4 versus 7.2, $p = 0.200$ in the low competitiveness environment; 4.9 versus 6.8, $p = 0.043$ in the high competitiveness environment). In sum, while not as pronounced as in stage 1, the opportunity cost of time in stage 2 is still higher under the queue rule than the lottery rule, especially when slots are scarcer.

[27]The lack of a performance incentive in the trial round precludes its use as a productivity indicator. Furthermore, if we use only stage 1 productivity as our measure of individual productivity, a significant portion of participants' productivity will not be measured since they have chosen to spend almost all their time queuing in stage 1 in almost all rounds.

Table 3: Random Effects Probit Regressions on the Likelihood of Obtaining a Slot

	Average marginal effects			
	Queue5		Queue7	
	Both stages	Stage 1	Both stages	Stage 1
Time valuation	-0.027	0.009	-0.009	-0.004
	(0.017)	(0.026)	(0.011)	(0.007)
Clusters	6	6	4	4
N	480	480	448	448

Notes: Standard errors clustered at the matching group level are in parentheses. We rescale the time valuation by dividing it by 100.

valuations and their likelihood of obtaining a slot in either Queue5 or Queue7. Therefore, it appears that the queue rule does not really harm the fairness of slot allocations.[28] Thus, we cannot reject the null hypothesis in favor of Hypothesis 2.

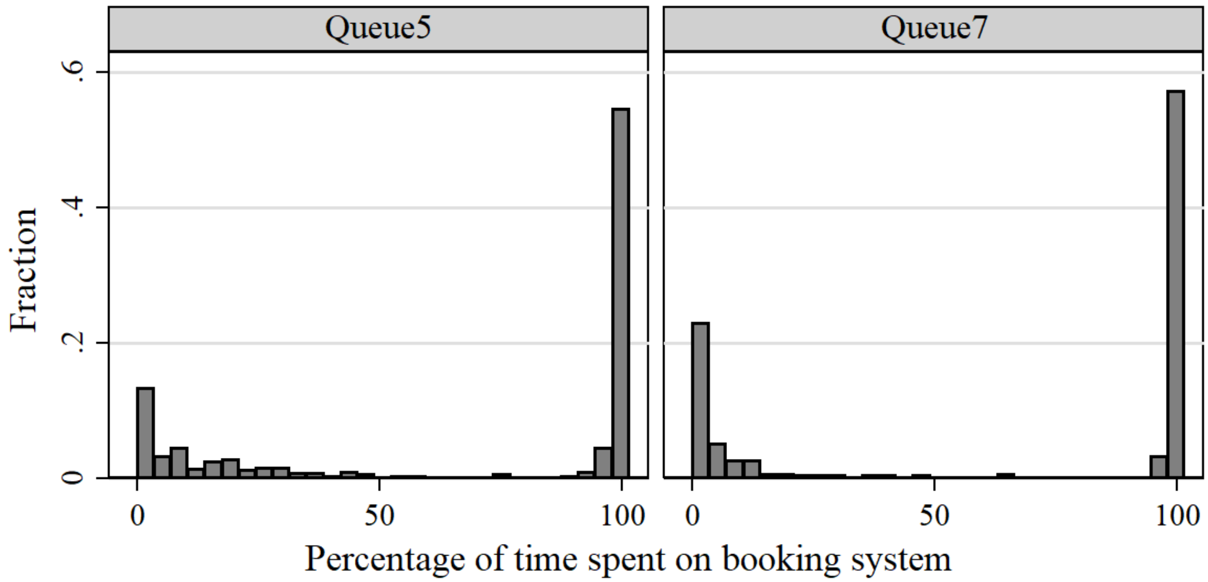
To explore this finding in greater depth, we compare participant time allocations in stage 1 for our two queue treatments. The findings in Figure 4 show a pattern of bimodal behavior in both queue treatments: participants spend either very little time or almost all of their time on the booking task. Specifically, we find that 66.0% (77.2%) of our observations in Queue5 (Queue7) consist of those at the extremes (either fewer than 5 seconds or more than 235 seconds).[29] In theory, this bimodal behavior would happen if participants' time valuation follows a bimodal distribution with peaks at both ends. However, this is clearly not what we observe in our experiment (see Figure B4 in Appendix B). Thus, the tenuous relationship between the time valuation and the likelihood of obtaining a slot is likely due to the weak relationship between the time valuation and the time spent on the booking system. We summarize these findings below.

Result 2. *There is no evidence of a positive correlation between the time valuation and the probability of obtaining a slot in the queue treatments, suggesting that the queue rule*

[28] Table B1 in Appendix B shows that these results are robust to controlling for individual characteristics that we collect in the post-experiment survey. We also observe that feeling more anxious in stage 2 is positively correlated with a higher likelihood of finally obtaining a slot. Moreover, females are less likely to obtain a slot, and more competitive individuals are more likely to obtain a slot, but only in the Queue7 treatment.

[29] A similar bimodal behavior is also observed in an experimental all-pay auction with incomplete information about individual marginal costs of bidding (Müller and Schotter, 2010).

Figure 4: Distribution of time spent on the booking system in stage 1 of the queue treatments



does not compromise allocation fairness. These results reflect a pattern of bimodal behavior: participants either drop out from the booking system or spend almost all their time on it.

4.2 Quantifying different sources of efficiency loss

The previous subsection shows that the lottery rule is superior to the queue rule in terms of productive efficiency. In this subsection, we compare the different types of efficiency losses at the group level across the two allocation rules. As discussed in Section 3.2, efficiency loss can stem from allocative (closely related to fairness) or productive efficiency loss, the latter of which can be further divided into strategic (due to wasted time) and behavioral (due to distraction taking away from the production task) efficiency loss.

Table 4 reports the quantified efficiency loss (in ECUs) of each type for each solo-track treatment. In addition, the table reports the total efficiency loss, which is the sum of the three types of efficiency losses. For strategic efficiency loss, we further distinguish our results by stage. For behavioral efficiency loss, we also separate our results by participant subgroups: one in which participants have already obtained a slot in stage 1 and the other in which participants have not obtained a slot in stage 1. We expect greater behavioral efficiency loss in this latter group.

From Table 4, we see that the strategic efficiency loss is approximately one order of magnitude

Table 4: The Breakdown of Efficiency Loss in Each Solo-track Treatment

	Queue5	Queue7	Lottery5	Lottery7
Allocative efficiency loss	85.115	166.406	59.229	147.313
	(9.603)	(13.853)	(5.454)	(9.752)
Productive efficiency loss				
Strategic	978.675	1242.640	53.892	101.092
	(31.433)	(40.859)	(7.034)	(10.133)
(a) Stage 1	844.057	965.542	24.578	43.712
	(30.721)	(34.088)	(3.743)	(7.664)
(b) Stage 2	134.618	277.098	29.314	57.380
	(7.705)	(19.205)	(3.605)	(5.144)
Behavioral	-16.045	-67.570	-46.570	-32.914
	(21.221)	(29.468)	(19.124)	(25.022)
(a) Slot	-5.569	-13.991	-32.353	-1.310
	(15.073)	(13.766)	(14.024)	(11.171)
(b) No slot	-10.476	-54.016	-14.217	-31.604
	(10.961)	(21.27)	(8.601)	(19.587)
Total efficiency loss	1047.745	1341.476	66.552	215.490
	(37.655)	(56.442)	(22.868)	(28.941)
Obs. (group × round)	96	64	96	64

Notes: Standard errors are in parentheses.

larger than the allocative efficiency loss under the queue rule, while the two are very similar under the lottery rule. Further, as expected, we see that the strategic efficiency loss under the queue rule is driven mainly by queuing in stage 1, although the amount of loss in stage 2 is not negligible. The strategic efficiency loss in either stage is significantly larger than the allocative efficiency loss ($p < 0.001$ in each comparison, Wilcoxon signed-rank test).

Moreover, we find that the allocative efficiency loss is slightly higher under the queue rule than the lottery rule, although this difference is only marginally significant in the low market competitiveness environment (Queue5 versus Lottery5, $p = 0.070$; Queue7 versus Lottery7, $p = 0.483$; Wilcoxon ranksum test).

Finally, and perhaps surprisingly, we find no evidence of behavioral efficiency loss under the queue rule. In Queue5, the behavioral efficiency loss does not differ from zero ($p =$

0.452, two-sided t-test). In Queue7, the behavioral efficiency loss is significantly lower than zero ($p = 0.025$), driven mainly by participants who have not obtained a slot in stage 1 ($p = 0.014$). These results may reflect greater production task effort, perhaps as a means of compensating for wasted time in both stages.[30] More importantly, however, the magnitude of the observed behavioral efficiency loss or gain is one order of magnitude lower than that of the strategic efficiency loss.

Overall, our results show that, under the queue rule, the productive efficiency loss from wasted time in stage 1 is much larger than the allocative or behavioral efficiency loss. The lottery rule is superior to the queue rule for each type of efficiency loss, and especially for strategic efficiency loss. From this set of results, we conclude that our data support Hypothesis 3(a) but we cannot reject the null hypothesis in favor of Hypothesis 3(b).

Result 3. *The lottery rule is superior to the queue rule in terms of productive efficiency, and not inferior in terms of allocative efficiency. The substantial productive efficiency loss under the queue rule reflects mainly strategic efficiency loss.*

4.3 Dual-track treatments

We have shown in our solo-track treatments that the lottery rule yields better efficiency than the queue rule in almost every aspect. In this section, we present our results from the two dual-track treatments which allow participants to choose between the queue rule and the lottery rule in stage 1.

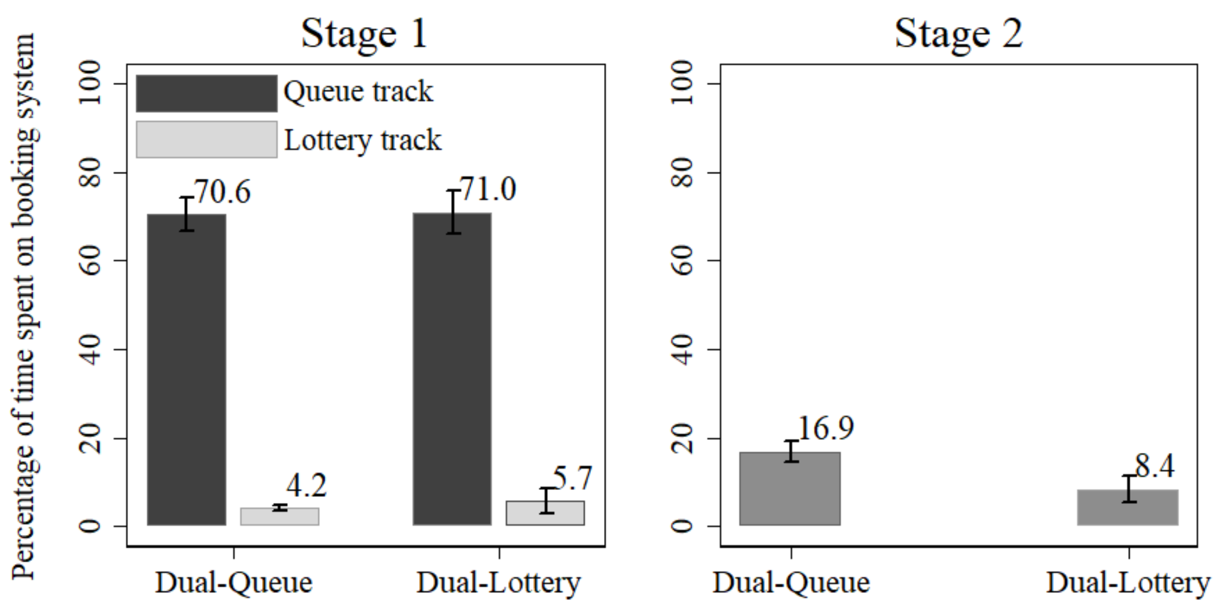
Before presenting our results related to Hypotheses 4 and 5, we briefly examine participants' behavior under each track. In general, we find similar behavior within an allocation rule across the solo- and dual-track systems.[31] From Figure 5, we see that dual-track queue

[30] In the post-experimental questionnaire, participants were asked to indicate their level of anxiety in both stages on a scale from 1 (not anxious at all) to 7 (extremely anxious). On average, the reported level of anxiety is significantly higher in each stage under the queue rule than the lottery rule. Further, the reported level of anxiety is higher in stage 2 than in stage 1 of both queue treatments ($p < 0.001$, Wilcoxon signed-rank test). In the questionnaire, participants were also asked to indicate how the booking task had disturbed their performance in the production task on a scale from 1 (not disturbing at all) to 7 (extremely disturbing). On average, the reported level of disturbance is about 3.85 ($s.d. < 2.00$) in both queue treatments, suggesting no evidence for such a disturbance. These results provide some suggestive evidence that the increased anxiety level in stage 2 is mainly performance-enhancing.

[31] While we derive a symmetric Bayesian NE in stage 1 for both the solo- and dual-track systems, without knowing the distribution of time valuation, it is hard to predict which system causes higher average wasted time under the queue rule.

(lottery) participants spend two-thirds (5%) of their time on the booking task in stage 1. We further see that queue participants spend significantly more time on the booking task in stage 2 compared to lottery participants (16.9% versus 8.4%, $p = 0.083$, Wilcoxon rank-sum test).[32] From Figure B5 in Appendix B, we see that queue participants yield substantially lower production outputs in stage 1 than lottery participants and slightly lower outputs in stage 2 (6.6 versus 5.7, $p = 0.083$). [33]

Figure 5: Percentage of time spent on the booking system in the dual-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level.

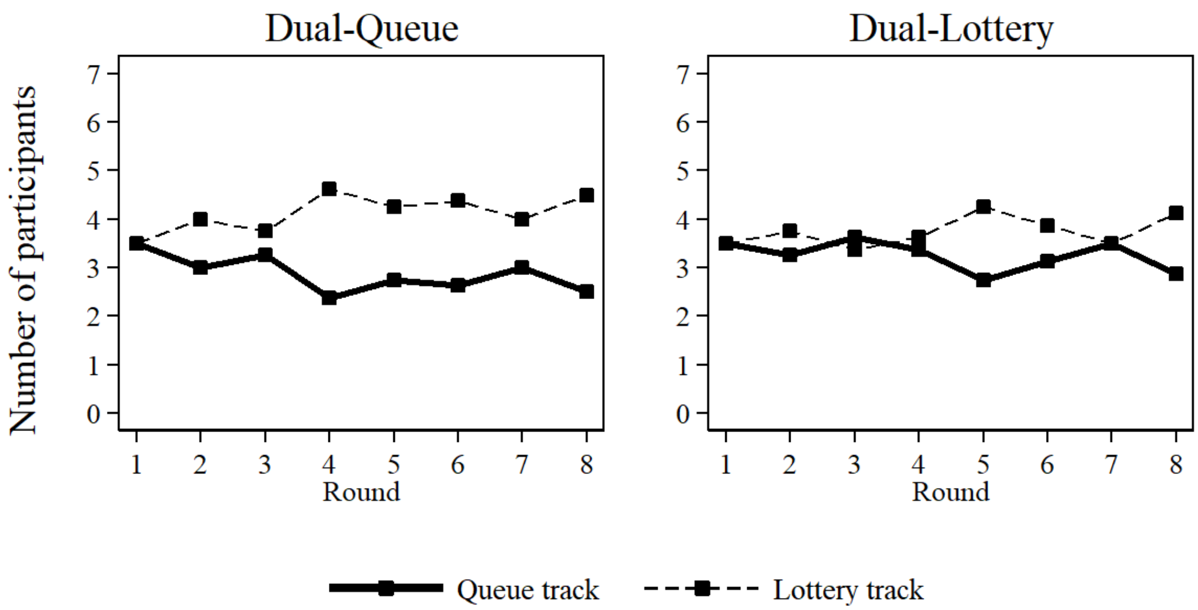
To test for Hypothesis 4, we examine the number of participants who chose each track over round. Figure 6 shows that, on average, participants are marginally significantly more likely to choose the lottery track than the queue track (Dual-Queue: 58.9% vs. 41.1%, $p = 0.066$;

[32]Figure B6 in Appendix B shows the time allocation behavior in stage 2 separately for those who need a slot and those who have already obtained a slot. The results are again similar to those for the solo-track treatments. Those seeking slots spend more time on the booking screen when stage 2 uses the queue rule versus the lottery rule. By contrast, the behavior is similar for those who have already obtained a slot.

[33]Figure B7 in Appendix B shows the stage 2 output separately for those who need a slot and those who have already obtained a slot. The results are again similar to those for the solo-track treatments. The average output of those seeking slots appears higher when stage 2 uses the lottery rule versus the queue rule, but the difference is not statistically significant ($p = 0.149$). By contrast, the output is similar for those who have already obtained a slot.

Dual-Lottery: 53.6% vs. 46.4%, $p = 0.068$, Wilcoxon signed-rank test). Thus, we reject the null hypothesis in favor of Hypothesis 4.

Figure 6: Track choices in the dual-track treatments



To test for Hypothesis 5, we conduct random-effects probit regression analysis of participants' likelihood of choosing the queue track on their time valuations.[34] The results in columns (1) and (3) of Table 5 show that those with higher time valuations are more likely to choose the queue track, though the effect is statistically significant only in the Dual-Lottery treatment. We conjecture that participants may be less sensitive to their productivity than their monetary valuation. Thus, we also investigate the separate effects of monetary valuation and productivity on a participant's choice of track. The results in columns (2) and (4) show that monetary valuation has a significantly positive influence on the likelihood of choosing the queue track: for example, in the Dual-Lottery treatment, a 100-ECUs increase in the monetary valuation increases the probability of choosing the queue track by 13.2%. On the contrary, productivity does not have a significant impact in either dual-track treatment.

Result 4. *In stage 1 of the dual-track treatments, participants are more likely to choose the*

[34] Table B3 in Appendix B tests for robustness by further controlling for individual characteristics we collect in the questionnaire. We also observe that feeling more anxious in stage 1 is correlated with a lower likelihood of choosing the queue track in the Dual-Queue treatment. Moreover, we find that females are more likely to choose the queue track in the Dual-Queue treatment.

Table 5: Random-effects Probit Regressions on the Likelihood of Choosing the Queue Track

	Average marginal effects			
	Dual-Queue		Dual-Lottery	
	(1)	(2)	(3)	(4)
Time valuation	0.037		0.027^{**}	
	(0.023)		(0.012)	
Monetary valuation		0.147^{***}		0.132^{***}
		(0.055)		(0.028)
Productivity		0.067		0.071
		(0.059)		(0.074)
Clusters	4	4	4	4
N	448	448	448	448

Notes: Standard errors clustered at the matching group level are in parentheses. We rescale the time valuation by dividing it by 100. *** $p < 0.01$.

lottery track instead of the queue track.

Result 5. *In the Dual-Lottery treatment, participants with higher time valuations are significantly more likely to choose the queue track instead of the lottery track.*

Recall that, in the solo-track treatments, the allocative efficiency loss tends to be larger under the queue rule than the lottery rule, as over 50% of participants spend all their time on the booking task regardless of their time valuation. The fact that the queue track in the dual-track treatments does attract some participants with higher time valuations implies that the overall allocative efficiency might be improved compared to the solo-track treatments: on the one hand, more than half of participants chose the lottery track, which should lead to a similar level of allocative efficiency as in the solo-track lottery treatments; on the other hand, participants with higher time valuations should have a higher likelihood of obtaining slots in the queue track.

This conjecture is partially supported. Table 6 reports the breakdown of efficiency loss in the dual-track treatments. We observe a very similar level of allocative efficiency loss in the Dual-Lottery treatment and Lottery7 treatment ($p = 0.703$, Wilcoxon ranksum test). Thus, the presence of the queue track does not increase the allocative efficiency loss in this treatment. However, in the Dual-Queue treatment, we observe a level of allocative efficiency loss comparable to that in the Queue7 treatment ($p = 0.768$). Overall, we conclude that the dual-track system may reduce the allocative efficiency loss seen in the solo-track queue system.

Table 6: The Breakdown of Efficiency Loss in Each Dual-track Treatment

	Dual-Queue	Dual-Lottery
Allocative efficiency loss	164.750	143.703
	(14.509)	(10.066)
Productive efficiency loss		
Strategic	814.570	657.927
	(48.037)	(40.011)
(a) Stage 1	521.376	564.597
	(38.569)	(38.011)
(b) Stage 2	293.194	93.330
	(22.560)	(8.168)
Behavioral	-58.612	-68.880
	(25.787)	(31.606)
(a) Slot	-13.789	-12.285
	(10.694)	(13.694)
(b) No slot	-44.823	-56.595
	(20.429)	(23.509)
Total efficiency loss	920.709	732.750
	(53.893)	(52.022)
Obs. (group $\times$ round)	64	64

Notes: Standard errors are in parentheses.

Examining productive efficiency loss in the dual-track treatments, we see that the strategic efficiency loss under the queue rule remains substantial although fewer participants leads to a lower total amount of loss compared to the solo-track queue treatment. Similar to the solo-track treatments, we find no evidence for behavioral efficiency loss, and again find that behavioral efficiency loss is significantly lower than zero ($p = 0.027$ and $p = 0.033$ for Dual-Queue and Dual-Lottery). Nevertheless, the efficiency gain is one order of magnitude lower than the strategic efficiency loss.

Overall, our results show that the dual-track system can substantially reduce the total amount of productive efficiency loss compared to the solo-track queue system. Moreover, the dual-track system can also reduce potential allocative efficiency loss by channeling some participants with higher time valuations to compete for slots in the queue track. We summarize these findings below.

Result 6. *The dual-track system partially restores the allocative and productive efficiency that has been lost in the solo-track queue system.*

5 Concluding Remarks

When scarce goods or services are provided for free or under price control, rationing rules must be used to allocate these resources in a manner that balances fairness and efficiency. This paper evaluates the fairness and efficiency of different rules for allocating goods such as appointment slots in public offices, tickets for entertainment and sporting events, and courses and on-campus housing at colleges and universities. One commonly-used allocation rule, the queuing system, is criticized for efficiency losses due to the opportunity cost of time spent on the queuing process.

Our study quantifies the different sources of efficiency loss under different allocation rules by developing and testing a flexible experimental framework. Specifically, we theoretically and experimentally compare the performance of a queue rule based on a first-come, first-served principle with that of a lottery rule that relies on a random selection process to allocate goods. Our experimental results show that the lottery rule yields superior fairness and efficiency in nearly every aspect. In particular, the queue rule creates efficiency losses due to opportunity costs of time spent waiting that are substantial enough to overwhelm other efficiencies, leading to lower participant welfare.

Acknowledging that participants may still prefer a queuing system, we also introduce a hybrid booking system that allows participants to choose between the two allocation rules. The results using our dual-track system show that a majority of participants select the lottery rule even when the ex-ante chance of obtaining a slot is exactly the same under both rules, since the significantly lower time cost under the lottery rule means a much higher payoff that more than offsets any lower ex-post chance of obtaining a slot.

While our study strongly supports the increased usage of the lottery rule in real-world applications, we note the practical implementation and design of the parameters of the lottery rule will depend on the context. For example, market designers will need to choose how early in advance appointment slots become available to participants as well as how late in the appointment cycle to allow cancellations and rebooking. For instance, public hospitals in China typically release appointment slots one week in advance, and cancellations are allowed one day before the service date. These design choices are important in relieving participant anxiety due to the uncertainty related to lottery results. If the initial application stage is too long and/or the date for the lottery draw is too close to the service date, participants will have less time to search for other possibilities if the lottery outcome turns out to be negative.

Our dual-track system may also hopefully inspire the development of other hybrid systems that may better cater to specific participant needs. For example, designers may implement a vertical type of hybrid system in which the lottery rule is used in the initial allocation stage and the queue rule is used in the cancellation and re-booking stage. Compared to a booking system that uses only the lottery rule, this hybrid system could allow cancellation and re-booking participants to know immediately whether they have successfully booked a slot. This hybrid system can be attractive especially when the cancellation rate is so low that the efficiency loss due to monitoring or searching is limited.

Finally, our flexible experimental framework is versatile enough to be applied to numerous settings. It can also form the basis for more complex booking scenarios, as when participants have preferences over different slots. For example, some patients visiting hospitals may prefer morning slots over afternoon ones. One potential solution to this issue is to borrow process steps from school choice matching algorithms. Participants might begin by submitting rank-order lists of slots to reveal their preferences. The process could then use a matching algorithm (which may include the use of lotteries to break ties) to find an initial allocation of slots. In the second step, participants would have the option to cancel assigned slots and re-book others. Depending on the number of remaining slots, all canceled and/or unassigned slots could be allocated either on a first-come, first-served basis or via lotteries.[35]

[35]We document a field application that contains some of the design features discussed. The application is the undergraduate course allocation process at a Chinese university. The allocation process for the next semester begins in the last few weeks of the current semester. All undergraduate courses for the next semester are allocated through three stages, each lasting from five days to two weeks. In the first stage, students can submit rank-order lists to indicate their course preferences. When the submission deadline is reached, the university runs the immediate acceptance algorithm to allocate courses and breaks priority ties for the same course first by credit hours earned and then by lotteries. The second stage includes two steps. In the first step, all remaining seats of courses are allocated using the same method as in the first stage. Most courses will be fully allocated after this step. In the second step, all remaining seats of courses are released at a certain date and allocated on a first-come, first-served basis, where students compete for the remaining

As long as the allocation of slots is independent of the timing of booking slots or submitting applications, this system can help to eliminate the efficiency loss due to the opportunity cost of time in a similar manner as the lottery rule.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 1998. "Random serial dictatorship and the core from random endowments in house allocation problems." *Econometrica*, 66(3): 689–701.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School choice: A mechanism design approach." *The American Economic Review*, 93(3): 729–747.

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan. 2020. "Thickness and information in dynamic matching markets." *Journal of Political Economy*, 128(3): 783–815.

Baccara, Mariagiovanna, SangMok Lee, and Leeat Yariv. 2020. "Optimal dynamic matching." *Theoretical Economics*, 15(3): 1221–1278.

Bartling, Björn, Ernst Fehr, and Holger Herz. 2014. "The Intrinsic Value of Decision Rights." *Econometrica*, 82(6): 2005–2039.

Barzel, Yoram. 1974. "A theory of rationing by waiting." *The Journal of Law and Economics*, 17(1): 73–95.

Bednar, Jenna, Yan Chen, Tracy Xiao Liu, and Scott Page. 2012. "Behavioral spillovers and cognitive load in multiple games: An experimental study." *Games and Economic Behavior*, 74(1): 12–31.

Chen, Yan, and Tayfun Sönmez. 2002. "Improving efficiency of on-campus housing: An experimental study." *American economic review*, 92(5): 1669–1686.

Che, Yeon-Koo, and Olivier Tercieux. 2021. "Optimal Queue Design." *working paper*.

seats by submitting applications as early as possible. The first two stages are completed before the end of the current semester. During the first two weeks of the next semester, students enter the third and last stage, where they may cancel some of their courses obtained in the previous semester and compete for other available courses on a first-come, first-served basis.

Dimakopoulos, Philipp D, and C-Philipp Heller. 2019. "Matching with waiting times: The german entry-level labor market for lawyers." *Games and Economic Behavior*, 115: 289–313.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2): 171–178.

Gavious, Arieh, Benny Moldovanu, and Aner Sela. 2002. "Bid costs and endogenous bid caps." *RAND Journal of Economics*, 33(4): 709–722.

Guillen, Pablo, and Onur Kesten. 2012. "Matching Markets with Mixed Ownership: The Case for a Real-life Assignment Mechanism." *International Economic Review*, 53(3): 1027–1046.

Hakimov, Rustamdjan, and Dorothea Kübler. 2021. "Experiments on centralized school choice and college admissions: a survey." *Experimental Economics*, 24(2): 434–488.

Hakimov, Rustamdjan, C.-Philipp Heller, Dorothea Kübler, and Morimitsu Kurino. 2021. "How to avoid black markets for appointments with online booking systems." *American Economic Review*, 111(7): 2127–51.

Heller, C, Johannes Johnen, Sebastian Schmitz, et al. 2019. "Congestion pricing: A mechanism design approach." *Journal of Transport Economics and Policy (JTEP)*, 53(1): 74–98.

Holt, Charles A, and Roger Sherman. 1982. "Waiting-line auctions." *Journal of Political Economy*, 90(2): 280–294.

Huang, Yangguang, and Quan Wen. 2019. "Auction–Lottery Hybrid Mechanisms: Structural Model and Empirical Analysis." *International Economic Review*, 60(1): 355–385.

Hugh-Jones, David, Morimitsu Kurino, and Christoph Vanberg. 2014. "An experimental study on the incentives of the probabilistic serial mechanism." *Games and Economic Behavior*, 87: 367–380.

Krishna, Vijay. 2009. *Auction Theory.* Academic Press.

Larson, Richard C. 1987. "OR forum – perspectives on queues: Social justice and the psychology of queueing." *Operations Research*, 35(6): 895–905.

Leroy, Sophie. 2009. "Why is it so hard to do my work? The challenge of attention residue when switching between work tasks." *Organizational Behavior and Human Decision Processes*, 109(2): 168–181.

Leshno, Jacob D. 2022. "Dynamic matching in overloaded waiting lists." *American Economic Review*, 112(12): 3876–3910.

Li, Shanjun. 2018. "Better lucky than rich? Welfare analysis of automobile licence allocations in Beijing and Shanghai." *The Review of Economic Studies*, 85(4): 2389–2428.

Müller, Wieland, and Andrew Schotter. 2010. "Workaholics and Dropouts in Organizations." *Journal of the European Economic Association*, 8(4): 717–743.

Naor, Pinhas. 1969. "The regulation of queue size by levying tolls." *Econometrica*, 15–24.

Nichols, Donald, Eugene Smolensky, and T Nicolaus Tideman. 1971. "Discrimination by waiting time in merit goods." *The American Economic Review*, 61(3): 312–323.

Owens, David, Zachary Grossman, and Ryan Fackler. 2014. "The Control Premium: A Preference for Payoff Autonomy." *American Economic Journal: Microeconomics*, 6(4): 138–161.

Parry, Ian W H, Margaret Walls, and Winston Harrington. 2007. "Automobile externalities and policies." *Journal of economic literature*, 45(2): 373–399.

Platz, Trine Tornøe, and Lars Peter Østerdal. 2017. "The curse of the first-in–first-out queue discipline." *Games and Economic Behavior*, 104: 165–176.

Roth, Alvin E. 2021. "Experiments in Market Design." In *Handbook of Experimental Economics*. Vol. 2, , ed. John H Kagel and Alvin E Roth, 290–346. Princeton, NJ:Princeton University Press.

Schummer, James. 2021. "Influencing waiting lists." *Journal of Economic Theory*, 195: 105263.

Schummer, James, and Azar Abizada. 2017. "Incentives in landing slot problems." *Journal of Economic Theory*, 170: 29–55.

Suen, Wing. 1989. "Rationing and rent dissipation in the presence of heterogeneous individuals." *Journal of Political Economy*, 97(6): 1384–1394.

Taylor, Grant A, Kevin KK Tsui, and Lijing Zhu. 2003. "Lottery or waiting-line auction?" *Journal of Public Economics*, 87(5-6): 1313–1334.

Tobin, James. 1952. "A Survey of the Theory of Rationing." *Econometrica*, 521–553.

Appendix

A Proof of Proposition 1 and 2

Proof of Proposition 1. For convenience, we first derive the equilibrium by ignoring the cap restriction T. For any participant i, to win a slot in equilibrium in stage 1 of the queue treatments, her time valuation must exceed the m-th largest of the remaining $n-1$ participants' time valuations. Let H and h respectively denote the cumulative distribution function and the density function of the m-th order statistics among $n-1$ independent draws from F:

$$H(y) = \sum_{\ell=n-m}^{n-1} \binom{n-1}{\ell} [F(y)]^{\ell}[1-F(y)]^{n-1-\ell}.$$

If a participant i bids b_i units of time in the queue for slots, her winning probability will be $H(t^{-1}(b_i))$. If she fails, she still has a probability of $\frac{1}{n-m}$ to win the canceled slot in stage 2. No matter she wins or not in stage 1, she needs to pay $b_i \cdot w_i$, the opportunity cost of time. Assuming that the other participants follow the equilibrium strategy $t(y)$, i needs to solve the following problem:

$$\max_{b_i} \left[H(t^{-1}(b_i))y_i w_i + (1 - H(t^{-1}(b_i)))\frac{1}{n-m} y_i w_i - b_i w_i\right]. \tag{1}$$

which is equivalent to

$$\max_{b_i} \left[H(t^{-1}(b_i))y_i \cdot \frac{n-m-1}{n-m} - b_i\right]. \tag{2}$$

The first-order condition to (2) is

$$h(t^{-1}(b_i))\frac{dt^{-1}(b_i)}{db_i} y_i \frac{n-m-1}{n-m} - 1 = 0. \tag{3}$$

In equilibrium, this equation holds when $b_i = t(y_i)$. So $t^{-1}(b_i) = y_i$ and $\frac{dt^{-1}(b_i)}{db_i} = \frac{1}{t'(y_i)}$.

Then

$$t'(y_i) = h(y_i)y_i \frac{n-m-1}{n-m}. \tag{4}$$

Given the boundary condition $t(\underline{y}) = 0$, we obtain

$$t(y_i) = \frac{n-m-1}{n-m} \int_{\underline{y}}^{y_i} h(s)s ds. \tag{5}$$

Now we analyze three cases of equilibrium.

Case 1: If $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds \leq T$, which means that the participant of the highest time valuation will bid no more than T when there is no cap, then the cap T is not a binding constraint. So $t(y_i)$ in equation (5) is the equilibrium bidding strategy under the queue rule.

Case 2: If $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds > T$, it means the cap T is a binding constraint. Then there exists a threshold $y^\diamond \in [\underline{y}, \overline{y}]$ such that all participants with time valuations weakly above $y^\diamond$ will pool by bidding the cap T. If $y^\diamond > \underline{y}$ (we will derive the condition for this to happen in Case 3), to derive the equilibrium strategy for participants of time valuations below $y^\diamond$, we note that for any such participant to win a slot in stage 1, her time valuation must exceed the m-th largest of the remaining $n-1$ participants' time valuations. So we can repeat the above steps to obtain the bidding strategy $t(y_i)$ in equation (5). To derive the threshold $y^\diamond$, note that a participant i of the time valuation $y^\diamond$ will be indifferent between bidding the cap T and bidding $t(y^\diamond)$. By bidding $t(y^\diamond)$, i's expected payoff is

$$\begin{aligned}&\big[H(y^\diamond)y^\diamond + (1-H(y^\diamond))\frac{1}{n-m}y^\diamond - t(y^\diamond)\big]\cdot w_i\\ =&\big[\frac{1}{n-m}y^\diamond + \frac{n-m-1}{n-m}H(y^\diamond)y^\diamond - \frac{n-m-1}{n-m}\int_{\underline{y}}^{y^\diamond} h(s)sds\big]\cdot w_i\\ =&\big[\frac{1}{n-m}y^\diamond + \frac{n-m-1}{n-m}\int_{\underline{y}}^{y^\diamond} H(s)ds\big]\cdot w_i.\end{aligned}$$

While by bidding T, i's probability of winning a slot in stage 1 is

$$\Pr(\text{win}|\text{T}) = \sum_{\ell=0}^{n-m-1}\binom{n-1}{\ell}[F(y^\diamond)]^\ell[1-F(y^\diamond)]^{n-1-\ell}\frac{m}{n-\ell} + \sum_{\ell=n-m}^{n-1}\binom{n-1}{\ell}[F(y^\diamond)]^\ell[1-F(y^\diamond)]^{n-1-\ell},$$

and therefore, i's expected payoff is

$$\big[\Pr(\text{win}|\text{T})y^\diamond + (1-\Pr(\text{win}|\text{T}))\frac{1}{n-m}y^\diamond - T\big]\cdot w_i.$$

So the threshold $y^\diamond$ is determined by the equation

$$\frac{n-m-1}{n-m}\big[\Pr(\text{win}|\text{T}) - \int_{\underline{y}}^{y^\diamond} H(s)ds\big] = T. \qquad (6)$$

Case 3: $y^\diamond = \underline{y}$ happens when it is profitable for a participant i of the lowest time valuation to pool with all of the others by bidding the cap T. By bidding T, her winning probability in stage 1 is $\frac{m}{n}$. If she fails, she still has a probability of $\frac{1}{n-m}$ to win the canceled slot in stage 2. So her total winning probability is $\frac{m}{n} + (1-\frac{m}{n})\frac{1}{n-m} = \frac{1}{n-m} + \frac{m(n-m-1)}{n(n-m)}$. If she does not bid T, she will fail for sure in stage 1 and have a probability of $\frac{1}{n-m}$ to win the canceled slot in stage 2. So by bidding T the net increase in winning probability is $\frac{m(n-m-1)}{n(n-m)}$. The cost of bidding T is the loss of the payoff $T \cdot w_i$ in the production task. So i will bid T if and only if $\frac{m(n-m-1)}{n(n-m)}\underline{y} \cdot w_i \geq T \cdot w_i$. Therefore, Case 3 happens when $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds > T$ and $\frac{m(n-m-1)}{n(n-m)}\underline{y} \geq T$, while Case 2 happens when $\frac{n-m-1}{n-m}\int_{\underline{y}}^{\overline{y}} h(s)sds > T$ and $\frac{m(n-m-1)}{n(n-m)}\underline{y} < T$. $\square$

Proof of Proposition 2. We first take the threshold y^* as given (we will derive y^* later). For any participant i, let $J(x; y^*)$ denote the probability that among the other $n-1$ participants x of them have time valuations weakly above y^* and therefore choose the queue track:

$$J(x; y^*) = \binom{n-1}{x}[1-F(y^*)]^x[F(y^*)]^{n-1-x}.$$

If i chooses the lottery track, in stage 1 she will win a slot for sure if $n-x \leq k$ (i.e., no more than k participants choose the lottery track) and with probability $\dfrac{k}{n-x}$ if otherwise. If i does not win a slot in the lottery track in stage 1, there are two cases in stage 2. If $x \geq k$, it means both tracks attract enough participants and in stage 2 there will be only a canceled slot. So i will win a slot in stage 2 with probability $\frac{1}{n-m}$. If $x < k$, then $k-x$ slots in the queue track will remain available in stage 2 and $n-x-k$ participants in the lottery track will not win slots in stage 1. So i will win a slot in stage 2 with probability $\frac{k-x+1}{n-x-k}$. Thus, i's probability of winning a slot by choosing the lottery track is

$$\begin{aligned} P_L(y_i; y^*) = &\underbrace{\sum_{x=n-k}^{n-1} J(x; y^*)}_{A} + \underbrace{\sum_{x=k}^{n-k-1} J(x; y^*)\Big[\frac{k}{n-x} + \frac{n-x-k}{n-x}\frac{1}{n-m}\Big]}_{B} \\ &+ \underbrace{\sum_{x=0}^{k-1} J(x; y^*)\Big[\frac{k}{n-x} + \frac{n-x-k}{n-x}\frac{k-x+1}{n-x-k}\Big]}_{C}, \end{aligned}$$

Note that $P_L(y_i; y^*)$ is independent of y_i. For convenience, we write it as $P_L(y^*)$.

If i chooses the queue track, i needs to choose a bidding strategy. We first consider the case that $t(y^*) = 0$. That is, if any participant i of the threshold time valuation chooses the queue track, she will wait for zero time in the queue. Then she will win a slot in stage 1 if and only if $x+1 \leq k$. If i does not win a slot in stage 1, there are two cases in stage 2. If $n-1-x \geq k$, both tracks attract enough participants and in stage 2 there will be only a canceled slot. So i will win a slot in stage 2 with probability $\frac{1}{n-m}$. If $n-1-x < k$, $k-(n-1-x)$ slots in the lottery track will remain available in stage 2 and $x+1-k$ participants in the queue track will not win slots in stage 1. So i will win a slot in stage 2 with probability $\frac{x+2+k-n}{x+1-k}$. Therefore, i's probability of winning a slot by choosing the queue track is

$$P_Q(y^*) = \underbrace{\sum_{x=0}^{k-1} J(x;y^*)}_{C'} + \underbrace{\sum_{x=k}^{n-k-1} J(x;y^*)\frac{1}{n-m}}_{B'} + \underbrace{\sum_{x=n-k}^{n-1} J(x;y^*)\frac{x+2+k-n}{x+1-k}}_{A'}.$$

Because a participant of the threshold time valuation y^* is indifferent between choosing the lottery track and the queue track, it means that y^* is determined by

$$P_L(y^*) = P_Q(y^*). \tag{7}$$

Before deriving the bidding strategy of the participants who choose the queue track, we first prove that in expectation more than half of the participants will choose the lottery track.

Claim 1. $F(y^*) > 1/2$.

Proof of Claim 1. It is easy to see that $B > B'$. We prove that if $F(y^*) \leq 1/2$, then $A + C > A' + C'$, which leads to a contradiction because $A + B + C = A' + B' + C'$.

Note that

$$A + C - (A' + C') = \sum_{x=n-k}^{n-1} J(x;y^*)\frac{n-(2k+1)}{x+1-k} - \sum_{x=0}^{k-1} J(x;y^*)\frac{n-(2k+1)}{n-x}.$$

Now for every $x \in \{n-k, \ldots, n-1\}$, define $x' = n-1-x$. Then $x' \in \{0, \ldots, k-1\}$. So

$$A + C - (A' + C') = \sum_{x=n-k}^{n-1} \left[J(x;y^*)\frac{n-(2k+1)}{x+1-k} - J(x';y^*)\frac{n-(2k+1)}{n-x'}\right].$$

If $F(y^*) \leq 1/2$, it is easy to see that for every $x \in \{n-k, \ldots, n-1\}$, $J(x;y^*) \geq J(x';y^*)$. Also, $x+1-k < n-x'$, because $n - x' - (x+1-k) = k > 0$. So $A + C > A' + C'$. □

To derive the bidding strategy of those who choose the queue track, we first assume that the cap T is not binding. Consider any participant i with $y_i > y^*$. She will win a slot for sure in the queue track if $x + 1 \leq k$, where x is the number of the other participants who choose the queue track. Otherwise she will win a slot if and only if y_i exceeds the k-th largest of the other x participants' time valuations. Note that the time valuations of participants who choose the queue track are distributed on $[y^*, \overline{y}]$ following the distribution function $\frac{F(y)-F(y^*)}{1-F(y^*)}$.

Let $H_{[k:x]}$ denote the cumulative distribution function of the k-th order statistics among x independent draws from $\frac{F(y)-F(y^*)}{1-F(y^*)}$:

$$H_{[k:x]}(y) = \sum_{\ell=x+1-k}^{x} \binom{x}{\ell} [\frac{F(y)-F(y^*)}{1-F(y^*)}]^{\ell} [\frac{1-F(y)}{1-F(y^*)}]^{x-\ell}.$$

Let the corresponding density function be $h_{[k:x]}(y)$.

When i does not win a slot in stage 1, she will obtain a slot in stage 2 with probability $\frac{1}{n-m}$ if $n-x-1 \geq k$ and with probability $\frac{x+2+k-n}{x+1-k}$ if $n-x-1 < k$. So i's probability of winning a slot by choosing the queue track is

$$\begin{aligned} P_Q(y_i) = \sum_{x=0}^{k-1} J(x;y^*) + \sum_{x=k}^{n-k-1} J(x;y^*) \big[H_{[k:x]}(y_i) + \big(1-H_{[k:x]}(y_i)\big)\frac{1}{n-m}\big] \\ + \sum_{x=n-k}^{n-1} J(x;y^*) \big[H_{[k:x]}(y_i) + \big(1-H_{[k:x]}(y_i)\big)\frac{x+2+k-n}{x+1-k}\big]. \qquad (8) \end{aligned}$$

Now suppose that i bids b_i units of time, then her expected payoff is $P_Q(t^{-1}(b_i))y_i w_i - b_i w_i$. To maximize $P_Q(t^{-1}(b_i))y_i w_i - b_i w_i$ it is equivalent to maximize

$$G(t^{-1}(b_i))y_i - b_i$$

where $G(y_i) = \sum_{x=k}^{n-k-1} \frac{n-m-1}{n-m} J(x;y^*) H_{[k:x]}(y_i) + \sum_{x=n-k}^{n-1} \frac{n-1-2k}{x+1-k} J(x;y^*) H_{[k:x]}(y_i)$.

The first-order condition to $\max_{b_i} \big[G(t^{-1}(b_i))y_i - b_i\big]$ is

$$G'(t^{-1}(b_i))\frac{dt^{-1}(b_i)}{b_i} y_i - 1 = 0,$$

and we know that it holds when $b_i = t(y_i)$. So

$$t'(y_i) = G'(y_i)y_i.$$

Given the boundary condition $t(y^*) = 0$, we obtain

$$\begin{aligned} t(y_i) &= \int_{y^*}^{y_i} G'(s)sds \\ &= \int_{y^*}^{y_i} [\sum_{x=k}^{n-k-1} \frac{n-m-1}{n-m} J(x;y^*)h_{[k:x]}(y_i) + \sum_{x=n-k}^{n-1} \frac{n-1-2k}{x+1-k} J(x;y^*)h_{[k:x]}(y_i)]sds. \end{aligned}$$

Now, there are two cases of equilibrium.

Case 1: If $t(\overline{y}) \leq T$, then the cap T will not be binding. So every i in the queue track will bid $t(y_i)$.

Case 2: Otherwise, as in Proposition 1, there will exist another threshold $y^{\diamond} \in (y^*, \overline{y})$ such that every i with $y_i \in [y^{\diamond}, \overline{y}]$ will bid T while every i with $y_i \in [y^*, y^{\diamond})$ will bid $t(y_i)$ in the queue track. The threshold $y^{\diamond}$ can be derived similarly as in Proposition 1 and we omit the detail.

The last case of equilibrium is as follows.

Case 3: All participants in the queue track pool their bid at T. So the queue track becomes another "lottery" track. For any participant i in the queue track, she will win a slot for sure in stage 1 if $x + 1 \leq k$. Otherwise, she will win a slot with probability $\frac{k}{x+1}$ in stage 1 and if she fails she still has a chance to win a slot in stage 2. Similar to before, i's probability of winning a slot by choosing the queue track is

$$\begin{aligned}\tilde{P}_Q(y_i) = \sum_{x=0}^{k-1} J(x; y^*) + \sum_{x=k}^{n-k-1} J(x; y^*)\Big[\frac{k}{x+1} + \big(1 - \frac{k}{x+1}\big)\frac{1}{n-m}\Big] \\ + \sum_{x=n-k}^{n-1} J(x; y^*)\Big[\frac{k}{x+1} + \big(1 - \frac{k}{x+1}\big)\frac{x+2+k-n}{x+1-k}\Big]. \qquad (9)\end{aligned}$$

Note that $\tilde{P}_Q(y_i)$ is independent of y_i.

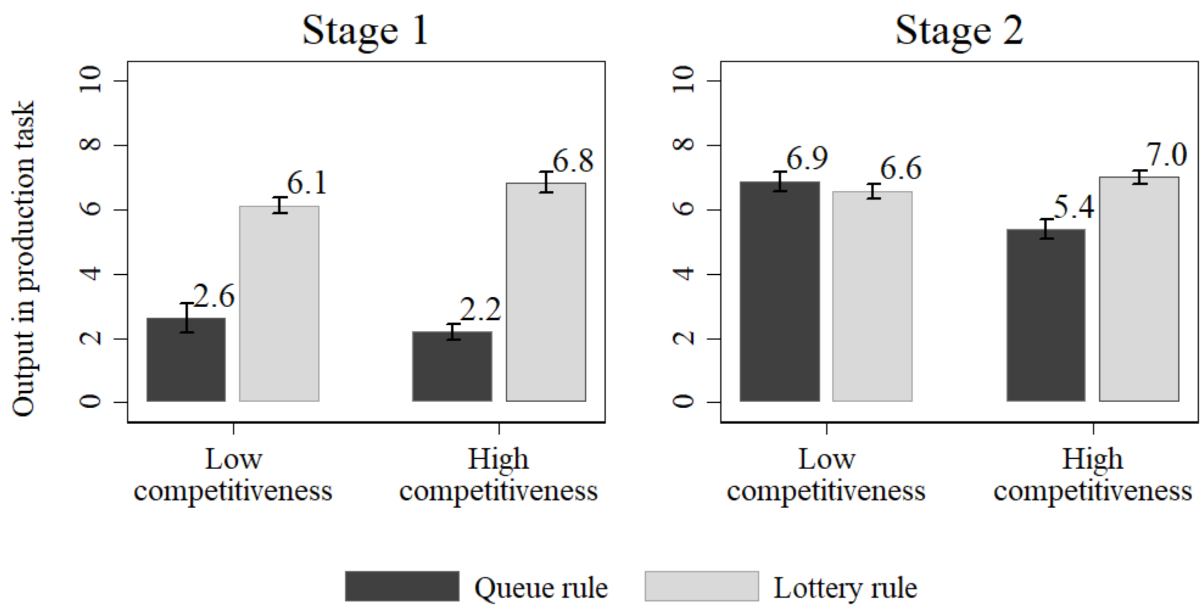
A participant of the threshold time valuation y^* is indifferent between choosing the lottery track and choosing the queue track with a bid T. So y^* is determined by

$$\big[\tilde{P}_Q(y^*) - P_L(y^*)\big] \cdot y^* = T. \qquad (10)$$

Moreover, it must be that $F(y^*) > 1/2$, because otherwise it is impossible for a participant of the threshold time valuation y^* to be indifferent between the lottery track and the queue track, which is another "lottery" track but requires participants to spend T units of time in the queue. $\square$

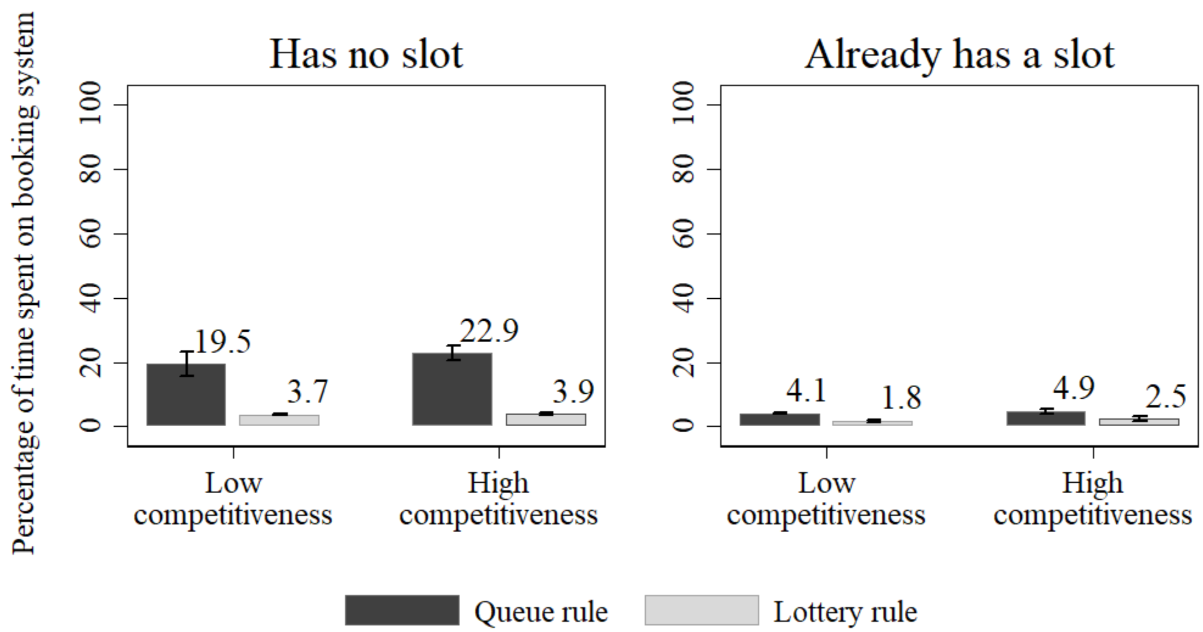
B Additional Figures and Tables

Figure B1: Average output in the production task in the solo-track treatments



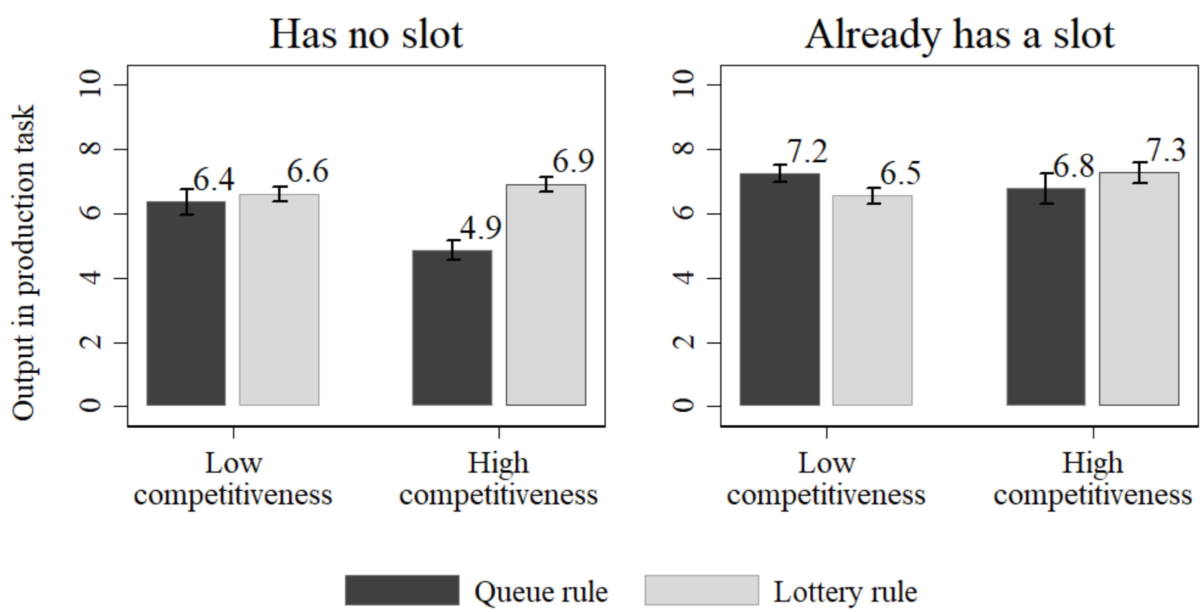
Notes: Error bars represent one standard error of means clustered at the matching group level. Low competitiveness stands for markets with 5 participants and 3 slots (Queue5 and Lottery5), while high competitiveness stands for markets with 7 participants and 2 slots (Queue7 and Lottery7).

Figure B2: Percentage of time spent on the booking system in stage 2 in the solo-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level. The graph is drawn both for participants who failed to obtain a slot in stage 1 (left) and for those who already have secured a slot in stage 1 including ones whose slots are later canceled (right). Low competitiveness stands for markets with 5 participants and 3 slots (Queue5 and Lottery5), while high competitiveness stands for markets with 7 participants and 2 slots (Queue7 and Lottery7).

Figure B3: Average output in the production task in stage 2 in the solo-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level. The graph is drawn both for participants who failed to obtain a slot in stage 1 (left) and for those who already have secured a slot in stage 1 including ones whose slots are later canceled (right). Low competitiveness stands for markets with 5 participants and 3 slots (Queue5 and Lottery5), while high competitiveness stands for markets with 7 participants and 2 slots (Queue7 and Lottery7).

Figure B4: Distribution of the computed time valuation across all treatments

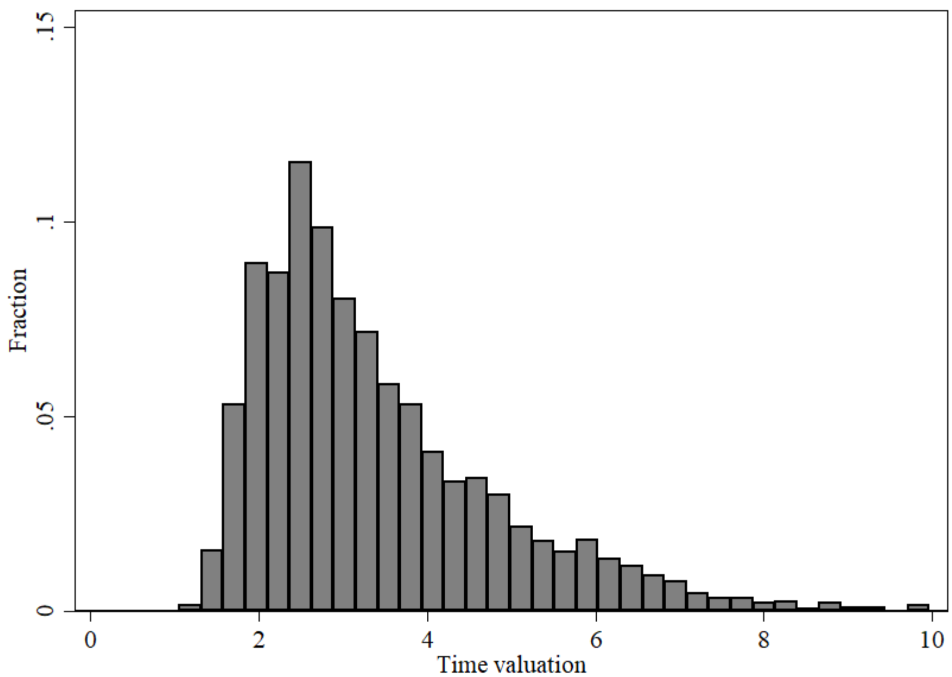
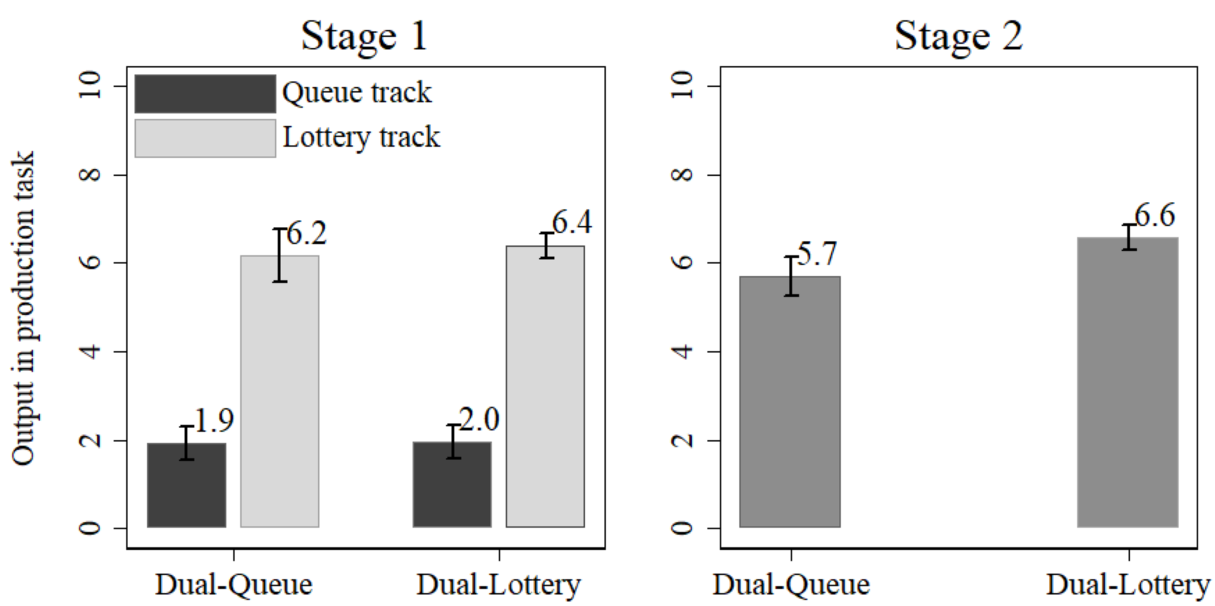
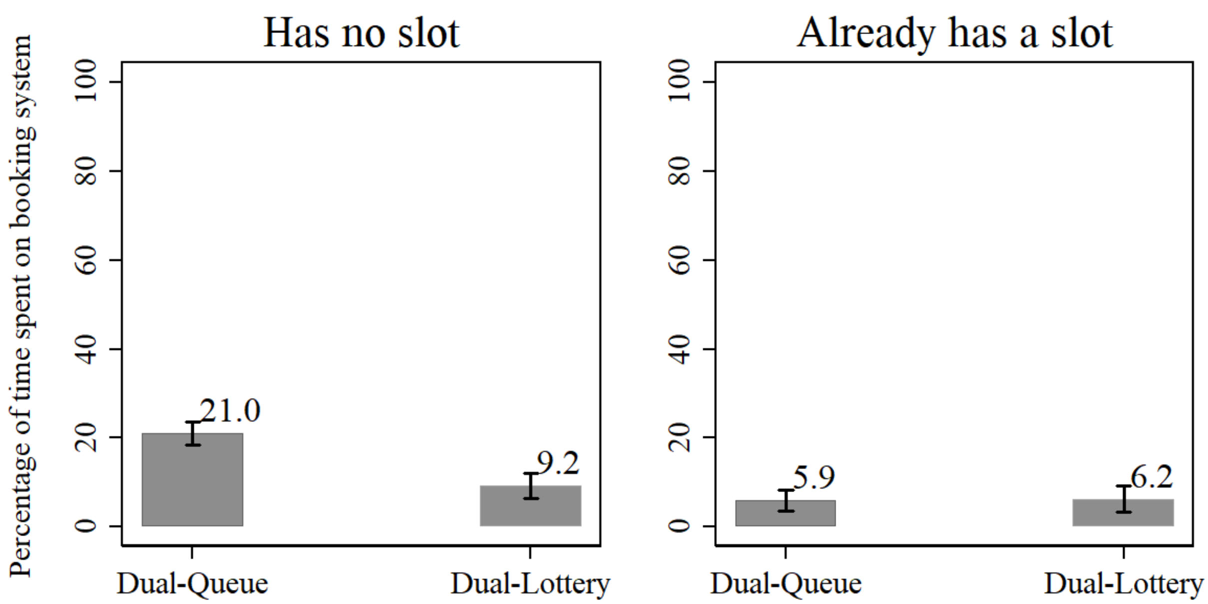


Figure B5: Average output in the production task in the dual-track treatments



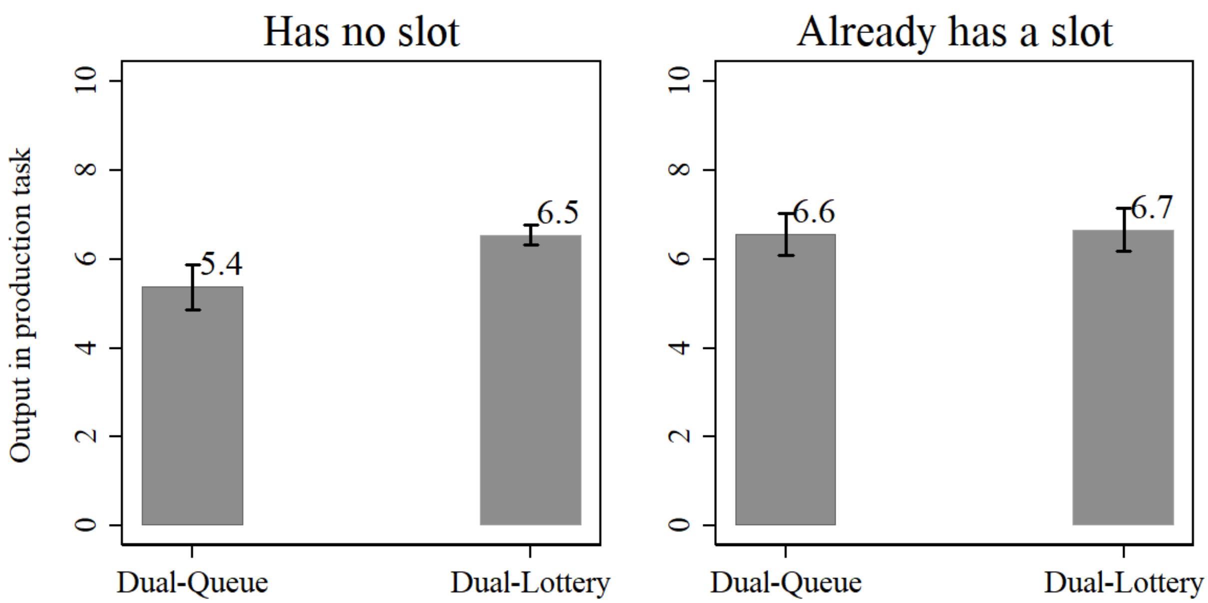
Notes: Error bars represent one standard error of means clustered at the matching group level.

Figure B6: Percentage of time spent on the booking system in stage 2 in the dual-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level. The graph is drawn both for participants who failed to obtain a slot in stage 1 (left) and for those who already have secured a slot in stage 1 including ones whose slots are later canceled (right).

Figure B7: Average output in the production task in stage 2 in the dual-track treatments



Notes: Error bars represent one standard error of means clustered at the matching group level. The graph is drawn both for participants who failed to obtain a slot in stage 1 (left) and for those who already have secured a slot in stage 1 including ones whose slots are later canceled (right).

Table B1: Random effects probit regressions on the likelihood of obtaining a slot

	Average marginal effects			
	Queue5		Queue7	
	Both stages	Stage 1	Both stages	Stage 1
Time valuation	-0.018	0.020	0.001	0.002
	(0.017)	(0.023)	(0.005)	(0.004)
Female	-0.117	-0.088	-0.201***	-0.159***
	(0.095)	(0.166)	(0.068)	(0.032)
Anxiety in Stage 1	-0.210	-0.007	-0.009	0.012
	(0.016)	(0.037)	(0.023)	(0.019)
Anxiety in Stage 2	0.027***	0.050	0.050**	0.036
	(0.009)	(0.035)	(0.021)	(0.026)
Risk-taking	0.011	0.027	0.042	0.027
	(0.017)	(0.058)	(0.035)	(0.027)
Competitiveness	-0.015	0.008	0.075***	0.026***
	(0.017)	(0.053)	(0.012)	(0.006)
Clusters	6	6	4	4
N	480	480	448	448

Notes: Standard errors clustered at the matching group level are in parentheses. We rescale the time valuation by dividing it by 100. Anxiety in Stage 1 and Stage 2 is the self-reported levels of anxiety on the scale from 1 (not anxious at all) to 7 (extremely anxious). Risk-taking is the self-reported general attitude toward taking risk on the scale from 1 (not risk-taking at all) to 7 (extremely risk-taking). Competitiveness is the self-reported general attitude toward competition on the scale from 1 (not competitive at all) to 7 (extremely competitive). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Random effects probit regressions on the % time spent on the booking system (dropping-out or full investment) in stage 1

	Average marginal effects	
	Queue5	Queue7
Time valuation	0.074*	0.076***
	(0.043)	(0.033)
Female	-0.146	-0.136
	(0.119)	(0.106)
Anxiety in Stage 1	0.006	0.049
	(0.014)	(0.036)
Anxiety in Stage 2	0.023	0.018
	(0.023)	(0.031)
Risk-taking	0.003	0.057**
	(0.023)	(0.028)
Competitiveness	-0.016	0.036
	(0.022)	(0.014)
Clusters	6	4
N	317	346

Notes: Standard errors clustered at the matching group level are in parentheses. The binary dependent variable is 1 if the total time spent on the booking system in stage 1 is no less than 235 seconds, and 0 if it is no more than 5 seconds. We rescale the time valuation by dividing it by 100. Anxiety in Stage 1 and Stage 2 is the self-reported levels of anxiety on the scale from 1 (not anxious at all) to 7 (extremely anxious). Risk-taking is the self-reported general attitude toward taking risk on the scale from 1 (not risk-taking at all) to 7 (extremely risk-taking). Competitiveness is the self-reported general attitude toward competition on the scale from 1 (not competitive at all) to 7 (extremely competitive). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Random effects probit regressions on the likelihood of choosing the queue track

	Average marginal effects			
	Dual-Queue		Dual-Lottery	
	(1)	(2)	(3)	(4)
Time valuation	0.036		0.033^{**}	
	(0.024)		(0.015)	
Monetary valuation		0.139^{***}		0.129^{***}
		(0.053)		(0.031)
Productivity		0.071		0.035
		(0.057)		(0.087)
Female	0.144^{**}	0.191^{***}	-0.002	-0.002
	(0.057)	(0.038)	(0.065)	(0.070)
Anxiety in Stage 1	-0.082^{***}	-0.071^{*}	-0.025	-0.021
	(0.028)	(0.038)	(0.030)	(0.031)
Anxiety in Stage 2	-0.001	-0.005	0.064^{***}	0.053^{***}
	(0.022)	(0.026)	(0.016)	(0.017)
Risk-taking	0.028	0.034	0.027	0.028
	(0.025)	(0.023)	(0.026)	(0.029)
Competitiveness	-0.023	-0.012	0.014	0.006
	(0.016)	(0.027)	(0.026)	(0.027)
Patience	-0.029	-0.043	-0.012	-0.020
	(0.021)	(0.018)	(0.028)	(0.031)
Clusters	4	4	4	4
N	448	448	448	448

Notes: Standard errors clustered at the matching group level are in parentheses. We rescale the time valuation and monetary valuation by dividing them by 100. Anxiety in Stage 1 and Stage 2 is the self-reported levels of anxiety on the scale from 1 (not anxious at all) to 7 (extremely anxious). Risk-taking is the self-reported general attitude toward taking risk on the scale from 1 (not risk-taking at all) to 7 (extremely risk-taking). Competitiveness is the self-reported general attitude toward competition on the scale from 1 (not competitive at all) to 7 (extremely competitive). Patience is the self-reported general attitude toward patience on the scale from 1 (not patient at all) to 7 (extremely patient). The question about patience was only asked in the two dual-track treatments. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C Experimental Instructions

In the following, we translate the original instructions in Chinese into English for the Queue7, Lottery7, Dual-Queue and Dual-Lottery treatments. The instructions for Queue5 and Lottery5 are omitted because they are exactly the same as Queue7 and Lottery7 except for the different number of group members and appointment slots

C.1 Instructions for Queue7

General Information

You are taking part in a decision-making experiment. Please read the instructions carefully. The instructions are the same for every participant. Please do not communicate with each other during the experiment. Turn off your mobile phone and put it into the envelop on your desk. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

You have earned 15 RMB for showing up on time. In addition, you can earn more money in this experiment. The amount of money you earn will depend upon the decisions you and other participants make. Your earnings in this experiment are expressed in EXPERIMENTAL CURRENCY UNITS, which we will refer to as ECUs. At the end of the experiment you will be paid using a conversion rate of 1 RMB for every 10 ECUs of earnings from the experiment.

Your final payment will be paid to you via bank transfer within 2-3 days on completion of today's experiment. All decisions are anonymous. That is, other participants will not know about your identity or your final payment.

Overview of the experiment

The experiment consists of 8 rounds. Each round lasts 8 minutes. A clock on the upper-right corner of the screen shows the time already past in each round. Each round has two tasks and each participant's earnings are sum of the earnings from the two tasks. At the beginning of each round, you can choose which task to start with: Task 1 or Task 2. The two tasks are displayed on different screens. You may freely switch between the two task screens at any time you want. But you cannot see both task screens at the same time.

Task 1 is about booking an appointment slot at a public office such as hospital. In each round, you will be randomly matched into groups of seven participants each. This means that your group members will most likely be different in each round. Each group has two appointment slots available and each participant can only book up to one slot. In each round, the private valuation for a slot is determined randomly and independently for each participant, and will be a natural number between (and including) 400 to 600 ECUs. Each participant is only informed about her own valuation, but not about other group members' valuations. A participant who books a slot will receive ECUs equivalent to her valuation and the one who does not will receive 0. We will discuss the booking procedure in Task 1 in more detail.

Task 2: In addition to the appointment booking task, in each round you can also work independently on a counting-dots task: counting the number of white dots in a series of dark-shaded squares. The figure below shows the task screen. You will enter the number of dots into the box next to the table. After you have entered the number, you can click the NEXT button. No matter whether the answer is correct or not, a new square will be generated. You will earn 35 ECUs for each square you solved correctly. If you enter a wrong number for a square, you will earn 0 for that square. Thus, in each round your earnings from the counting-dots task = the number of correct answers × 35 ECUs.

We will now describe in more detail the appointment booking system in Task 1, which consists of two steps.

Step 1: Start from 00:00 (minutes: seconds), end at 04:00. The 2 slots will become immediately available at 04:00 and will be assigned on a first-come-first-served basis in the following way: Each participant can choose to switch to and stay on the booking screen to reserve a position in a queue. Those who switch to the booking screen earlier reserve a front position. However, if a participant switches to the counting-dots task screen and then back to the booking screen, he will have to go to the back of the queue. When 4:00 is reached, the slots will be assigned as follows:

1. If the number of participants staying in the queue > 2, each of the first two in the queue will obtain a slot.
2. If the number of participants staying in the queue ≤ 2, each of them obtains a slot. Any remaining slot(s) will become available in Step 2.

Step 2: Start from 04:00, end at 08:00. Every participant who has not obtained a slot in Step 1 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. The number of available slots will be shown on the screen and updated in real time. (Note that you will need to wait for a slot to become available to book.) Those who have obtained a slot in Step 1 can still switch to the booking screen. But they will not see the BOOKING button as they do not need to book again.

The available slot(s) in Step 2 can be any number from 1 to 2 and have two sources:

1. Cancelled slot: To mimic the real-world situation in which people may choose to cancel their appointments, one participant who has obtained a slot in Step 1 will be randomly chosen by the computer to cancel her slot at some point in Step 2 (the timing is again randomly chosen by the computer). The cancelled slot will be released and become available in the booking system. The participant whose slot is cancelled will still receive ECUs equal to her valuation of the slot. But he will not be allowed to book another slot in Step 2. Furthermore, he will be notified of the cancellation whenever he switches to the booking screen.

2. Unassigned slot(s): The unassigned slot(s) may come from Step 1 if there are fewer than 2 participants staying in the queue in Step 1 before the end of 04:00. There will be no unassigned slot if the two slots have been assigned in Step 1.

Until 08:00 is reached, the available slots will be assigned on a first-come-first-served basis by pressing the BOOKING button. In particular, the cancelled slot will become available on the book screen at the moment when the cancellation happens. The unassigned slot(s) will become immediately available at 04:00.

Payoff

At the end of each round, your round payoff is the sum of the payoffs from the two tasks. At the end of the experiment, one randomly chosen round will be paid out, in addition to the show-up fee.

This completes the instructions. To ensure every participant understand the instructions, please answer the quiz on your screen. If there is any question, please raise your hand. Once everyone correctly answers the quiz, we will start the experiment. Also, in order to help participants familiarize themselves with the counting-dots task, there will be one practice (non-paying) round consisting of a 5-minute counting-dots task only.

C.2 Instructions for Lottery7

General Information

You are taking part in a decision-making experiment. Please read the instructions carefully.

The instructions are the same for every participant. Please do not communicate with each other during the experiment. Turn off your mobile phone and put it into the envelop on your desk. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

You have earned 15 RMB for showing up on time. In addition, you can earn more money in this experiment. The amount of money you earn will depend upon the decisions you and other participants make. Your earnings in this experiment are expressed in EXPERIMENTAL CURRENCY UNITS, which we will refer to as ECUs. At the end of the experiment you will be paid using a conversion rate of 1 RMB for every 10 ECUs of earnings from the experiment.

Your final payment will be paid to you via bank transfer within 2-3 days on completion of today's experiment. All decisions are anonymous. That is, other participants will not know about your identity or your final payment.

Overview of the experiment

The experiment consists of 8 rounds. Each round lasts 8 minutes. A clock on the upper-right corner of the screen shows the time already past in each round. Each round has two tasks and each participant's earnings are sum of the earnings from the two tasks. At the beginning of each round, you can choose which task to start with: Task 1 or Task 2. The two tasks are displayed on different screens. You may freely switch between the two task screens at any time you want. But you cannot see both task screens at the same time.

Task 1 is about booking an appointment slot at a public office such as hospital. In each round, you will be randomly matched into groups of seven participants each. This means that your group members will most likely be different in each round. Each group has two appointment slots available and each participant can only book up to one slot. In each round, the private valuation for a slot is determined randomly and independently for each participant, and will be a natural number between (and including) 400 to 600 ECUs. Each participant is only informed about her own valuation, but not about other group members' valuations. A participant who books a slot will receive ECUs equivalent to her valuation and the one who does not will receive 0. We will discuss the booking procedure in Task 1 in more detail.

Task 2: In addition to the appointment booking task, in each round you can also work independently on a counting-dots task: counting the number of white dots in a series of dark-shaded squares. The figure below shows the task screen. You will enter the number of dots into the box next to the table. After you have entered the number, you can click the NEXT button. No matter whether the answer is correct or not, a new square will be generated. You will earn 35 ECUs for each square you solved correctly. If you enter a wrong number for a square, you will earn 0 for that square. Thus, in each round your earnings from the counting-dots task = the number of correct answers $\times$ 35 ECUs.

We will now describe in more detail the appointment booking system in Task 1, which consists of two steps.

Step 1: Start from 00:00 (minutes: seconds), end at 04:00. Every participant can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. When 4:00 is reached, the slots will be assigned as follows:

1. If the number of applicants > 2, all applications will be put into a virtual urn. Then, one by one, applications are randomly drawn from the urn to fill the two slots.
2. If the number of applicants ≤ 2, each applicant obtains a slot. Any remaining slot(s) will become available in Step 2.

Step 2: Start from 04:00, end at 08:00. Every participant who has not obtained a slot in Step 1 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. The number of available slots will be shown on the screen and updated in real time. (Note that you do not need to wait for a slot to become available to apply.) Those who have obtained a slot in Step 1 can still switch to the booking screen. But they will not see the BOOKING button as they do not need to apply again.

The available slot(s) in Step 2 can be any number from 1 to 2 and have two sources:

1. Cancelled slot: To mimic the real-world situation in which people may choose to cancel their appointments, one participant who has obtained a slot in Step 1 will be randomly chosen by the computer to cancel her slot at some point in Step 2 (the timing is again randomly chosen by the computer). The cancelled slot will be released and become available in the booking system. The participant whose slot is cancelled will still receive ECUs equal to her valuation of the slot. But he will not be allowed to book another slot in Step 2. Furthermore, he will be notified of the cancellation whenever he switches to the booking screen.
2. Unassigned slot(s): The unassigned slot(s) may come from Step 1 if there are fewer than 2 applicants in Step 1 before the end of 04:00. There will be no unassigned slot

if the two slots have been assigned in Step 1.

When 08:00 is reached, the slots will be assigned as follows:

1. If in Step 2 the number of applicants > the number of available slots, all applications will be put into a virtual urn. Then, one by one, applications are randomly drawn from the urn to fill the available slots.

2. If in Step 2 the number of applicants ≤ the number of available slots, each applicant obtains a slot.

Payoff

At the end of each round, your round payoff is the sum of the payoffs from the two tasks. At the end of the experiment, one randomly chosen round will be paid out, in addition to the show-up fee.

This completes the instructions. To ensure every participant understand the instructions, please answer the quiz on your screen. If there is any question, please raise your hand. Once everyone correctly answers the quiz, we will start the experiment. Also, in order to help participants familiarize themselves with the counting-dots task, there will be one practice (non-paying) round consisting of a 5-minute counting-dots task only.

C.3 Instructions for Dual-Queue

General Information

You are taking part in a decision-making experiment. Please read the instructions carefully. The instructions are the same for every participant. Please do not communicate with each other during the experiment. Turn off your mobile phone and put it into the envelop on your desk. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

You have earned 15 RMB for showing up on time. In addition, you can earn more money in this experiment. The amount of money you earn will depend upon the decisions you and other participants make. Your earnings in this experiment are expressed in EXPERIMENTAL CURRENCY UNITS, which we will refer to as ECUs. At the end of the experiment you will be paid using a conversion rate of 1 RMB for every 10 ECUs of earnings from the experiment.

Your final payment will be paid to you via bank transfer within 2-3 days on completion of today's experiment. All decisions are anonymous. That is, other participants will not know about your identity or your final payment.

Overview of the experiment

The experiment consists of 8 rounds. Each round lasts 8 minutes. A clock on the upper-right corner of the screen shows the time already past in each round. Each round has two tasks and each participant's earnings are sum of the earnings from the two tasks. At the beginning of each round, you can choose which task to start with: Task 1 or Task 2. The two tasks are displayed on different screens. You may freely switch between the two task screens at any time you want. But you cannot see both task screens at the same time.

Task 1 is about booking an appointment slot at a public office such as hospital. In each round, you will be randomly matched into groups of seven participants each. This means that your group members will most likely be different in each round. Each group has two appointment slots available and each participant can only book up to one slot. In each round, the private valuation for a slot is determined randomly and independently for each participant, and will be a natural number between (and including) 400 to 600 ECUs. Each participant is only informed about her own valuation, but not about other group members' valuations. A participant who books a slot will receive ECUs equivalent to her valuation and the one who does not will receive 0. We will discuss the booking procedure in Task 1 in more detail.

Task 2: In addition to the appointment booking task, in each round you can also work independently on a counting-dots task: counting the number of white dots in a series of dark-shaded squares. The figure below shows the task screen. You will enter the number of dots into the box next to the table. After you have entered the number, you can click the NEXT button. No matter whether the answer is correct or not, a new square will be generated. You will earn 35 ECUs for each square you solved correctly. If you enter a wrong number for a square, you will earn 0 for that square. Thus, in each round your earnings from the counting-dots task = the number of correct answers × 35 ECUs.

We will now describe in more detail the appointment booking system in Task 1, which consists of two steps.

Step 1: At the beginning of each round, each participant must choose one of the following two tracks for booking slots.

- Track 1: This track has one slot which will be assigned on a first-come-first-served basis.

- Track 2: This track has one slot which will be assigned based on applications.

Note: You can only choose one of the tracks. Your choice cannot be changed during a round. You will be informed about how many group members choose Track 1 and how many choose Track 2.

Rules in Track 1: Start from 00:00 (minutes: seconds), end at 04:00. One slot will become immediately available at 04:00 and will be assigned on a first-come-first-served basis in the following way: Each participant who chooses Track 1 can choose to switch to and stay on the booking screen to reserve a position in a queue. Those who switch to the booking screen earlier reserve a front position. However, if a participant switches to the counting-dots task screen and then back to the booking screen, he will have to go to the back of the queue. When 4:00 is reached, the slot will be assigned as follows:

1. If the number of participants staying in the queue > 1, the first one in the queue will obtain the slot.

2. If the number of participants staying in the queue ≤ 1, the one in the queue obtains the slot. If no one waited in the queue, the slot will become available in Step 2.

Rules in Track 2: Start from 00:00 (minutes: seconds), end at 04:00. Every participant who chooses Track 2 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. When 4:00 is reached, the slot will be assigned as follows:

1. If the number of applicants > 1, all applications will be put into a virtual urn. Then, one application is randomly drawn from the urn to fill the slot.

2. If the number of applicants ≤ 1, this applicant obtains the slot. If no one applied for the slot, the slot will become available in Step 2.

Step 2: Start from 04:00, end at 08:00. Every participant who has not obtained a slot in Step 1 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. The number of available slots will be shown on the screen and updated in real time. (Note that you will need to wait for a slot to become available to book.) Those who have obtained a slot in Step 1 can still switch to the booking screen. But they will not see the BOOKING button as they do not need to book again.

The available slot(s) in Step 2 can be any number from 1 to 2 and have two sources:

1. Cancelled slot: To mimic the real-world situation in which people may choose to cancel their appointments, one participant who has obtained a slot in Step 1 will be randomly chosen by the computer to cancel her slot at some point in Step 2 (the timing is again randomly chosen by the computer). The cancelled slot will be released and become available in the booking system. The participant whose slot is cancelled will still receive ECUs equal to her valuation of the slot. But he will not be allowed to book another slot in Step 2. Furthermore, he will be notified of the cancellation whenever he switches to the booking screen.
2. Unassigned slot(s): The unassigned slot(s) may come from Step 1. There will be no unassigned slot if the two slots have been assigned in Step 1.

Until 08:00 is reached, the available slots will be assigned on a first-come-first-served basis by pressing the BOOKING button. In particular, the cancelled slot will become available on the book screen at the moment when the cancellation happens. The unassigned slot(s) will become immediately available at 04:00.

Payoff

At the end of each round, your round payoff is the sum of the payoffs from the two tasks. At the end of the experiment, one randomly chosen round will be paid out, in addition to the show-up fee.

This completes the instructions. To ensure every participant understand the instructions, please answer the quiz on your screen. If there is any question, please raise your hand. Once everyone correctly answers the quiz, we will start the experiment. Also, in order to help participants familiarize themselves with the counting-dots task, there will be one practice (non-paying) round consisting of a 5-minute counting-dots task only.

C.4 Instructions for Dual-Lottery

General Information

You are taking part in a decision-making experiment. Please read the instructions carefully. The instructions are the same for every participant. Please do not communicate with each other during the experiment. Turn off your mobile phone and put it into the envelop on your desk. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

You have earned 15 RMB for showing up on time. In addition, you can earn more money in this experiment. The amount of money you earn will depend upon the decisions you and other participants make. Your earnings in this experiment are expressed in EXPERIMENTAL CURRENCY UNITS, which we will refer to as ECUs. At the end of the experiment you will be paid using a conversion rate of 1 RMB for every 10 ECUs of earnings from the experiment.

Your final payment will be paid to you via bank transfer within 2-3 days on completion of today's experiment. All decisions are anonymous. That is, other participants will not know about your identity or your final payment.

Overview of the experiment

The experiment consists of 8 rounds. Each round lasts 8 minutes. A clock on the upper-right corner of the screen shows the time already past in each round. Each round has two tasks and each participant's earnings are sum of the earnings from the two tasks. At the beginning of each round, you can choose which task to start with: Task 1 or Task 2. The two tasks are displayed on different screens. You may freely switch between the two task screens at any time you want. But you cannot see both task screens at the same time.

Task 1 is about booking an appointment slot at a public office such as hospital. In each round, you will be randomly matched into groups of seven participants each. This means that your group members will most likely be different in each round. Each group has two appointment slots available and each participant can only book up to one slot. In each round, the private valuation for a slot is determined randomly and independently for each participant, and will be a natural number between (and including) 400 to 600 ECUs. Each participant is only informed about her own valuation, but not about other group members' valuations. A participant who books a slot will receive ECUs equivalent to her valuation and the one who does not will receive 0. We will discuss the booking procedure in Task 1 in more detail.

Task 2: In addition to the appointment booking task, in each round you can also work independently on a counting-dots task: counting the number of white dots in a series of dark-shaded squares. The figure below shows the task screen. You will enter the number of dots into the box next to the table. After you have entered the number, you can click the NEXT button. No matter whether the answer is correct or not, a new square will be generated. You will earn 35 ECUs for each square you solved correctly. If you enter a wrong number for a square, you will earn 0 for that square. Thus, in each round your earnings from the counting-dots task = the number of correct answers × 35 ECUs.

We will now describe in more detail the appointment booking system in Task 1, which consists of two steps.

Step 1: At the beginning of each round, each participant must choose one of the following two tracks for booking slots.

- Track 1: This track has one slot which will be assigned on a first-come-first-served basis.
- Track 2: This track has one slot which will be assigned based on applications.

Note: You can only choose one of the tracks. Your choice cannot be changed during a round. You will be informed about how many group members choose Track 1 and how many choose Track 2.

Rules in Track 1: Start from 00:00 (minutes: seconds), end at 04:00. One slot will become immediately available at 04:00 and will be assigned on a first-come-first-served basis in the following way: Each participant who chooses Track 1 can choose to switch to and stay on the booking screen to reserve a position in a queue. Those who switch to the booking screen earlier reserve a front position. However, if a participant switches to the counting-dots task screen and then back to the booking screen, he will have to go to the back of the queue. When 4:00 is reached, the slot will be assigned as follows:

1. If the number of participants staying in the queue > 1, the first one in the queue will obtain the slot.
2. If the number of participants staying in the queue ≤ 1, the one in the queue obtains the slot. If no one waited in the queue, the slot will become available in Step 2.

Rules in Track 2: Start from 00:00 (minutes: seconds), end at 04:00. Every participant who chooses Track 2 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. When 4:00 is reached, the slot will be assigned as follows:

1. If the number of applicants > 1, all applications will be put into a virtual urn. Then, one application is randomly drawn from the urn to fill the slot.
2. If the number of applicants ≤ 1, this applicant obtains the slot. If no one applied for the slot, the slot will become available in Step 2.

Step 2: Start from 04:00, end at 08:00. Every participant who has not obtained a slot in

Step 1 can switch to the booking screen at any time and apply for one slot by pressing the BOOKING button. The number of available slots will be shown on the screen and updated in real time. (Note that you do not need to wait for a slot to become available to apply.) Those who have obtained a slot in Step 1 can still switch to the booking screen. But they will not see the BOOKING button as they do not need to apply again.

The available slot(s) in Step 2 can be any number from 1 to 2 and have two sources:

1. Cancelled slot: To mimic the real-world situation in which people may choose to cancel their appointments, one participant who has obtained a slot in Step 1 will be randomly chosen by the computer to cancel her slot at some point in Step 2 (the timing is again randomly chosen by the computer). The cancelled slot will be released and become available in the booking system. The participant whose slot is cancelled will still receive ECUs equal to her valuation of the slot. But he will not be allowed to book another slot in Step 2. Furthermore, he will be notified of the cancellation whenever he switches to the booking screen.

2. Unassigned slot(s): The unassigned slot(s) may come from Step 1. There will be no unassigned slot if the two slots have been assigned in Step 1.

When 08:00 is reached, the slots will be assigned as follows:

1. If in Step 2 the number of applicants $>$ the number of available slots, all applications will be put into a virtual urn. Then, one by one, applications are randomly drawn from the urn to fill the available slots.

2. If in Step 2 the number of applicants $\leq$ the number of available slots, each applicant obtains a slot.

Payoff

At the end of each round, your round payoff is the sum of the payoffs from the two tasks. At the end of the experiment, one randomly chosen round will be paid out, in addition to the show-up fee.

This completes the instructions. To ensure every participant understand the instructions, please answer the quiz on your screen. If there is any question, please raise your hand. Once everyone correctly answers the quiz, we will start the experiment. Also, in order to help participants familiarize themselves with the counting-dots task, there will be one practice (non-paying) round consisting of a 5-minute counting-dots task only.