Cooperation in Queueing Systems

This Version: May 30th, 2023
First Version: February 24th, 2020

Yaroslav Rosokha[†] Chen Wei[‡]

Abstract

We study a social dilemma in a single-queue system in which human servers have discretion over the effort with which to process orders that arrive stochastically. We show theoretically that the efficient outcome in the form of high effort can be sustained in the subgame perfect equilibrium if the interactions are long-term (even when each server has a short-term incentive to free ride and provide low effort). In addition, we show that queue visibility plays an important role in the type of strategies that can sustain a high-effort equilibrium. In particular, we show that limiting feedback about the current state of the queue may be beneficial if the expected duration of interaction is long. We conduct two controlled lab experiments to test the theoretical predictions, and find that effort increases with the expected duration of an interaction. We also find that visibility has a strong impact on the strategies that human subjects use to provide effort in a dynamic setting. We discuss implications for managers and firms that are trying to improve service systems.

Keywords: Behavioral Operations, Single-Queue Systems, Stochastic Dynamic Games, Indefinitely Repeated Prisoner's Dilemma, Finite Mixture Models

---

[†] Purdue University • yrosokha@purdue.edu

[‡] Washington University in St. Louis • chen.wei@wustl.edu

1 Introduction

Queueing systems composed of servers that carry out a sequence of (randomly) arriving tasks underlie most economic activity. Examples abound and include the restaurant and food service industry, in which human workers prepare and serve food that customers demand; the manufacturing industry, in which a combination of human and non-human workers transform raw materials into finished products; and the healthcare industry, in which providers deliver services to patients. Not surprisingly then, queueing theory has had a vibrant history across many domains including mathematics (Erlang, 1909; Kolmogorov, 1931; Kendall, 1951), operations research (Cobham, 1954; Little, 1961), management (Kao and Tung, 1981; Graves, 1982), and economics (Sah, 1987; Polterovich, 1993). Although most of the early research assumed servers process orders at fixed rates — a reasonable assumption when one deals with machines — more recently, the field has seen a push to understand the implications of servers having discretion over work speed (George and Harrison, 2001; Hopp, Iravani, and Yuen, 2007), being utility maximizing (Gopalakrishnan, Doroudi, Ward, and Wierman, 2016), or being susceptible to behavioral biases (Bendoly, Croson, Goncalves, and Schultz, 2010).

Multi-server queueing systems have two important but largely understudied features. The first is that servers interact repeatedly. For example, in the car repair industry, technicians work in groups to provide services for a queue of cars (Wergin, 2003). In this situation, managers need to decide how frequently to rotate the workers and change group composition. The repeated interactions provide room for reputation-building and reciprocity, which may result in more complex strategies on the part of the decision-makers (i.e., human servers). Although such strategies have been studied in the theoretical and experimental literature on repeated games (Dal Bó and Fréchette, 2018), to the best of our knowledge, these topics have not been investigated in the context of queueing systems. What makes the queueing setting distinct is the stochastic nature of customer arrivals and the dynamic implication of servers' decisions. Specifically, when servers exert high effort, more customer orders are being processed and the length of the queue is likely to decrease. On the other hand, when servers exert low effort, fewer customer orders are being processed and the length of the queue is likely to increase. In this paper, we focus on a scenario in which human servers have discretion over effort, and the compensation depends on the total number of customers processed by the group, which creates an incentive to free ride.[1] We formalize the queueing environment as a dynamic stochastic game (henceforth *stochastic game*) with multiple states and transition probabilities determined partly by the actions of the players and partly by the arrival process. We show theoretically that even when individuals face incentives to free ride, high effort can be supported in the subgame perfect Nash equilibrium (henceforth SPE) of the stochastic game if the

---

[1] Group-based payment schemes are frequently observed in the real world. For example, Ortega (2009) finds that, according to the European Working Conditions Survey, group-based performance pay is the third most frequent type of performance pay among employees. In the queuing context, examples include Hamilton, Nickerson, and Owan (2003), who document team-based incentives in the garment-industry setting with a group of workers facing a queue of cloth pieces that need to be sewn together into garments (the team then receives a piece rate for the entire garment), and Tan and Netessine (2019), who document that restaurant workers face, at least in part, team-based incentives.

expected length of the interaction is long enough.

The second feature that is critical for managing multi-server queuing systems is how much information about the state of the queue to provide to the servers. For example, in the car repair industry, customers typically make appointments ahead of time. In this situation, managers need to decide whether to tell technicians how many new customers will visit the next day.[2] Although the topic of queue visibility has received some attention in the recent literature (e.g., Buell, Kim, and Tsay, 2017; Shunko, Niederhoff, and Rosokha, 2018; Hathaway, Kagan, and Dada, 2022), the role of common knowledge about the state of the queue has not been considered from the game-theoretic perspective. In particular, visibility of the queue directly impacts the types of strategies that can be used to enforce high effort in the equilibrium of the stochastic game. We show theoretically that when the queue is visible, there exist equilibria in which players play a class of state- and history-contingent trigger strategies that provide high effort when the queue is long and low effort when the queue is short – dynamics that have been documented in the empirical studies (e.g., Kc and Terwiesch, 2009). We also show that when the queue is not visible, sustaining high effort across all states is possible if the interactions are long-term.

We conduct two controlled laboratory experiments to test our theoretical predictions for a simplified two-server three-state queueing system. In both experiments, we vary two factors – the expected duration of an interaction (i.e., probability of continuing an interaction to the next period) and queue visibility (i.e., whether servers know the state of the queue). The two experiments differ with respect to the micro-foundations and the resulting payoff parameters as well as the manner in which the decision context is presented. In both experiments, we find clear evidence that the average effort increases with the expected duration of repeated interactions. We also find that when the queue is visible, the average effort is higher when the queue is long than when it is short. Arguably, the most interesting result is that we find support for the theoretical predictions regarding the effect of queue visibility; namely, when the expected duration of future interactions is short, on average, servers provide higher effort when the queue is visible than when it is not. When the expected duration of future interactions is long, the opposite is true. In addition to the analysis of (observable) effort, we carry out an econometric estimation of (unobservable) strategies that subjects use. When the queue is not visible, we find that subjects either provide low effort in all states of the queue regardless of the actions of the other server (i.e., play Always Defect) or provide high or low effort conditional on the behavior of the other server (i.e., play Tit-for-Tat or Grim Trigger). When the queue is visible, we find a significant proportion of subjects use sophisticated state-contingent versions of these strategies, in that subjects respond to the behavior observed the last time the queue was in a particular state.

---

[2] Other examples include the restaurant industry where customers make reservations either via calls or online platforms and form a virtual queue (Kimes, 2009). In this scenario, managers may have better information about the number of reservations and need to decide whether to share this information with their staff. Similar features are present in the airline services industry. In particular, (i) aircraft cleaning crew work as a team to clean the aircraft interior and (ii) airlines have better forecasts than the cleaning crew about the number of flights a cleaning crew may need to process during a shift (Belcastro, Marozzo, Talia, and Trunfio, 2016; Sternberg, Soares, Carvalho, and Ogasawara, 2017). In this setting, airlines have the option to share these forecasts with their cleaning crews.

The rest of the paper is organized as follows: In section 2, we review related literature in operations management and economics. In section 3, we develop the notation and set up theoretical characterization of an SPE. In section 4, we present the experimental design for a simplified environment with two servers and three states of the queue, and provide theoretical predictions for the chosen parameters. In section 5, we carry out the analysis of the data. In particular, we first analyze effort choices and then conduct an econometric estimation of underlying repeated-game strategies. In section 6, we present the details of the second study, with the results presented in section 7. We conclude in section 9.

2 Related Literature

Our work contributes to four broad streams of research across operations management and economics. The first stream includes papers that investigate queueing systems with human servers.[3] Our contribution to this stream can be organized along two dimensions. The first dimension includes the theoretical analysis of the effort provision when servers are utility-maximizing (e.g., Zhan and Ward, 2018, 2019; Armony, Roels, and Song, 2021). Among the most relevant theoretical papers along this dimension is Gopalakrishnan, Doroudi, Ward, and Wierman (2016), who study strategic servers in multi-server systems and the impact of scheduling policies on the equilibrium of the one-shot game among the servers. Our project contributes to this dimension by theoretically investigating the impact of long-term relationships and queue visibility on the servers' effort provision. In particular, we focus on the strategies that servers can use to enforce high effort in the SPE of the stochastic game underlying the queueing system. The second dimension includes experimental papers on human-server behavior in queueing systems. The most relevant papers along this dimension include Schultz, Juran, Boudreau, McClain, and Thomas (1998), Schultz, Juran, and Boudreau (1999), Powell and Schultz (2004), and Staats and Gino (2012), who consider behavioral factors that influence effort provision in a variety of queueing systems. As mentioned above, recent work along this dimension also includes Buell, Kim, and Tsay (2017), who find that operational transparency increases customers' perceptions of service quality and reduces throughput times; Shunko, Niederhoff, and Rosokha (2018), who find that the visibility of the queue may speed up servers' service rate; Hathaway, Kagan, and Dada (2022), who find that servers incorporate the state of the queue into their decisions; and Kremer and de Véricourt (2022), who find that decision-makers are insufficiently sensitive to congestion when facing an accuracy/speed trade-off. Our work is distinct in that we provide a game-theoretic foundation for the servers' behavior and highlight that visibility of the queue may have different consequences on servers' effort provision depending on the expected duration of an interaction among servers. In addition, although a large body of literature has considered empirical regularities associated with human-server behavior (e.g., a review

---

[3] For a thorough discussion of issues studied within the stream of literature that considers servers as decision-makers, we refer the reader to section 9.3 of the recent review by Allon and Kremer (2018). The review also encompasses related streams that consider the effect of the customer (section 9.2) and the manager (section 9.4) having discretion over the respective decisions.

by Delasay, Ingolfsson, Kolfal, and Schultz, 2019), our paper is the first to conduct econometric investigation of the repeated-game strategies that human servers may use in queueing systems.

The second stream of research that we contribute to includes papers in operations and supply-chain management that investigate the impact of long-term relational incentives.[4] Papers in this stream of literature include Nosenzo, Offerman, Sefton, and van der Veen (2016), who investigate the threat of punishment and power of rewards in the repeated inspection game; Davis and Hyndman (2018), who investigate the efficacy of relational incentives for managing the quality of a product in a two-tier supply chain; Beer, Ahn, and Leider (2018), who show that the benefits of buyer-specific investments for both suppliers and buyers are strengthened when firms interact repeatedly; and Hyndman and Honhon (2020), who investigate indefinitely binding and temporarily binding contracts in the repeated two-person newsvendor game. Taken together, the findings from these papers suggest that long-term relationships can be effective in enforcing more efficient outcomes. Our project contributes to this stream of research by highlighting the role of repeated interactions on the behavior of servers in the queueing setting.

The third stream of literature that we contribute to is the experimental and theoretical work in economics that investigates behavior in the indefinitely repeated Prisoners' Dilemma (henceforth PD) game (see Dal Bó and Fréchette, 2018, for a review). Papers in this stream of literature have shown that cooperation is sensitive to the probability of continuation and payoffs, and that cooperation may not always be sustained even if theoretically possible (e.g., Dal Bó and Fréchette, 2011; Blonski, Ockenfels, and Spagnolo, 2011). Regarding the strategies that human subjects use in PD experiments, recent papers, including Dal Bó and Fréchette (2011, 2019) and Romero and Rosokha (2018, 2019b), show that simple strategies such as Grim Trigger, Always Defect, and Tit-for-Tat are prevalent. We contribute to this stream of literature by investigating the extent to which these strategies will be played in a stochastic environment with a transition between the PD and non-PD games. In particular, in our setting, cooperation (i.e., high effort) in the PD game leads to a higher likelihood that the non-PD game in which low effort is both the Nash equilibrium and the socially optimal outcome will be played next. These transitions create room for spillover effects related to Knez and Camerer (2000), Peysakhovich and Rand (2016), and Cason, Lau, and Mui (2019), and path dependence in equilibrium selection studied by Romero (2015).

The fourth stream of literature that we contribute to explores dynamic and stochastic repeated games. Early papers in this literature include Green and Porter (1984) and Rotemberg and Saloner (1986), who theoretically show that collusion among firms can be supported in the presence of stochastic demand shocks that are independent of firms' decisions.[5] Recent experimental work by

---

[4] In this paper, we focus on the repeated nature of an interaction and abstract away from settings with end-of-shift or temporary workforce considerations, which may be better captured with a finite-horizon model. For an early discussion of finite versus indefinite horizon, see Cox and Oaxaca (1989). For a recent study of finitely-repeated PD games, see Embrey, Fréchette, and Yuksel (2018).

[5] Our work is also related to the study of dynamic common-pool resource games (Walker, Gardner, and Ostrom, 1990; Gardner, Ostrom, and Walker, 1990). Recent papers that experimentally study common-pool resource games by Vespa (2020) find that although efficiency can be supported with history-contingent strategies, in practice, subjects find it difficult to cooperate and rely on state-contingent Markov strategies.

Rojas (2012) confirms that collusion in such environments can arise in the lab.[6] In an experimental study of the dynamic oligopoly game, Salz and Vespa (2020) point out that restricting attention to Markov strategies, when decision-makers can use a richer class of state- and history-contingent strategies to support cooperation in the SPE of the repeated game, may lead to systematic biases in estimation of strategies. Our work is also closely related to the dynamic (Vespa and Wilson, 2015, 2019) and stochastic (Kloosterman, 2019) variations of the repeated PD game. Vespa and Wilson (2015) find that subjects are conditionally cooperative and adjust their behavior not only in response to the state, but also to the history. Vespa and Wilson (2019) test the extent to which subjects internalize the incentives of changing the transition rule from endogenous to stochastic. Kloosterman (2019) focuses on the beliefs about the future in a two-state stochastic PD and finds that subjects cooperate when beliefs about the future support a large scope for punishment. Our work is distinct in that the queueing problem that we study combines both the dynamic and the stochastic components. In particular, the dynamic implications of decisions are different from the environments studied in previous work. In addition, we focus on the common knowledge about the underlying state. We find evidence that when the queue is visible, a significant portion of subjects rely on history-contingent repeated-game strategies to sustain high-effort cooperation.

3 Theoretical Background

Consider a single-queue system with $N = \{1, 2\}$ identical servers and a finite buffer of size B. In particular, suppose that in each time period $t \in \{1, ..., \infty\}$, λ_t customer orders arrive to the queue, and servers discount the future according to the common discount factor δ. Further suppose λ_t is a random variable that is distributed according to G, where G is a distribution with integer support on $[\lambda_{min}, \lambda_{max}]$. Then, let $\Theta = \left\{ \theta \in \mathbb{N} \;\middle|\; \lambda_{min} \leq \theta \leq B \right\}$ denote the set of states of the queueing system. That is, $\theta_t \in \Theta$ denotes the number of customers in line (including customers who are in service) in period t.

In this paper, we are interested in scenarios in which servers face a social dilemma in at least one state of the queue. To set up such a dilemma, we consider an environment in which (i) servers have discretion over effort and (ii) free-riding incentives exist for each of the servers. Regarding the discretion over effort, we assume that each server $i \in N$ can choose whether to provide high effort or low effort, $e_i \in E_i := \{h, l\}$, with high effort resulting in more capacity within a period. We further assume that the cost of processing orders to server i, $c(e_i, e_{-i}, \theta)$, increases with own effort, e_i, and is convex in the number of orders processed within a period by the server, $m(e_i, e_{-i}, \theta)$.[7] Regarding the free-riding incentives, we assume that individual effort choices are not observed

---

[6] In the intermediate treatment of Rojas (2012), two firms play a repeated Cournot game in which the demand of each period is stochastic. Firms only know the distribution for next period's demand state, and they need to select the amount of quantities to produce in each period. In that environment, firms' choices in a certain period do not affect the demand realizations in the next period.

[7] The convex-cost assumption is a standard component across the theoretical, empirical, and experimental streams of literature (e.g., Mas and Moretti, 2009; Ortega, 2009; Clark, Masclet, and Villeval, 2010; Gill and Prowse, 2012). Note that, the cost function depends on the effort by the other server, because when work is available, a server would rather split the workload than do the majority of it alone.

by the managers and that individual payoff, $r(e_i, e_{-i}, \theta)$, is a function of the total number of customers processed by the group, $M(e_i, e_{-i}, \theta)$. That is, the net payoff to server i within a period is $u(e_i, e_{-i}, \theta) = r(e_i, e_{-i}, \theta) - c(e_i, e_{-i}, \theta)$.

3.1 Stage Game

Let $g(\theta) = \langle N, E, U(\theta) \rangle$ denote the stage game played in state θ, where the set of players is given by N, the set of strategy profiles is given by $E = \prod E_i$, and the set of payoffs is given by $U(\theta) = \{u(e, \theta) : e \in E\}$. In this paper, we consider scenarios in which providing low effort is a dominant action of $g(\theta)\ \forall\ \theta \in \Theta$, but $\theta' \in \Theta$ exists for which a high effort profile is socially optimal:

$$u_i(l, e_j, \theta) > u_i(h, e_j, \theta)\ \forall\ e_j \in E_j, \theta \in \Theta \text{ and} \tag{1}$$

$$\exists\ \theta' \in \Theta : 2r(h, h, \theta') - 2c(h, h, \theta') > 2r(l, l, \theta') - 2c(l, l, \theta'). \tag{2}$$

Inequality (1) means that regardless of what the other player does, each player receives a higher payoff for providing low effort than for providing high effort. In particular, (1) implies that effort profile $e^d = (l, l)$ is the unique Nash equilibrium of the stage game $g(\theta)\ \forall \theta \in \Theta$. Inequality (2) means that there exists a state θ' in which both players receive a lower payoff if both provide low effort than if both provide high effort. Note that (1) and (2) imply that the game played in state θ' is a two-player PD.

3.2 Stochastic Game

Let $\Gamma = \langle N, E, U, \Theta, \mathcal{P} \rangle$ denote a stochastic game implied by the queueing environment above. In particular, in addition to sets N, E, U, and Θ, let $\mathcal{P}$ denote the transition probability across the states. Specifically, let $\mathcal{P}_{\theta\theta'}^{e_i e_{-i}} := \mathcal{P}(\theta' | \theta, e)$ denote the probability that the next state is θ' given the current state θ and the effort profile $e \in E$. Notice that the transition probability is fully determined by the current state, the action profile by the servers, and the arrival process. Next, we consider some of the strategies the players may use to sustain high effort in equilibrium for the two cases – when the queue is visible and when it is not.

3.3 Strategies and Subgame Perfect Equilibrium When Queue Is Visible

When the queue is visible, players may utilize three types of repeated-game strategies. The first type is the state-contingent *Markov strategies*. These strategies condition only on the realization of θ. For example, a player may always provide high effort in one particular state θ and always provide low effort in all other states. We refer to this strategy as AC^{θ}. The second type is the *history-contingent strategies*. These strategies condition only on the realized history of actions but not on the current state or the history of states. An example of this type of strategy is the well-known *Grim Trigger* strategy (henceforth GT), which begins by providing high effort in the first period and continues to provide high effort until one of the players provides low effort. The third type is the *state- and history-contingent strategies*. These strategies condition on both the state realization and the history of actions. An example of this type of strategy is a strategy that plays GT in a particular state θ but always provides low effort in all states $\theta' \neq \theta$. We refer to such a strategy as GT^{θ}.

To check whether a strategy profile s is an SPE, we have to check whether, for each player i and each subgame, no single deviation would increase player i's payoff in the subgame. For example, to find conditions under which strategy profile $s^{GT} = (GT, GT)$ is an SPE, we have to check single deviations in two kinds of contingencies: (i) after histories in which all players provided high effort and (ii) after histories in which at least one of the players provided low effort at some point. To evaluate whether a player has a profitable deviation in state θ for the first type of contingency, we need to compare the total value from continuing to provide high effort, which we denote as V^c_θ, and the total value of deviating, which we denote as V^{dev}_θ. Formally,

$$V^c_\theta = u(e^c, \theta) + \delta \sum_{\theta'} \mathcal{P}^{hh}_{\theta\theta'} V^c_{\theta'} \tag{3}$$

$$V^{dev}_\theta = u(e^{dev}, \theta) + \delta \sum_{\theta'} \mathcal{P}^{lh}_{\theta\theta'} V^d_{\theta'} \tag{4}$$

$$V^d_\theta = u(e^d, \theta) + \delta \sum_{\theta'} \mathcal{P}^{ll}_{\theta\theta'} V^d_{\theta'}. \tag{5}$$

The second type of contingency in which one of the players has deviated is satisfied because the best course of action given that the other is going to provide low effort is to provide low effort oneself. Thus, a strategy profile s^{GT} is an SPE of Γ if

$$V^c_\theta - V^{dev}_\theta \geq 0 \ \forall \ \theta. \tag{6}$$

3.4 Strategies and Subgame Perfect Equilibrium When Queue Is Not Visible

When the queue is not visible, the state of the game is not known with certainty. Therefore, to enforce high effort in equilibrium, the servers cannot condition on the current state θ, but instead are limited to the history-contingent strategies. For example, when the queue is not visible, a strategy profile s^{GT} is an SPE if

$$\mathbf{E}_\theta \left[V^c_\theta - V^{dev}_\theta \right] \geq 0. \tag{7}$$

Note the difference between (6) and (7) is that the former has to hold for each state (including states with high incentives to deviate), whereas the latter has to hold in expectation. In section 4.1, we show that this feature means that not knowing the state of the queue may lead to higher effort provision across all states. By contrast, knowing the queue length means that servers may more easily sustain high effort in a subset of states.

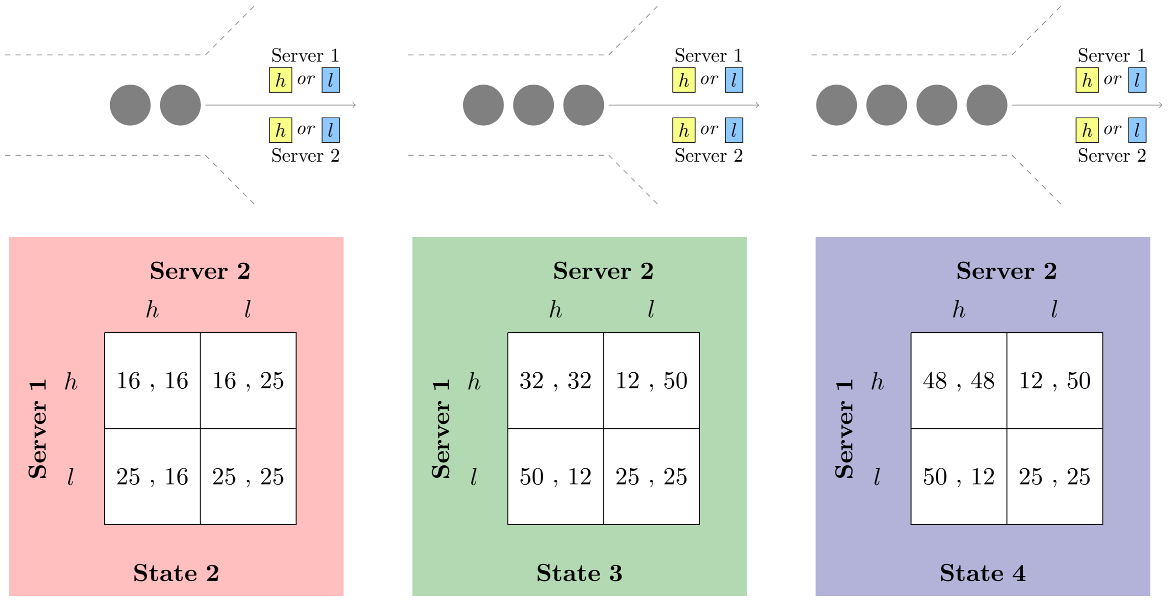
4 Study 1: Experimental Design and Theoretical Predictions

In this section, we describe the environment, provide the theoretical predictions, and formulate the hypotheses for the first set of experiments. In particular, we set buffer $B = 4$, distribution G uniform, and the minimum and maximum number of new arriving customers $\lambda_{min} = 2$, $\lambda_{max} = 4$, which means that the set of states $\Theta = \{2, 3, 4\}$. We chose these parameters so that the number of states is small (so can be reasonably implemented in the lab) yet provides room for queueing dynamics, with the queue being shorter or longer than the average arrival rate. In terms of the payoffs, we picked the parameters of the convex cost function, $c(.)$, and the parameters of the compensation function, $r(.)$, so that in addition to creating an environment with desired properties (1) and (2), the payoffs in certain states match studies in the literature that have been shown to yield a range of cooperative behavior (e.g., Dal Bó and Fréchette, 2011). In particular, we assume that with low effort, each server will always process one task, but with high effort, each can process up to two tasks, if available. We further assume that the cost of processing $m_i(.)$ tasks with low effort is $c(l, e_j, \theta) = am_i(l, e_j, \theta)^2 + bm_i(l, e_j, \theta) + c$ and the cost of processing $m_i(.)$ tasks with high effort is $c(h, e_j, \theta) = am_i(h, e_j, \theta)^2 + bm_i(h, e_j, \theta) + z$ with $a = 22$, $b = -37$, $c = 40$, and $z = 49$ (see Appendix A.1 for more details). Note that these cost functions imply that, *ceteris paribus*, providing high effort is more costly.

Regarding the compensation function, we assume that the compensation to the individual server depends on the number of total tasks, $M(.)$, processed by the group: $r(e_i, e_j, \theta) = kM(e_i, e_j, \theta) + \mathbf{1}_{M(e_i,e_j,\theta)=4} bonus$ with $k = 25$, and $bonus = 11$. The interpretation is that the server is compensated based on the total number of tasks processed by the group in two ways. The first is per-unit compensation that depends on $M(.)$, with $M(l, l, \theta) = 2 \, \forall \theta$, $M(l, h, \theta) = min(3, \theta)$, $M(h, h, \theta) = min(4, \theta)$. Therefore, the queue length and the chosen effort naturally influence the server's compensation (i.e., when the queue is shorter, fewer customers can be processed; when servers reduce effort, fewer customers can be processed). The second is a bonus that is paid when the favorable output is observed (e.g., four tasks can be processed only if both servers provide high effort). Bonus payment based on the observable outcomes is a common feature of many compensation structures (e.g., Hashimoto, 1979; Blakemore, Low, and Ormiston, 1987; Bell and Van Reenen, 2014; Hathaway, Kagan, and Dada, 2022).

In Figure 1, we present resulting payoffs for each combination of effort choices (see Appendix A.1.3 for calculation details). The consequence of these payoffs is that when two customer orders are available, the dominant action –to provide low effort– is also socially optimal from the servers' perspective. However, when three or four customer orders are available, the socially optimal outcome is for both servers to provide high effort even though, individually, each would prefer to provide low effort. In a repeated-interaction context, when three or four customers are in line, short-term incentives to free ride but long-term incentives to cooperate exist. In terms of the difference between states 3 and 4, the free-riding incentives are larger when three customer orders are in line, because each player would prefer that the other provide high effort and process two of the three customer orders.

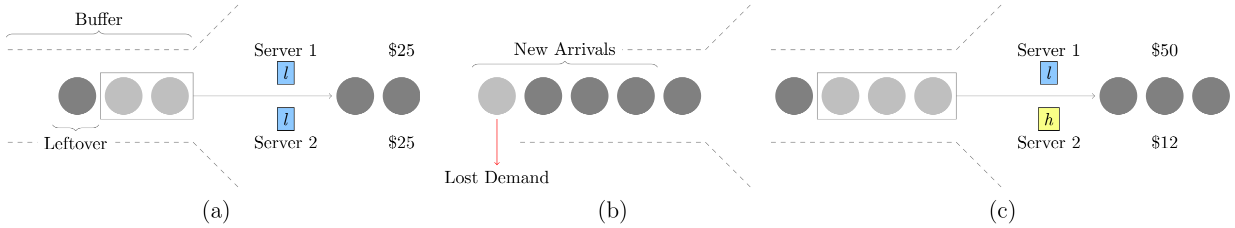
Figure 1: Stage-Game Payoffs in Each State (Study 1)



Notes: The three columns present stage games played in the three possible states. State $\theta \in \{2, 3, 4\}$ corresponds to θ customer orders in line. Each player chooses low effort (l) or high effort (h) to process customer orders. With low effort, each player can process up to one order; with high effort, each player can process up to two orders. Matrices present normal form representation of the stage game played in each state. Stage games played in states 3 and 4 are PD games studied in Dal Bó and Fréchette (2011). The stage game played in state 2 is non-PD.

Figure 2 presents example dynamics in our experimental environment. In particular, panel (a) presents an example in which three customer orders are in line, and both servers select low effort. In such a case, one order will be left over for the next period. Panel (b) presents an example in which four new customer orders arrive. In this case, the total number of customer orders will exceed the buffer size, and as a result, one order will be lost. Panel (c) presents the outcome if one server provides high effort and the other server provides low effort. Note that the state of the queue is partly endogenous (via effort choices) and partly exogenous (via stochastic arrivals). If, for example, three customers arrived every period and the buffer size was limited to three, we would have a non-stochastic version of the indefinitely repeated PD, with payoffs identical to the $R = 32$ treatment studied in Dal Bó and Fréchette (2011).

Figure 2: Example Dynamics



Notes: Panel (a) presents an example decision in period t. In particular, suppose three customer orders are in the queue and each server selects low effort; then, two orders are processed in period t and each server earns 25 points. The payoffs are determined from the stage-game payoff matrix in Figure 1 corresponding to state 3. Panel (b) presents example arrivals in period $t+1$. For this example, four orders are arriving in period $t+1$, and because the new orders together with the leftover orders from period t exceed the buffer size, one order is considered lost demand. Panel (c) presents an example decision in period $t+1$ whereby server 1 chooses l and server 2 chooses h. The payoffs are determined from the stage-game payoff matrix in Figure 1 corresponding to state 4.

An important experimental design choice that we made concerns how to vary queue visibility. In particular, to vary queue visibility, we chose to modify the timing of when the number of new order arrivals was revealed within the decision period. Specifically, for the treatments in which the queue was visible, the number of new orders was revealed *before* subjects made their decisions for that period. Therefore, subjects knew the number of leftover orders from the previous period and the number of new arrivals at the time of making their decision. Furthermore, we clearly stated which stage-game payoff matrix determined the payoffs for the period. For the treatments in which the queue was not visible, the number of new orders was revealed *after* subjects made their decisions for that period. Thus, although at the time of their decision, subjects knew the number of leftover orders from the previous period, they did not know the exact number of new arrivals nor the exact stage-game payoff matrix. We decided to show the previous period's outcomes for several reasons. First, doing so allowed for minimal change between the treatments (i.e., the only difference in the instructions was regarding the timing of arrivals). Second, by providing the outcome of the previous period, we removed any uncertainty and confusion about the payoffs obtained in the previous period. Finally, the design has clear practical implications. For instance, in the car service industry example that we mentioned in the introduction, technicians may observe how many cars are left unfinished when they get off work, but may not know how many cars will arrive for service the next day.

4.1 Theoretical Predictions

In this section, we provide conditions under which cooperation in the form of high effort may arise in the stochastic game underlying the queueing experiment. Our setup has a nice feature that both the Nash equilibrium of all stage games and the Markov perfect equilibrium of the overall stochastic game is to provide low effort in all three states. Thus, high effort can *only* be sustained using strategies that condition on the past history of play. Specifically, we follow the typical approach in the game-theoretic literature and focus on GT-type strategies described in section 3. The reliance on GT-type strategies is in part due to their analytical tractability, and in part because they us allow to find the minimal discount factor (i.e., the probability that an interaction continues to the next period) that supports cooperation. In addition, recent experimental studies (see Dal Bó and Fréchette, 2018, for a review) show GT is one of the five most popular strategies that human subjects use when playing indefinitely repeated PD. In what follows, we outline how to derive the condition on the discount factor to ensure high effort could be supported in equilibrium of the infinitely repeated stochastic game (see Appendix A.1.5 for more details).

To determine whether GT is an equilibrium strategy of the stochastic game, we first find the transition probability matrix implied by the strategy profile s^{GT}. In our case, as long as both players provide high effort, they process all of the customer orders. Therefore, the transition probability $\mathcal{P}^{hh}_{\theta\theta'}$ is determined by the arrival process (i.e., uniform distribution). However, if one of the players decides to deviate and plays low effort, then some of the states will have leftover customer orders, which, together with the arrival process, implies that the transition to states with more customers is more likely. In addition, a deviation to the low effort will cause all players to play low effort in all future periods, meaning that the transition probability $\mathcal{P}^{ll}_{\theta\theta'}$ will determine the evolution of states from next period onward. In this way, a single deviation has an immediate and a long-term impact on the rewards and transition probabilities.

To determine the critical threshold $\delta^*_v(GT)$ for which s^{GT} is an SPE when the queue is visible, we solve equation (6). We find that $\delta^*_v(GT) = 0.72$, meaning that as long as the probability of continuing an interation is at least 0.72, a strategy profile s^{GT} is an SPE of the stochastic game. An important point is that GT does not distinguish among the states. However, we expect that a human participant would do so if the queue is visible. Therefore, we consider two state-contingent variations of this strategy. The first, which we term GT^{34}, plays GT across states 3 and 4 and always provides low effort in state 2. We find that $\delta^*_v(GT^{34}) = 0.64$. The second, which we term GT^4, plays GT in state 4 only and provides low effort in both states 2 and 3. Solving for $\delta^*_v(GT^4)$, we get 0.19. These results highlight that GT^4 is the easiest to sustain, followed by GT^{34}, and GT is the most difficult to sustain when the queue is visible. It is important to note that when $\delta > 0.72$, all of these equilibria are possible. However, some equilibria will lead to higher payoffs from the server's perspective (i.e., be more efficient). In particular, the highest joint payoff for the servers will be reached in the GT^{34} equilibrium.

When the queue is not visible, we solve equation (7) to determine the critical threshold $\delta^*_{nv}(GT)$ for which s^{GT} is an SPE and find that $\delta^*_{nv}(GT) = 0.58$. This finding means that high effort in all states of the queue can be supported at a lower discount factor when the queue is not visible.[8] The reason is that when the queue is visible, players know the exact state they are in, so they know the exact benefit of providing high or low effort in the current period. However, if players do not know

---

[8] In Appendix A.1.4, we investigate how these results depend on the parameters of the cost and compensation functions.

the exact state, they can only consider the expected benefit. Once δ is large enough, servers are willing to provide high effort in all states because the expected value of doing so is larger than the expected value of unilateral deviation. Thus, theoretically, if the goal is to achieve high effort in all states of the queue, reducing visibility may be helpful. We summarize these results in Table 1. In addition to the critical thresholds, the table presents servers' effort provision in the symmetric SPE associated with each strategy.

Table 1: Summary of Symmetric SPE (Study 1)

Strategy	Critical Threshold When Queue is Visible (δ^*_v)	Critical Threshold When Queue is Not Visible (δ^*_{nv})	High Effort (%)				Comment
			State 2	State 3	State 4	Overall	
GT	0.72	0.58	100	100	100	100	History contingent
GT^{34}	0.64	-	0	100	100	67	State and history contingent
GT^4	0.19	-	0	0	100	44	State and history contingent
$D.AlT^4$	0.40	0.40	0	0	50	31	Relies on inference
AD	0	0	0	0	0	0	-

Notes: The percentage of high effort is calculated by assuming the other server uses the same strategy. The calculation is for the steady state.

We would like to reiterate few points. First, providing low effort is always an SPE. Second, when cooperation is supported, infinitely many SPEs exist, so analyzing GT-type strategies is useful to identify the maximum amount of high effort that could be sustained. Third, for the case of a visible queue, when $\delta > 0.72$, we need to distinguish between the maximum efficient and maximum absolute proportions of high effort that can be sustained in SPE. We do so because even though, theoretically, 100% of high effort can be sustained in SPE, we expect that human subjects may recognize that this sustained effort is not beneficial to them, because they can be better off by limiting their effort when the queue is short. Lastly, for the case in which a queue is not visible, equilibria exist that rely on players' inference regarding the probability of being in a particular state given a sequence of action profiles. For example, given our design, even if the number of new arrivals (λ_t) has not been revealed in period t, if players observe that $\theta_{t-1} = 4$ and $e_{t-1} = (l, l)$, they should conclude $Pr(\theta_t = 4) = 1$, and providing high effort will be beneficial as long as the other player does so too. We define a trigger strategy $D.AlT^4$ that provides low effort until mutual low effort has been observed in state 4 and then provides high effort in the subsequent period. Then, after one period of high effort, this strategy immediately reverts back to low effort until another mutual low effort is observed in state 4 in the past. We calculate that $\delta^*_{nv}(D.AlT^4) = 0.40$. The strategy is also a trigger strategy in that it prescribes low effort forever if one of the players did not provide high effort after mutual low effort has been observed in state 4. Whether and to what extent subjects actually use these strategies and provide high effort is an empirical question we address with behavioral experiments.

4.2 Treatments and Hypotheses

For the first study, we implement a 3 × 2 factorial design in which we vary the expected length of the interaction and queue visibility. We chose the three values of δ so that, in combination with the variation in visibility, we obtain distinct predictions regarding the maximum high effort sustained in an SPE across the three states. To induce long-term relationships, we implemented a random termination protocol of Roth and Murnighan (1978). In particular, we described this protocol to subjects as the computer rolling a 12-sided die each period of the match, with the match continued if the number was below 7 (9; 11) for the $\delta = \frac{3}{6}$ ($\delta = \frac{4}{6}$; $\delta = \frac{5}{6}$) treatment. To ensure subjects were comfortable with this procedure, we included a testing phase in which we required them to roll the computerized dice to simulate a duration of 10 matches. The rolls in the actual experiment were pre-drawn so that different visibility treatments had the same match-length realizations. The match-length realizations for each treatment are presented in Table A1 in the Appendix A.2.

Next, we provide three general hypotheses based on our theoretical predictions summarized in Table 1. Our first hypothesis deals with the effect of the expected duration of the interactions. In particular, we expect to observe the highest effort when $\delta = \frac{5}{6}$ and the lowest when $\delta = \frac{3}{6}$. Specifically, the table shows that when $\delta = \frac{5}{6}$, 100% effort could supported in an SPE for both the visible and non-visible treatments (critical thresholds of .72 and .58 are below $\frac{5}{6}$). Similarly, when $\delta = \frac{3}{6}$, at most 44% and 31% of cooperation could be supported in an SPE for the visible and non-visible treatments, respectively. The theoretical predictions regarding the expected duration of interaction are consistent with the existing experimental evidence on cooperation in the indefinitely repeated Prisoner's Dilemma settings (e.g., see Result 1 in the survey by Dal Bó and Fréchette, 2018). In addition to calculating critical thresholds, in Appendix A.1.5, we calculate a measure of strategies uncertainty called the size of the basin of attraction of Always Defect that has received considerable empirical support in the experimental literature (e.g., see Result 4 in the survey by Dal Bó and Fréchette, 2018). The results of these calculations show that strategic uncertainty faced by the player who chooses to provide high effort reduces across the three treatments (see Figure A4). Combined, we expect that in both the visible and not-visible treatments, a longer expected duration of interaction will lead to greater effort.

Hypothesis 1 *Effort increases in the expected duration of an interaction (δ).*

Our second hypothesis deals with the effect of queue visibility on effort provision. In particular, when $\delta = \frac{3}{6}$, Table 1 shows that GT^4 can be sustained as SPE when the queue is visible (critical thresholds of .19 is below $\frac{3}{6}$). This implies that up to 44% of high effort can be sustained when the queue is visible because subjects should find it easy to provide high effort when the queue is long ($\theta = 4$). However, if the queue is not visible, subjects will not be able to play any state-contingent strategies. Instead, they might learn to play strategies that rely on the inference of the state such as $D.AlT^4$ (critical thresholds of .40 is below $\frac{3}{6}$), which leads to a lower percentage of high effort. When $\delta = \frac{5}{6}$, the full effort can be supported in equilibrium when the queue is visible and not visible (critical thresholds of .72 and .58 are below $\frac{4}{6}$). However, there is an important distinction based on efficiency. Namely, when the queue is visible, an efficient outcome is to provide low effort when the queue is short ($\theta = 2$) and high effort when the queue is medium or long ($\theta = 3$ and $\theta = 4$). Whereas when the queue is not visible, the efficient outcome is to provide full effort. Thus, we expect that when $\delta = \frac{5}{6}$ the highest effort will be achieved with a non-visible queue.

Hypothesis 2 (a) *When the expected duration of future interactions is short ($\delta = \frac{3}{6}$), effort is higher when the queue is visible.*

(b) *When the expected duration of future interactions is long ($\delta = \frac{5}{6}$), effort is higher when the queue is not visible.*

The third hypothesis deals with the type of strategies that we should observe across the treatments. Specifically, we expect that subjects will use state- and history-contingent strategies when the queue is visible, but use only history-contingent strategies when the queue is not visible. Although this expectation may seem trivial, the multiplicity of equilibria in repeated-game settings like the ones studied in this paper means the theory does not provide a sharp prediction. In fact, subjects could rely only on history-contingent strategies or could provide low effort across all states in both visible and not-visible treatments. Thus, without conducting lab experiments, whether subjects will learn to play strategies that differ in effort provision across different states and whether any differences will exist among the visibility treatments is not clear.

Hypothesis 3 *When the queue is visible, subjects use state-contingent strategies.*

Note that Hypothesis 3 is important because it provides an insight regarding the mechanism behind the main hypothesis of the paper (Hypothesis 2). In particular, Hypothesis 3 highlights the reason visibility matters. Namely, by changing the visibility of the queue, a manager can influence the type of strategies that servers use to support high effort in equilibrium.

4.3 Experiment Details and Administration

For study 1, we used ORSEE software (Greiner, 2015) to recruit 280 students on the campus of a large public US university between January and February of 2020. We ran 24 sessions with the experimental interface programmed in oTree (Chen, Schonger, and Wickens, 2016) (see Appendix A.2.2 for screenshots of the interface). For each session, we invited 14 subjects; however, because of the no-shows, the actual number of participants in each session varied between 10 and 14. Instructions used in the experiment consisted of a set of interactive screens that explained all aspects of the experiment, and we provided a printed copy that subjects could use for reference during the experiment (see Appendix A.2.3). At the end of the instructions, subjects completed a 10-question quiz (see Appendix A.2.4). Subjects were asked to describe their strategies in a survey right after the experiment. They were also asked if their decisions in a period depended on what happened in the previous periods and if their strategy was different between the initial matches and the later matches. We report their answers in Table A2 in Appendix A.3.

We used a between-subjects design whereby each participant took part in only one experimental treatment. Table A3 of the Appendix A.3 presents a summary of the six treatments. Each treatment consisted of four sessions, and each session consisted of either 80, 50, or 25 matches depending on the probability of continuation. At the beginning of each match, subjects were randomly paired with one other subject and remained paired with that subject for the duration of the match. Subjects remained anonymous throughout the session. Throughout the experiment, we used experimental points as the currency, with 250 points equaling \$1. Subjects were paid in cash at the end of the experiment. The average earning in our experiment was \$22.60 (including the \$5 show-up fee).

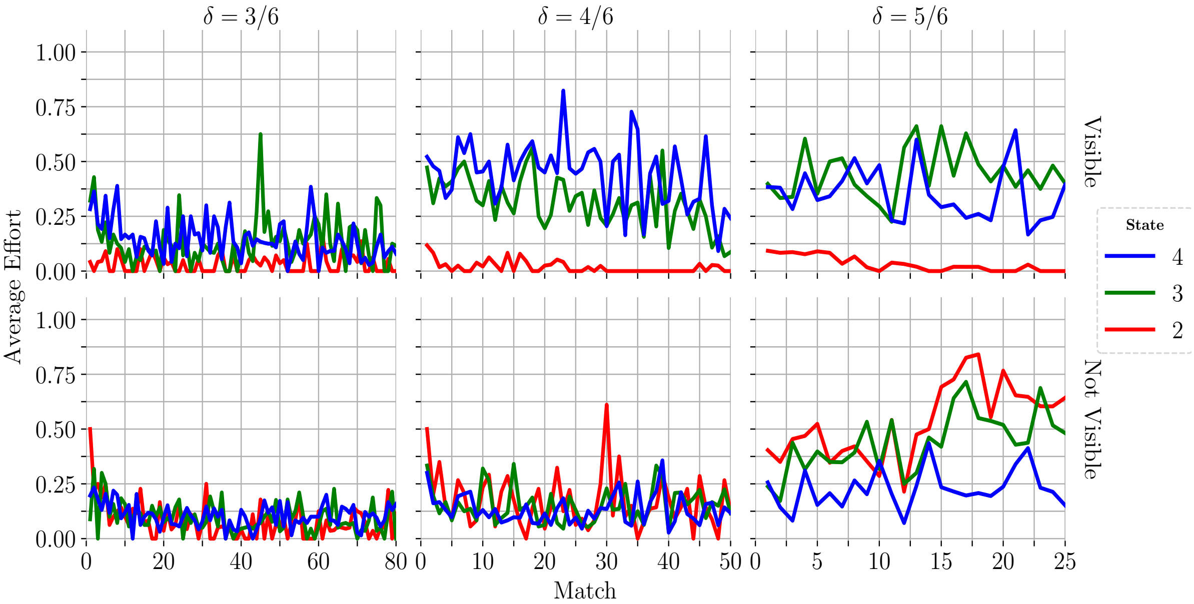
5 Study 1: Results

Figure 3 presents the evolution of effort across all matches in our experiment. Several clear patterns emerge. First, the average effort is higher for higher values of δ. Second, when the queue is visible, the average effort across states differs, indicating that subjects make choices contingent on states. As expected, this pattern is more salient for the high δ treatments. Third, in the treatment in which the queue is not visible and $\delta = \frac{5}{6}$, the average effort increases as subjects gain experience. Furthermore, the figure also shows that some subjects provide high effort in state 2 when the queue is not visible. This finding is more salient in the second half of matches of the $\delta = \frac{5}{6}$ treatment. Although this result is good news for a manager who wants servers to exert a homogeneous and fast processing speed, it is inefficient from the subjects' perspective because in state 2, both the Pareto-optimal and the Nash-equilibrium is to provide low effort for both servers. In fact, the efficiency is higher when the queue is visible than when it is not visible across all three δ values.[9]

---

[9] Table A4 in Appendix A.3 presents the efficiency of outcomes for each of the six treatments. The effect of visibility on efficiency is significant at the 5 percent level when $\delta = \frac{4}{6}$ and $\delta = \frac{5}{6}$ ($p = 0.01$ and 0.03, respectively), but not significant when $\delta = \frac{3}{6}$ ($p = 0.45$).

Figure 3: Evolution of Effort (Study 1)



Notes: Average effort in a match is calculated as the percentage of high effort chosen by all participants in that match. High effort is coded as 1, and low effort is coded as 0.

Table 2 presents the percentage of high effort observed in the second half of our experiment. We focus on the second half of the experiment because, at that point, subjects have become familiar with the strategic environment and the learning has stabilized. The table breaks down actions by effort in the first period and effort across all periods. The first-period effort is important because it provides clear evidence of the subject's intention for the match. In addition, the first-period effort provides an unbiased view of the decisions across the three states. Finally, combined with effort across all periods, the first-period effort provides indirect evidence on the dynamics within the interaction. For example, when $\delta = \frac{5}{6}$ and the queue is visible, effort in state 4 across all periods is approximately half the effort in state 4 in the first period (29.0% vs. 65.0%). This finding suggests that subjects have used strategies that punished deviations from high effort. The end-of-experiment survey also reveals this dynamic (see Table A2 of the Appendix A.3).

Table 2: Percentage of High Effort (Study 1)

Treatment		First Period			All Periods			
Visibility	δ	State 2	State 3	State 4	State 2	State 3	State 4	All States
Yes	$\frac{3}{6}$	3.7 (2.6)	11.4 (4.0)	15.0 (4.2)	3.9 (2.2)	17.0 (5.0)	10.7 (2.9)	10.9 (2.8)
Yes	$\frac{4}{6}$	0.3 (0.3)	28.6 (5.4)	66.5 (6.0)	0.7 (0.3)	23.8 (4.7)	34.8 (4.4)	25.0 (2.6)
Yes	$\frac{5}{6}$	0.0 (0.0)	41.3 (6.2)	65.0 (6.5)	1.1 (0.4)	50.5 (5.7)	29.0 (4.1)	28.4 (3.0)
No	$\frac{3}{6}$	3.6 (1.2)	3.5 (1.8)	3.8 (1.4)	6.1 (1.7)	8.9 (2.3)	9.8 (1.8)	8.9 (1.7)
No	$\frac{4}{6}$	9.7 (4.0)	8.7 (3.5)	8.9 (3.9)	15.5 (3.3)	16.0 (3.5)	12.9 (2.3)	14.0 (2.4)
No	$\frac{5}{6}$	50.5 (6.3)	36.7 (6.2)	44.3 (5.5)	64.7 (4.8)	54.2 (4.9)	23.5 (3.1)	36.8 (4.3)

Notes: Standard errors (in parentheses) are calculated by taking one subject as a unit of observation.

Several observations from Table 2 are noteworthy. The first observation concerns effort provision in the first period of interactions in each treatment. As expected, no difference exists across the three states when the queue is not visible.[10] This finding is reassuring in that subjects could not distinguish among the states in the first period. When the queue is visible, however, we find a clear trend – higher effort in the states with more customer orders in line. Specifically, when the discount factor is $\frac{3}{6}$, the percentage of high effort increases from 3.7% in state 2, to 11.4% in state 3, to 15.0% in state 4. When the discount factor is $\frac{4}{6}$, the percentage of high effort increases from 0.3% in state 2, to 28.6% in state 3, to 66.5% in state 4. When the discount factor is $\frac{5}{6}$, the percentage of high effort increases from 0.0% in state 2, to 41.3% in state 3, to 65.0% in state 4. All of the increases from state 2 to state 4 are significant at the 1 percent level using a matched-pairs t-test.[11]

Table 2 shows that the percentage of high effort increases with δ. The difference is present across all states when the queue is not visible and across states 3 and 4 when the queue is visible. The difference is particularly noticeable in the first-period outcomes because outcomes after the first period largely depend on what happened initially. For example, if half of the subjects play the *AD* strategy and the other half play the *GT* strategy, we would observe a cooperation rate close to 25%, even though 50% of subjects play cooperative strategies. To formally test whether

---

[10] The p-value for a matched-pairs t-test of choosing high effort between state 2 and state 3 when the discount factor is $\delta = \frac{3}{6}(\frac{4}{6}, \frac{5}{6})$ is 0.97 (0.68, 0.61). The p-value for a matched-pairs t-test of choosing high effort between state 3 and state 4 when the discount factor is $\delta = \frac{3}{6}(\frac{4}{6}, \frac{5}{6})$ is 0.34 (0.65, 0.58).

[11] As a robustness check, we run a probit regression of the choice of high effort in the first period on the dummy for whether the state is 2 or 4 (clustering standard errors at the individual level). We find the increase from state 2 to state 4 is significant at the 5 percent level.

the difference between the first-period outcomes is significant, we first compare treatments between $\delta = \frac{3}{6}$ and $\delta = \frac{5}{6}$. We run a probit regression of the choice of high effort in the first period on the dummy for whether the discount factor is high ($\delta = \frac{5}{6}$), with standard errors clustered at the session level. We find a significant difference ($p < 0.01$) both when the queue is visible and when it is not. We also consider other comparisons. In particular, if the dummy variable is 0 when $\delta = \frac{3}{6}$ and 1 otherwise, the difference is still significant ($p < 0.01$) both when the queue is visible and when it is not. If the dummy variable is 0 when $\delta = \frac{3}{6}$ and 1 when $\delta = \frac{4}{6}$, the difference is significant ($p < 0.01$) when the queue is visible but not significant ($p = 0.33$) when the queue is not visible. If the dummy variable is 0 when $\delta = \frac{4}{6}$ and 1 when $\delta = \frac{5}{6}$, the difference is not significant ($p = 0.73$) when the queue is visible but is significant ($p < 0.01$) when the queue is not visible. We summarize these observations as Result 1.1.

Result 1.1 *Hypothesis 1 is supported: servers provide higher effort when the expected duration of future interactions is longer.*

Next, Table 2 shows that when the discount factor is low ($\delta = \frac{3}{6}$), servers provide higher effort when the queue is visible than when it is not. To formally test whether the difference is significant, we run a probit regression of high effort in the first period on the dummy for whether the state is visible, with standard errors clustered at the session level. We find that the difference is significant at the 10 percent level ($p = 0.05$). When the discount factor is medium ($\delta = \frac{4}{6}$), the theory suggests that more cooperation is possible when the queue is not visible; however, in the data, we find that cooperation is significantly higher when the queue is visible ($p = 0.06$). One reason for this finding might be that although *GT* is an SPE strategy, it is not risk dominant when $\delta = \frac{4}{6}$.[12] Taken together, the two treatments provide evidence that servers choose significantly higher effort when the queue is visible ($p = 0.03$).[13] When the discount factor is high ($\delta = \frac{5}{6}$), the overall effort is higher when the queue is not visible, which is consistent with the theoretical prediction of the impact of queue visibility. However, the difference is not significant ($p = 0.33$).[14] We summarize these findings as Result 1.2.

Result 1.2 (a) *Hypothesis 2.a is supported: when the expected duration of future interactions is short, servers provide higher effort when the queue is visible.*

(b) *The evidence is directionally consistent with Hypothesis 2.b: when the expected duration of*

---

[12] To check whether a strategy is risk dominant, we convert the indefinitely repeated PD game into a coordination game (Dal Bó and Fréchette, 2011; Blonski and Spagnolo, 2015). When the queue is not visible, *GT* is a risk-dominant strategy when $\delta = \frac{5}{6}$ but not when $\delta = \frac{4}{6}$ or $\delta = \frac{3}{6}$. To apply this approach to the case when the queue is visible, we do this exercise for GT^{34}. We find that GT^{34} is risk dominant when $\delta = \frac{5}{6}$ and $\delta = \frac{4}{6}$, and it is not risk dominant when $\delta = \frac{3}{6}$.

[13] A probit regression of high effort in the first period on two dummy variables – visibility and discount factor – yields *p*-values of 0.01 and less than 0.01 when clustering standard errors at the session level.

[14] This result holds if we include the state dummy variables in the probit regression. For example, when the discount factor is low ($\delta = \frac{3}{6}$), the difference is significant at the 10 percent level ($p = 0.06$). When the discount factor is medium ($\delta = \frac{4}{6}$), the difference is marginally not significant ($p = 0.10$). Taken together, the difference is significant at the 5 percent level ($p = 0.04$). When the discount factor is high ($\delta = \frac{5}{6}$), we do not find such a difference ($p = 0.34$).

future interactions is long, servers provide higher effort when the queue is not visible, but the difference is not significant.

The fact that effort provision depends on the state realization when the queue is visible leads us to believe that subjects use state-contingent strategies. The fact that effort in the first period is greater than the effort across all periods leads us to believe that subjects use history-contingent strategies. Therefore, we use a finite-mixture estimation approach to formally estimate the strategies underlying choices in our experiment. This approach has advantages over other estimation methods, and evidence shows that it performs well (Dal Bó and Fréchette, 2018; Romero and Rosokha, 2018; Dal Bó and Fréchette, 2019). The method works by first specifying the set of K strategies considered by the modeler. Then, for each subject $n \in N$, and each strategy $k \in K$, the method prescribes comparing subject n's actual play with how strategy k would have played in her place. Let $X(k,n)$ denote the number of periods in which subject n's play correctly matches the play of strategy k. Then, let X denote a $K \times N$ matrix of the number of correct matches for all combinations of subjects and strategies. Similarly, let Y denote a $K \times N$ matrix of the number of mismatches when comparing subjects' play with what the strategies would do in their place. Then, define a Hadamard-product P:

$$P = \beta^X \circ (1-\beta)^Y, \tag{8}$$

where β is the probability that a subject plays according to a strategy and $(1-\beta)$ is the probability that the subject deviates from that strategy. Thus, each entry $P(k,n)$ is the likelihood that strategy k generated the observed choices by subject n. Then, using the matrix dot product, we define the log-likelihood function $\mathcal{L}$:

$$\mathcal{L}(\beta, \phi) = ln\big(\phi' \cdot P\big) \cdot \mathbf{1}, \tag{9}$$

where ϕ is a vector of strategy frequencies.

For our estimation, the set of strategies encompasses the five most common strategies found in the literature on repeated games as well as state-contingent variations of those strategies. In particular, we include Always Cooperate (AC), Always Defect (AD), Grim Trigger (GT), Tit-for-Tat (TFT), and Suspicious Tit-for-Tat ($D.TFT$) – the five memory-1 strategies that account for the majority of the strategies in 16 out of 17 treatments reviewed by Dal Bó and Fréchette (2018). We also include modified versions of these strategies that condition on either state 4 or both states 3 and 4. Notably, we include GT^{34} and GT^4, which were analyzed theoretically. In addition, we include the $D.AlT^4$ strategy that could sustain some amount of high effort when the queue is not visible (as well as the corresponding $D.AlT^{34}$ and $D.AlT$ strategies).

Table 3: Estimated Percentage of Strategies (Study 1)

Visibility	δ	AD	AC	TFT	GT	$D.TFT$	$D.AlT$	AC^{34}	TFT^{34}	GT^{34}	$D.TFT^{34}$	$D.AlT^{34}$	AC^{4}	TFT^{4}	GT^{4}	$D.TFT^{4}$	$D.AlT^{4}$	β (%)	$\mathcal{L}$
Yes	$\frac{3}{6}$	59.4				6.5		2.4			22.5		2.4	2.4		4.3		93.7	-895.9
		(9.9)				(5.4)		(2.5)			(9.0)		(2.3)	(2.6)		(5.8)		(1.7)	
Yes	$\frac{4}{6}$	37.9							11.2	14.2	2.5			15.6	18.6			92.0	-1053.6
		(7.5)							(4.6)	(6.3)	(2.2)			(12.8)	(12.9)			(0.9)	
Yes	$\frac{5}{6}$	32.7							32.1	13.7	4.5		5.5	5.4	6.2			93.7	-904.9
		(7.1)							(7.5)	(6.4)	(2.8)		(3.6)	(4.4)	(5.2)			(0.7)	
No	$\frac{3}{6}$	65.8	2.1			17.0	4.1					2.1				6.8	2.1	93.1	-1096.1
		(7.2)	(2.2)			(6.9)	(3.0)					(2.3)				(4.4)	(2.1)	(1.3)	
No	$\frac{4}{6}$	56.7		6.3	2.1	21.4						4.1				3.2	6.3	92.2	-1082.3
		(8.6)		(3.5)	(2.2)	(6.4)						(3.4)				(3.1)	(3.9)	(1.4)	
No	$\frac{5}{6}$	36.6	2.1	37.2	6.2	15.8											2.1	89.5	-1244.4
		(7.1)	(2.0)	(7.7)	(3.5)	(5.5)											(2.2)	(1.3)	

Notes: For ease of reading, estimated percentages < 0.1 are not displayed. Strategy superscripts denote states in which this strategy is played; in states that are not included in the superscripts, the strategy specifies playing AD. The first five strategies are viewed as history-contingent strategies; AC^{34} and AC^{4} are state-contingent strategies; and the rest are state- and history-contingent strategies. The value of $(1 - \beta)$ can be interpreted as the amount of noise not captured by the specified strategies. Bootstrapped standard errors are in parentheses. The unit of observation is one subject.

Table 3 presents the estimation results for the second half of matches (see Table A5 of the Appendix A.3 for the estimation results for the first half of matches). We find that the most common strategy across all six treatments is the AD strategy. This finding is not surprising given the prevalence of the AD strategy in the literature on the indefinitely repeated PD with parameters similar to ours (e.g., Dal Bó and Fréchette, 2011). Even so, we find a clear pattern in the type of other strategies used across the treatments. In particular, when the queue is visible, the frequency of state- and history-contingent strategies is 29.2% ($\delta = \frac{3}{6}$), 62.1% ($\delta = \frac{4}{6}$), and 61.9% ($\delta = \frac{5}{6}$). When the queue is not visible, the frequency of state- and history-contingent strategies is 11% ($\delta = \frac{3}{6}$), 13.6% ($\delta = \frac{4}{6}$), and 2.1% ($\delta = \frac{5}{6}$).[15] The difference for the $\delta = \frac{3}{6}$ ($\delta = \frac{4}{6}$, $\delta = \frac{5}{6}$) treatment is significant at the 5 percent (1 percent, 1 percent) level using a non-parametric permutation test.[16] We summarize this finding as Result 1.3.

---

[15] The low proportion of state-contingent *D.AlT*-type strategies is supported by the end-of-experiment unincentivized survey. In particular, in the period following both subjects choosing low effort in state 4, the probability of being in state 4 is 100%, but we only observe 1.0% (2.0%, 1.0%) of mutual high effort when $\delta = \frac{3}{6}$ ($\delta = \frac{4}{6}$, $\delta = \frac{5}{6}$). Regarding the end-of-experiment survey, we asked subjects to describe strategies that they used. We find that 0.0% (8.0%, and 2.0%) of subjects described strategies close to *D.AlT* when $\delta = \frac{3}{6}$ ($\delta = \frac{4}{6}$, $\delta = \frac{5}{6}$). And among those who did describe *D.AlT* strategy, several mentioned difficulty getting others to play this strategy. The end-of-experiment survey results are summarized in Table A2 of Appendix A.3.

[16] For the permutation test (Good, 2013), the null hypothesis is that no difference exists between the proportion of strategies across the visibility treatments for a fixed value of δ. Given the null hypothesis, the distribution of the test statistic is obtained by randomly permuting the treatment labels among subjects (unit of observation).

Result 1.3 *Hypothesis 3 is supported: when the queue is visible, servers use state-contingent strategies.*

In terms of the particular strategies played, we find that when the queue is visible, subjects play sophisticated TFT- and GT-like strategies that provide high effort when the queue is long but low effort when the queue is short. These strategies are different from the TFT and GT strategies studied in the repeated-game literature in that they respond to the opponent's behavior conditional on the state of the queue. Importantly, the proportion of these strategies observed in the data predictably varies by treatment and is comparable to the existing literature on repeated games. In particular, we observe an initial increase and then decrease of the proportion of strategies that cooperate in state 4 but not in state 3 (TFT^4 and GT^4 accounted for 2.4% when $\delta = \frac{3}{6}$; 34.1% when $\delta = \frac{4}{6}$; and, 11.6% when $\delta = \frac{5}{6}$).[17] The decrease is due to the switch to strategies that cooperated across more states (e.g., TFT^{34} and GT^{34}). For example, when $\delta = \frac{5}{6}$, 45.8% of subjects use cooperative strategies that provide high effort in states 3 and 4 as opposed to 25.4% when $\delta = \frac{4}{6}$, and 0.0% when $\delta = \frac{3}{6}$. When the queue is not visible, subjects could not play any of these strategies; as a result, when δ is increased, we observe a switch from non-cooperative strategies (AD and DTFT account for 82.8% when $\delta = \frac{3}{6}$) to cooperative strategies (TFT and GT comprise 43.4% when $\delta = \frac{5}{6}$).

Regarding the comparison to the existing literature, consider the non-stochastic version of our study with $\theta = 3$ every period (i.e., servers always face three customers in every period) studied in Dal Bo and Frechette (2011). In particular, the authors find the proportion of TFT in $\delta = \frac{3}{4}$ treatment to be 35.2%. Thus, our estimation of 37.2% of TFT in the not-visible treatment and 37.5% of state-contingent TFT-like strategies is very much in line with the existing literature.

6 Study 2: Experimental Design and Theoretical Predictions

We had two main objectives for the second study. The first was to provide a more natural queueing framing for the participants (instead of a context-neutral presentation in the first study), including details of the cost and revenue functions. The second was to explore a wider parameter space with a streamlined microfoundation. In particular, we explicitly assume that servers work together to process tasks (e.g., a team of nurses/doctors treating a patient in a healthcare setting) and the cost incurred by each server depends on her utilization, which in turn depends on the effort by the other server.

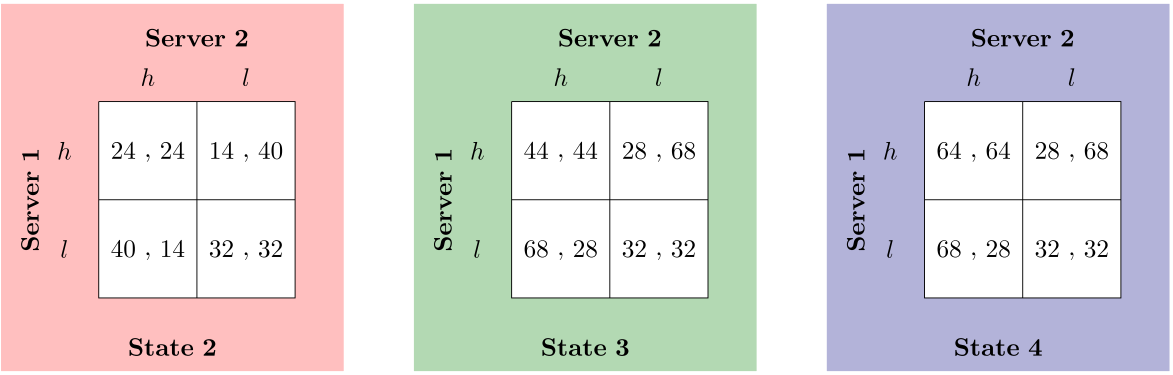
Appendix B.2.3 presents instructions that were used for the second study. To highlight some of the details, we told the subjects, *"You and the participant you are paired with will work together to process the task queue. Specifically, in each round, you will choose how much capacity to allocate (either 1 or 2 units). The participant you are paired with will also choose how much capacity to*

---

[17] The high standard errors of the estimates for TFT^4 and GT^4 are the results of these two strategies being very similar in behavior for the considered duration of interactions. When we estimate the joint proportion, we obtain 34.1 with a standard error of 6.8.

allocate (either 1 or 2 units).” We chose to present the decision as the choice of capacity rather than effort or speed of processing to minimize any negative associations with “low effort” or “slow speed.” We then went on to describe the implications of the capacity choices on the dynamics (i.e., leftover tasks) and how payoffs are determined. In terms of the payoffs, we explicitly provided the cost function, the revenue function, and the payoff calculations associated with the capacity choices. Finally, the decision screen was different from the one in Study 1, in that we presented the state of the queue in text, and we provided a button for decision support that, when clicked, presented the summary of the payoffs similar to how the payoff tables were presented in study 1.

Figure 4: Stage-Game Payoffs in Each State (Study 2)



State 2

Server 1 \ Server 2	h	l
h	24 , 24	14 , 40
l	40 , 14	32 , 32

State 3

Server 1 \ Server 2	h	l
h	44 , 44	28 , 68
l	68 , 28	32 , 32

State 4

Server 1 \ Server 2	h	l
h	64 , 64	28 , 68
l	68 , 28	32 , 32

Notes: The three columns present stage-games played in the three possible states. State $\theta \in \{2, 3, 4\}$ corresponds to θ customer orders in line. Each player chooses low capacity (l) or high capacity (h) to process customer orders. Matrices present normal form representation of the stage game played in each state.

In terms of the payoffs, we picked another cost function, $c(.)$, and different parameters for the compensation function, $r(.)$. Specifically, for study 2, the cost is a function of the server's capacity choice and the realized utilization (i.e., the fraction of the time that a server ends up working in a period), $T = min(1, \frac{\theta}{total\ capacity})$. Intuitively, the fraction of the time that server ends up working in a period increases with the number of customer orders but decreases with the sum of servers' capacity. At the same time, this fraction cannot exceed 1. We set the cost of choosing one unit of capacity $c(1, T) = aT^2 + bT + c$ with $a = 2$, $b = 18$, and $c = 20$ and the cost of choosing two units of capacity $c(2, T) = xT^2 + yT + z$ with $x = 6$, $y = 54$, and $z = 20$. Note that the cost function in study 2 relies on the aggregate measure of servers' performance T, because servers work together to process a single task. This foundation is different from study 1, where servers were assumed to work separately. In Appendix B.1, we provide more details on the cost function, including the visual representation (Figure B1) and comparative statics regarding theoretical calculations (Figure B2). Regarding the compensation, we simplify the function relative to Study 1 and only include the per-unit compensation. Specifically, when the group processed $M(.)$ tasks, $r(e_i, e_j, \theta) = kM(e_i, e_j, \theta) + \mathbf{1}_{M(e_i,e_j,\theta)=4} bonus$ with $k = 36$, and $bonus = 0$. In Figure 4, we present resulting payoffs for each combination of effort choices (see Appendix B.1.3 for calculation details).

Using the approach described in section 3, we find that when the queue is visible, $\delta_v^*(GT) = 0.67$, $\delta_v^*(GT^{34}) = 0.62$, and $\delta_v^*(GT^4) = 0.26$. When the queue is not visible, $\delta_{nv}^*(GT) = 0.55$ and $\delta_{nv}^*(D.AlT^4) = 0.46$ (see Appendix B.1.5 for more details). We summarize these results in Table 4. Importantly, the patterns on cooperation across treatments are the same as that in study 1, suggesting that our theoretical results from section 4.2 are robust to the alternative specification of the compensation and cost functions.

Table 4: Summary of Symmetric SPE (Study 2)

Strategy	Critical Threshold When Queue is Visible (δ_v^*)	Critical Threshold When Queue is Not Visible (δ_{nv}^*)	High Effort (%)				Comment
			State 2	State 3	State 4	Overall	
GT	0.67	0.55	100	100	100	100	History contingent
GT^{34}	0.62	-	0	100	100	67	State and history contingent
GT^4	0.26	-	0	0	100	44	State and history contingent
$D.AlT^4$	0.46	0.46	0	0	50	31	Relies on inference
AD	0	0	0	0	0	0	-

Notes: δ_v^* is the critical threshold that a strategy is an SPE when the queue is visible. δ_{nv}^* is the critical threshold that a strategy is an SPE when the queue is not visible. The percentage of high effort is calculated by assuming the other server uses the same strategy.

In terms of the experiment, we implement a 2×2 factorial between-subjects design in which we vary the expected duration of an interaction (i.e., $\delta = \frac{3}{6}$ and $\delta = \frac{5}{6}$) and whether servers know the state of the queue. The theoretical predictions and hypotheses remain the same as in Study 1.

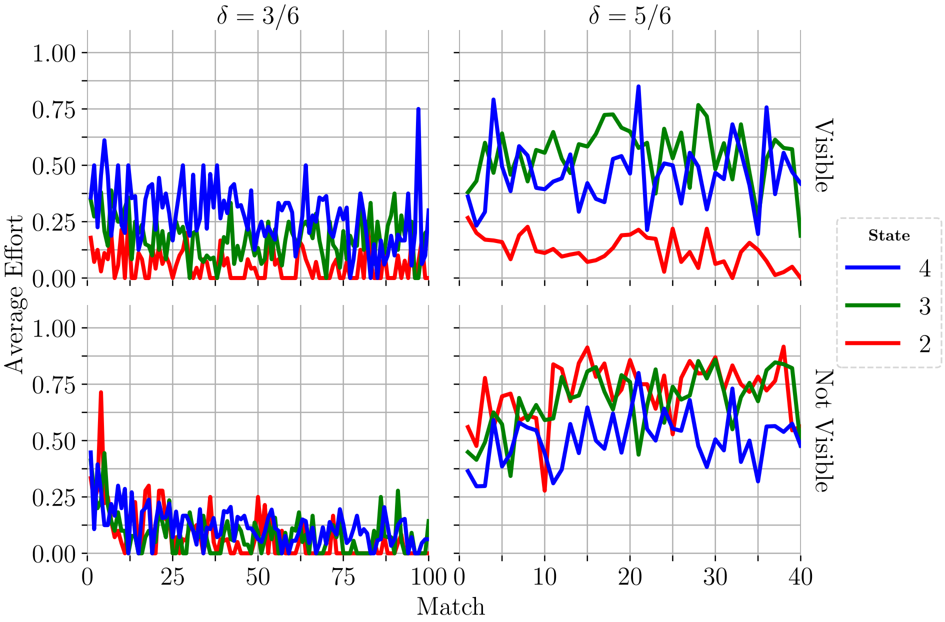
Regarding the experimental administration, we used ORSEE software (Greiner, 2015) to recruit 142 students in April 2021. We ran 12 sessions with the experimental interface programmed in oTree (Chen, Schonger, and Wickens, 2016) (see Appendix B.2.2 for screenshots of the interface and Appendix B.2.4 for the quiz). Each treatment consisted of three sessions, and each session consisted of either 100 matches (for $\delta = \frac{3}{6}$) or 40 matches (for $\delta = \frac{5}{6}$). Note that we increased the number of matches to ensure subjects had enough time to learn, because subjects may still have been learning in the second half of study 1 (see Figure A5 in Appendix A.3). Throughout the experiment, we used experimental points as the currency, with 400 points equaling \$1. Subjects were paid in cash at the end of the experiment. The average earning in our experiment was \$27.67 (including the \$5 show-up fee).

7 Study 2: Results

Figure 5 presents the evolution of effort across matches in the second study, and Table 5 provides more details on the average percentage of high effort observed in the second half of the study. Similar to study 1, we find that high effort increases with δ. In particular, we run a probit regression of the choice of high effort in the first period on the dummy for whether the discount factor is high

($\delta = 0.83$), with standard errors clustered at the session level. We find the difference is significant ($p < 0.01$) both when the queue is visible and when it is not.

Figure 5: Evolution of Effort (Study 2)



Notes: Average effort in a match is calculated as the percentage of High effort chosen by all participants in that match. High effort is coded as 1, and low effort is coded as 0.

Table 5: Percentage of High Effort (Study 2)

Treatment		First Period			All Periods			
Visibility	δ	State 2	State 3	State 4	State 2	State 3	State 4	All States
Yes	$\frac{3}{6}$	3.8 (2.8)	16.3 (5.2)	34.1 (7.6)	4.4 (2.9)	16.8 (4.6)	21.6 (5.0)	16.5 (3.4)
Yes	$\frac{5}{6}$	6.8 (3.7)	50.4 (7.1)	66.1 (7.8)	9.2 (3.5)	54.7 (4.6)	41.6 (5.1)	38.5 (3.2)
No	$\frac{3}{6}$	1.4 (0.8)	1.4 (1.0)	0.6 (0.4)	2.9 (1.0)	6.7 (1.9)	10.4 (2.3)	8.1 (1.8)
No	$\frac{5}{6}$	63.2 (8.3)	66.4 (8.0)	65.6 (7.7)	75.6 (3.7)	73.9 (3.6)	49.0 (4.6)	62.4 (4.1)

Notes: The table presents the data in the second half of the experiment: 50-100 matches for $\delta = 0.5$ and 20-40 matches for $\delta = 0.83$. Standard errors (in parentheses) are calculated by taking one subject as a unit of observation.

Result 2.1 *Hypothesis 1 is supported: servers provide higher effort when the expected duration of future interactions is longer.*

Regarding part (a) of the second hypothesis, we find that when $\delta = \frac{3}{6}$, servers provide significantly more effort when the queue is visible than when it is not ($p < 0.01$). Regarding part (b) of the second hypothesis, we find that when $\delta = \frac{5}{6}$, servers provide significantly less effort when the queue is visible than when it is not. However, unlike study 1, this difference is now significant ($p < 0.01$). The difference becomes significant for two possible reasons. The first is that for study 2, we increased the number of matches. Therefore, subjects had more opportunities to learn, and the trend noted in Figure 3 (and Figure A5 in the Appendix A.3) led to the difference becoming more pronounced. The second is that the parameters chosen for the second study are more favorable for cooperation, which led to an increase in the proportion of high effort.[18] However, although the increase occurred for all three states in the not-visible treatment, the increase is minimal in state 2 of the visible treatment (in which we did not expect any cooperation). These results validate our second hypothesis.

Result 2.2 (a) *Hypothesis 2.a is supported: when the expected duration of future interactions is short, servers provide higher effort when the queue is visible.*

(b) *Hypothesis 2.b is supported: when the expected duration of future interactions is long, servers provide higher effort when the queue is not visible.*

Table 6: Estimated Percentage of Strategies (Study 2)

Visibility	δ	AD	AC	TFT	GT	$D.TFT$	$D.AlT$	AC^{34}	TFT^{34}	GT^{34}	$D.TFT^{34}$	$D.AlT^{34}$	AC^{4}	TFT^{4}	GT^{4}	$D.TFT^{4}$	$D.AlT^{4}$	β (%)	$\mathcal{L}$
Yes	$\frac{3}{6}$	53.1 (8.7)	2.8 (3.0)					2.8 (3.0)	3.3 (3.2)	7.8 (4.9)			8.4 (4.6)	8.3 (4.8)	4.5 (4.7)	9.0 (5.3)		93.7 (1.1)	-755.2
Yes	$\frac{5}{6}$	13.9 (6.1)		2.8 (2.5)		11.1 (5.2)		9.2 (4.8)	22.2 (7.1)	16.8 (6.7)	5.6 (3.7)		2.5 (2.2)		16.0 (7.5)			89.6 (1.1)	-1640.4
No	$\frac{3}{6}$	56.3 (7.4)				12.7 (6.8)	2.8 (2.8)				4.5 (4.9)	8.4 (3.9)				4.1 (3.9)	11.2 (4.4)	95.1 (1.2)	-708.7
No	$\frac{5}{6}$	5.9 (3.8)	14.6 (5.7)	37.1 (7.7)	16.0 (7.1)	20.6 (6.8)											5.9 (3.9)	89.1 (1.1)	-1578.2

Notes: For ease of reading, estimated percentages < 0.1 are not displayed. Strategy superscripts denote states in which this strategy is played; in states not included in the superscripts, the strategy specifies playing Always Defect (AD). The first five strategies are viewed as history-contingent strategies; AC^{34} and AC^{4} are state-contingent Markov strategies; and the rest are state- and history-contingent strategies. Bootstrapped standard errors are in parentheses. The unit of observation is one subject.

---

[18] One way to compare the extent to which cooperation is favorable is to use the size of the basin of attraction of AD against GT – $SizeBAD$. For $\delta = \frac{5}{6}$, in study 1, $SizeBAD = 0.32$, in study 2, $SizeBAD = 0.16$. See detailed discussions of $SizeBAD$ and its relationship with δ in Figure A4 of Appendix A.1.5 and Figure B4 of Appendix B.1.5.

Lastly, we conduct the finite-mixture estimation to uncover the strategies used by subjects in the second study. Table 6 presents the results for the second half of matches. We find that when the queue is visible, 44.1% (72.3%) of subjects rely on state-contingent strategies for $\delta = \frac{3}{6}$ ($\delta = \frac{5}{6}$), which is greater than 28.2% (5.9%) observed when the queue is not visible.[19]

Result 2.3 *Hypothesis 3 is supported: when the queue is visible, servers use state-contingent strategies.*

Comparing the strategies used across the two studies, we find many similarities. For example, when $\delta = \frac{3}{6}$, subjects primarily rely on uncooperative strategies (AD and versions of $D.TFT$), whereas when $\delta = \frac{5}{6}$, the most popular strategies include the TFT-like strategies. One notable difference is that for $\delta = \frac{5}{6}$, the proportion of AD is smaller in study 2 (13.9 and 5.9 in study 2 vs. 32.7 and 36.6 in study 1). This difference is consistent with overall higher levels of cooperation in the second study noted above.

To summarize, we find that the results of the second set of experiments provide clear support for the three hypotheses derived from theoretical predictions. The results also complement findings from the first set of experiments by establishing the robustness to the decision-making context and functional specification of individual cost and compensation functions.

8 System Performance

System performance metrics, such as the average customer waiting time and the percentage of demand lost due to the system being full, are important for managers. In this section, we analyze these metrics in three ways. First, extending the analysis of effort choice, we consider theoretical predictions based on the efficient equilibrium summarized in Table 1 (see Appendix C.1 for more details). We summarize the results in the column labeled *Theoretical Calculation* of Table 7 and find that theoretical predictions regarding the waiting time are consistent with the predictions regarding the effort choice – waiting time is decreasing in the expected duration of interaction, and the benefit of (non-)visibility depends on the expected duration of interaction. Predictions regarding the lost demand are not as straightforward, because multiple equilibria exist with no lost demand (e.g., symmetric SPEs involving either GT^{34} or GT have no lost demand), and therefore, the theory does not provide a prediction regarding the impact of visibility for treatments with $\delta = 5/6$.

---

[19] The p-value for the non-parametric randomization test for $\delta = \frac{3}{6}$ is .22. The p-value for the non-parametric randomization test for $\delta = \frac{5}{6}$ is $< .01$.

Table 7: System Performance

Treatment		Customer Waiting Time						Lost Demand (%)					
		Study 1			Study 2			Study 1			Study 2		
Visibility	δ	Theoretical Calculation	Raw Data	Data-Driven Simulation	Theoretical Calculation	Raw Data	Data-Driven Simulation	Theoretical Calculation	Raw Data	Data-Driven Simulation	Theoretical Calculation	Raw Data	Data-Driven Simulation
Yes	$\frac{3}{6}$	0.96	1.11 (0.03)	1.90 (0.00)	0.79	0.83 (0.05)	1.47 (0.00)	3.7	5.4 (0.7)	29.7 (0.2)	3.7	3.6 (1.1)	24.4 (0.2)
Yes	$\frac{5}{6}$	0.78	1.08 (0.09)	1.57 (0.00)	0.61	0.71 (0.11)	1.23 (0.00)	0.0	7.3 (2.4)	21.2 (0.1)	0.0	5.5 (1.7)	17.7 (0.1)
No	$\frac{3}{6}$	1.21	1.12 (0.01)	1.94 (0.00)	1.01	0.92 (0.01)	1.72 (0.00)	12.8	5.1 (0.3)	31.4 (0.2)	12.8	4.6 (0.5)	32.2 (0.2)
No	$\frac{5}{6}$	0.67	1.09 (0.06)	1.58 (0.00)	0.53	0.59 (0.05)	0.93 (0.00)	0.0	8.5 (1.1)	22.1 (0.1)	0.0	4.0 (0.6)	10.4 (0.1)

Notes: The theoretical predictions are for the steady state. The system performance in the "Raw Data" columns is calculated using the raw experimental data from all periods. The system performance in the "Data-Driven Simulation" columns is calculated using simulated data from the steady states. See Appendix C.2 for the simulation steps. Standard errors for simulations are calculated using bootstrapping. In the case that a standard error is below 0.005, we denote it as 0.00.

Second, we consider the two metrics calculated based on the data obtained in our experiment. The columns labeled *Raw Data* in Table 7 show realized metrics across various treatments in the two studies. The table shows the impact of the expected duration of interaction on the customer waiting time is consistent with the analysis of cooperation. For example, in study 2, when the queue is visible, the average waiting time when the discount factor is low is 0.92 versus 0.59 when the discount factor is high. However, the impact of visibility on the system performance is mixed. For example, in study 2, when the discount factor is high, the customer waiting time is shorter when the queue is not visible than when it is (0.59 vs. 0.71). To formally test the effect of visibility and expected duration of interaction on the system performance, we conduct regression analysis on the pooled data in which we include dummy variables for the study, visibility, and discount factor. In addition, we include an interaction term to capture the differential impact of the visibility depending on the discount factor (see Table C2 in Appendix C.2). The results show that the customer waiting time significantly reduces with the expected duration of interaction ($p < .01$). However, we find that the effect of the expected duration of interaction on the lost demand is positive but not significant ($p = .60$). This finding is not surprising, because a longer interaction is associated with more periods and thus may cause more lost demand. Lastly, although the effects of visibility and the interaction term are not significant, the magnitudes of the coefficients are directionally consistent with the theoretical predictions.

The third type of analysis that we carry out is based on estimated strategies from the experiment. In particular, columns labeled *Data-Driven Simulation* present steady-state metrics from strategy distributions estimated in Tables 3 and 6. This analysis goes beyond the analysis of raw data by allowing the impact of the initial conditions to be reduced, as well as by allowing direct comparison across various values of δ. Specifically, because the buffer is empty at the start of each match, the waiting time for the customers arriving in period 1 and the initial lost demand in period 1 would be underestimated. Furthermore, when $\delta = 3/6$, 50% of interactions lasted one period, which means the bias associated with the initial conditions could be substantial. Using the estimated strategies, we can carry out simulations that reach steady states, which makes the metrics more comparable to theoretical predictions (see Appendix C.2 for simulation steps). The simulation results in Table 7 show that our results on the beneficial impact of the expected duration of interaction are robust. In addition, in study 2, we find the beneficial impact of non-visibility when the expected duration of interaction is long, and in both studies, we find the detrimental impact of non-visibility when the expected duration of interactions is short. We also carry out regression analysis using the simulated data in which we include dummy variables for the study, visibility, discount factor, and the interaction term between the visibility and discount factor. Table C2 in Appendix C.2 shows the coefficients of visibility, discount factor, and the interaction term are all significant at the 1 percent level.

To summarize, the goal of this section was to take a closer look at system performance variables that are of importance to managers. Consistent with our results on effort provision discussed in sections 5 and 7, we find evidence regarding the beneficial impact of the expected duration of interaction on both the customer waiting time and the lost demand. We also find that the impact of visibility depends on the duration of the interactions. In particular, we find evidence that when the expected duration of interaction is long, a benefit arises from hiding the queue from the servers.

9 Discussion

In this paper, we theoretically and experimentally investigated the effort provision in a single-queue two-server system when compensation is based on the group performance. To the best of our knowledge, we are the first to focus on the repeated nature of interaction among servers in queueing systems and show theoretically that high effort can be sustained in equilibrium even when short-term incentives to free-ride are present for each server in each of the possible states of the queue (i.e., number of customers in line). Furthermore, if servers' interactions are long term, high effort can be sustained regardless of whether the queue is visible. However, as interactions get shorter, visibility becomes an important determinant of the types of strategies that can support high effort in equilibrium. In particular, we theoretically show that providing less visibility of the queue may be better because players will average incentives across multiple states, which will lead them to provide high effort even in states corresponding to the short queue. We also show that if the queue is visible, sustaining high effort when the queue is long is much easier than when the queue is short.

We conducted two laboratory studies to test the theoretical predictions. In particular, in both studies, we varied the expected length of interaction among the servers and the visibility of the queue. We have three main findings. First, longer expected interactions lead to higher effort. This finding establishes the results from the repeated-games literature in the stochastic game underlying the queueing setting. The second and arguably most interesting finding is that queue visibility may have a different impact depending on the expected duration of the interaction. In particular, if the expected duration of the interaction is short, subjects provide higher effort when the queue is visible than when it is not. However, if the expected duration of the interaction is long, subjects provide lower effort when the queue is visible than when it is not. Our third finding concerns the strategies that human servers use in the two visibility treatments, which establishes the mechanism behind the second finding. Specifically, following the repeated-games literature, we carried out finite-mixture model estimation of the strategies. We find that when the queue is not visible, subjects primarily relied on history-contingent strategies. Among them, the most common ones are tit-for-tat and suspicious tit-for-tat. When the queue is visible, however, a significant proportion of subjects rely on state- and history-contingent versions of these strategies. These state- and history-contingent strategies are sophisticated in that they keep track of what players did the last time they were in the current state.

Our results have several implications for managers who are trying to design more efficient queueing systems. First, in the presence of a group-based incentive scheme, emphasizing the long-term nature of the interaction among the servers is important. The emphasis on repeated interactions should encourage reputation building and provide room for the threat of future punishment. This implication also suggests that managers face trade-offs when implementing policies regarding rotating human servers among groups. Although regularly rotating workers among teams could limit the collusion (Bandiera, Barankay, and Rasul, 2005), reduce the ratchet effect (Wei, 2020), and promote innovation (Dearden, Ickes, and Samuelson, 1990), it may intensify free-riding and thus impair the teamwork. Therefore, when fast working speed and cooperation among group members are more valued in an organization, managers may place more weight on a fixed group composition than on a fluid one.

Second, when the interactions are long, hiding information about the state of the queue may be beneficial if the manager would like to instill fast processing speeds across all of the states of the queue and improve system performance. For example, we find that when $\delta = \frac{5}{6}$, the system has better performance in steady state (i.e., shorter customer waiting time and less lost demand) when the queue is not visible than when it is visible. Fast processing speed and shorter customer waiting times are critical in industries in which service speed plays an important role in customer satisfaction (e.g., banking industry) (Davis and Maggard, 1990; Kara, Kaynak, and Kucukemiroglu, 1995; Mathe-Soulek, Slevitch, and Dallinger, 2015). Because customer satisfaction also correlates with customer loyalty and profitability (Hallowell, 1996; Kandampully and Suhartanto, 2000), lower effort with few customers in line might not be in a company's interest.

Our paper opens many exciting avenues for future research on understating the behavior of servers and customers on both the theoretical and experimental fronts. First and foremost, in this paper, we focused on the strategic implications of repeated interactions among servers. Previous work has shown that the visibility of the queue may also influence customers' decisions to join/leave the queue (for a review of the literature that considers the impact of information about the queue on customers' decisions and the resulting system properties, see Chapter 3 of Hassin, 2016). Extending the equilibrium analysis to include strategic customers or scheduling policies would be of great importance. Second, we analyzed the case of discrete effort levels, discrete states, and discrete timelines. Given recent advances in running (near-) continuous-time experiments (e.g., Friedman and Oprea, 2012), considering a similar setting with continuous variables along each of those dimensions would be interesting. Third, we considered a case of identical customers and servers, but introducing heterogeneity in worker ability and customer orders (and thus an asymmetry in the dynamic game) would add more realism to the environment. Fourth, we focused on the case of an indefinite duration of the interaction. Understanding the implications of the finite horizon in this setting would be important. Lastly, the extent to which communication among servers and different matching mechanisms (such as the one studied in Honhon and Hyndman, 2020) can improve effort provision in the queueing setting would be of great interest.

References

Allon, G., and M. Kremer (2018): "Behavioral foundations of queueing systems," *The Handbook of Behavioral Operations*, pp. 325–359.

Armony, M., G. Roels, and H. Song (2021): "Pooling Queues with Strategic Servers: The Effects of Customer Ownership," *Operations Research*, 69(1), 13–29.

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social Preferences and the Response to Incentives: Evidence from Personnel Data*," *The Quarterly Journal of Economics*, 120(3), 917–962.

Beer, R., H.-S. Ahn, and S. Leider (2018): "Can trustworthiness in a supply chain be signaled?," *Management science*, 64(9), 3974–3994.

Belcastro, L., F. Marozzo, D. Talia, and P. Trunfio (2016): "Using Scalable Data Mining for Predicting Flight Delays," *ACM Trans. Intell. Syst. Technol.*, 8(1), 1–20.

Bell, B., and J. Van Reenen (2014): "Bankers and Their Bonuses," *The Economic Journal*, 124(574), 1–21.

Bendoly, E., R. Croson, P. Goncalves, and K. Schultz (2010): "Bodies of knowledge for research in behavioral operations," *Production and Operations Management*, 19(4), 434–452.

Blakemore, A. E., S. A. Low, and M. B. Ormiston (1987): "Employment Bonuses and Labor Turnover," *Journal of Labor Economics*, 5(4, Part 2), S124–S135.

Blonski, M., P. Ockenfels, and G. Spagnolo (2011): "Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence," *American Economic Journal: Microeconomics*, 3(3), 164–92.

Blonski, M., and G. Spagnolo (2015): "Prisoners' other Dilemma," *International Journal of Game Theory*, 44(1), 61–81.

Buell, R. W., T. Kim, and C.-J. Tsay (2017): "Creating Reciprocal Value Through Operational Transparency," *Management Science*, 63(6), 1673–1695.

Cason, T., and V.-L. Mui (2019): "Individual versus group choices of repeated game strategies: A strategy method approach," *Games and Economic Behavior*, 114, 128–145.

Cason, T. N., S.-H. P. Lau, and V.-L. Mui (2019): "Prior interaction, identity, and cooperation in the Inter-group Prisoner's Dilemma," *Journal of Economic Behavior & Organization*, 166, 613–629.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Clark, A. E., D. Masclet, and M. C. Villeval (2010): "Effort and Comparison Income: Experimental and Survey Evidence," *ILR Review*, 63(3), 407–426.

Cobham, A. (1954): "Priority assignment in waiting line problems," *Journal of the Operations Research Society of America*, 2(1), 70–76.

Cox, J. C., and R. L. Oaxaca (1989): "Laboratory Experiments with a Finite-Horizon Job-Search Model," *Journal of Risk and Uncertainty*, 2(3), 301–329.

Dal Bó, P., and G. R. Fréchette (2011): "The evolution of cooperation in infinitely repeated games: Experimental evidence," *The American Economic Review*, 101(1), 411–429.

——— (2018): "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," *Journal of Economic Literature*, 56(1), 60–114.

——— (2019): "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma," *American Economic Review*, 109(11), 3929–52.

Dal Bó, P., and G. R. Fréchette (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101(1), 411–29.

Davis, A. M., and K. Hyndman (2018): "An experimental investigation of managing quality through monetary and relational incentives," *Management Science*, 64(5), 2345–2365.

Davis, M. M., and M. J. Maggard (1990): "An analysis of customer satisfaction with waiting times in a two-stage service process," *Journal of Operations Management*, 9(3), 324 – 334.

Dearden, J., B. W. Ickes, and L. Samuelson (1990): "To Innovate or Not to Innovate: Incentives and Innovation in Hierarchies," *The American Economic Review*, 80(5), 1105–1124.

Delasay, M., A. Ingolfsson, B. Kolfal, and K. Schultz (2019): "Load effect on service times," *European Journal of Operational Research*, 279(3), 673–686.

Embrey, M., G. R. Fréchette, and S. Yuksel (2018): "Cooperation in the finitely repeated prisoner's dilemma," *The Quarterly Journal of Economics*, 133(1), 509–551.

Erlang, A. K. (1909): "Sandsynlighedsregning og telefonsamtaler," *Nyt tidsskrift for Matematik*, 20, 33–39.

Friedman, D., and R. Oprea (2012): "A continuous dilemma," *American Economic Review*, 102(1), 337–63.

Fudenberg, D., D. G. Rand, and A. Dreber (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review*, 102(2), 720–49.

Gardner, R., E. Ostrom, and J. M. Walker (1990): "The Nature of Common-Pool Resource Problems," *Rationality and Society*, 2(3), 335–358.

George, J. M., and J. M. Harrison (2001): "Dynamic control of a queue with adjustable service rate," *Operations research*, 49(5), 720–731.

Gill, D., and V. Prowse (2012): "A Structural Analysis of Disappointment Aversion in a Real Effort Competition," *American Economic Review*, 102(1), 469–503.

Good, P. (2013): *Permutation tests: a practical guide to resampling methods for testing hypotheses.* Springer Science & Business Media.

Gopalakrishnan, R., S. Doroudi, A. R. Ward, and A. Wierman (2016): "Routing and Staffing When Servers Are Strategic," *Operations Research*, 64(4), 1033–1050.

Graves, S. C. (1982): "The application of queueing theory to continuous perishable inventory systems," *Management Science*, 28(4), 400–406.

Green, E. J., and R. H. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52(1), 87–100.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125.

Hallowell, R. H. (1996): "The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study," *International Journal of Service Industry Management*, 7, 27–42.

Hamilton, B., J. Nickerson, and H. Owan (2003): "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111, 465–497.

Hashimoto, M. (1979): "Bonus Payments, on-the-Job Training, and Lifetime Employment in Japan," *Journal of Political Economy*, 87(5, Part 1), 1086–1104.

Hassin, R. (2016): *Rational queueing.* Chapman and Hall/CRC.

Hathaway, B. A., E. Kagan, and M. Dada (2022): "The Gatekeeper's Dilemma: 'When Should I Transfer This Customer?'," *Operations Research*, forthcoming.

Honhon, D., and K. Hyndman (2020): "Flexibility and Reputation in Repeated Prisonerâ€™s Dilemma Games," *Management Science*, 66(11), 4998–5014.

Hopp, W. J., S. M. Iravani, and G. Y. Yuen (2007): "Operations systems with discretionary task completion," *Management Science*, 53(1), 61–77.

Hyndman, K., and D. Honhon (2020): “Flexibility in long-term relationships: An experimental study,” *Manufacturing & Service Operations Management*, 22(2), 273–291.

Kandampully, J., and D. Suhartanto (2000): “Customer loyalty in the hotel industry: The role of customer satisfaction and image,” *International Journal of Contemporary Hospitality Management*, 12(6), 346–351.

Kao, E. P., and G. G. Tung (1981): “Bed allocation in a public health care delivery system,” *Management Science*, 27(5), 507–520.

Kara, A., E. Kaynak, and O. Kucukemiroglu (1995): “Marketing strategies for fast-food restaurants: a customer view,” *International journal of contemporary hospitality management*, 7(4), 16–22.

Kc, D. S., and C. Terwiesch (2009): “Impact of workload on service time and patient safety: An econometric analysis of hospital operations,” *Management Science*, 55(9), 1486–1498.

Kendall, D. G. (1951): “Some problems in the theory of queues,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 13(2), 151–173.

Kimes, S. (2009): “Customer attitudes towards restaurant reservations policies,” *Journal of Revenue and Pricing Management*, 10.

Kloosterman, A. (2019): “Cooperation in stochastic games: a prisoner’s dilemma experiment,” *Experimental Economics*, 23, 447–467.

Knez, M., and C. Camerer (2000): “Increasing cooperation in prisoner’s dilemmas by establishing a precedent of efficiency in coordination games,” *Organizational Behavior and Human Decision Processes*, 82(2), 194–216.

Kolmogorov, A. (1931): “Sur le probleme d’attente,” *Mathematicheskiy Sbornik*, 38(1-2), 101–106.

Kremer, M., and F. de Véricourt (2022): “Mismanaging Diagnostic Accuracy Under Congestion,” *Operations Research*, forthcoming.

Little, J. D. (1961): “A proof for the queuing formula: L= λ W,” *Operations research*, 9(3), 383–387.

Mas, A., and E. Moretti (2009): “Peers at Work,” *American Economic Review*, 99(1), 112–45.

Mathe-Soulek, K., L. Slevitch, and I. Dallinger (2015): “Applying mixed methods to identify what drives quick service restaurant’s customer satisfaction at the unit-level,” *International Journal of Hospitality Management*, 50, 46 – 54.

Nosenzo, D., T. Offerman, M. Sefton, and A. van der Veen (2016): “Discretionary sanctions and rewards in the repeated inspection game,” *Management Science*, 62(2), 502–517.

Ortega, J. (2009): “Employee Discretion and Performance Pay,” *The Accounting Review*, 84(2), 589–612.

Peysakhovich, A., and D. G. Rand (2016): “Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory,” *Management Science*, 62(3), 631–647.

Polterovich, V. (1993): “Rationing, queues, and black markets,” *Econometrica*, 61(1), 1–28.

Powell, S. G., and K. L. Schultz (2004): “Throughput in serial lines with state-dependent behavior,” *Management Science*, 50(8), 1095–1105.

Rojas, C. (2012): “The role of demand information and monitoring in tacit collusion,” *The RAND Journal of Economics*, 43(1), 78–109.

Romero, J. (2015): “The effect of hysteresis on equilibrium selection in coordination games,” *Journal of Economic Behavior & Organization*, 111, 88–105.

Romero, J., and Y. Rosokha (2018): “Constructing strategies in the indefinitely repeated prisoner’s dilemma game,” *European Economic Review*, 104, 185 – 219.

——— (2019a): “The Evolution of Cooperation: The Role of Costly Strategy Adjustments,” *American Economic Journal: Microeconomics*, 11(1), 299–328.

——— (2019b): “Mixed Strategies in the Indefinitely Repeated Prisoner’s Dilemma,” *Purdue Working Paper.*

Rotemberg, J. J., and G. Saloner (1986): “A Supergame-Theoretic Model of Price Wars during Booms,” *The American Economic Review*, 76(3), 390–407.

Roth, A. E., and J. K. Murnighan (1978): “Equilibrium behavior and repeated play of the prisoner’s dilemma,” *Journal of Mathematical psychology*, 17(2), 189–198.

Sah, R. K. (1987): “Queues, rations, and market: comparisons of outcomes for the poor and the rich,” *The American Economic Review*, 77(1), 69–77.

Salz, T., and E. Vespa (2020): “Estimating dynamic games of oligopolistic competition: an experimental investigation,” *The RAND Journal of Economics*, 51(2), 447–469.

Schultz, K. L., D. C. Juran, and J. W. Boudreau (1999): “The effects of low inventory on the development of productivity norms,” *Management Science*, 45(12), 1664–1678.

Schultz, K. L., D. C. Juran, J. W. Boudreau, J. O. McClain, and L. J. Thomas (1998): “Modeling and Worker Motivation in JIT Production Systems,” *Management Science*, 44(12-part-1), 1595–1607.

Shunko, M., J. Niederhoff, and Y. Rosokha (2018): “Humans are not machines: The behavioral impact of queueing design on service time,” *Management Science*, 64(1), 453–473.

Staats, B., and F. Gino (2012): “Specialization and Variety in Repetitive Tasks: Evidence from a Japanese Bank,” *Management Science*, 58, 1141–1159.

Sternberg, A., J. Soares, D. Carvalho, and E. Ogasawara (2017): “A Review on Flight Delay Prediction,” *Working Paper.*

Tan, T. F., and S. Netessine (2019): “When You Work with a Superman, Will You Also Fly? An Empirical Study of the Impact of Coworkers on Performance,” *Management Science*, 65(8), 3495–3517.

Vespa, E. (2020): “An Experimental Investigation of Cooperation in the Dynamic Common Pool Game,” *International Economic Review*, 61(1), 417–440.

Vespa, E., and A. J. Wilson (2015): “Experimenting with Equilibrium Selection in Dynamic Games,” *Working Paper.*

Vespa, E., and A. J. Wilson (2019): “Experimenting with the transition rule in dynamic games,” *Quantitative Economics*, 10(4), 1825–1849.

Walker, J. M., R. Gardner, and E. Ostrom (1990): “Rent dissipation in a limited-access common-pool resource: Experimental evidence,” *Journal of Environmental Economics and Management*, 19(3), 203 – 211.

Wei, C. (2020): “Can job rotation eliminate the Ratchet effect: Experimental evidence,” *Journal of Economic Behavior & Organization*, 180, 66 – 84.

Wergin, N.-E. (2003): “Teamwork in the Automobile Industry - an Anglo-German Comparison,” *European Political Economy Review*, 1, 152–190.

Zhan, D., and A. Ward (2019): “Staffing, Routing, and Payment to Trade off Speed and Quality in Large Service Systems,” *Operations Research*, 67(6), 1738–1751.

Zhan, D., and A. R. Ward (2018): “The M/M/1+ M queue with a utility-maximizing server,” *Operations Research Letters*, 46(5), 518–522.

Appendices

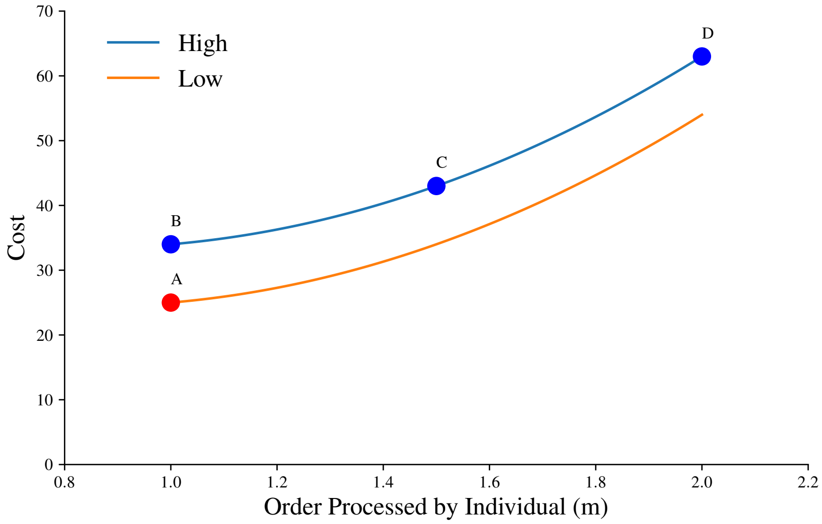
A Study 1: Additional Details

A.1 Microfoundations

A.1.1 Cost Function

As described in the first paragraph of section 4, the cost of processing $m_i(.)$ orders with low effort is $c(l, e_j, \theta) = am_i(l, e_j, \theta)^2 + bm_i(l, e_j, \theta) + c$ with $a = 22$, $b = -37$, and $c = 40$; and the cost of processing $m_i(.)$ orders with high effort is $c(h, e_j, \theta) = xm_i(h, e_j, \theta)^2 + ym_i(h, e_j, \theta) + z$ with $x = 22$, $y = -37$, and $z = 49$. Figure A1 presents the two cost functions in the same graph given $m_i(.)$ ranges between 1 and 2. Figure A1 also labels all possible costs that could realize in the setup used for the experiment. In particular, point A in the figure corresponds to the cost of providing low effort. Note that for the setup used in study 1, the cost of providing low effort is independent of the effort provided by the other server. Points B, C, and D correspond to the cost of providing high effort. The cost varies depending on the state and the actions of the other server. For example, if the state is 2, the server will only process one order, whereas if the state is 4, the server will process two orders.

Figure A1: Cost Function



Notes: **A:** (l,l,2), (l,h,2), (l,l,3), (l,h,3), (l,l,4), and (l,h,4); **B:** (h,h,2) and (h,l,2); **C:** (h,h,3); **D:** (h,l,3), (h,h,4), and (h,l,4). The first item in the tuple (e_i,e_j,θ) represents the server's own effort; the second item represents the other server's effort; the last item represents the state of the queue (i.e., the number of orders in the queue).

A.1.2 Compensation Function

As described in the second paragraph of section 4, an individual server is compensated based on the group performance. Specifically, if the group processes $M(.)$ orders, the compensation function is $r(e_i, e_j, \theta) = kM(e_i, e_j, \theta) + \mathbf{1}_{M(e_i,e_j,\theta)=4} bonus$ with $k = 25$ and $bonus = 11$.

A.1.3 Examples

Here, we provide two examples of how the cost and compensation function maps the payoff matrices.

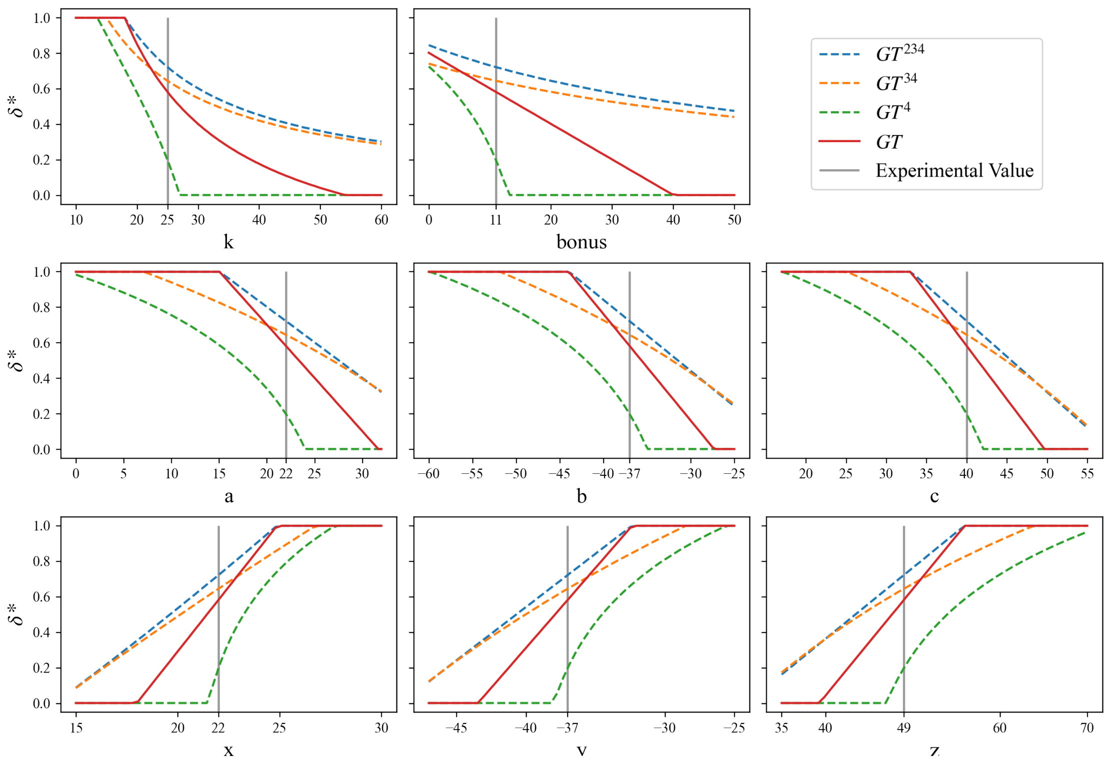
- **Example 1**: Suppose 2 tasks are in the queue in a given period. If a server chooses high effort and her teammate also chooses high effort, both tasks are processed (each server processes 1 task). This scenario leads to the server's cost 34(=22×1^2-37×1+49) and the revenue 50(=25×2). Therefore, this server's payoff is 16(=50-34).
- **Example 2**: Suppose 4 tasks are in the queue in a given period. If a server chooses low effort and her teammate chooses high effort, the team is able to process 3 tasks. This scenario leads to the server's cost 25(=22×1^2-37×1+40). The revenue is 75(=25×3). Therefore, this server's payoff is 50(=75-25).

A.1.4 Comparative Statics

Combined, the compensation function and the cost function have 8 parameters. Figure A2 presents comparative statics of the critical thresholds of four GT strategies for each of the eight parameters. In particular, for each subfigure, we vary only one parameter and hold all other parameters constant at the experimental value. For example, in the top-right figure, we study the relationship between the critical thresholds of the

GT (solid red), GT^4 (dashed green), GT^{34} (dashed orange), and GT^{234} (dashed blue) and the value of k while keeping $bonus = 11, a = 22, b = -37, c = 40, x = 22, y = -37, z = 49$.

Figure A2: Comparative Statics and Parameters



Notes: For each graph, we hold the other parameters constant as the experimental value when varying a given parameter. The experimental values are $k = 25, bonus = 11, a = 22, b = -37, c = 40, H = 1, d = 9$.

The eight graphs in Figure A2 share two common patterns. The first pattern is that $\delta_v^*(GT^4) \leq \delta_v^*(GT^{34}) \leq \delta_v^*(GT^{234})$ when $\delta_v^*(GT^4) > 0$.[20] The interpretation of this result is that when the queue is visible, sustaining cooperation is easier when the queue is long than when it is short. The second pattern is that $\delta_v^*(GT^4) \leq \delta_{nv}^*(GT) \leq \delta_v^*(GT^{234})$. In other words, although sustaining high effort is easier across all states when the queue is not visible, sustaining high effort is easier across a subset of the states (e.g., $\theta = 4$) when the queue is visible. These two common patterns are consistent with our hypotheses discussed in section 4.2.

A.1.5 Theoretical Predictions

To determine whether GT is an equilibrium strategy, we first find the transition-probability matrix implied by the strategy profile s^{GT}. In particular, if both players provide high effort, the transition-probability

---

[20] $\delta_v^*(GT^4)$=0 indicates the incentives to provide the high effort are large. If the incentives to provide the high effort are large enough (or the cost to provide the high effort is small enough), servers may play a PD game when the state is 2. Therefore, it is possible that $\delta_v^*(GT^{34}) > \delta_v^*(GT^{234})$.

matrix is given by $\mathcal{P}^{hh}$; if one player deviates from high effort, the transition-probability matrix is given by $\mathcal{P}^{lh}$; and if both players provide low effort, the transition-probability matrix is given by $\mathcal{P}^{ll}$:

$$\mathcal{P}^{hh} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 1/3 & 1/3 & 1/3 \\ 4 & 1/3 & 1/3 & 1/3 \end{array} \qquad \mathcal{P}^{lh} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 1/3 & 1/3 & 1/3 \\ 4 & 0 & 1/3 & 2/3 \end{array} \qquad \mathcal{P}^{ll} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 0 & 1/3 & 2/3 \\ 4 & 0 & 0 & 1 \end{array}. \tag{10}$$

Thus, if both players provide high effort, they process all of the customer orders, and therefore, the transition probability $\mathcal{P}^{hh}$ is determined by the arrival process (i.e., uniform distribution). However, if one or both players provide low effort, some of the states will have leftover customer orders, which, together with the arrival process, implies that a transition to states with more customers is more likely. Vectors u^c, u^{dev}, and u^d specify payoffs obtained in each of the states:

$$u^c = \begin{pmatrix} u^c_2 \\ u^c_3 \\ u^c_4 \end{pmatrix} = \begin{pmatrix} 16 \\ 32 \\ 48 \end{pmatrix} \qquad u^{dev} = \begin{pmatrix} u^{dev}_2 \\ u^{dev}_3 \\ u^{dev}_4 \end{pmatrix} = \begin{pmatrix} 25 \\ 50 \\ 50 \end{pmatrix} \qquad u^d = \begin{pmatrix} u^d_2 \\ u^d_3 \\ u^d_4 \end{pmatrix} = \begin{pmatrix} 25 \\ 25 \\ 25 \end{pmatrix}. \tag{11}$$

Lastly, the total values for the three cases in matrix notation are

$$V^c = [I - \delta\mathcal{P}^{hh}]^{-1}u^c, \qquad V^{dev} = u^{dev} + \delta\mathcal{P}^{lh}V^d, and \qquad V^d = [I - \delta\mathcal{P}^{ll}]^{-1}u^d, \tag{12}$$

where $V^c = \left(\begin{smallmatrix} V^c_2 \\ V^c_3 \\ V^c_4 \end{smallmatrix}\right)$, $V^d = \left(\begin{smallmatrix} V^d_2 \\ V^d_3 \\ V^d_4 \end{smallmatrix}\right)$, and I is the identity matrix. To show that s^{GT} is an SPE when the queue is visible, we need to find δ so that each element of V^c is at least as large as the corresponding element of V^{dev}. We find that s^{GT} is an SPE when δ is at least 0.72. We denote this critical threshold as $\delta^*_v(GT)$.

Note GT does not distinguish between the states. However, we expect that a human participant would. Therefore, we consider two trigger strategies that do. In particular, the first strategy, which we term GT^{34}, plays GT across states 3 and 4 and always provides low effort in state 2. The only difference in the analysis above is that $u^c = \left(\begin{smallmatrix} 25 \\ 32 \\ 48 \end{smallmatrix}\right)$, which leads to $\delta^*_v(GT^{34}) = 0.64$.

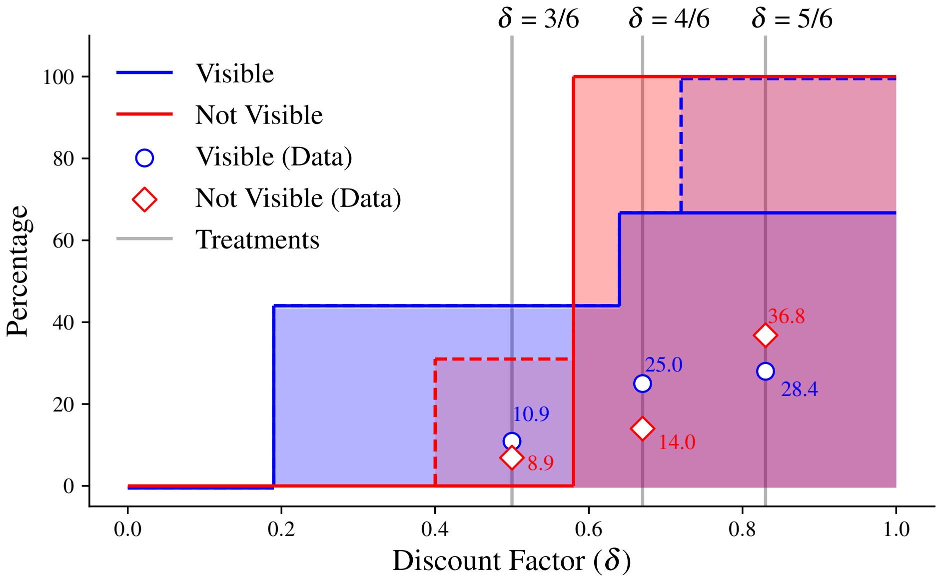
The second state- and history-contingent strategy, which we term GT^4, plays GT in state 4 only and provides low effort in both states 2 and 3. The implied transition-probability matrices and the payoff vectors for this strategy are

$$\mathcal{P}^{c} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 0 & 1/3 & 2/3 \\ 4 & 1/3 & 1/3 & 1/3 \end{array} \qquad \mathcal{P}^{dev} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 0 & 1/3 & 2/3 \\ 4 & 0 & 1/3 & 2/3 \end{array} \qquad \mathcal{P}^{ll} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 0 & 1/3 & 2/3 \\ 4 & 0 & 0 & 1 \end{array} \tag{13}$$

$$u^c = \begin{pmatrix} 25 \\ 25 \\ 48 \end{pmatrix} \qquad u^{dev} = \begin{pmatrix} 25 \\ 25 \\ 50 \end{pmatrix} \qquad u^d = \begin{pmatrix} 25 \\ 25 \\ 25 \end{pmatrix}. \tag{14}$$

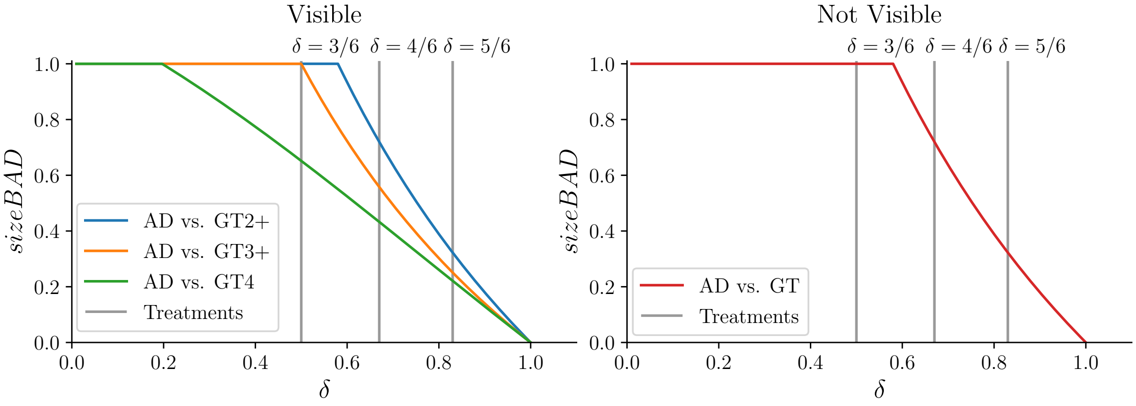
Note that we chose to label the transition probabilities as $\mathcal{P}^c$ instead of $\mathcal{P}^{hh}$ and $\mathcal{P}^{dev}$ instead of $\mathcal{P}^{lh}$, because the cooperative path of s^{GT^4} involves low effort in states 2 and 3. Solving for $\delta^*_v(GT^4)$, we get 0.19.

Figure A3: High Effort, Discount Factor, and Data



Notes: Data are the percentage of high effort chosen by subjects across all periods in the second half of matches. Servers are assumed to use GT-type strategies in this figure ($\delta^*_{nv}(GT) = 0.58$; $\delta^*_v(GT^4) = 0.19$; $\delta^*_v(GT^{34}) = 0.64$ and $\delta^*_v(GT) = 0.72$). Even GT is an equilibrium strategy when the queue is visible and $\delta > 0.72$, we assume servers provide low effort when the state is 2, because it is the efficient choice. $D.AlT^4$ is also an SPE when the queue is not visible and δ is large enough; it does not change the main hypotheses ($\delta^*_{nv}(D.AlT^4)$=0.40).

Figure A4: SizeBAD and Discount Factor



Notes: $SizeBAD$ represents the limit of the basin of attraction of Always Defect (AD) versus cooperative strategies such as Grim Trigger. When a subject is considering which of these two types of strategies is more profitable in expectation, she might believe that the other subject will choose the cooperative strategy. If this belief is less than $SizeBAD$, the subject will maximize the expected payoff by choosing AD (Dal Bó and Fréchette, 2018). The larger $SizeBAD$ is, the more likely she is to choose AD. Dal Bó and Fréchette (2011) show that cooperation rates decrease in $sizeBAD$ in their experiment, especially when cooperation is risk dominant.

A.2 Experimental Design

A.2.1 Match-Length Realizations

Table A1: Match Lengths

(a) $\delta = \frac{3}{6}$

Supergame Number:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Realization #1:	4	2	1	5	1	1	1	1	2	1	1	1	3	1	1	9	5	3	1	2	2	2	2	7	1
Realization #2:	2	2	1	4	3	1	1	1	2	1	1	1	2	1	1	1	1	2	1	1	4	1	2	2	1
Realization #3:	1	1	2	2	2	2	2	1	4	2	2	1	3	2	2	1	1	3	3	6	1	1	2	2	1
Realization #4:	2	2	4	1	1	2	3	2	2	4	1	1	2	3	1	2	2	1	1	3	1	3	3	2	1
Supergame Number:	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
Realization #1:	4	1	1	1	3	1	1	2	1	1	2	4	1	2	2	6	3	1	6	12	2	2	2	2	1
Realization #2:	2	1	2	1	1	2	2	2	1	7	1	1	1	1	2	3	1	1	1	1	5	3	3	5	5
Realization #3:	3	2	2	2	1	1	1	1	1	1	1	1	2	1	1	1	1	2	2	2	3	2	1	4	1
Realization #4:	2	2	1	5	1	2	4	3	1	1	1	1	1	2	2	2	1	2	1	1	1	2	1	2	2
Supergame Number:	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
Realization #1:	1	1	2	3	7	1	1	3	1	1	1	1	5	2	2	2	1	1	4	2	1	2	1	2	6
Realization #2:	1	1	4	2	13	2	1	1	1	2	1	2	3	2	1	3	1	1	2	1	2	2	2	4	1
Realization #3:	1	4	1	1	4	2	1	1	1	2	1	3	7	7	1	1	1	2	1	1	5	3	3	4	1
Realization #4:	2	4	1	6	3	2	2	1	1	1	4	4	2	1	2	2	2	2	2	2	2	4	1	1	3
Supergame Number:	**76**	**77**	**78**	**79**	**80**																				
Realization #1:	1	1	1	1	2																				
Realization #2:	1	1	1	3	1																				
Realization #3:	1	4	3	1	3																				
Realization #4:	2	1	4	1	1																				

(b) $\delta = \frac{4}{6}$

Supergame Number:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Realization #1:	4	3	7	1	1	5	3	1	1	9	8	1	2	4	10	4	1	1	1	4	1	4	2	4	1
Realization #2:	1	2	1	2	6	2	1	1	2	2	2	5	3	1	1	4	1	1	2	2	2	3	1	4	6
Realization #3:	3	2	1	3	1	1	2	1	4	2	4	1	1	1	4	3	2	4	2	2	7	3	3	2	3
Realization #4:	2	4	7	7	1	1	1	3	1	5	3	3	5	1	7	1	3	2	2	5	6	8	1	1	5
Supergame Number:	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
Realization #1:	2	8	3	7	12	4	2	2	1	2	2	3	7	1	1	3	2	1	1	9	2	1	1	4	2
Realization #2:	2	1	2	2	2	2	2	9	1	1	3	3	1	1	7	3	3	5	5	1	1	4	15	4	6
Realization #3:	1	3	3	6	1	6	3	2	2	2	1	1	4	1	3	2	1	1	6	3	3	6	4	2	9
Realization #4:	1	2	3	1	3	4	5	1	2	2	1	5	3	7	1	2	4	4	1	2	1	4	3	4	4

(c) $\delta = \frac{5}{6}$

Supergame Number:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Realization #1:	4	12	30	4	15	1	10	2	5	2	11	19	6	2	1	2	5	7	2	3	2	2	9	8	6
Realization #2:	8	3	15	1	11	10	6	2	3	2	4	16	4	1	13	11	20	4	6	9	2	3	4	8	1
Realization #3:	7	6	4	7	8	3	1	3	9	1	6	3	2	2	2	38	22	7	2	1	4	5	3	3	6
Realization #4:	11	15	8	1	7	6	3	4	5	6	16	2	4	4	3	1	7	8	12	2	1	10	11	6	4

A.2.2 Screenshots

Queue-Visible-Treatment Decision Screen

Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

Round	1	2
Number of New Tasks	2	3
Table #	2	3
My Choice	2	?
Other's Choice	1	
My Payoff	16	
Other's Payoff	25	
Dice Roll	6	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 7 or greater.

Please select your choice for Round **2** of Match #2

Queue-Visible-Treatment Waiting Screen

Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

Round	1	2
Number of New Tasks	2	3
Table #	2	3
My Choice	2	1
Other's Choice	1	
My Payoff	16	
Other's Payoff	25	
Dice Roll	6	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 7 or greater.

Please wait while the other participant makes a choice.

Queue-Not-Visible-Treatment Decision Screen

Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

Round	1	2
Number of New Tasks	4	
Table #	4	
My Choice	2	?
Other's Choice	1	
My Payoff	12	
Other's Payoff	50	
Dice Roll	1	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Please select your choice for Round **2** of Match #1

Queue-Not-Visible-Treatment Waiting Screen

Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

Round	1	2
Number of New Tasks	4	
Table #	4	
My Choice	2	1
Other's Choice	1	
My Payoff	12	
Other's Payoff	50	
Dice Roll	1	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Please wait while the other participant makes a choice.

A.2.3 Instructions

Experiment Overview

Today's experiment will last about 60 minutes.

You will be paid a show-up fee of $5 together with any money you accumulate during this experiment. The amount of money you accumulate will depend partly on your actions and partly on the actions of other participants. This money will be paid at the end of the experiment in private and in cash.

It is important that during the experiment you remain **silent**. If you have a question or need assistance of any kind, please **raise your hand, but do not speak** - and an experiment administrator will come to you, and you may then whisper your question.

In addition, please **turn off your cell phones and put them away now**.

Anybody that breaks these rules will be asked to leave.

Agenda

1. Instructions
2. Quiz
3. Experiment

How Matches Work

The experiment is made up of **80 matches**.

At the start of each match you will be randomly paired with another participant in this room.

You will then play a number of rounds with that participant (this is what we call a "match").

Each match will last for a random number of **rounds**:

- At the end of each round the computer will roll a twelve-sided fair dice.
- If the computer rolls a number less than 7, then the **match continues** for at least one more round (**50% probability**).
- If the computer rolls a 7 or greater, then the **match ends** (**50% probability**).

To test this procedure, click 'Test' button below. You will need to test this procedure 10 times.

Choices and Payoffs

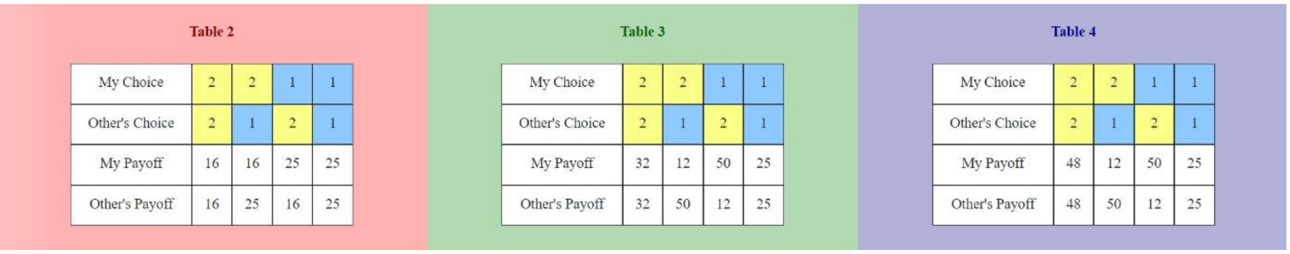


Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

In each round of a match, you will choose whether to complete **1** or **2** tasks. The participant you are paired with will also choose whether to complete **1** or **2** tasks.

In each round of a match, your payoff will be according to **one** of the three tables (labeled **Table 2**, **Table 3**, and **Table 4**). Each table presents payoffs from the four pairs of choices that are possible. These payoffs are in **points**.

The **Table #** is determined based on the number of total tasks available in that round. Thus, when there are 2 tasks available, the payoff is based on **Table 2**; when there are 3 tasks available, the payoff is based on **Table 3**; and when there are 4 tasks available, the payoff is based on **Table 4**.

For example, if you choose **2** and the participant you are paired with chooses **2** and if the payoff

- is according to **Table 2**, then your payoff for the round will be 16 points, and the other's payoff will be 16 points.
- is according to **Table 3**, then your payoff for the round will be 32 points, and the other's payoff will be 32 points.
- is according to **Table 4**, then your payoff for the round will be 48 points, and the other's payoff will be 48 points.

At the end of the experiment, your total points will be converted into cash at the exchange rate of 250 points = $1.

Which Table Will be Used

In each round, a random number of new tasks will become available. This number will be drawn at random from a set of numbers $\{2, 3, 4\}$, with each number equally likely. We will refer to this random number as the **Number of New Tasks**.

To determine the **Table #** in a round, we will use the **Number of New Tasks** together with any leftover tasks from the previous round as follows:

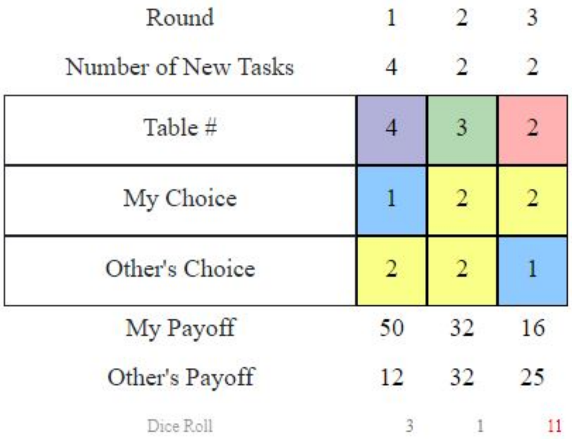
- In Round 1, there are no previous rounds and, therefore, **Table #** will be equal to the **Number of New Tasks**.
- In Round > 1, **Table #** will be determined in two steps

  - First, we will determine the **Number of Leftover Tasks** from the previous round. Notice that if (**Table #** in the previous round) is less than the sum of (My Choice in the previous round) and (Other's Choice in the previous round) then there will be no leftover tasks and, therefore, **Number of Leftover Tasks** will be equal to 0.
  - Second, we will determine the **Table #** in the current round by adding the **Number of Leftover Tasks** from the previous round to the **Number of New Tasks** in the current round. Importantly, the number of tasks available in each round could be at most 4, so any tasks beyond 4 will be discarded.

For example:

- Suppose that in Round 1 the **Number of New Tasks** is randomly drawn to be **4**, then the payoff in Round 1 will be determined by **Table 4**.
- If you choose to complete **1** task while the participant you are paired with chooses to complete **2** tasks, then your payoff for Round 1 will be 50 points, and the other's payoff will be 12 points.
- Suppose that in Round 2 the **Number of New Tasks** is **2**, then your payoff in Round 2 will be determined by **Table 3**.
  - Specifically, we first determine that the **Number of Leftover Tasks** from the first round is **1**(=4-[1+2]). Second, we add the **Number of Leftover Tasks** to the **Number of New Tasks** and determine that **Table #** for the second round is **3** (=**1**+**2**).

How History Will be Recorded



Round	1	2	3
Number of New Tasks	4	2	2
Table #	4	3	2
My Choice	1	2	2
Other's Choice	2	2	1
My Payoff	50	32	16
Other's Payoff	12	32	25
Dice Roll	3	1	11

The history of all variables will be recorded in a history table like the one presented above. In this table you can see an example history of a match in which the computer picked actions at random. The recorded variables include:

- Round –– round number.
- Number of New Tasks –– a random draw in that round (one number is drawn from $\{2, 3, 4\}$ with each number is equally likely).
- Table –– table that is used to determined the payoffs for that round (either **Table 2**, **Table 3**, or **Table 4** depending on the number of tasks available in that round).

- My choice −− your choice (either **1** or **2**).
- Other's Choice −− the choice by the participant that you are paired with (either **1** or **2**).
- My Payoff −− your payoff in that round.
- Other's Payoff −− payoff of the participant that you are paired with.

Reminder, your earnings will be the sum of your points across all matches converted into cash at the exchange rate of 250 points = $1. In addition, you will be paid your show-up fee of $5.

Quiz

Next, there will be a quiz with 10 questions.

You have to answer each question correctly in order to proceed to the next question.

If you answer a question incorrectly, you will see a hint. At that point you will have an opportunity to answer again.

Throughout the quiz, you may refer to the printed instructions.

Matches 1–80

During today's experiment, the **Number of New Tasks** will be randomly drawn **after** you and the participant with whom you are matched make decisions.

This means that in each round, you and the participant with whom you are matched make decisions **without knowing the Number of New Tasks** for that round.

The above instructions were used for the no visibility and $\delta = .5$ *treatment. The number of matches, probability of continuation, and the information about the timing of the decisions relative to the revelation of the Number of New Tasks were adjusted for each treatment.*

A.2.4 Quiz

The top third of each screen contains the three payoff tables

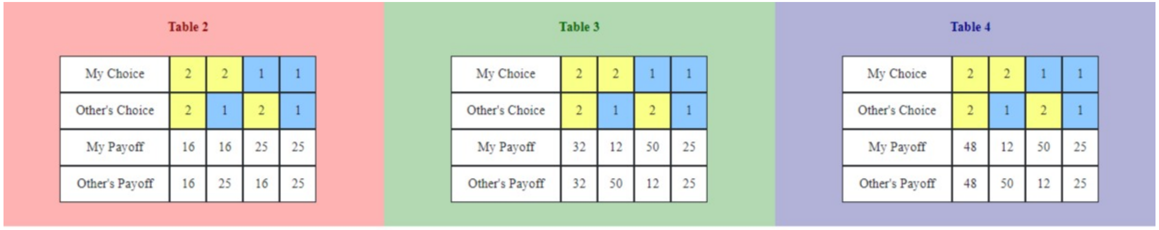


Table 2

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	16	16	25	25
Other's Payoff	16	25	16	25

Table 3

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	32	12	50	25
Other's Payoff	32	50	12	25

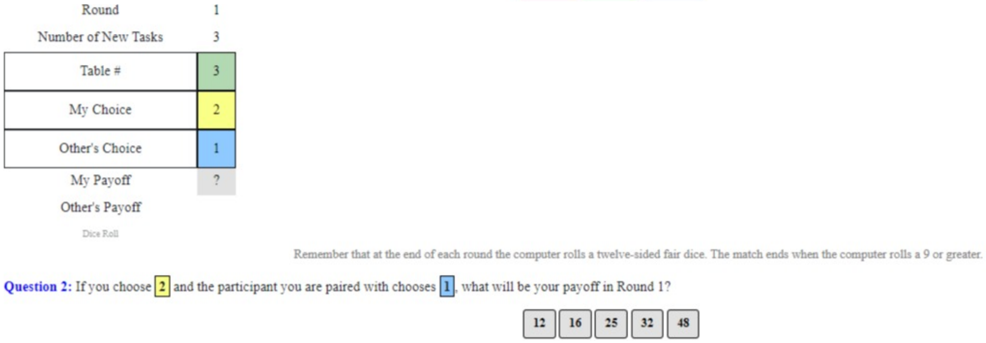
Table 4

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	48	12	50	25
Other's Payoff	48	50	12	25

Round	1
Number of New Tasks	3
Table #	?
My Choice	
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 1: If the **Number of New Tasks** is 3, what **Table #** will be used to determine payoffs in Round 1?



Round	1
Number of New Tasks	3
Table #	3
My Choice	2
Other's Choice	1
My Payoff	?
Other's Payoff	
Dice Roll	

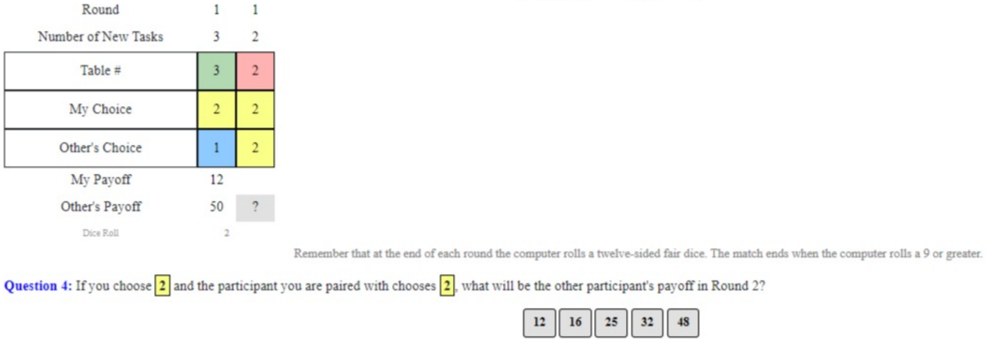
Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 2: If you choose 2 and the participant you are paired with chooses 1, what will be your payoff in Round 1?

Round	1	2
Number of New Tasks	3	2
Table #	3	?
My Choice	2	
Other's Choice	1	
My Payoff	12	
Other's Payoff	50	
Dice Roll	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

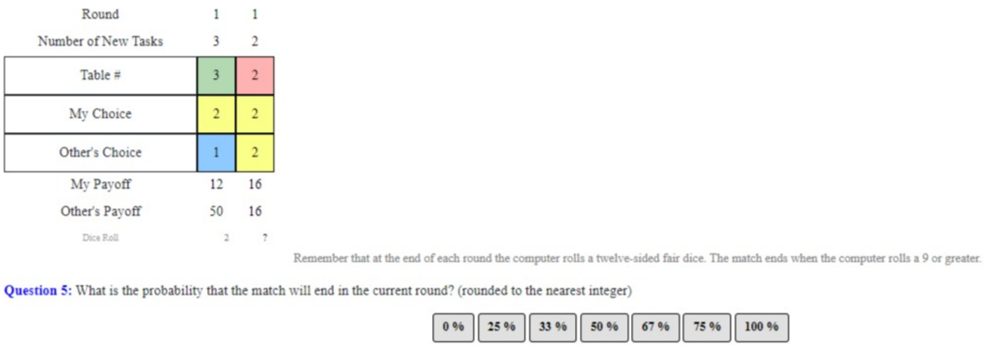
Question 3: If the **Number of New Tasks** is 2, what **Table #** will be used to determine payoffs in Round 2?



Round	1	1
Number of New Tasks	3	2
Table #	3	2
My Choice	2	2
Other's Choice	1	2
My Payoff	12	
Other's Payoff	50	?
Dice Roll	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 4: If you choose 2 and the participant you are paired with chooses 2, what will be the other participant's payoff in Round 2?



Round	1	1
Number of New Tasks	3	2
Table #	3	2
My Choice	2	2
Other's Choice	1	2
My Payoff	12	16
Other's Payoff	50	16
Dice Roll	2	?

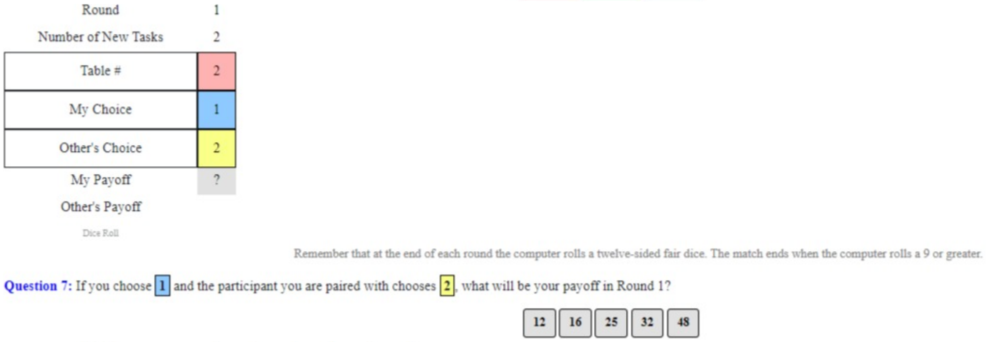
Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 5: What is the probability that the match will end in the current round? (rounded to the nearest integer)

Round	1
Number of New Tasks	2
Table #	?
My Choice	
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

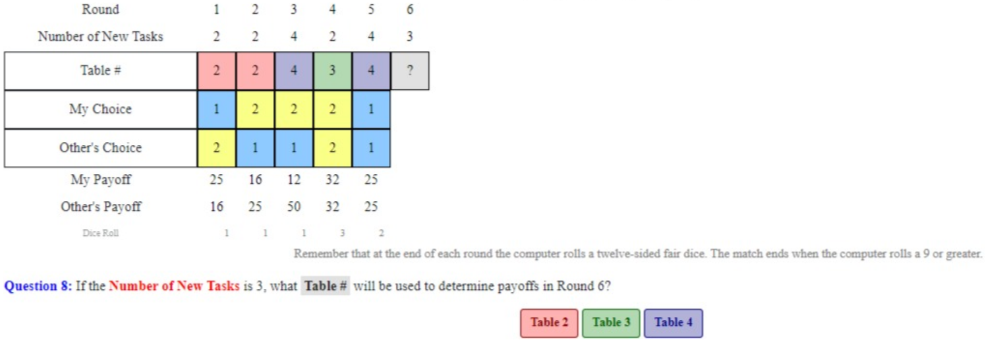
Question 6: If the **Number of New Tasks** is 2, what **Table #** will be used to determine payoffs in Round 1?



Round	1
Number of New Tasks	2
Table #	2
My Choice	1
Other's Choice	2
My Payoff	?
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

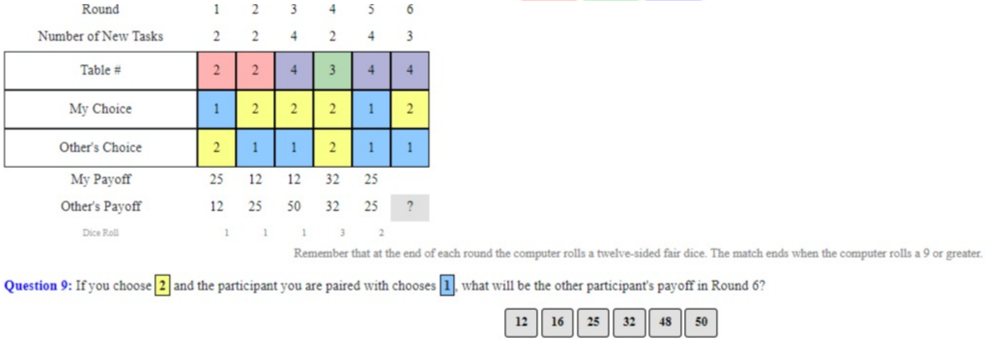
Question 7: If you choose **1** and the participant you are paired with chooses **2**, what will be your payoff in Round 1?



Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
Table #	2	2	4	3	4	?
My Choice	1	2	2	2	1	
Other's Choice	2	1	1	2	1	
My Payoff	25	16	12	32	25	
Other's Payoff	16	25	50	32	25	
Dice Roll	1	1	1	3	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

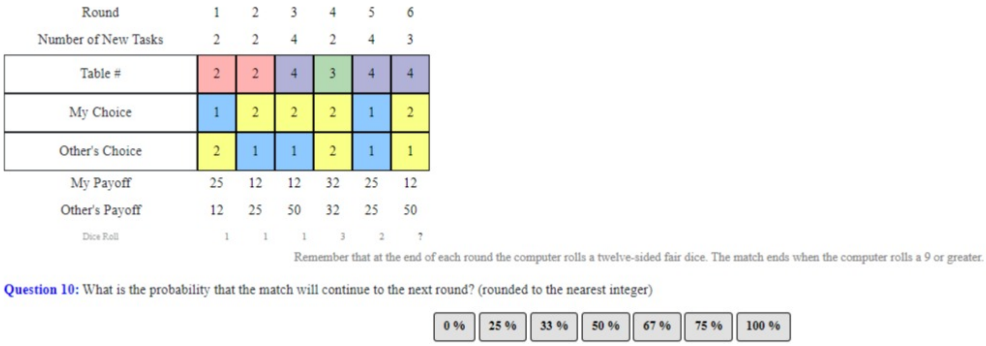
Question 8: If the **Number of New Tasks** is 3, what **Table #** will be used to determine payoffs in Round 6?



Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
Table #	2	2	4	3	4	4
My Choice	1	2	2	2	1	2
Other's Choice	2	1	1	2	1	1
My Payoff	25	12	12	32	25	
Other's Payoff	12	25	50	32	25	?
Dice Roll	1	1	1	3	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 9: If you choose **2** and the participant you are paired with chooses **1**, what will be the other participant's payoff in Round 6?



Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
Table #	2	2	4	3	4	4
My Choice	1	2	2	2	1	2
Other's Choice	2	1	1	2	1	1
My Payoff	25	12	12	32	25	12
Other's Payoff	12	25	50	32	25	50
Dice Roll	1	1	1	3	2	7

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 9 or greater.

Question 10: What is the probability that the match will continue to the next round? (rounded to the nearest integer)

A.3 Additional Tables and Figures

Table A2: Self-reported Strategies and Decisions

	Visibility	$\delta = \frac{3}{6}$	$\delta = \frac{4}{6}$	$\delta = \frac{5}{6}$	Average
Dependence on the previous periods	Yes	52.3	50.0	58.3	53.5
	No	50.0	58.3	56.3	54.9
Difference from the earlier matches	Yes	42.9	54.3	75.0	57.4
	No	50.0	60.4	50.0	53.5
AD strategies	Yes	35.7	21.7	20.8	26.1
	No	56.3	56.3	35.4	49.3
GT-like strategies	Yes	16.7	30.4	25.0	24.0
	No	2.1	6.2	25.0	11.1
TFT-like strategies	Yes	11.9	2.2	8.3	7.5
	No	12.5	8.3	2.1	7.6
DALT-like strategies	Yes	0	0	0	0
	No	0	8.0	2.0	3.3

Notes: The first row reports the percentage of subjects who answered "yes" for the question, "Did your decision in a round depend on what happened in the previous rounds?" The second row reports the percentage of subjects who answered "yes" for the question, "Was your strategy different between the initial matches and the later matches of the experiment?" The next four rows report the answers for the question, "What was your strategy during the experiment? (Please be specific)". The third row shows the percentage of subjects whose description of their strategy is an AD strategy.

Table A3: Summary of Experiment Administration

Treatment		Administration				Demographics		
Visibility	δ	Sessions	Subjects	Matches	Earnings	% Male	% STEM	% US HS
Yes	$\frac{3}{6}$	4	42	80	22.6 (0.3)	50.0 (7.9)	61.9 (7.5)	69.0 (7.4)
Yes	$\frac{4}{6}$	4	46	50	22.9 (0.3)	56.5 (7.2)	65.2 (7.0)	56.5 (7.5)
Yes	$\frac{5}{6}$	4	48	25	24.1 (0.2)	54.2 (7.2)	64.6 (7.0)	75.0 (6.0)
No	$\frac{3}{6}$	4	48	80	22.2 (0.2)	66.7 (6.8)	62.5 (6.9)	68.8 (6.9)
No	$\frac{4}{6}$	4	48	50	20.8 (0.2)	50.0 (7.0)	68.8 (6.6)	72.9 (6.5)
No	$\frac{5}{6}$	4	48	25	22.8 (0.1)	60.4 (7.4)	60.4 (7.0)	66.7 (6.7)

Notes: Earnings are reported in USD and include a $5 show-up fee. Standard errors are in parentheses.

Table A4: Efficiency

Treatment		First Period			All Periods			
Visibility	δ	State 2	State 3	State 4	State 2	State 3	State 4	All States
Yes	$\frac{3}{6}$	98.7 (1.0)	82.2 (2.1)	56.5 (1.4)	98.6 (0.8)	83.4 (1.7)	55.6 (0.8)	66.1 (0.8)
Yes	$\frac{4}{6}$	99.9 (0.1)	87.2 (3.2)	80.4 (2.7)	99.8 (0.1)	85.2 (1.8)	66.5 (1.8)	74.9 (1.5)
Yes	$\frac{5}{6}$	100.0 (0.0)	91.5 (4.3)	79.2 (2.8)	99.6 (0.1)	91.5 (2.1)	64.6 (1.6)	73.3 (1.8)
No	$\frac{3}{6}$	98.7 (0.4)	79.4 (1.2)	53.0 (0.6)	97.8 (0.6)	81.2 (1.1)	55.0 (0.4)	65.0 (0.5)
No	$\frac{4}{6}$	96.5 (1.4)	81.2 (2.1)	55.2 (0.9)	94.4 (1.1)	83.0 (1.5)	56.4 (0.7)	64.7 (0.8)
No	$\frac{5}{6}$	81.8 (2.3)	90.0 (4.1)	68.0 (2.4)	75.8 (1.9)	92.0 (1.7)	61.2 (1.2)	67.7 (1.2)

Table A5: SFEM Estimates – First Half of Matches

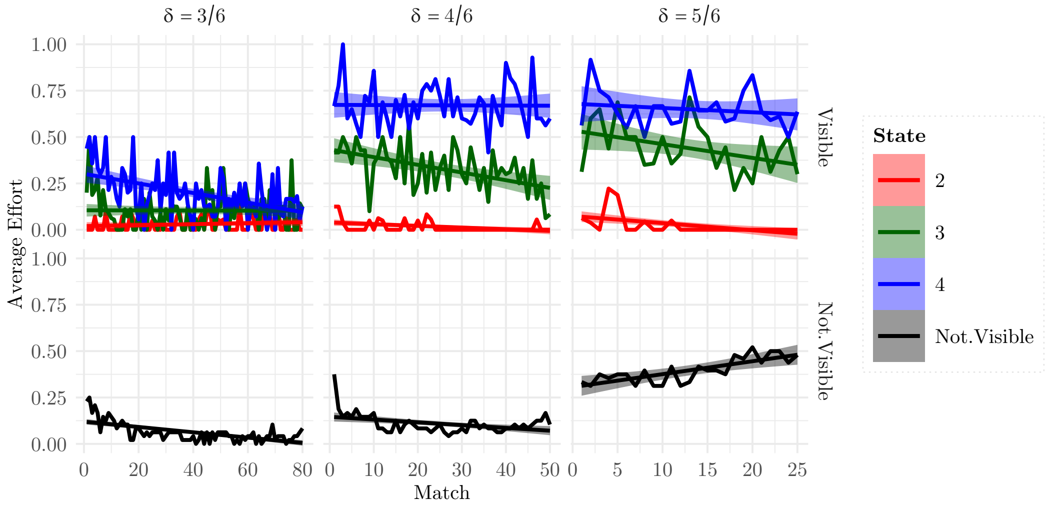
Visibility	δ	AD	AC	TFT	GT	$D.TFT$	$D.AlT$	AC^{34}	TFT^{34}	GT^{34}	$D.TFT^{34}$	$D.AlT^{34}$	AC^{4}	TFT^{4}	GT^{4}	$D.TFT^{4}$	$D.AlT^{4}$	β (%)	$\mathcal{L}$
Yes	$\frac{3}{6}$	63.2							0.2	2.1	20.6			9.5		4.2		91.3	-1026.9
		(8.1)							(1.0)	(2.0)	(7.0)			(5.1)		(5.1)		(1.2)	
Yes	$\frac{4}{6}$	29.0			1.9				13.5	15.0			16.4	15.8	4.8	3.6		87.7	-1325.0
		(6.6)			(2.0)				(6.0)	(6.5)			(6.2)	(8.5)	(6.5)	(3.5)		(1.2)	
Yes	$\frac{5}{6}$	30.0			4.1				33.1	3.4	2.0		2.1	18.6	5.0	1.6		86.3	-1768.8
		(6.7)			(3.2)				(7.4)	(2.7)	(1.9)		(2.5)	(6.2)	(4.3)	(2.0)		(1.1)	
No	$\frac{3}{6}$	59.1			2.1	18.5	4.1				7.8	4.2					4.1	90.9	-1199.5
		(8.5)			(1.9)	(8.8)	(2.8)				(6.4)	(2.4)					(2.6)	(1.4)	
No	$\frac{4}{6}$	68.1		5.5		20.2					1.0						5.2	88.3	-1292.3
		(7.4)		(3.6)		(7.3)					(3.0)						(3.5)	(1.6)	
No	$\frac{5}{6}$	49.8	2.1	13.2	10.6	17.4				2.4		2.1					2.5	86.2	-1788.7
		(7.5)	(1.9)	(4.2)	(5.0)	(5.1)				(2.3)		(2.0)					(2.9)	(1.3)	

Table A6: SFEM Estimates – Set of Strategies from Fudenberg et al. (2012)

Visibility	δ	AC	AD	TFT	$DTFT$	$TF2T$	$TF3T$	$2TFT$	$2TF2T$	$T2$	$GRIM$	$GRIM2$	$GRIM3$	$WSLS$	$2WSLS$	$CtoD$	$DTF2T$	$DTF3T$	$DGRIM2$	$DGRIM3$	$DCAlt$	β (%)	$\mathcal{L}$
No	$\frac{3}{6}$	2.1	70.6		17.2												8.1			2.0		92.9	-1107.3
		(1.9)	(7.6)		(5.8)												(4.4)			(1.7)		(1.4)	
No	$\frac{4}{6}$		60.1	6.2	26.0											2.1		5.6				91.7	-1112.5
			(7.5)	(3.5)	(7.2)											(1.8)		(3.5)				(1.3)	
No	$\frac{5}{6}$	2.1	39.0	19.4	10.6	5.3			5.6		2.5	5.3	3.6				2.5		4.1			90.6	-1176.8
		(2.4)	(7.3)	(5.6)	(4.6)	(3.4)			(4.6)		(2.5)	(3.5)	(3.0)				(2.1)		(2.5)			(1.1)	

Notes: The value of $(1-\beta)$ can be interpreted as the amount of noise not captured by the specified strategies. When the queue is visible, our estimates of $(1-\beta)$ in Table 3 are similar to the estimates in Romero and Rosokha (2018) and Romero and Rosokha (2019a). However, when the queue is not visible, the values are somewhat lower, suggesting that the set of strategies may be missing relevant strategies. Table A6 presents the estimates when using an expanded set of strategies. In particular, we use 20 commonly studied strategies in the indefinitely repeated PD literature (Fudenberg, Rand, and Dreber, 2012; Cason and Mui, 2019).

Figure A5: Evolution of Effort



(a) First Period

Notes: High effort is coded as 1, and low effort is coded as 0. We find evidence that subjects' choices have a time-trend effect. For example, when $\delta = 4/6$ and the queue is visible, we run a probit regression of subjects' first-period choice in state 3 on the match number and find the effect of the match number is significant at the 5% level if standard errors are clustered at the individual level (p-value=0.035). When $\delta = 5/6$ and the queue is not visible, we run a probit regression of subjects' first-period choice on the match number and find the effect of the match number is also significant at the 5% level if standard errors are clustered at the individual level (p-value=0.019). For these two cases, the time-trend effect is not statistically significant when the standard errors are clustered at the session level (p-value=0.359 for the first case, and p-value=0.142 for the second case).

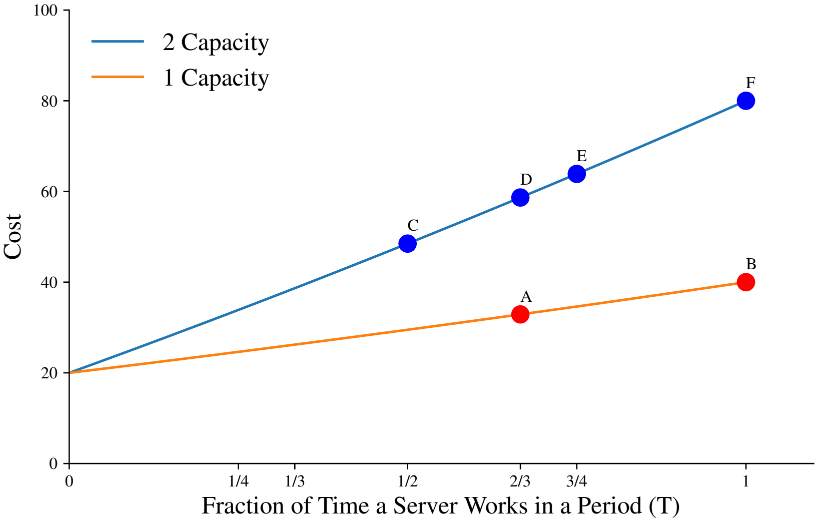
B Study 2: Additional Details

B.1 Microfoundations

B.1.1 Cost Function

As described in the third paragraph of section 6, the cost of working T fraction of a period with 1 unit of effort (i.e., capacity) is $c(1,T) = aT^2 + bT + c$ with $a = 2$, $b = 18$, and $c = 20$. The cost of choosing 2 units of capacity is $c(2,T) = xT^2 + yT + z$ with $x = 6$, $y = 54$, and $z = 20$. Figure B1 draws the two cost functions in the same graph given T ranges between 0 and 1. Figure B1 also labels all possible costs that relate to the three payoff matrices in the second study. In particular, points A and B in Figure B1 correspond to the cost of providing 1 unit of capacity. Points C, D, E, and F correspond to the cost of providing 2 units of capacity. Note that in study 2, because $T = min(1, \frac{\theta}{total\ capacity})$ is the fraction of time that person ends up working in a period, the individual's cost varies depending on the state and the actions of both servers.

Figure B1: Cost Function



Notes: A: (1,2,2); B: (1,1,2), (1,2,3), (1,1,3), (1,2,4) and (1,1,4); C: (2,2,2); D: (2,1,2); E: (2,2,3); F: (2,1,3), (2,2,4), (2,1,4). The first item in the tuple(e_i,e_j,θ) represents a server's own capacity choice; the second item represents the other server's capacity choice; the last item represents the state of the queue (i.e., the number of orders in the queue).

B.1.2 Compensation Function

As described in the third paragraph of section 6, an individual server is compensated based on the group performance. Specifically, if the group processes $M(.)$ orders, the compensation function is $r(e_i, e_j, \theta) = kM(e_i, e_j, \theta) + \mathbf{1}_{M(e_i,e_j,\theta)=4} bonus$ with $k = 36$ and $bonus = 0$.

B.1.3 Examples

Here, we provide two examples of how the cost and compensation function maps the payoff matrices.

- **Example 1**: Suppose 2 tasks are in the queue in a given period. If a server chooses 2 units of capacity and her teammate chooses 2 units of capacity, the total capacity is 4(=2+2) and 2 tasks are processed. The fraction of the time that the team ends up working T is 2/4, which leads to the server's cost 48(=6×T^2+54×T+20). The revenue is 72(=36×2). Therefore, this server's payoff is 24(=72-48).
- **Example 2**: Suppose 4 tasks are in the queue in a given period. If a server chooses 1 unit of capacity and her teammate chooses 2 units of capacity, the total capacity is 3(=1+2) and 3 tasks are processed. The fraction of the time that the team ends up working T is 1(=$min(1, \frac{4}{3})$), which leads to the server's cost 40(=2×T^2+18×T+20). The revenue is 108(=36×3). Therefore, this server's payoff is 68(=108-40).

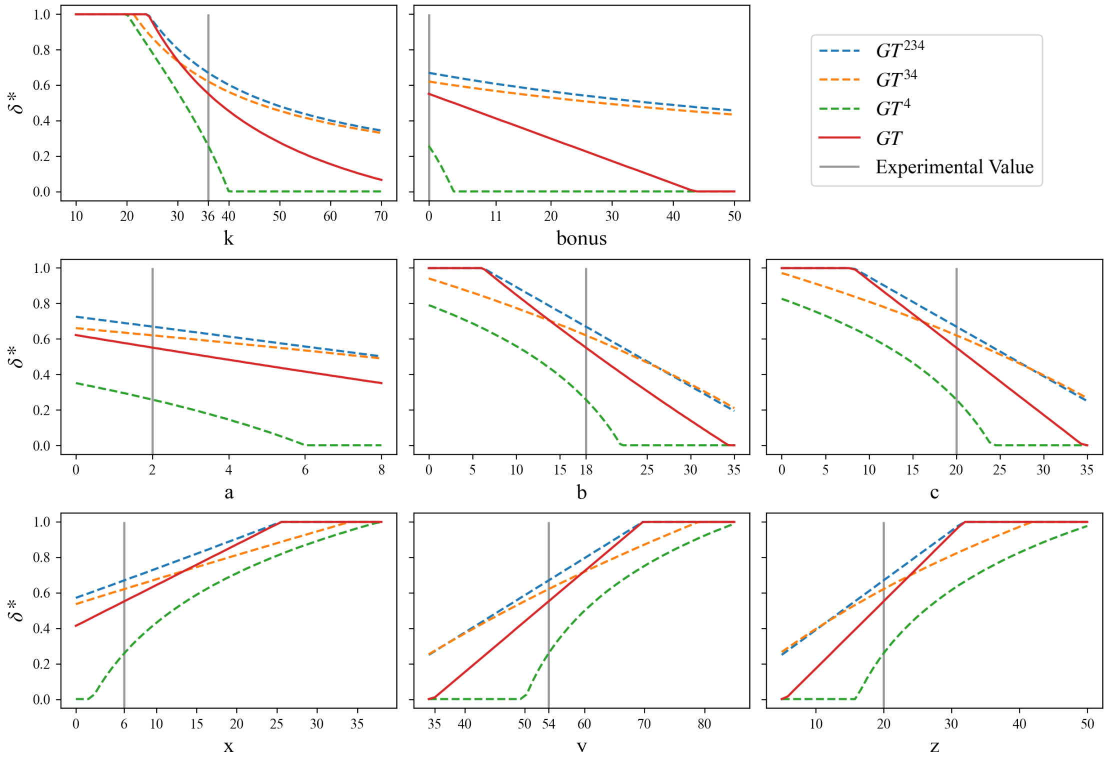
B.1.4 Comparative Statics

As in study 1, the compensation function and the cost function have 8 parameters in study 2. Figure B2 presents comparative statics of the critical thresholds of four GT strategies for each of the eight parameters.

In particular, for each subfigure, we vary one parameter while holding all other parameters constant at the experimental value. For example, in the top-right figure, we study the relationship between the critical thresholds of the GT (solid red), GT^4 (dashed green), GT^{34} (dashed orange), and GT^{234} (dashed blue) and the value of k while keeping $bonus = 0, a = 2b = 18, c = 20, x = 6, y = 54, z = 20$.

The patterns observed in Figure A2 can also be observed in Figure B2. Specifically, when $\delta_v^*(GT^4) > 0$, we find $\delta_v^*(GT^4) \leq \delta_v^*(GT^{34}) \leq \delta_v^*(GT^{234})$. This result suggests that when the queue is visible, sustaining cooperation is easier when the queue is long than when it is short. Then, $\delta_v^*(GT^4) \leq \delta_{nv}^*(GT) \leq \delta_v^*(GT^{234})$. In other words, sustaining high effort across all states is easier when the queue is not visible, and sustaining high effort across a subset of the states (e.g., $\theta = 4$) is easier when the queue is visible. These two common patterns are consistent with the hypotheses discussed in section 6.

Figure B2: Comparative Statics and Parameters



Notes: For each graph, we hold the other parameters constant as the experimental value when varying a given parameter. The experimental values are $k = 36, bonus = 0, a = 2, b = 18, c = 20, x = 6, y = 54$, and $z = 20$.

B.1.5 Theoretical Predictions

As in Study 1, we first find the transition-probability matrix implied by the strategy profile s^{GT} to find its critical threshold. If both players provide high effort, the transition-probability matrix is given by $\mathcal{P}^{hh}$; if one player deviates from high effort, the transition-probability matrix is given by $\mathcal{P}^{lh}$; and if both players provide low effort, the transition-probability matrix is given by $\mathcal{P}^{ll}$:

$$\mathcal{P}^{hh} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \end{array} \quad \mathcal{P}^{lh} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array} \quad \mathcal{P}^{ll} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \end{array}. \tag{15}$$

Vectors u^c, u^{dev}, and u^d specify payoffs obtained in each of the states:

$$u^c = \begin{pmatrix} u_2^c \\ u_3^c \\ u_4^c \end{pmatrix} = \begin{pmatrix} 24 \\ 44 \\ 64 \end{pmatrix} \qquad u^{dev} = \begin{pmatrix} u_2^{dev} \\ u_3^{dev} \\ u_4^{dev} \end{pmatrix} = \begin{pmatrix} 40 \\ 68 \\ 68 \end{pmatrix} \qquad u^d = \begin{pmatrix} u_2^d \\ u_3^d \\ u_4^d \end{pmatrix} = \begin{pmatrix} 32 \\ 32 \\ 32 \end{pmatrix}. \tag{16}$$

Lastly, the total values for the three cases in matrix notation are

$$V^c = [I - \delta\mathcal{P}^{hh}]^{-1}u^c, \qquad V^{dev} = u^{dev} + \delta\mathcal{P}^{lh}V^d, and \qquad V^d = [I - \delta\mathcal{P}^{ll}]^{-1}u^d, \tag{17}$$

where $V^c = \begin{pmatrix} V_2^c \\ V_3^c \\ V_4^c \end{pmatrix}$, $V^d = \begin{pmatrix} V_2^d \\ V_3^d \\ V_4^d \end{pmatrix}$, and I is the identity matrix. To show that s^{GT} is an SPE when the queue is visible, we need to find δ so that each element of V^c is at least as large as the corresponding element of V^{dev}. We find that s^{GT} is an SPE when δ is at least 0.67. We denote this critical threshold as $\delta_v^*(GT)$.

GT^{34} plays GT across states 3 and 4 and always provides low effort in state 2. The only difference in the analysis above is that $u^c = \begin{pmatrix} 32 \\ 44 \\ 64 \end{pmatrix}$, which leads to $\delta_v^*(GT^{34}) = 0.62$.

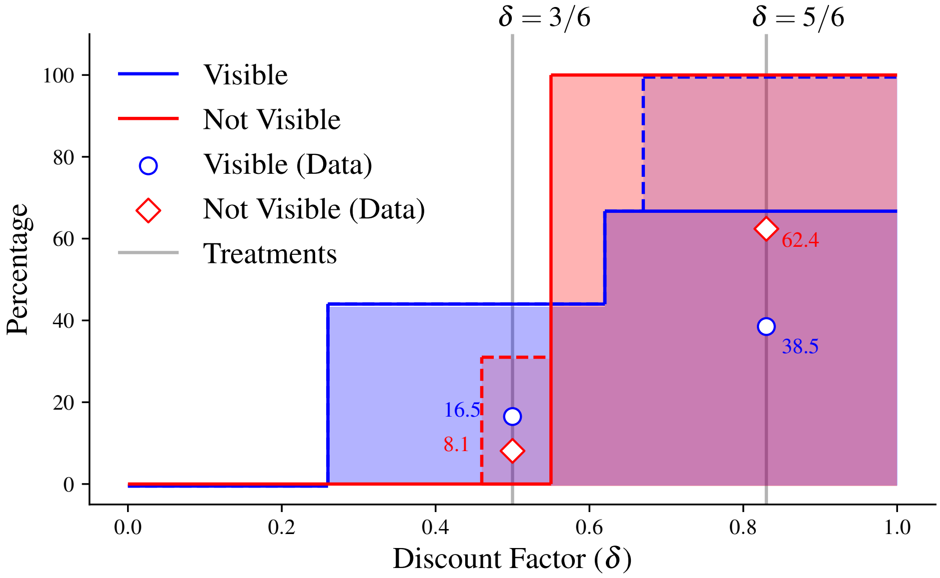
GT^4 plays GT in state 4 only and provides low effort in both states 2 and 3. The implied transition-probability matrices and the payoff vectors for this strategy are

$$\mathcal{P}^{c} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \end{array} \quad \mathcal{P}^{dev} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array} \quad \mathcal{P}^{ll} = \begin{array}{cc} \theta & \begin{array}{ccc} 2 & 3 & 4 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{18}$$

$$u^c = \begin{pmatrix} 32 \\ 32 \\ 64 \end{pmatrix} \qquad u^{dev} = \begin{pmatrix} 32 \\ 32 \\ 68 \end{pmatrix} \qquad u^d = \begin{pmatrix} 32 \\ 32 \\ 32 \end{pmatrix}. \tag{19}$$

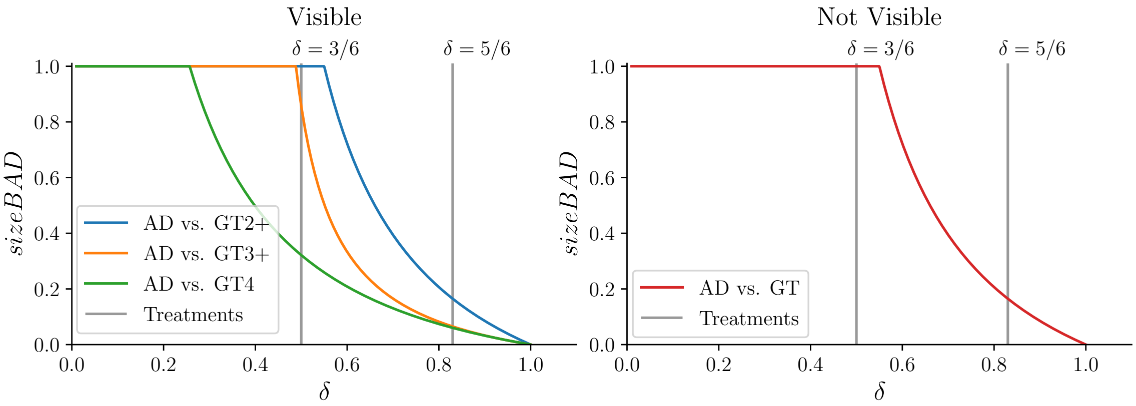
Note that we chose to label the transition probabilities as $\mathcal{P}^c$ instead of $\mathcal{P}^{hh}$ and $\mathcal{P}^{dev}$ instead of $\mathcal{P}^{lh}$, because the cooperative path of s^{GT^4} involves low effort in states 2 and 3. Solving for $\delta_v^*(GT^4)$, we get 0.26.

Figure B3: High Effort, Discount Factor, and Data



Notes: Data are the percentage of high capacity chosen by subjects across all periods in the second half of matches. Servers are assumed to use GT-type strategies in this figure ($\delta^*_{nv}(GT) = 0.55$; $\delta^*_v(GT^4) = 0.26$; $\delta^*_v(GT^{34}) = 0.62$ and $\delta^*_v(GT) = 0.67$). Even GT is an equilibrium strategy when the queue is visible and $\delta > 0.67$, we assume servers provide low capacity when the state is 2 because it is the efficient choice. $D.AlT^4$ is also an SPE when the queue is not visible and δ is large enough; it does not change the main hypotheses ($\delta^*_{nv}(D.AlT^4)$=0.46).

Figure B4: SizeBAD and Discount Factor



Notes: *SizeBAD* represents the limit of the basin of attraction of Always Defect (AD) versus cooperative strategies such as Grim Trigger. When a subject is considering which of these two types of strategies is more profitable in expectation, she might believe that the other subject will choose the cooperative strategy. If this belief is less than *SizeBAD*, the subject will maximize the expected payoff by choosing AD. The larger *SizeBAD* is, the more likely she is to choose AD.

B.2 Experimental Design

B.2.1 Match-Length Realizations

Table B1: Match Lengths

(a) $\delta = \frac{3}{6}$

Supergame Number:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Realization #1:	2	2	4	1	1	2	3	2	2	4	1	1	2	3	1	2	2	1	1	3	1	3	3	2	1
Realization #2:	2	2	1	4	3	1	1	1	2	1	1	1	2	1	1	1	1	2	1	1	4	1	2	2	1
Realization #3:	4	2	1	5	1	1	1	1	2	1	1	1	3	1	1	9	5	3	1	2	2	2	2	7	1

Supergame Number:	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
Realization #1:	2	2	1	5	1	2	4	3	1	1	1	1	1	2	2	2	1	2	1	1	1	2	1	2	2
Realization #2:	2	1	2	1	1	2	2	2	1	7	1	1	1	1	2	3	1	1	1	1	5	3	3	5	5
Realization #3:	4	1	1	1	3	1	1	2	1	1	2	4	1	2	2	6	3	1	6	12	2	2	2	2	1

Supergame Number:	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
Realization #1:	2	4	1	6	3	2	2	1	1	1	4	4	2	1	2	2	2	2	2	2	2	4	1	1	3
Realization #2:	1	1	4	2	13	2	1	1	1	2	1	2	3	2	1	3	1	1	2	1	2	2	2	4	1
Realization #3:	1	1	2	3	7	1	1	3	1	1	1	1	5	2	2	2	1	1	4	2	1	2	1	2	6

Supergame Number:	**76**	**77**	**78**	**79**	**80**	**81**	**82**	**83**	**84**	**85**	**86**	**87**	**88**	**89**	**90**	**91**	**92**	**93**	**94**	**95**	**96**	**97**	**98**	**99**	**100**
Realization #1:	2	1	4	1	1	1	1	1	3	1	2	4	1	2	1	1	1	4	1	5	2	1	2	2	1
Realization #2:	1	1	1	3	1	1	2	1	1	1	2	3	3	2	1	4	4	2	1	1	1	1	1	1	2
Realization #3:	1	1	1	1	2	2	1	1	1	5	1	1	6	1	3	1	1	2	2	1	1	1	1	1	4

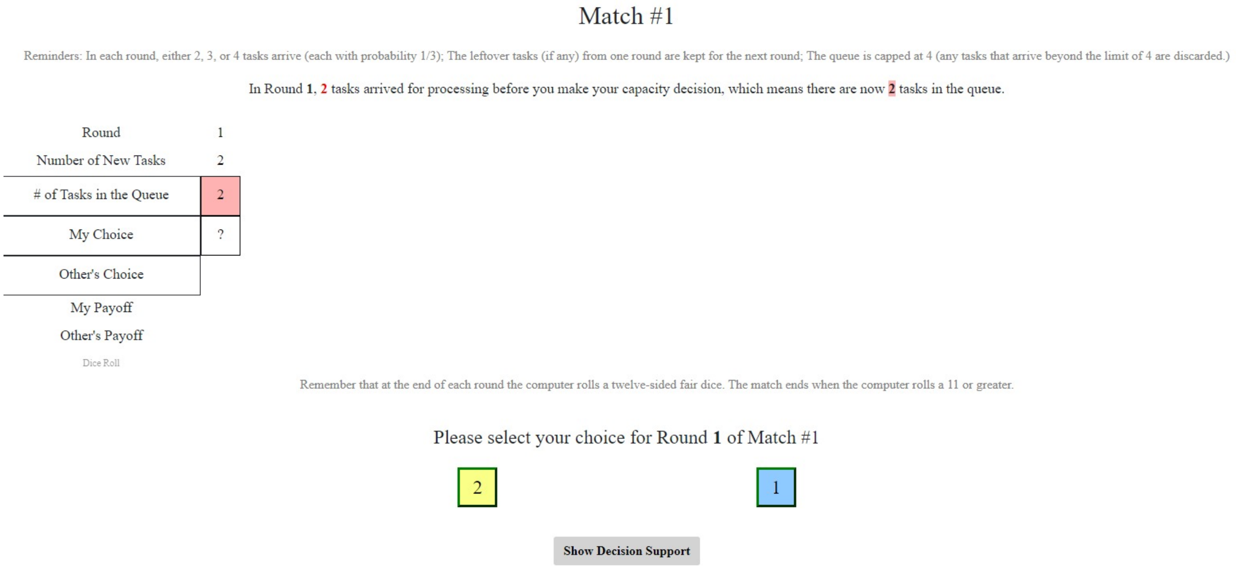
(b) $\delta = \frac{5}{6}$

Supergame Number:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Realization #1:	11	15	8	1	7	6	3	4	5	6	16	2	4	4	3	1	7	8	12	2	1	10	11	6	4
Realization #2:	4	12	30	4	15	1	10	2	5	2	11	19	6	2	1	2	5	7	2	3	2	2	9	8	6
Realization #3:	8	3	15	1	11	10	6	2	3	2	4	16	4	1	13	11	20	4	6	9	2	3	4	8	1

Supergame Number:	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
Realization #1:	6	1	4	18	2	9	1	6	10	20	1	1	1	3	5
Realization #2:	2	3	13	1	21	13	1	12	9	8	6	15	2	3	1
Realization #3:	10	6	12	14	11	14	3	2	11	6	5	12	12	2	3

B.2.2 Screenshots

Queue-Visible-Treatment Decision Screen



Match #1

Reminders: In each round, either 2, 3, or 4 tasks arrive (each with probability 1/3); The leftover tasks (if any) from one round are kept for the next round; The queue is capped at 4 (any tasks that arrive beyond the limit of 4 are discarded.)

In Round **1**, 2 tasks arrived for processing before you make your capacity decision, which means there are now 2 tasks in the queue.

Round	1
Number of New Tasks	2
# of Tasks in the Queue	2
My Choice	?
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Please select your choice for Round **1** of Match #1

Queue-Visible-Treatment Waiting Screen

Match #1

Reminders: each round, either 2, 3, or 4 tasks arrive (each with probability 1/3); the leftover tasks (if any) from one round are kept for the next round; the queue is capped at 4 (any tasks that arrive beyond the limit of 4 are discarded.)

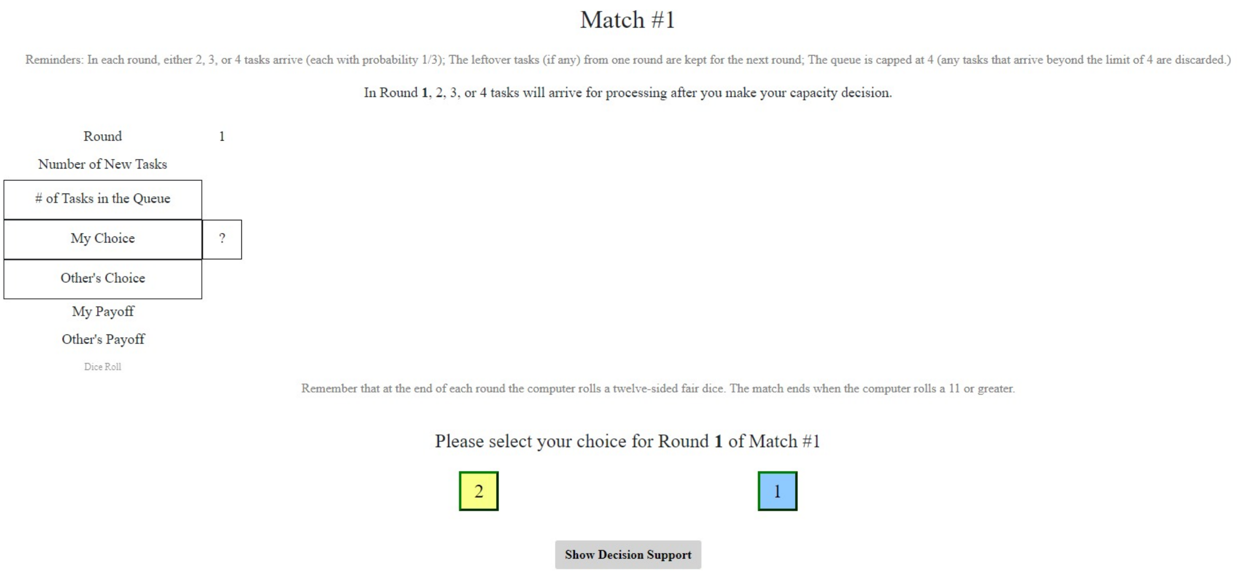
In Round **1**, 2 tasks arrived for processing before you make your capacity decision, which means there are now 2 tasks in the queue.

Round	1
Number of New Tasks	2
# of Tasks in the Queue	2
My Choice	2
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Please wait while the other participant makes a choice.

Queue-Not-Visible-Treatment Decision Screen



Match #1

Reminders: In each round, either 2, 3, or 4 tasks arrive (each with probability 1/3); The leftover tasks (if any) from one round are kept for the next round; The queue is capped at 4 (any tasks that arrive beyond the limit of 4 are discarded.)

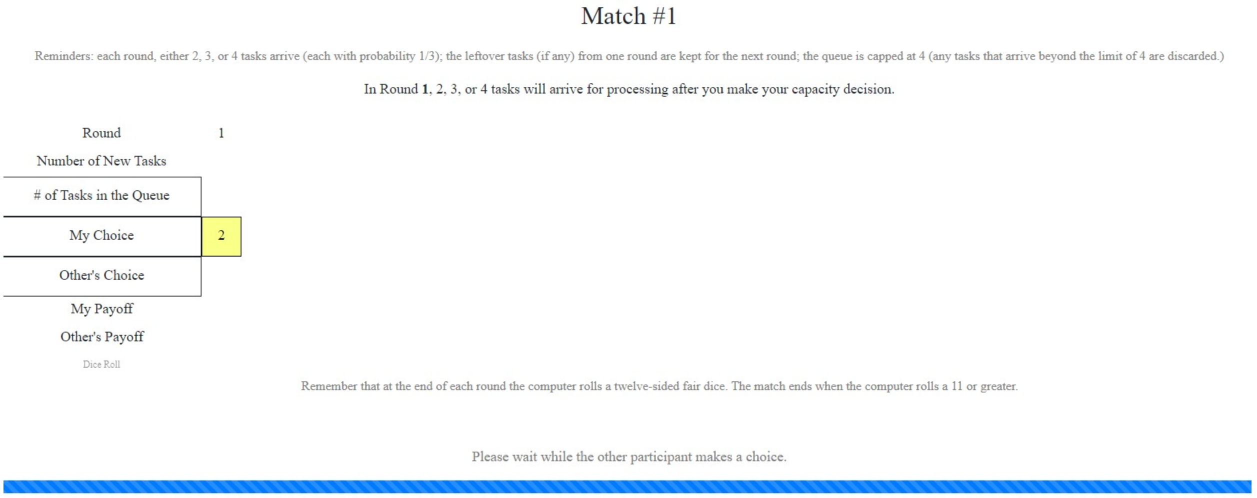
In Round **1**, 2, 3, or 4 tasks will arrive for processing after you make your capacity decision.

Round	1
Number of New Tasks	
# of Tasks in the Queue	
My Choice	?
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Please select your choice for Round **1** of Match #1

Queue-Not-Visible-Treatment Waiting Screen



Match #1

Reminders: each round, either 2, 3, or 4 tasks arrive (each with probability 1/3); the leftover tasks (if any) from one round are kept for the next round; the queue is capped at 4 (any tasks that arrive beyond the limit of 4 are discarded.)

In Round **1**, 2, 3, or 4 tasks will arrive for processing after you make your capacity decision.

Round	1
Number of New Tasks	
# of Tasks in the Queue	
My Choice	2
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Please wait while the other participant makes a choice.

B.2.3 Instructions

Experiment Overview

Today's experiment will last about 60 minutes.

You will be paid a show-up fee of $5 together with any money you accumulate during this experiment. The amount of money you accumulate will depend partly on your actions and partly on the actions of other participants. This money will be paid at the end of the experiment in private and in cash.

It is important that during the experiment you remain **silent**. If you have a question or need assistance of any kind, please **raise your hand, but do not speak -** and an experiment administrator will come to you, and you may then whisper your question.

In addition, please **turn off your cell phones and put them away now**.

Anybody that breaks these rules will be asked to leave.

Agenda

1. Instructions
2. Quiz
3. Experiment

How Matches Work

The experiment is made up of **100 matches**.

At the start of each match you will be randomly paired with another participant in this room.

You will then play a number of rounds with that participant (this is what we call a "match").

Each match will last for a random number of **rounds**:

- At the end of each round the computer will roll a twelve-sided fair dice.
- If the computer rolls a number less than 7, then the **match continues** for at least one more round (**50% probability**).
- If the computer rolls a 7 or greater, then the **match ends** (**50% probability**).

To test this procedure, click 'Test' button below. You will need to test this procedure 10 times.

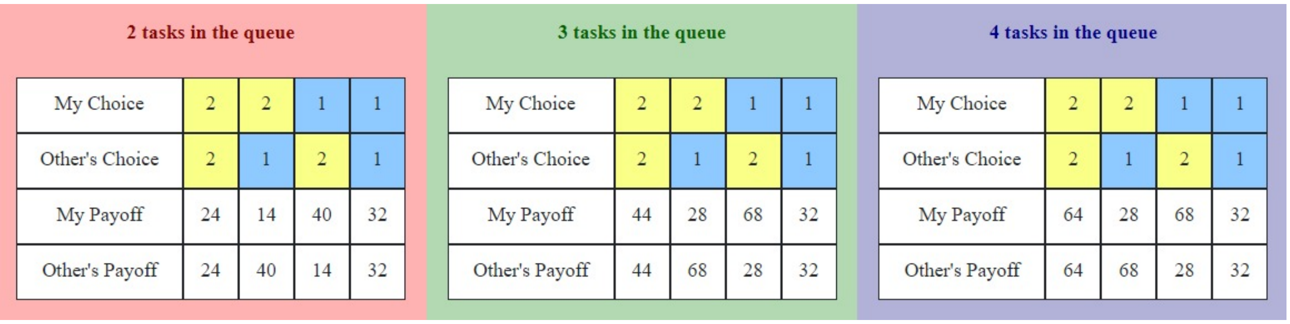
Round Description

In each round of a match, "tasks" will arrive for processing and join the **Task Queue**.

You and the participant you are paired with will work together to process the task queue. Specifically, in each round, you will choose how much capacity to allocate (either **1** unit or **2** units). The participant you are paired with will also choose how much capacity to allocate (either **1** unit or **2** units).

We denote the **sum** of your choice and the choice of the participant you are paired with as **Total Capacity** in that round. For example, if you choose **2** and the participant you are paired with chooses **2** then the Total Capacity is 4.

The number of tasks that can be processed in a given round is the smaller of the **Total Capacity** and the **# of Task in the Queue**. For example, if the total capacity is 4 but there are 3 tasks in the queue, then 3 tasks will be processed in that round. Another example – if the total capacity is 2 but there are 3 tasks in the queue, then 2 tasks will be processed in that round and ***1 Leftover Task*** will remain in the queue for the next round. Thus, the two of you can process available tasks in a round up to the total capacity in that round.



2 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	24	14	40	32
Other's Payoff	24	40	14	32

3 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	44	28	68	32
Other's Payoff	44	68	28	32

4 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	64	28	68	32
Other's Payoff	64	68	28	32

Possible combinations of choices and the resulting payoffs are summarized in the three tables above (labeled **2 Tasks in the Queue**, **3 Tasks in the Queue**, and **4 Tasks in the Queue**). These payoffs are determined based on the revenue and costs associated with the capacity choices and the tasks in the queue. The details of how the payoffs are determined will be described next, for now, we will provide a few examples of how to read these summary tables. In particular, if you choose **2** and the participant you are paired with chooses **2** and there are

- **2 tasks in the queue**, then your payoff for the round will be 24 points, and the other's payoff will be 24 points.
- **3 tasks in the queue**, then your payoff for the round will be 44 points, and the other's payoff will be 44 points.
- **4 tasks in the queue**, then your payoff for the round will be 64 points, and the other's payoff will be 64 points.

At the end of the experiment, your total points (accumulated across all rounds and matches) will be converted into cash at the exchange rate of 250 points = $1. In addition, you will be paid your show-up fee of $5.

How Round Payoffs are Determined

Your **Payoff** in a given round is the difference between the **Revenue** that you get and the **Cost** that you incur from processing tasks.

The **Revenue** is a function of the tasks processed by you and the participant you are paired with. Specifically, you will receive **36 points per task** processed.

The **Cost** is a function of your capacity choice and the fraction of the time that you end up working in that round (which we denote T). We calculate T as the smaller between 1 and the ratio of (# of Tasks in the Queue) and (Total Capacity). That is, if the total capacity in a given round is less than or equal to the # of tasks in the queue, then you work the whole round (T=1). However, if the Total Capacity in a given round is larger than the # of tasks in the queue, then the fraction of the time that you work is T=(# of Tasks in the Queue) / (Total Capacity). For example, if in a given round the total capacity is 4, and there are

3 tasks in the queue, then 3 tasks will be processed and the fraction of time you work in that round will be $^3/_4$.

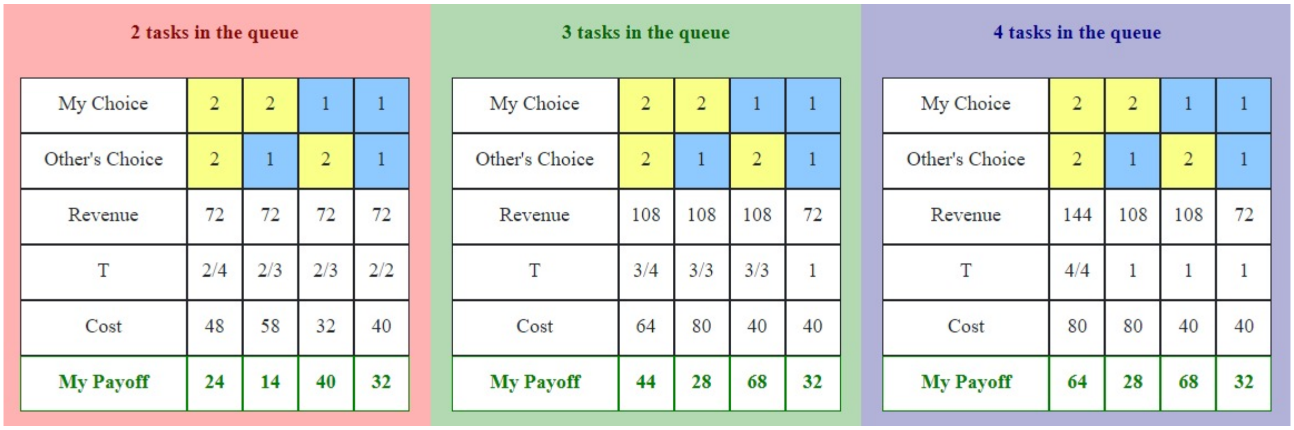
The cost function is increasing in your capacity choice and the fraction of the time you work as follows:

- if you choose to allocate **1** unit, your cost is $(2 \times T^2 + 18 \times T + 20)$
- if you choose to allocate **2** units, your cost is $(6 \times T^2 + 54 \times T + 20)$

The payoffs for the participant that you are paired with are calculated in the same way.

Next, we will consider a specific example.

Payoff Calculations Example



2 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
Revenue	72	72	72	72
T	2/4	2/3	2/3	2/2
Cost	48	58	32	40
My Payoff	**24**	**14**	**40**	**32**

3 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
Revenue	108	108	108	72
T	3/4	3/3	3/3	1
Cost	64	80	40	40
My Payoff	**44**	**28**	**68**	**32**

4 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
Revenue	144	108	108	72
T	4/4	1	1	1
Cost	80	80	40	40
My Payoff	**64**	**28**	**68**	**32**

Possible combination of choices, tasks in the queue, and the resulting payoffs are presented in the tables above. In addition, we present revenue, cost, and the fraction of the time you will work associated with those combinations.
For example, if in a given round there are **2 Tasks in the Queue**, you choose **2** and the participant that you are paired with chooses **2**, then

- the total capacity is 4(=**2**+**2**).
- **2** tasks are processed (the smaller between **2 Tasks in the Queue** and the total capacity of 4).
- the **Revenue** is **72**(=36 × 2).
- the fraction of the time you end up working T is 2/4 (=(2 tasks in the queue)/(total capacity of 4)).
- the **Cost** is **48**$(= 6 \times (^2/_4)^2 + 54 \times (^2/_4) + 20$, rounded to the nearest integer)
- your **Payoff** is **24**(=72-48)

The payoffs for the participant that you are paired with are calculated in the same way.

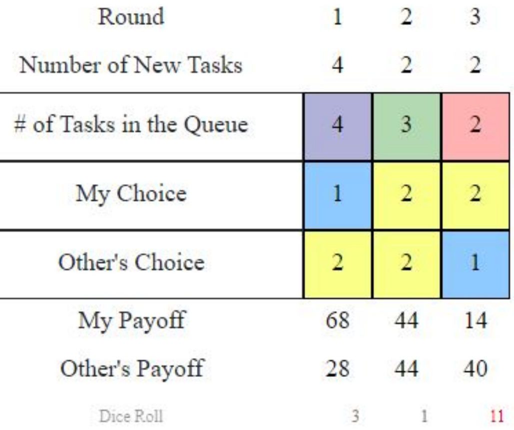
How # of Tasks in the Queue are Determined

In each round, a random number of new tasks will become available to join the queue. This number will be drawn at random from a set of numbers $\{2, 3, 4\}$, with each number equally likely. We will refer to this random number as the **Number of New Tasks**.

To determine the **# of Task in the Queue** in a given round, we will use the **Number of New Tasks** together with any ***leftover tasks*** from the previous round as follows:

- In Round 1, there are no previous rounds and, therefore, the **# of Task in the Queue** will be equal to the **Number of New Tasks**.
- In Round > 1, **# of Task in the Queue** will be determined in two steps:
  - First, we will determine the number of ***leftover tasks*** from the previous round.
    - If the **# of Task in the Queue** in the previous round is less than the **Total Capacity** in the previous round, then the number of ***leftover tasks*** is 0.
    - If the **# of Task in the Queue** in the previous round is greater than the **Total Capacity** in the previous round, then the number of ***leftover tasks*** is (**# of Task in the Queue** −**Total Capacity**)
  - Second, we will determine the **# of Task in the Queue** in the current round by adding the ***leftover tasks*** to the **Number of New Tasks** in the current round.
- Importantly, the Queue is capped at 4 tasks. That is, any tasks that arrive beyond the limit of 4 will be discarded. For example, if in a given round there are 2 leftover tasks in the queue and 3 new tasks arrive that round, then the queue will contain 4 tasks and 1 will be discarded.

How History Will be Recorded



Round	1	2	3
Number of New Tasks	4	2	2
# of Tasks in the Queue	4	3	2
My Choice	1	2	2
Other's Choice	2	2	1
My Payoff	68	44	14
Other's Payoff	28	44	40
Dice Roll	3	1	11

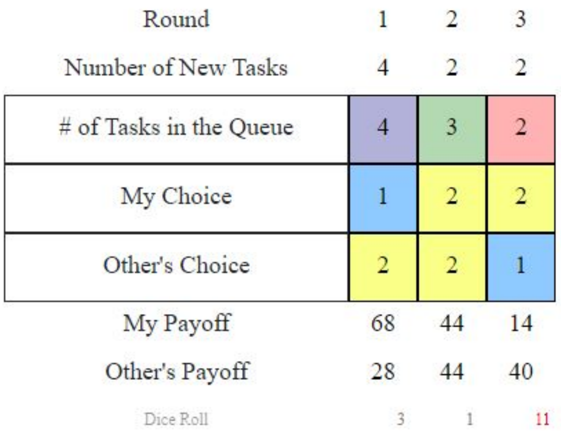
The history of all variables will be recorded in a history table like the one presented above. In this table you can see an example history of a match in which the computer picked actions at random. The recorded variables include:

- Round −− round number.
- Number of New Tasks −− a random draw in that round (one number is drawn from $\{2, 3, 4\}$ with each number is equally likely).

- # of Tasks in the Queue −− number of tasks in the queue in that round (either **2**, **3**, or **4**).
- My choice −− your choice (either **1** or **2**).
- Other's Choice −− the choice by the participant that you are paired with (either **1** or **2**).
- My Payoff −− your payoff in that round.
- Other's Payoff −− payoff of the participant that you are paired with in that match.

Reminder, your earnings will be the sum of your points across all matches converted into cash at the exchange rate of 250 points = $1. In addition, you will be paid your show-up fee of $5.

Example



Round	1	2	3
Number of New Tasks	4	2	2
# of Tasks in the Queue	4	3	2
My Choice	1	2	2
Other's Choice	2	2	1
My Payoff	68	44	14
Other's Payoff	28	44	40
Dice Roll	3	1	11

The following example is visualized in the above history:

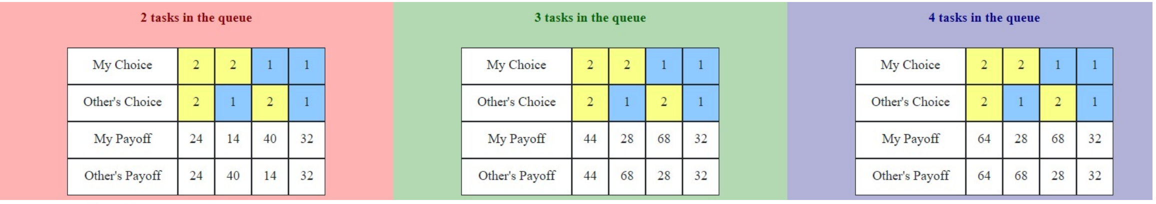
- Suppose that in Round 1, the **Number of New Tasks** is **4**, then there are **4 tasks in the queue** in Round 1.
- If you choose **1** while the participant you are paired with chooses **2**, then your payoff for Round 1 will be 68 points, and the other's payoff will be 28 points.
- Suppose that in Round 2 the **Number of New Tasks** is **2**, then there are **3 tasks in the queue** in Round 2.
  - First, we determine that the ***1***(=4-[1+2])***leftover task*** remains in the queue at the end of Round 1.
  - Second, we add this ***leftover task*** to the **Number of New Tasks** and determine that for Round 2, # of tasks in the queue is **3** (=***1***+**2**).
- If you choose **2** and the participant you are paired with chooses **2**, then your payoff for Round 2 will be 44 points, and the other's payoff will be 44 points.
- Suppose that in Round 3 the **Number of New Tasks** is **2**, then there are **2 Tasks in the Queue** in Round 3.
  - First, we determine that there are ***no leftover task*** remains in the queue at the end of Round 2 (Total Capacity ≧3).

  - Therefore, the # of tasks in the queue is equal to the **Number of New Tasks** in Round 3, which means that # of tasks in the queue is **2** (=*0*+2)

- If you choose **2** and the participant you are paired with chooses **1**, then your payoff for Round 3 will be 14 points, and the other's payoff will be 40 points.

B.2.4 Quiz

The top third of each screen contains the three payoff tables



2 tasks in the queue

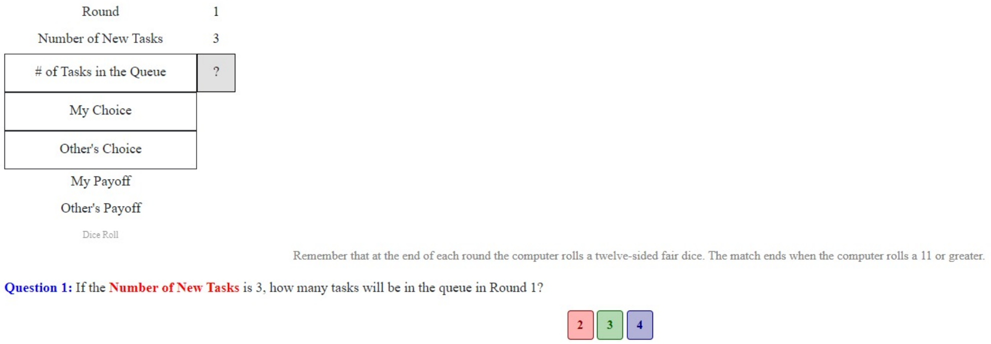
My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	24	14	40	32
Other's Payoff	24	40	14	32

3 tasks in the queue

My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	44	28	68	32
Other's Payoff	44	68	28	32

4 tasks in the queue

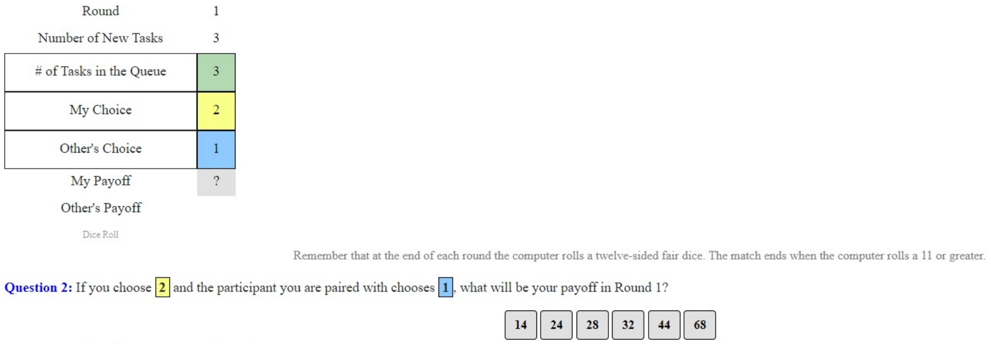
My Choice	2	2	1	1
Other's Choice	2	1	2	1
My Payoff	64	28	68	32
Other's Payoff	64	68	28	32



Round	1
Number of New Tasks	3
# of Tasks in the Queue	?
My Choice	
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

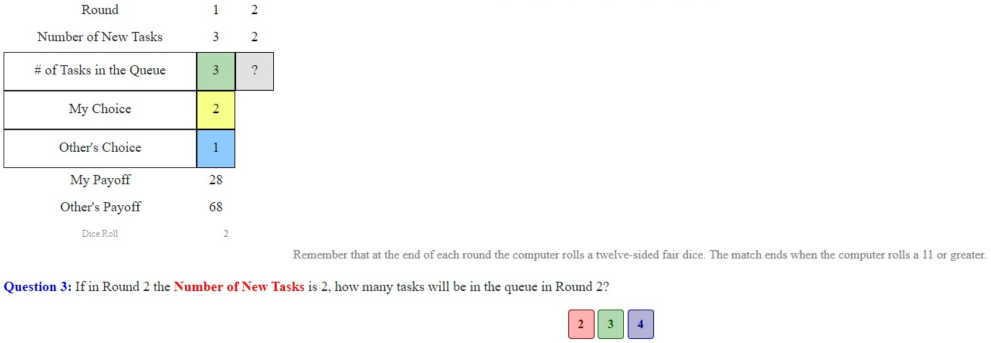
Question 1: If the **Number of New Tasks** is 3, how many tasks will be in the queue in Round 1?



Round	1
Number of New Tasks	3
# of Tasks in the Queue	3
My Choice	2
Other's Choice	1
My Payoff	?
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

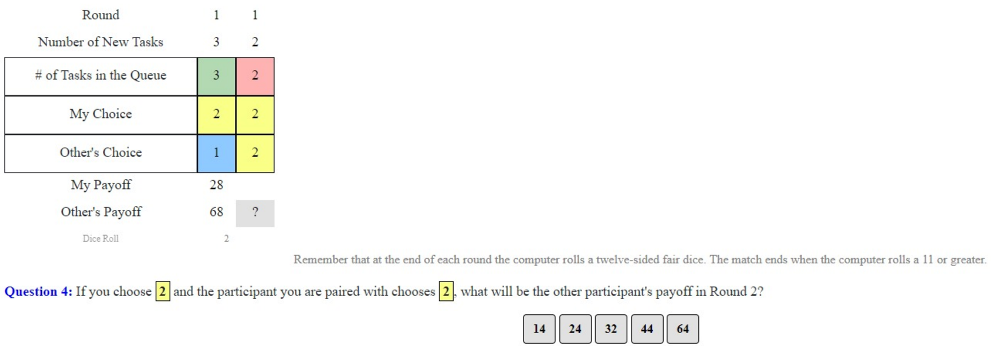
Question 2: If you choose 2 and the participant you are paired with chooses 1, what will be your payoff in Round 1?



Round	1	2
Number of New Tasks	3	2
# of Tasks in the Queue	3	?
My Choice	2	
Other's Choice	1	
My Payoff	28	
Other's Payoff	68	
Dice Roll	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 3: If in Round 2 the **Number of New Tasks** is 2, how many tasks will be in the queue in Round 2?

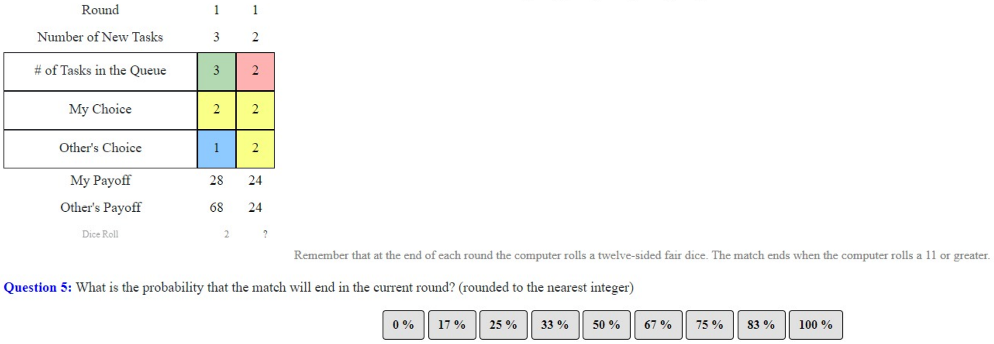


Round	1	1
Number of New Tasks	3	2
# of Tasks in the Queue	3	2
My Choice	2	2
Other's Choice	1	2
My Payoff	28	
Other's Payoff	68	?
Dice Roll	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 4: If you choose 2 and the participant you are paired with chooses 2, what will be the other participant's payoff in Round 2?

14 24 32 44 64

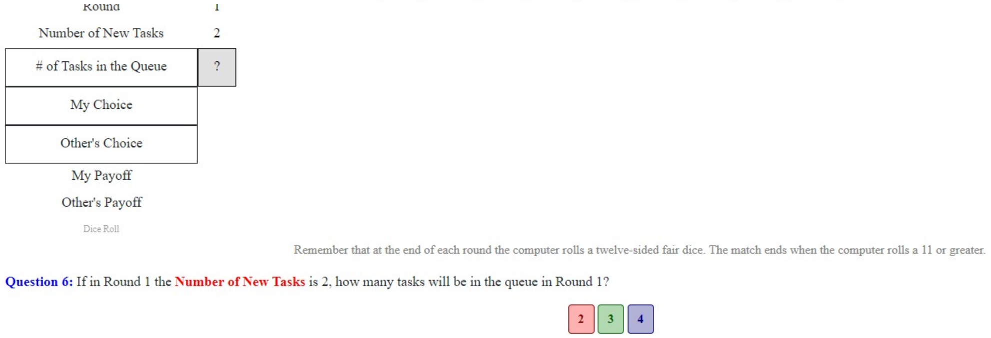


Round	1	1
Number of New Tasks	3	2
# of Tasks in the Queue	3	2
My Choice	2	2
Other's Choice	1	2
My Payoff	28	24
Other's Payoff	68	24
Dice Roll	2	?

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 5: What is the probability that the match will end in the current round? (rounded to the nearest integer)

0 % 17 % 25 % 33 % 50 % 67 % 75 % 83 % 100 %

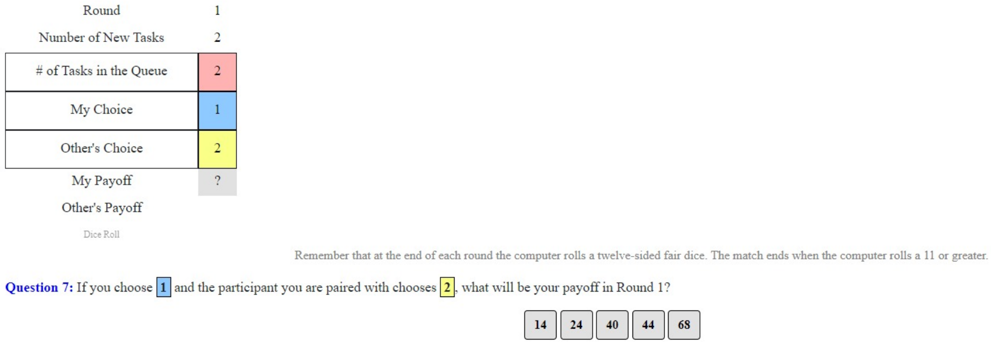


Round	1
Number of New Tasks	2
# of Tasks in the Queue	?
My Choice	
Other's Choice	
My Payoff	
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 6: If in Round 1 the **Number of New Tasks** is 2, how many tasks will be in the queue in Round 1?

2 3 4

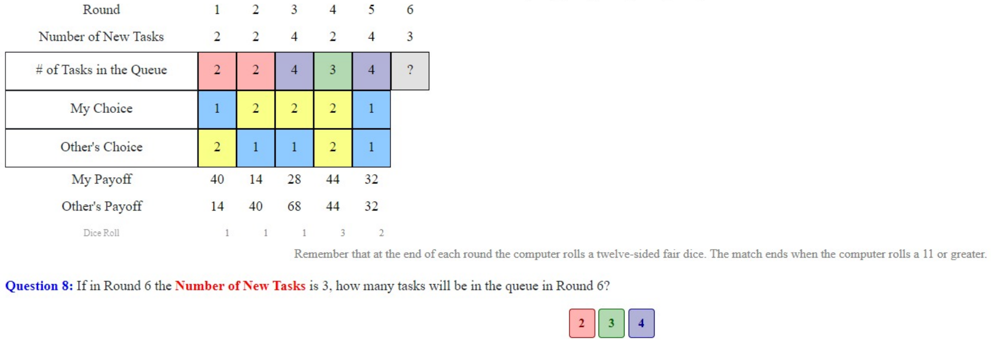


Round	1
Number of New Tasks	2
# of Tasks in the Queue	2
My Choice	1
Other's Choice	2
My Payoff	?
Other's Payoff	
Dice Roll	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 7: If you choose 1 and the participant you are paired with chooses 2, what will be your payoff in Round 1?

14 24 40 44 68

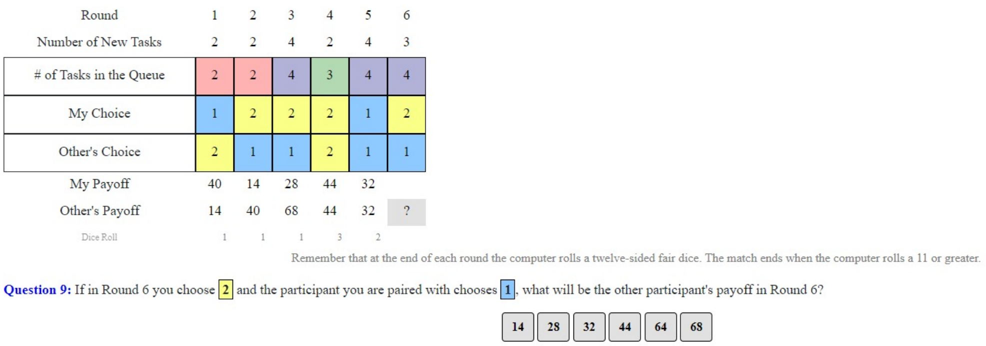


Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
# of Tasks in the Queue	2	2	4	3	4	?
My Choice	1	2	2	2	1	
Other's Choice	2	1	1	2	1	
My Payoff	40	14	28	44	32	
Other's Payoff	14	40	68	44	32	
Dice Roll	1	1	1	3	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 8: If in Round 6 the **Number of New Tasks** is 3, how many tasks will be in the queue in Round 6?

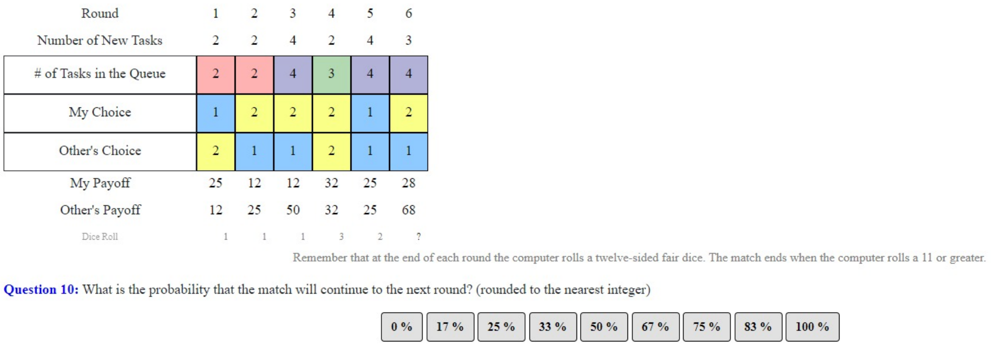
2 3 4



Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
# of Tasks in the Queue	2	2	4	3	4	4
My Choice	1	2	2	2	1	2
Other's Choice	2	1	1	2	1	1
My Payoff	40	14	28	44	32	
Other's Payoff	14	40	68	44	32	?
Dice Roll	1	1	1	3	2	

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 9: If in Round 6 you choose 2 and the participant you are paired with chooses 1, what will be the other participant's payoff in Round 6?



Round	1	2	3	4	5	6
Number of New Tasks	2	2	4	2	4	3
# of Tasks in the Queue	2	2	4	3	4	4
My Choice	1	2	2	2	1	2
Other's Choice	2	1	1	2	1	1
My Payoff	25	12	12	32	25	28
Other's Payoff	12	25	50	32	25	68
Dice Roll	1	1	1	3	2	?

Remember that at the end of each round the computer rolls a twelve-sided fair dice. The match ends when the computer rolls a 11 or greater.

Question 10: What is the probability that the match will continue to the next round? (rounded to the nearest integer)

B.3 Additional Tables and Figures

Table B2: Summary of Experiment Administration for the Second Study

Treatment		Administration				Demographics		
Visibility	δ	Sessions	Subjects	Matches	Earnings	% Male	% STEM	% US HS
Yes	$\frac{3}{6}$	3	36	100	22.7	56.5	60.9	73.9
					(0.2)	(10.5)	(10.3)	(9.4)
Yes	$\frac{5}{6}$	3	36	40	33.0	38.9	61.1	72.2
					(0.6)	(8.3)	(8.2)	(7.2)
No	$\frac{3}{6}$	3	36	100	22.4	61.9	61.9	66.7
					(0.1)	(10.5)	(10.7)	(10.4)
No	$\frac{5}{6}$	3	34	40	32.8	55.9	55.9	70.6
					(0.5)	(8.7)	(9.0)	(7.8)

Notes: Earnings are reported in USD and include a $5 show-up fee. Standard errors are in parentheses. One session in (Yes, $\frac{3}{6}$) and one session in (No, $\frac{3}{6}$) treatments were stopped before match 100 due to time constraints. The demographics data for these two sessions are missing.

Table B3: Study 2: Efficiency

Treatment		First Period			All Periods			
Visibility	δ	State 2	State 3	State 4	State 2	State 3	State 4	All States
Yes	$\frac{3}{6}$	98.8 (1.6)	76.9 (1.9)	67.1 (1.6)	98.6 (1.8)	76.2 (1.3)	60.8 (1.3)	69.8 (1.2)
Yes	$\frac{5}{6}$	97.9 (2.3)	89.8 (3.1)	83.0 (2.4)	97.3 (1.7)	86.8 (0.9)	70.8 (2.1)	78.0 (1.7)
No	$\frac{3}{6}$	99.6 (0.5)	67.6 (0.3)	50.2 (0.2)	99.1 (0.5)	70.4 (0.5)	55.2 (0.5)	63.5 (0.5)
No	$\frac{5}{6}$	82.6 (4.6)	93.2 (3.5)	82.8 (2.3)	80.3 (1.6)	91.4 (1.1)	74.5 (2.1)	79.8 (1.3)

Table B4: Study 2: SFEM Estimates – First Half of Matches

Visibility	δ	AD	AC	TFT	GT	$D.TFT$	$D.AlT$	AC^{34}	TFT^{34}	GT^{34}	$D.TFT^{34}$	$D.AlT^{34}$	AC^{4}	TFT^{4}	GT^{4}	$D.TFT^{4}$	$D.AlT^{4}$	β (%)	$\mathcal{L}$
Yes	$\frac{3}{6}$	53.0 (9.1)	2.8 (2.3)			1.2 (2.3)		2.8 (2.7)	6.4 (4.5)	2.3 (3.0)			8.2 (4.8)		11.3 (6.4)	9.6 (5.1)	2.4 (2.7)	86.1 (1.8)	-1566.3
Yes	$\frac{5}{6}$	13.9 (5.5)		5.6 (3.4)		8.3 (4.3)		2.8 (2.6)	49.5 (8.8)	8.3 (5.2)	2.8 (3.1)				8.8 (4.2)			86.4 (1.4)	-2043.2
No	$\frac{3}{6}$	66.2 (7.3)		2.8 (2.8)			1.6 (2.2)					9.4 (4.2)				3.2 (3.5)	16.8 (5.9)	89.0 (1.5)	-1334.5
No	$\frac{5}{6}$	6.0 (4.2)	3.0 (3.1)	47.2 (9.7)	16.1 (6.1)	19.8 (7.4)					2.1 (2.4)						5.9 (4.7)	82.5 (1.5)	-2235.2

C System Performance

C.1 Theoretical Predictions

This section shows the steps of conducting theoretical predictions of the system performance (i.e., customer waiting time and lost demand) in steady states. In each treatment, we assume subjects use the efficient SPE strategy suggested in Table 1. Given this strategy and the transition probability matrix $\mathcal{P}$ suggested by this strategy, we can calculate the stationary distribution of states π by setting $\pi\mathcal{P} = \pi$. Using the stationary distribution of states, the average customer waiting time in a steady state is defined as the ratio of the expected total orders' waiting time and the expected total order processed in a period.[21] The average percentage of lost demand is defined as the ratio of the expected total lost orders and the expected total new

---

[21] An order's waiting time includes the processing time and depends on the servers' effort provision. In study 1, an order is processed in 0.5 periods if the server chooses high effort and 1.0 periods if the server chooses low effort. In study 2, an order is processed in 0.25 periods if both servers choose high effort, in $\frac{1}{3}$ periods if one server chooses high effort and the other server chooses low effort, and in 0.5 periods if both servers choose low effort.

arrival orders. The calculation is trivial for the treatment where $\delta = \frac{5}{6}$ because there are no leftovers and the transition probabilities are determined by the arrival process (i.e., uniform distribution). The calculations for the other treatments are subtle. Table C1 presents the results of these calculations. We use the treatment where the queue is visible and $\delta = \frac{3}{6}$ as an example. In that treatment, servers are using $GT4$, so the transition probability matrix is given by $\mathcal{P}$:

$$\mathcal{P} = \begin{array}{cccc} \theta & 2 & 3 & 4 \\ 2 & 1/3 & 1/3 & 1/3 \\ 3 & 0 & 1/3 & 2/3 \\ 4 & 1/3 & 1/3 & 1/3 \end{array}$$

Therefore, the stationary distribution of states is $[\frac{2}{9}, \frac{3}{9}, \frac{4}{9}]$. In the first study, when the current state is 2, the total waiting time for those 2 orders is $1 + 1 = 2$. When the current state is 3, there are 4 situations that should be considered. With probability $\frac{1}{9}$ (using the Bayes rule and the transition probability), the previous state is 3, and the next state is 3. We only consider the last 2 orders in the queue when calculating the average waiting time because the first order (leftover order) should be counted in the previous period. Therefore, the total waiting time for the last 2 orders is $1 + 2 = 3$. With probability $\frac{2}{9}$, the previous state is 3, and the next state is 4. The total waiting time for the last 2 orders is $1 + 1.5 = 2.5$. With probability $\frac{2}{9}$, the previous state is not 3, and the next state is 3. There are no leftovers, and the total waiting time for the 3 orders is $1 + 1 + 2 = 4$. Lastly, With probability $\frac{4}{9}$, the previous state is not 3, and the next state is 4. The total waiting time for the 3 orders is $1 + 1 + 1.5 = 3.5$. When the current state is 4, with probability $\frac{1}{2}$, the previous state is 3, and there is a leftover order. We only take the last 3 orders into account, and the total waiting time for these 3 orders is $0.5 + 1 + 1 = 2.5$. With probability $\frac{1}{2}$, there is no leftover order, and the total waiting time for those 4 orders is $0.5 + 0.5 + 1 + 1 = 3$. As a result, the expected total waiting time $\frac{2}{9} \times 2 + \frac{3}{9} \times (\frac{1}{9} \times 3 + \frac{2}{9} \times 2.5 + \frac{2}{9} \times 4 + \frac{4}{9} \times 3.5) + \frac{4}{9} \times (\frac{1}{2} \times 2.5 + \frac{1}{2} \times 3) = 2.78$, and the expected total order processed is $\frac{2}{9} \times 2 + \frac{3}{9} \times (\frac{1}{9} \times 2 + \frac{2}{9} \times 2 + \frac{2}{9} \times 3 + \frac{4}{9} \times 3) + \frac{4}{9} \times (\frac{1}{2} \times 3 + \frac{1}{2} \times 4) = 2.89$. Thus, the average waiting time is 0.96.

For the average lost demand, the only possible scenario in which we may see a lost demand is when the current state is 4 and the previous state is 3. In this scenario, we can observe a lost demand only when the new arrival customer is 4. Therefore, the expected lost demand in steady states is $\frac{4}{9} \times \frac{6}{12} \times \frac{1}{2} \times 1$. The expected number of new arrival orders is 3. Thus, the average percentage of lost demand is 3.7%.

Additionally, we conduct theoretical simulations to validate the theoretical predictions. Specifically, for each treatment, two servers interact with each other for 1100 periods using the efficient SPE strategy. The average waiting time and lost demand in steady states are calculated using the last 1000 periods. We repeat this process 1000 times and present the results in the table C1. The theoretical simulation results are perfectly consistent with the theoretical calculations.

Table C1: System Performance in Theory

Treatment		**Customer Waiting Time**				**Lost Demand (%)**			
		Study 1		Study 2		Study 1		Study 2	
Visibility	δ	Theoretical Calculation	Theoretical Simulation	Theoretical Calculation	Theoretical Simulation	Theoretical Calculation	Theoretical Simulation	Theoretical Calculation	Theoretical Simulation
Yes	$\frac{3}{6}$	0.96	0.96 (0.01)	0.79	0.79 (0.01)	3.7	3.7 (0.1)	3.7	3.7 (0.1)
Yes	$\frac{5}{6}$	0.78	0.78 (0.00)	0.61	0.61 (0.00)	0.0	-	0.0	-
No	$\frac{3}{6}$	1.21	1.21 (0.00)	1.01	1.01 (0.00)	12.8	12.8 (0.2)	12.8	12.8 (0.2)
No	$\frac{5}{6}$	0.67	0.67 (0.00)	0.53	0.53 (0.00)	0.0	-	0.0	-

Notes: Waiting times include processing times. The theoretical predictions are for the steady state. Standard errors for theoretical simulations are calculated using bootstrapping. In the case that a standard error is below 0.005, we denote it as 0.00.

C.2 Raw Data and Data-Driven Simulations

In this section, we first present the regression analysis of the system performance using the raw experimental data. We run a linear regression of the customer waiting time on three dummy variables (i.e., study, visibility, and discount factor) and the interaction variable between visibility and discount factor. The first column of Table C2 presents the results of this regression. We find it significantly decreases with the discount factor, which is consistent with the theoretical prediction. Although we do not find significant effects of queue visibility and the interaction variable, the magnitudes of these two coefficients are directionally consistent with the theoretical predictions. Our second regression considers the percentage of lost demand as the independent variable. The second column of Table C2 presents the results of this regression. We find that the magnitudes of the coefficients of queue visibility and the interaction variable are directionally consistent with the theoretical predictions. However, the lost demand increases with the discount factor. This is not surprising as a high discount factor is associated with more periods and thus may cause more lost demand.

Table C2: Regression Results

Variables	Raw Data Waiting Time	 Lost Demand	Data-Driven Simulation Waiting Time	 Lost Demand
study	-0.29***	-1.7**	-0.41***	-0.05***
	(0.03)	(0.6)	(0.01)	(0.00)
visibility	-0.05	-0.3	-0.15***	-0.05***
	(0.03)	(0.7)	(0.01)	(0.00)
discount	-0.18***	1.4	-0.58***	-0.16***
	(0.06)	(0.9)	(0.01)	(0.00)
visibility × discount	0.11	0.5	0.29***	0.08***
	(0.09)	(1.6)	(0.01)	(0.00)
constant	1.16***	5.7***	2.03***	0.34***
	(0.02)	(0.4)	(0.01)	(0.00)
Observations	4,736	4,736	800	800

Notes: For the regression result from the raw data, standard errors (in parentheses) are clustered at the session level, and we only consider the data in the second half of matches. The dummy variable "study" is 0 if the observation is in Study 1 and is 1 if the observation is in Study 2. The dummy variable "visibility" is 0 if the queue is not visible and is 1 if the queue is visible. The dummy variable "discount" is 0 if $\delta = \frac{3}{6}$ and is 1 if $\delta = \frac{5}{6}$.

Because the system performance metrics in the raw experimental data are differentially impacted by the initial and terminal conditions of the experiment (e.g., the first period has no lost demand, and the waiting time for the leftover customers in the last period is ambiguous), we believe that analyzing the system performance in the steady state using simulations based on the estimates of the strategies that subjects use is beneficial. That is, we can simulate servers' actions for each treatment based on the estimated strategies in that treatment (see Table 3 for study 1 and Table 6 for study 2). We performed the simulation using the following steps:

- For each combination of 2 (out of 16) strategies, we simulate 2 strategies' actions in a 20-period match.
- To focus on the periods that have reached the steady state, we (i) exclude the first 4 periods (in the simulations, we noticed that customer waiting time and lost demand need at least 4 periods to reach the steady state in our simulation); and (ii) exclude the last period because the average waiting time of the leftover customers in the last period is ambiguous.
- Next, we calculate the average customer waiting time and the percentage of lost demand for the selected range of periods.
- We then create a 16×16 matrix where each row represents a strategy that a server can choose, and each column represents a strategy that the other server can choose. Each entry in this matrix represents the metric of interest (i.e., average customer waiting time or the percentage of lost demand) when the row strategy is played against the column strategy.
- We calculate the average customer waiting time and the percentage of lost demand for a treatment based on this matrix and the estimated percentage of the 16 strategies for that treatment.

We repeat the above process 100 times. The regression results are presented in the last columns in Table C2. We find that the effects of queue visibility, the discount factor, interaction variable are all significant at 1 percent level. The coefficients of these three variables are consistent with our theoretical predictions.