Queueing to Learn*

Chiara Margaria†

January 12, 2017

Abstract

I study the design of a queue when agents learn from past service experiences and can be served multiple times. A continuum of forward-looking agents compete for a constant flow of a resource and decide whether and when to engage in costly queueing to be served. Valuations fluctuate over time, independently across agents; each agent faces an experimentation problem because payoffs are informative about the prevailing valuation. To maximize efficiency, the designer offers a simple binary menu of queues (i.e., two customer classes): service is rendered on a first-come first-served basis in one queue, and in random order in the other. Surprisingly, allowing for strategic reneging, the designer can implement the optimal menu with a single queueing discipline.

Keywords: Queues; Experimentation; Reneging; Congestion; Mechanism Design.

JEL Codes: C73, D47, D82

1 Introduction

Queues and waiting lists are common tools for allocating scarce resources. Designing a queue is inherently a dynamic problem. First, the resource to be allocated becomes available over time, and decisions must be made online. Second, the pool of potential customers changes as a result of agents joining and leaving the queue. Third, agents acquire information when they are allocated the resource, and thus their private information evolves.

*I am indebted to Johannes Hörner and Larry Samuelson for their invaluable help and support throughout this project. I would like to thank Dirk Bergemann, Eduardo Faingold, Mira Frick, and Ryota Iijima for their useful comments and suggestions, and Gad Allon, Barış Ata, Achal Bassamboo for insightful discussions. I benefited from many conversations with visitors at Yale, in particular V. Bhaskar, Laura Doval, Caroline Thomas, and Takuo Sugaya, and from excellent comments from Allen I. K. Vong and various seminar participants.

†Yale University, 28 Hillhouse Ave., New Haven, CT 06520, USA, chiara.margaria@yale.edu.

In the context of queues, and more broadly when a resource is allocated to a pool of agents that changes over time, the assumption of static private information has largely been maintained. In contrast, I study the design of a queue to maximize efficiency when strategic agents require service repeatedly, and learn their valuation when served. Learning and repeated consumption raise new questions. Are common queueing disciplines well suited to address congestion when agents learn about their valuation over time? In a dynamic setting with repeated service, can the designer resort to the last-come first-served discipline to induce agents to follow socially desirable behavior? Does the opportunistic behavior of agents, who might renege, harm or help the designer?

To address these questions, I analyze the problem of a designer who seeks to maximize efficiency by offering a queueing discipline (or a menu thereof). The contribution of the paper is twofold. First, I show that the optimal menu is remarkably simple and involves well-known queueing disciplines: service is rendered in a first-come first-served manner in one queue, and in random order (possibly with a minimum wait requirement) in the other. Second, I demonstrate that strategic reneging enables the designer to implement the optimal menu with a single queue. The possibility of strategic reneging (and rejoining) complicates the designer's problem, as the welfare from a queueing discipline depends on the induced agents' behavior. However, the opportunity to join and leave the queue allows agents to "engineer" personalized waiting-time distributions. Surprisingly, this enhanced flexibility suffices to implement the optimal menu.

In the model, a constant flow of a resource is to be allocated to a continuum of forward-looking agents. Capacity is limited: over any interval of time, a fixed mass of agents can be served. Over time, agents decide whether and when to engage in costly queueing to be served. Each of them can be served multiple times. Valuations fluctuate independently across agents: each agent faces an experimentation problem because the lump-sum payoff collected at each service reveals the prevailing valuation. Rivalry generates an externality problem, as agents ignore the fact that their actions affect overall congestion.

A designer devises a queueing discipline (or a menu thereof) to maximize efficiency in a steady-state equilibrium. In other words, she commits to a dynamic allocation rule that differentiates between agents based on the time that each has waited since joining the queue. Even if the mechanism determines the times when new information is revealed, its content remains wholly private. The designer's choice shapes the experimentation problem faced by the agents. The value of the information that an agent acquires when he consumes is determined by his ability to act on it, which in turns depends on the equilibrium level of congestion.

The analysis in divided into two parts. First, I analyze the problem of the designer when she is constrained to disciplines that deter reneging. The optimal queueing discipline is either first-come first-served, or service-in-random-order, with a possible minimum waiting requirement.

The result stems from a trade-off between thickness and congestion. Roughly, first-come first-served minimizes the queue length and is optimal when the waiting cost is high. However, when served in order of arrival, agents may join the line only to acquire the right to receive service at some point in the future, if they expect their valuation to be high by the time they are served. This is detrimental to welfare: agents joining the line impose negative externalities on future arrivals, particularly if arriving agents are likely to have a higher expected valuation. Service-in-random-order alleviates this problem. Agents with lower beliefs have no incentive to join the line if the probability of being served early on is high enough. Moreover, allocating the resource to agents irrespective of the time at which they join the queue allows for serving some of those who have just learned that their prevailing valuation is high. This is beneficial because beliefs are mean-reverting, and the information acquired at the last service is the most valuable in the short run.

Second, I adopt an indirect approach to solve the unconstrained problem in which the designer is free to select any feasible and incentive-compatible queueing discipline, potentially inducing reneging. I first solve for the optimal menu of queueing disciplines. When pooling different types of agents is optimal—that is, when the optimal menu is a singleton—the designer's choice coincides with the solution to the constrained problem. When screening is optimal, the optimal menu is binary; service is rendered in order of arrival in one queue and in random order in the other. Agents who join the line to guarantee future service join the first-come first-served queue; agents who rejoin after having received good news about their valuation join the queue in which they are served in random order.

Finally, I show that the optimal menu can be (virtually) implemented by offering a single queueing discipline by having agents renege and, in particular, restart (renege and immediately rejoin) the queue. Informally, when an agent joins the queue, he is offered a waiting-time distribution, which describes the (random) time at which he is served if he does not renege. An agent is able to "engineer" for himself a different waiting-time distribution by leaving and joining the queue repeatedly over time. In principle, the class of distributions that can be generated by this behavior is limited. However, I show that the designer can tailor the waiting-time distribution (i.e., the queueing discipline) such that agents are incentivized to play an appropriate "restarting" strategy, and the efficiency of the optimal menu is virtually achieved.

These results are important for understanding the performance of different disciplines in a dynamic environment. For example, congestion problems arise in the allocation of specialized resources, such as prototype labs, within organizations; it has been noted[1] that queues are also costly because they postpone the time at which agents (in the example, developer teams) receive feedback, which is valuable since it informs their future decisions. In other setups, such as processing jobs on capacity-constrained computers and services provided via the Internet, agents can easily conceal their identity, meaning that the designer can only differentiate between them on the basis of their waiting-time.[2] Finally, institutional constraints or practical considerations sometimes limit the ability to discriminate across agents: for example, the scheduling of shared research facilities such as lab instruments is usually implemented via automated systems that identify each agent with a job to be

[1] See, for example, Thomke and Reinertsen (2012).

[2] This assumption is implicit in operations research models in which customers can cancel their service request and resubmit a new one. See Hassin (2016, Sec. 1.4, Chap. 4).

processed. A better understanding of the tradeoffs arising when consumers require service repeatedly and have an exploration motive can inform the design of priorities for service allocation in these settings.

Related Literature. Queue management motivated one of the earliest contributions to dynamic mechanism design, Dolan (1978). Since then, dynamic mechanism design has developed well beyond the context of queues, analyzing both the case of dynamic allocations and the case of dynamic private information,[3] with and without transfers. However, the vast majority of existing models studying the allocation of a resource to a pool of agents that changes over time, assumes that private information is static. On the contrary, in my model the agents' private information is endogenous, and the agents strategically choose the timing of their arrivals in the queue.

The paper contributes to three streams of literature. First, congestion externalities have been widely studied in the literature on strategic behavior in queues.[4] While the novel ingredient with respect to that literature is learning, my model also differs from the majority of rational queueing models in that agents are served multiple times. The idea that in a first-come first-served (FCFS) queue rational agents adopt suboptimal behavior dates back to Naor (1969); in that framework, Hassin (1985) shows that last-come first-served (LCFS) achieves the social optimum without the need for transfers. More recently, Platz and Østerdal (2016) find that in a concert queueing game,[5] FCFS and LCFS achieve the minimal and the maximal aggregate equilibrium payoff among all queueing disciplines, respectively.

With learning, the cost of congestion is twofold: beyond increasing the total queueing cost, congestion affects the value of the information collected at each service. Consequently, the ranking of different queueing disciplines depends on which of these two effects is stronger, and FCFS is sometimes optimal. Further, the results in Naor (1969) and Platz and Østerdal (2016) rely on the designer's ability to prevent restarting, whereas in my model, the designer cannot punish or detect reneging and restarting. As a result, LCFS cannot be part of an equilibrium.[6] In fact, my paper contributes to the scarce existing literature on strategic retrials and restarting.[7,8] The motivation and modeling choices of my paper are close to

[3]Krähmer and Strausz (2015) distinguish two approaches to introducing dynamics in mechanism design problems: dynamic private information and dynamic allocations. The overview by Bergemann and Pavan (2015) provides a finer classification of existing models when transfers are allowed.

[4]Hassin and Haviv (2003) and Hassin (2016) survey the literature on rational queueing in great detail.

[5]In a concert queueing game, agents choose their arrival time. Service is provided starting at a known opening time. In Platz and Østerdal (2016) service value decreases over time, waiting is costly, and agents cannot queue up before the opening time.

[6]As in Platz and Østerdal (2016), it is understood that under LCFS if agents arrive at a rate greater than capacity, at check-in each agent faces a lottery: he is served either immediately or at some later point in time.

[7]Hassin (2016, Chap. 4) discusses this literature; notice, however, that the chapter primarily concerns non-stationary models, while I focus on the steady state.

[8]However, in the operations research literature, queueing systems in which a fixed number of *non-strategic* agents are served repeatedly are common. See Bassamboo and Randhawa (2016) for a recent application to

those of Bassamboo and Randhawa (2016) who study scheduling policies in a queueing system with abandoning customers, abstracting from strategic considerations.

Second, the dynamic allocation of objects to agents arriving over time through waiting lists has been studied in the context of public housing and organ transplants. In these applications, however, agents have unit demand, and thus there is no scope for individual experimentation, and the flow of agents joining the pool is exogenous. Both Leshno (2015) and Bloch and Cantala (2016) consider the problem of allocating a sequence of heterogenous items to agents on a waiting list. Leshno (2015) assumes that agents' valuations are constant over time and the waiting is overloaded, and thus, there is always an agent who is assigned the object. Hence, the designer cannot affect aggregate waiting costs, and her objective is to minimize mismatches. Leshno (2015) examines different priority rules to serve waiting agents by focusing on mechanisms in which the designer asks an agent to report his preference only when she considers allocating him the good.[9] Contrary to Leshno (2015), Bloch and Cantala (2016) assume that agents' valuations evolve over time independently across periods. They consider a constant-size waiting list, and thus, wasting occurs when all agents reject the object. They show that first-come first-served always outperforms service-in-random-order. In my setting, the fact that the length of the queue is endogenously determined gives rise to a tradeoff between thickness and congestion that is absent from these two models.

Third, the interaction among agents who engage in individual experimentation has been studied by the strategic experimentation literature. Even if the general ideas are similar, owing to the continuum of agents, my model markedly differs from those in this literature. While most of the literature has focused on information externalities, which are absent in my model, Thomas (2014) analyzes a game of experimentation with congestion externalities. Cripps and Thomas (2016) investigate the interaction between information externalities and congestion externalities in a queueing model. Their paper is, however, only tangentially related to mine. In their model, agents arrive over time and there is a common source of uncertainty, the service rate of a server. Observational learning arises because the length of the queue and other agents' reneging decisions reveal their private information.

2 The Model

2.1 Setup

Time is continuous, indexed by $t \geq 0$, and the horizon is infinite. A designer (she) wishes to allocate a perishable and indivisible good to a unit mass of long-lived agents indexed by $i \in [0,1]$. Units of the good arrive at rate λ, so that a mass $\lambda (t'' - t')$ is to be assigned over any interval of time $[t', t'']$, $t'' > t'$.

the online DVD rental business. Formally, these models fall into the class of closed Jackson networks (see Chen and Yao (2013, Ch. 7) or Haviv (2013, Ch. 11)).

[9] The Disjoint Queues (DQ) mechanism analyzed by Leshno (2015) (see his Appendix E) relaxes this assumption. The DQ Mechanism is similar to a menu of queueing disciplines, but each of the queues is assumed to serve agents according to the FCFS discipline.

At all times, each agent chooses whether to queue or not. Namely, agent i's action at time t, a_t^i, is equal to 0 or 1, with the interpretation that $a_t^i = 1$ when he queues, and $a_t^i = 0$ when he does not. Letting N_t^i denote the total number of times agent $i \in [0,1]$ has been allocated the good by time t (formally, N_t^i is a counting process), feasibility requires that, for all $t' \geq 0$ and all $t'' > t'$,

$$\int_0^1 \left(\int_{t'}^{t''} \mathrm{d}N_t^i \right) \mathrm{d}i \leq \lambda \left(t'' - t' \right). \tag{1}$$

In addition, only agents who queue can be served. Formally, $\mathrm{d}N_t^i > 0$ only if $a_t^i = 1$.

Queueing entails a flow cost of $c \geq 0$, relative to not queueing. When allocated the good, agent i receives a lump-sum payoff equal to some state θ^i. The state θ^i can take two possible values θ_0 and θ_1, with $\theta_0 < 0 < \theta_1$. Agent i's state evolves unbeknown to him according to a continuous-time Markov chain $\left(\theta_t^i\right)_{t \geq 0}$ with state space[10] $\{\theta_0, \theta_1\}$, transition matrix $\left(\left(-\rho_0, \rho_0\right), \left(\rho_1, -\rho_1\right)\right)$, and initial probability of state θ_1 given by $\rho_0 / \left(\rho_0 + \rho_1\right)$. For any pair of agents, their individual state processes are assumed to be independent.

Given some integrable action process a_t^i, the realization of the state process θ_t^i, and individual allocation process N_t^i, the realized payoff of agent i is given by the long-run average

$$\limsup_{T \to \infty} \frac{1}{T} \left(\int_0^T \theta_t^i \, \mathrm{d}N_t^i - \int_0^T c \, a_t^i \, \mathrm{d}t \right).$$

This payoff has two components. The first is the sum of lump-sum payoffs collected at each consumption experience. The second is the total cost borne by the agent while queueing. Note the absence of discounting.

2.1.1 Information and Strategies

Each agent i observes his payoff, and hence he perfectly learns his prevailing state θ_t^i at each consumption experience.[11] Agents do not observe the queue.[12] Let $\left(\mathcal{F}_t^i\right)_{t \geq 0}$ be the filtration corresponding to the information of agent i.[13]

An agent's strategy specifies when to join and leave the queue. Because time is continuous, the formal definition requires some care. In particular, an agent who leaves the queue at time t, either because he reneges[14] or because he is served at that time, may want to

[10] Because this paper focuses on steady-state analysis, using the subscript t to denote calendar time, so that $\theta_t^i \in \{\theta_0, \theta_1\}$, will hopefully not create any confusion.

[11] I use the terms *service* (being served) and *consumption experience* (being allocated the good) interchangeably.

[12] Appendix A.2 shows that in this setup one can dispense with this assumption without affecting the results. However, most models with observable queues focus on observational learning, unlike my paper. Hence, I maintain the assumption in the main body of the paper.

[13] Formally, for any $t \geq 0$, $\mathcal{F}_t^i$ is the σ-algebra generated by $\left(N_s^i\right)_{0 \leq s \leq t}$, $\left\{\theta_s^i, s \leq t : \mathrm{d}N_s^i > 0\right\}$, and $\left(a_s^i\right)_{0 \leq s \leq t}$.

[14] In the operations research terminology, *reneging* refers to the act of leaving a queue after joining it.

rejoin the queue with no delay. To allow for this possibility, strategies are defined as impulse controls.[15]

Informally, if at time t, the agent is not queueing, a (pure) strategy dictates the time τ at which he joins the line. If queueing at t, a (pure) strategy is a time $\tau > t$ with the interpretation that the agent leaves the queue at τ if he has not been served earlier. That is, a pure strategy specifies the time at which the agent reneges conditional on the event $\left\{T^i(t) \geq \tau\right\}$, where $T^i(t) := \inf\left\{t' > t : \mathrm{d}N^i_{t'} > 0\right\}$ is the (random) time of the first service after t. In this case, the strategy also prescribes the action to be taken when reneging at τ, rejoining or not. Finally, a strategy prescribes whether to rejoin the queue or not at any time the agent is served, that is, at any time t such that $\mathrm{d}N^i_t > 0$.

Formally, an agent's strategy is a double sequence of random times (the action times) and random variables:

$$\sigma = \left\{(\tau_k, a_k)\right\}_{k=1}^{\infty},$$

where

(i) $0 \leq \tau_1 < \tau_2 < \ldots,$

(ii) for any $k \in \mathbb{N}$,[16]

$$\tau_k = \left\{t > \tau_{k-1} : \mathrm{d}N^i_t > 0\right\} \wedge \tilde{\tau}_k,$$

where $\tilde{\tau}_k$ is a predictable[17] stopping time adapted to the filtration $\left(\mathcal{F}^i_t\right)_{t\geq 0}$,

(iii) for any $k \in \mathbb{N}$, $a_k \in \{0, 1\}$ is a $\mathcal{F}^i_{\tau_k}$-measurable random variable,

(iv) for any $k \in \mathbb{N}$, if $a_k = 0$, then $a_{k+1} = 1$ a.s. (the action at τ_k impacts the outcome).

With some abuse of notation, a strategy defines an action process $\left(a^i_t\right)_{t\geq 0}$ taking value in $\{0, 1\}$ which is piecewise constant, as well as a piecewise continuous process $t - \sup_{\tau_k \leq t} \tau_k$ describing the amount of time elapsed since the agent last changed his action.

2.1.2 Queueing Discipline

At any t, the designer observes the aggregate distribution of the agents' actions and, for each queueing agent, the time elapsed since he last joined the queue. Crucially, I assume

[15] Stokey (2008) summarizes economic applications of impulse control models when the decision maker faces a lump sum cost to adjust the state variable under his control, so that the optimal adjustments made are lumpy. In my setup, lumpiness arises from the finite action space; in this sense, even if strategies are formally impulse controls, the idea is closer to the literature on timing games (see footnote 16).

[16] Because of condition (ii), my definition of strategy bears some similarity that of Rosenberg et al. (2013). In a two-player exit game with private learning, the authors model the idea that each player has two information nodes at each instant, by allowing a player to exit with *no* delay when the opponent exits.

[17] Predictability enforces the informational restriction that, when queueing, an agent chooses the stopping time τ at which he reneges conditional on the event $\left\{T^i(t) \geq \tau\right\}$.

that agents joining the queue at the same time cannot be distinguished by the designer: the time spent in line since each agent last started queueing is the only information about agents' histories available to her. Specifically, at any time t and for any agent i, the designer observes the current action a_t^i and, if the agent is queueing the time in queue, $t - \sup_{\tau_k \leq t} \tau_k$. She does not observe either the individual allocation $\left(N_s^i\right)_{s<t}$ or the past realized payoffs. It is convenient to define the time-in-queue w_t^i as the product between the time elapsed since agent $i \in [0,1]$ last adjusted his action and the current action,

$$w_t^i := a_t^i \cdot \left(t - \sup_{\tau_k \leq t} \tau_k \right),$$

so that the information about agent i available to the designer at time t is summarized by a_t^i and w_t^i.[18]

The designer commits to a queueing discipline. Informally, a queueing discipline dictates how the good is allocated across agents in the queue on the basis of the designer's information. Given a queueing discipline, agents engage in a queueing game; the goal of the designer is to choose a queueing discipline to maximize aggregate payoffs in some Nash equilibrium of the induced game. Given an allocation process N_t^i, the realization of the state process θ_t^i and some integrable action process a_t^i for any $i \in [0,1]$, the realized aggregate payoff is given by

$$\int_0^1 \left(\limsup_{T \to \infty} \frac{1}{T} \left(\int_0^T \theta_t^i \, \mathrm{d}N_t^i - \int_0^T c \, a_t^i \, \mathrm{d}t \right) \right) \mathrm{d}i \, .$$

There are no transfers. Since I focus on steady-state equilibria, the general definition of queueing discipline is relegated to Appendix A.1. In fact, in steady state, the problem of the designer can be stated as a choice over offered waiting-time distributions.

The offered waiting-time distribution $H : \mathbb{R}_+ \to [0,1]$ is the distribution of the time an agent that checks in and never reneges would wait before being served, in steady state. It is an equilibrium object: it relates to the queueing discipline and to the aggregate distribution of other agents' behavior. The problem of the designer can be stated in standard fashion: she chooses a distribution H and a strategy profile σ^i, $i \in [0,1]$, subject to some incentive compatibility and feasibility constraints. With an abuse of terminology, I sometimes refer to H as the queueing discipline. I posit that the probability distributions that can be induced by some queueing discipline together with a profile of strategies σ^i, $i \in [0,1]$, belong to $\mathcal{H}$, as defined below.

Definition 1. *The set $\mathcal{H}$ is the set of cumulative distribution functions, $H : \mathbb{R}_+ \to [0,1]$, such that $H(0) = 0$ and $\int t \, \mathrm{d}H(t) < \infty$.*

The requirement that H should not have atoms at 0 is necessary to guarantee that the allocation processes N_t^i, $i \in [0,1]$, are well defined for all strategy profiles of the agents. As I discuss in Section 4 and Section 5, there is no loss in restricting attention to distributions having finite mean. However, this assumption is necessary for the results of Section 3.

[18] The information available to the designer is described by the filtration $(\mathcal{G}_t)_{t \geq 0}$ which is formally defined in Appendix A.1.

2.2 Equilibrium

A given queueing discipline defines a game between agents.[19] Here, I define a (steady-state) equilibrium of this game, given such a queueing discipline.

Without loss, attention is restricted to (non-stationary) Markov strategies. In fact, the payoff of an agent is affected by the behavior of other agents only through the aggregate distribution of their actions. Therefore, as explained in Appendix A.2, an agent's best-reply problem can be stated as a Markov decision problem. I now introduce the relevant state variables.

First, at any time, an agent's information about his current state is summarized by the belief that he attaches to it being equal to θ_1. Specifically, at any t, agent $i \in [0,1]$ entertains a posterior belief

$$p_t^i := \Pr\left[\theta_t^i = \theta_1 \mid \mathcal{F}_t^i\right].$$

As long as he is not served, the belief evolves according to (to the first order)

$$\mathrm{d}p_t^i = \left(\left(1 - p_t^i\right)\rho_0 - p_t^i \rho_1\right)\mathrm{d}t.$$

Specifically, for all $t \geq 0$ and all $t' > t$, such that no service occurs from t to t',

$$p_{t'}^i = e^{-(\rho_0+\rho_1)(t'-t)} p_t^i + \left(1 - e^{-(\rho_0+\rho_1)(t'-t)}\right) \frac{\rho_0}{\rho_0 + \rho_1}. \tag{2}$$

Equation (2) makes plain that the belief of agent i is a convex combination of his past belief p_t and the invariant probability of θ_1, $\rho_0/(\rho_0+\rho_1)$. Along the history with no service, the posterior belief that the state is θ_1 converges to $\rho_0/(\rho_0+\rho_1)$. As soon as the agent is served, his belief jumps to 1 or 0.

Because the probability of being served depends on the agent's current action and, potentially, on how long he has queued, the current action and the time-in-queue are the other state variables.

Appendix A.2 shows that it is without loss to assume that agents' strategies are Markov in $\left(t, p_t^i, a_t^i, w_t^i\right)$ (calendar time, posterior belief, current action, and time-in-queue). Specifically, at any time t such that $t \geq \tau_{k-1}$, the next action time τ_k and the action taken at that point a_k are independent of what occurred at $t' < t$. A stationary Markov strategy is a Markov strategy that does not depend on calendar time. Let Σ be the set of stationary Markov strategies.

At any time t, the joint distribution of state variables (p, a, w) for the unit mass of agents is described by $M_t : [0,1] \times \{0,1\} \times \mathbb{R}_+ \to [0,1]$;[20] the queueing discipline and agents' strategies jointly determine the evolution of M_t. Let $\mathcal{M}$ denote the set of distributions over $[0,1] \times \{0,1\} \times \mathbb{R}_+$.

[19] Specifically, the game is an anonymous sequential game (Jovanovic and Rosenthal (1988)).

[20] That is, for any vector $(p, a, w) \in [0,1] \times \{0,1\} \times \mathbb{R}_+$ the fraction of agents $i \in [0,1]$ such that $p_t^i \leq p, a_t^i \leq a$, and $w_t^i \leq w$ equals $M_t(p, a, w)$.

I focus on symmetric steady-state equilibria. In a steady state, each agent $i \in [0,1]$ chooses his strategy σ^i to maximize

$$\mathbb{E}_{\sigma^i,H}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t\right)\right],$$

where $H \in \mathcal{H}$ is the offered waiting-time distribution. Note that an agent's payoff depends on his opponents' actions only through their effect on the waiting-time distribution, H.

Definition 2. *A symmetric steady-state equilibrium is a tuple $(\sigma, H, M) \in \Sigma \times \mathcal{H} \times \mathcal{M}$, such that*

1. *(Optimality) The strategy $\sigma^i = \sigma$ maximizes*

$$\mathbb{E}_{\sigma^i,H}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t\right)\right];$$

2. *(Stationarity) The distribution $M_t = M$ is independent of t.*

Note that, by definition, agents play stationary Markov strategies, and the equilibrium is symmetric.

3 Best Replies

I begin by analyzing the best-reply problem. In steady state, given a queueing discipline and the strategies of the other agents, an agent faces a time-independent offered waiting-time distribution. Hence, throughout the section, I fix a distribution $H \in \mathcal{H}$ and an agent i, although the index i is usually omitted. Throughout the section, I assume that given $H \in \mathcal{H}$, the agent has a best reply that yields strictly positive payoffs. As it turns out, there is no need to discuss the case in which this assumption fails. Not only can an agent assure himself a payoff of 0 by never queueing, but the designer can always guarantee strictly positive payoffs in equilibrium. (See Lemma 13 in Appendix B.)

Section 3.1 shows that one state variable suffices to describe stationary Markov strategies.Given this observation, in Section 3.2, I show that the payoff from any stationary Markov strategy can be written as a function of a few sufficient statistics. In Section 3.3 I first introduce two classes of strategies, and then show that any best reply is in one of these two classes.

3.1 Reducing to One State Variable

The agent faces a Markov decision problem with state variable $(p, w, a) \in [0,1]\times\mathbb{R}_+\times\{0,1\}$, his belief, his time-in-queue, and his action. In this section, I show that a stationary Markov strategy can be described in terms of one state variable only, the posterior belief.

First, by definition, the time-in-queue equals zero at any action time τ_k, i.e., $w_{\tau_k} = 0$. Hence, it is without loss to assume that the action taken at τ_k is only a function of the posterior belief at that time. Second, I argue that at any $t \geq \tau_{k-1}$, the next action time τ_k can be written as an exit time of the posterior belief from an open set.

To see this, I distinguish two cases: either $\tau_k = T(\tau_{k-1})$, or $\tau_k < T(\tau_{k-1})$, where $T(\tau_{k-1})$ is the first time the agent is served after τ_{k-1}. In the first case, at τ_k, the posterior belief jumps to either 0 or 1 and hence exits from any open subset of the unit interval. In the second case, what occurs between τ_{k-1} and τ_k provides no relevant information: p_t and w_t evolve deterministically and monotonically over $[\tau_{k-1}, \tau_k]$. As a result, τ_k can be expressed as the exit time of the belief process from some open set. Further, as $w_{\tau_{k-1}} = 0$, it is without loss to take this set to be only a function of the belief at $p_{\tau_{k-1}}$.

This observation has an immediate implication: the belief process induced by a strategy is a Markov process. As will become clear, it is a regenerative process.[21] A regenerative process has the property that there exist points in time, called regeneration points, when the process probabilistically restarts itself. The successive return times to a fixed belief constitute the regeneration points of the belief process. Specifically, for a fixed $p \in [0,1]$, define $\tau_p(n) \in (0, \infty]$, $n \in \mathbb{N}$, recursively as

$$\tau_p(0) = \inf\{t \geq 0 : p_t = p\}, \qquad \tau_p(n+1) = \inf\{t > \tau_p(n) : p_t = p\}, \text{ for } n \in \mathbb{N}.$$

Given a fixed p, and any $n \in \mathbb{N}$, the belief process after $\tau_p(n)$, $(p_t)_{t \geq \tau_p(n)}$, has the same distribution as the belief process after $\tau_p(0)$, $(p_t)_{t \geq \tau_p(0)}$. Hence, one can appeal to renewal theory to compute payoffs, as explained in Section 3.2.

Notice that, strictly speaking, this description of strategies (in terms of the posterior belief only) does not specify the behavior of an agent following off-path histories.[22] With respect to payoffs, this is innocuous in my environment. Even if the agent were to start the game with an arbitrary posterior belief, the transient component of the payoff (that is, the payoff collected before reaching the recurrent path) does not affect the (long-run average) payoff.[23]

3.2 Sufficient Statistics

In this section, I identify a finite number of sufficient statistics that determine the payoff from any stationary Markov strategy. Let $\sigma \in \Sigma$ be fixed throughout.

The payoff process induced by σ can be expressed as a function of the belief process. This follows from two observations. First, as discussed in Section 3.1, without loss, the action is only a function of the belief, as is the flow cost incurred by the agent. Second, the times at which the belief jumps coincide with the times at which the agent collects lump sums. As a result, the payoff process (as the belief) can be decomposed into cycles, and

[21] A good reference on renewal theory is Asmussen (2008).

[22] That is, the strategy does not specify behavior after one's own deviation. For a discussion of equilibrium refinement, see Section A.2.1.

[23] There is a small caveat: there exist strategies such that, given a starting belief p_0, $\mathbb{E}[\tau_{p_0}(0)] = \infty$. I discuss these "absorbing" strategies at the end of the proof of Lemma 1, in Appendix B. Absorbing strategies cannot be part of an equilibrium, and hereafter I focus on non-absorbing strategies.

heuristically, the long-run average payoff is equal to

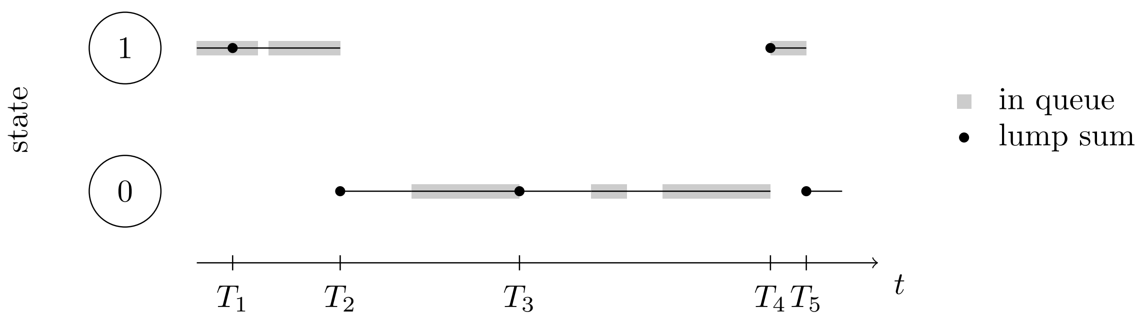
$$\frac{\mathbb{E}\left[\text{total payoff within a cycle}\right]}{\mathbb{E}\left[\text{cycle length}\right]}.$$

I make this intuition precise in Lemma 1. The idea is to construct a semi-Markov process by sampling the belief process $(p_t)_{t\geq 0}$ at appropriately chosen time points. The natural choice of sampling points are the service times for agent i, defined as

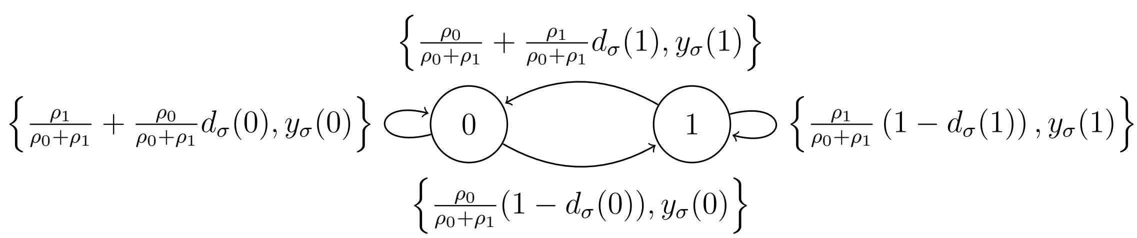
$$\{T_n\}_{n=1,2,\ldots} \coloneqq \{t \in [0,\infty) : \mathrm{d}N_t > 0\}.$$

Set $\pi_t \coloneqq p_{T_n}$ for $t \in [T_n, T_{n+1})$. I refer to π_t as the minimal semi-Markov process.

The top panel of Figure 1 depicts a path for the state evolution of the process π_t, while the bottom panel provides its symbolic representation. The semi-Markov process π_t takes values in $\{0,1\}$. It is a convenient tool to compute the payoff from $\sigma \in \Sigma$. The strategy σ determines the transition probabilities for the semi-Markov process, and the total payoff collected by an agent at each visit to the two states.



(a) Typical sample path. Dots represent service times, and hence correspond to transitions of the semi-Markov process. Notice that the state of the semi-Markov process does not necessarily change at each transition.



(b) Semi-Markov Representation. The expression in brackets represents the transition probability, and the average sojourn time. The computation of these variables can be found in the proof of Lemma 1.

Figure 1: Semi-Markov process.

Specifically, the transition probabilities for the minimal semi-Markov process are determined by the strategy σ and the waiting-time distribution H via two pairs of statistics, defined next. I defer the interpretation of these quantities to Section 3.3, where I focus on a particular class of strategies. For $j = 0,\, 1$, given $\tau \in \{\tau_j(n), n \geq 0\}$, define the functions $y_\sigma(j)$, $y_\sigma : \{0,1\} \to (0,\infty]$, and $d_\sigma(j)$, $d_\sigma : \{0,1\} \to [0,1)$,

$$y_\sigma(j) \coloneqq \mathbb{E}_{\sigma,H}\left[T(\tau) - \tau\right], \qquad d_\sigma(j) \coloneqq \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)}\right].$$

Turning to the reward structure, given $\tau \in \{\tau_j(n), n \geq 1\}$, let $q_\sigma(j)$, $q_\sigma : \{0,1\} \to (0,\infty]$,

$$q_\sigma(j) \coloneqq \mathbb{E}_{\sigma,H}\left[\int_\tau^{T(\tau)} a_t \,\mathrm{d}t\right],$$

be the expected time that an agent spends in queue before the next service, starting from the moment the belief reaches j. Plainly, at each visit in state $j = 0, 1$, the reward collected equals $\theta_j - c\, q_\sigma(j)$.

Having defined the transition and reward structure of the semi-Markov process induced by a strategy σ, the computation of payoffs is standard. Lemma 1 summarizes the result. The proof is in Appendix B.

Lemma 1. *Fix $H \in \mathcal{H}$, and $\sigma \in \Sigma$. The payoff*

$$\mathbb{E}_{\sigma,H}\left[\limsup_{T\to\infty} \frac{1}{T}\left(\int_0^T \theta_t \,\mathrm{d}N_t - \int_0^T c\, a_t \,\mathrm{d}t\right)\right],$$

and the individual service rate[24]

$$\lim_{t\to\infty} \frac{1}{t} N_t,$$

are a function of y_σ, d_σ, q_σ only.

Notice that the law of the counting process N_t is jointly determined by H and σ, but for notational simplicity, I keep such dependence implicit.

3.3 Reneging and Abandoning

Depending on the waiting-time distribution, an agent may benefit from leaving the queue before being served. To take a specific example, consider a waiting-time distribution with a piecewise constant hazard rate and one downward jump at some $t > 0$. If the agent were to renege and immediately rejoin whenever his time-in-queue equals t, he would effectively be served according to an exponential distribution with rate equal to the hazard rate before t. More generally, it is intuitive to expect that if a waiting-time distribution has a decreasing

[24] As shown in the proof, $\lim_{t\to\infty} \frac{1}{t} N_t$ converges a.s.

hazard rate, the best reply involves reneging.[25] Note that this reasoning hinges on the fact that the designer is unable to detect restarting (the combined action of reneging and rejoining).[26]

Two classes of stationary Markov strategies are of particular interest: the class of non-reneging strategies and the (larger) class of non-abandoning strategies. As they play a crucial role in the analysis, this section defines non-reneging and non-abandoning strategies, and describes their properties. As explained, in this case, the sufficient statistics are easy to interpret. Proposition 1 states the main result of the section: any best reply is necessarily a non-abandoning strategy. Recall that $T(\tau)$ is the first time the agent is served after τ.

Definition 3. *A strategy $\sigma \in \Sigma$ is non-reneging if $a_k = 1$ implies $\tau_{k+1} = T(\tau_k)$.*

In other words, according to a non-reneging strategy, an agent never voluntarily leaves the queue after joining it; he leaves the queue only when he is dismissed upon service. A strategy according to which the agent never joins the queue is trivially a non-reneging strategy.

Definition 4. *A strategy $\sigma \in \Sigma$ is non-abandoning if $a_k = 1$ implies either $\tau_{k+1} = T(\tau_k)$ or $a_{k+1} = 1$, or both.*

In other words, an agent playing a non-abandoning strategy rejoins the queue immediately whenever he reneges: he restarts, but he never abandons the queue.

Let Σ^{NR} and Σ^{NA} denote the set of non-reneging and non-abandoning strategies, respectively. Note that $\Sigma^{NR} \subset \Sigma^{NA}$.

Non-abandoning strategies have the feature that after joining the line, an agent queues uninterruptedly until he is served. This property simplifies the computation of the sufficient statistics, as I explain in the remainder of the section.

To understand the belief dynamics induced by non-abandoning strategies, it is useful to construct a finer semi-Markov representation of the belief process. Let $\underline{p}_\sigma$ and $\overline{p}_\sigma$ be the lowest and highest long-run belief at which the agent joins the queue when playing σ.

$$\underline{p}_\sigma := \liminf_{k\to\infty} \{p_{\tau_k} : a_k = 1\}, \qquad \overline{p}_\sigma := \limsup_{k\to\infty} \{p_{\tau_k} : a_k = 1\}.$$

Proceeding as in the previous section, I construct a semi-Markov process by sampling the belief process $(p_t)_{t\geq 0}$ at $\bigcup_{j\in\{0,\underline{p}_\sigma,\overline{p}_\sigma,1\}} \{\tau_j(n)\}_{n=0}^{\infty}$, the return times to the beliefs 0, $\underline{p}_\sigma$, $\overline{p}_\sigma$, and 1. Figure 2 provides a symbolic representation of the semi-Markov process generated by this sampling.[27] According to the properties of non-abandoning strategies, this representation divides each of the states of the minimal process into two states according

[25] This intuition is formalized in Proposition 2.

[26] Restarting is discussed in, for example, Hassin (1985), who observes that last-come first-served leads to socially optimal behavior in Naor's model. An important assumption of that paper is that the designer is able to detect and prevent this behavior. See Platz and Østerdal (2016) for a similar result.

[27] The semi-Markov process in Figure 2 can be constructed for any $\sigma \in \Sigma$, not only for $\sigma \in \Sigma^{NA}$. The transition structure in Figure 2 is valid more generally. However, the reward structure is specific to non-abandoning strategies.

to the action played, queueing or not. When playing $\sigma \in \Sigma^{NA}$, the agent queues at any $p_t \in [\underline{p}_\sigma, \overline{p}_\sigma]$, that is, when the process occupies states $\underline{p}_\sigma$ and $\overline{p}_\sigma$, and does not queue otherwise.

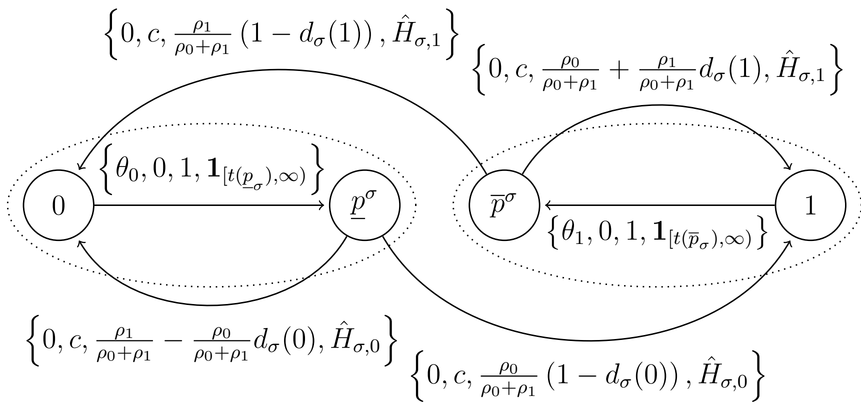


Figure 2: Semi-Markov representation. The expression in brackets represents the lump-sum payoff, the cost rate, the transition probability and the sojourn time distribution. This representation divides each of the states of the minimal semi-Markov process: each of the dotted ellipses corresponds to a state in the minimal process.

An immediate implication is that, focusing on non-abandoning strategies, the reward structure complies with the standard formulation of semi-Markov decision processes (e.g., see Puterman, 1994, Ch. 11). The agent receives lump-sum payoffs in states 0 and 1, and incurs a cost at rate c as long as the semi-Markov process occupies state $\underline{p}_\sigma$ or $\overline{p}_\sigma$.

Before computing the value of the sufficient statics for non-abandoning strategies, I define two distributions that are important for the sequel and provide additional insights into the best-reply problem faced by the agents.

Let $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ be the distributions of the sojourn times in states $\underline{p}_\sigma$ and $\overline{p}_\sigma$ for the semi-Markov process in Figure 2. Formally,

$$\begin{aligned} \hat{H}_{\sigma,0}(t) &:= \Pr_{\sigma,H}\left[T(\tau) - \tau < t \mid p_\tau = \underline{p}_\sigma\right], \\ \hat{H}_{\sigma,1}(t) &:= \Pr_{\sigma,H}\left[T(\tau) - \tau < t \mid p_\tau = \overline{p}_\sigma\right]. \end{aligned} \tag{3}$$

To be clear, these distributions are statistics of the belief process induced by the strategy σ. I shall refer to $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ as the induced waiting-time distributions.

When σ is a non-abandoning strategy, conditional on not being served, the agent potentially restarts the queue many times. As a consequence, $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ are convolutions of countably many truncated versions of H. In the special case in which σ is a non-reneging strategy, both induced distributions are equal to the offered waiting-time distribution H.

I now explain how the three pairs of sufficient statistics in Lemma 1 can be expressed in terms of the distributions $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ and the threshold beliefs. For notational convenience, let $t(p)$, $t : (0,1) \to \mathbb{R}_+$ be the time required for the belief to increase from 0 to p, if

$p \leq \rho_0/(\rho_0+\rho_1)$, or to decrease from 1 to p if $p > \rho_0/(\rho_0+\rho_1)$; explicitly,

$$t\left(p\right) = \begin{cases} \frac{1}{\rho_0+\rho_1}\ln\left(\frac{\rho_0}{\rho_0(1-p)-\rho_1 p}\right) & \text{if } p \leq \frac{\rho_0}{\rho_0+\rho_1}, \\ \frac{1}{\rho_0+\rho_1}\ln\left(\frac{\rho_1}{-\rho_0(1-p)+\rho_1 p}\right) & \text{if } p > \frac{\rho_0}{\rho_0+\rho_1}. \end{cases}$$

First, the expected time that an agent spends in queue before the next service starting from the moment the belief reaches 0 and 1 equals the expected sojourn time in state $\underline{p}_\sigma$ and $\overline{p}_\sigma$, respectively. Formally,

$$q_\sigma(j) = \int t \, \mathrm{d}\hat{H}_{\sigma,j}(t), \text{ for } j = 0, 1.$$

Second, the statistic $y_\sigma(0)$ ($y_\sigma(1)$) is the sum of the expected sojourn times in states 0 and $\underline{p}_\sigma$ ($\overline{p}_\sigma$ and 1, respectively). This is a consequence of having divided each of the states of the minimal semi-Markov process. Because the semi-Markov process spends a deterministic amount of time in states 0 and 1, it follows that

$$\begin{aligned} y_\sigma(0) &= t(\underline{p}_\sigma) + q_\sigma(0), \\ y_\sigma(1) &= t\left(\overline{p}_\sigma\right) + q_\sigma(1). \end{aligned}$$

Finally, by the same reasoning, it holds that

$$\begin{aligned} d_\sigma(0) &= \int e^{-(\rho_0+\rho_1)(t(\underline{p}_\sigma)+t)} \, \mathrm{d}\hat{H}_{\sigma,0}(t), \\ d_\sigma(1) &= \int e^{-(\rho_0+\rho_1)(t(\overline{p}_\sigma)+t)} \, \mathrm{d}\hat{H}_{\sigma,1}(t). \end{aligned} \tag{4}$$

These observations imply that, when an agent plays a non-abandoning strategy, the sufficient statistics, and hence the payoffs, are functions of two threshold beliefs and of some summary statistics of each induced waiting-time distribution.

I now clarify the computation of the transition functions for the two semi-Markov processes and provide further interpretation of the sufficient statistics. Intuitively, given the law of motion of the belief in (2), an agent who joins the queue with a belief $\underline{p}_\sigma$ realizes a payoff of θ_1 at the next service with probability

$$\left(\int_0^\infty e^{-(\rho_0+\rho_1)t} \, \mathrm{d}\hat{H}_{\sigma,0}(t)\right)\underline{p}_\sigma + \left(1 - \int_0^\infty e^{-(\rho_0+\rho_1)t} \, \mathrm{d}\hat{H}_{\sigma,0}(t)\right)\frac{\rho_0}{\rho_0+\rho_1}. \tag{5}$$

By definition, this is also the transition probability from state $\underline{p}_\sigma$ to state 1 or, equivalently, the transition probability from state 0 to state 1 in the minimal semi-Markov representation. Combining (5) and (4), it is straightforward to show that the probability of transiting from state $\underline{p}_\sigma$ to state 1 equals

$$\frac{\rho_0}{\rho_0+\rho_1}\left(1 - d_\sigma(0)\right).$$

This formula is valid for any $\sigma \in \Sigma$ (see the proof of Lemma 1). It can be understood as follows. Starting from a belief of 0, the average time before the next service equals $y_\sigma(0)$. Because the belief evolves over time according to (2), the probability of receiving a lump-sum θ_1 at the next service (equivalently, of transiting to state 1) is a convex combination of the initial belief 0 and of the invariant probability of state θ_1, $\rho_0/(\rho_0+\rho_1)$. The weight attached to the initial belief in the convex combination is the expectation of the exponential function that appears in (2), and by definition it is equal to $d_\sigma(0)$. Because service never occurs at beliefs below $\underline{p}_\sigma$, the transition probability can be equivalently computed using equation (5), which has a similar interpretation.

By a symmetric argument, the transition probability from state $\overline{p}_\sigma$ to state 1 equals

$$\int_0^\infty \left(e^{-(\rho_0+\rho_1)t}\overline{p}_\sigma + \left(1-e^{-(\rho_0+\rho_1)t}\right)\frac{\rho_0}{\rho_0+\rho_1}\right)\mathrm{d}\hat{H}_{\sigma,1}(t) = \frac{\rho_0}{\rho_0+\rho_1} + \frac{\rho_1}{\rho_0+\rho_1}d_\sigma(1).$$

The distributions $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ affect the payoffs via two summary statistics. First, when joining the queue, the agent expects to bear a total waiting cost before the next service that is proportional to the first moment of one of these two distributions. Second, because beliefs evolve non-linearly, the expected valuation conditional on service is again a function of the moment generating function of $\hat{H}_{\sigma,0}$ or $\hat{H}_{\sigma,1}$ evaluated at $-(\rho_0+\rho_1)$. Focusing on non-abandoning strategies, when best-replying, the agent selects a pair of threshold beliefs, and tailors his restarting times in an attempt to affect the two summary statistics of each of the induced waiting-time distributions.

As will become clear, it suffices to restrict attention to a subclass of non-abandoning strategies. Proposition 1 states that, after reneging, an agent never finds it optimal to wait before rejoining. Moreover, whenever the belief jumps to 1, the agent always finds it optimal to immediately rejoin the queue. The proof is relegated to the Appendix.

Proposition 1. *If $\sigma \in \Sigma$ maximizes*

$$\mathbb{E}_{\sigma,H}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t\,\mathrm{d}N_t - \int_0^T c\,a_t\,\mathrm{d}t\right)\right],$$

then $\sigma \in \Sigma^{NA}$, and $\overline{p}_\sigma = 1$.

Given Proposition 1, when choosing a feasible and incentive-compatible pair $(\sigma, H) \in \Sigma \times \mathcal{H}$, the designer is restricted to $\sigma \in \Sigma^{NA}$. I employ an indirect approach to solve the designer's problem that allows me to restrict attention, in most of the analysis, to the more tractable set of non-reneging strategies Σ^{NR}.

4 Optimal Queueing Discipline Without Reneging

I address the designer's problem in two steps. First, in this section, I solve a constrained problem in which the designer is further restricted to select $\sigma \in \Sigma^{NR}$. The solution to the constrained problem illustrates the main tradeoffs and guides the subsequent analysis. Then, in Section 5, I adopt an indirect approach to solve the unconstrained problem.

The objective of the designer is to maximize aggregate equilibrium payoffs. Because, by definition, the equilibrium is symmetric, each agent achieves the same realized payoff. Hence, the aggregate payoff equals the payoff of a representative agent, denoted by i hereafter.

The designer faces the aggregate capacity constraint (1). As a result of the law of large numbers, this constraint can be written as a bound on the aggregate (long-run) service rate,

$$\int_0^1 \left(\lim_{t\to\infty} \frac{1}{t} N_t^i \right) \mathrm{d}i \le \lambda,$$

where the limit inside the integrand is understood in the sense of almost sure convergence. (For conciseness, this clarification is omitted hereafter.) More formally, an agent's belief is a Markov process, in steady state. The stationary distribution of the belief process coincides with the cross-sectional distribution of beliefs.[28] By the nature of the learning process, in steady state, the aggregate service rate equals the total "birth rate" of agents with extreme beliefs, 0 or 1.

The bulk of the analysis, in Sections 4.1, 4.2, and 4.3, consists of restating the optimization problem in a two-dimensional space and identifying the set of implementable pairs $(H, \sigma) \in \mathcal{H} \times \Sigma$ in this space. As explained in Section 4.4, solving for the optimal queueing discipline is then simply a matter of characterizing the designer's preferences over the implementable pairs.

4.1 Constrained Problem

As made precise below, the designer selects a waiting-time distribution and a strategy to maximize the aggregate payoff, subject to feasibility and incentive compatibility constraints. In the constrained problem, she is further restricted to choosing a non-reneging strategy. Formally, in the constrained problem, the designer solves

$$\text{[P]} \qquad \sup \mathbb{E}_{\sigma,H} \left[\limsup_{T\to\infty} \frac{1}{T} \left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t \right) \right],$$

over distribution functions $H \in \mathcal{H}$ and strategies $\sigma \in \Sigma$, subject to[29]

$$\sigma \in \arg\max_{\sigma^i \in \Sigma} \mathbb{E}_{\sigma^i,H} \left[\limsup_{T\to\infty} \frac{1}{T} \left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t \right) \right], \tag{IC}$$

$$\lim_{t\to\infty} \frac{1}{t} N_t^i \le \lambda, \tag{C}$$

$$\sigma \in \Sigma^{NR}. \tag{R}$$

[28] This is a result of the law of large numbers as formalized by Sun (2006). The construction of an appropriate product space is needed to overcome measure-theoretic problems arising from having a continuum of (almost everywhere pairwise independent) random variables.

[29] Recall that σ and H jointly determine the law of the counting process N_t^i, so that (C) places a constraint on the set of implementable pairs (σ, H).

Notice that even if the designer is constrained to non-reneging strategies, each agent, when best-replying, is not restricted to this class, as is clear from the incentive compatibility constraint (IC). The constraint (R) restricts the designer to waiting-time distributions H and strategies σ such that an agent finds it optimal not to renege; specifically, he must find it optimal to play the non-reneging strategy σ.

4.2 Non-reneging Constraint

Seemingly, requiring the designer to select $\sigma \in \Sigma^{NR}$ makes the incentive problem more difficult: she must provide incentives to play threshold strategies, and agents' incentives to do so depend on fine details of the distribution H. However the constraint (R) can be expressed as a simple condition on the distribution H. Proposition 2 formalizes this.

Definition 5. *A distribution $H \in \mathcal{H}$ is called NBUE (new better than used in expectation) if for all $w > 0$:*

$$\int_w^\infty (t - w)\, \mathrm{d}H(t \mid t > w) \leq \int_0^\infty t\, \mathrm{d}H(t) \, .$$

Let $\mathcal{H}^N \subset \mathcal{H}$ denote the set of NBUE distributions.[30]

Proposition 2.

1. *If $\sigma \in \Sigma^{NR}$ is optimal within Σ given $H \in \mathcal{H}$, then $H \in \mathcal{H}^N$.*[31]
2. *If $\sigma \in \Sigma^{NR}$ is optimal within Σ^{NR} given $H \in \mathcal{H}^N$, then there exists $H' \in \mathcal{H}^N$ such that*
   - *(i) σ is optimal within Σ given H';*
   - *(ii) (σ, H') and (σ, H) induce the same payoffs and the same service rate.*

Proposition 2 ensures that there is no loss in focusing on non-reneging strategies if the constraint (R) is replaced by the requirement that H be an NBUE distribution. In the proof, contained in Appendix B, I show that if the offered waiting-time distribution does not satisfy the NBUE property, an agent's best reply involves reneging.

An NBUE waiting-time distribution has the property that an agent, when queueing, faces an expected residual wait no longer that the average waiting time offered to "newcomers." Because expected valuations change over time, the failure of this property does not automatically imply that the agent wants to renege irrespective of his current belief. Queueing agents with beliefs below the invariant probability of state θ_1, $\rho_0/(\rho_0 + \rho_1)$, may benefit from being served later in the future, as their belief is increasing. However, it turns out that

[30] Strictly speaking, the concept of NBUE (see Shaked and Shanthikumar, 2007) is defined for any nonnegative distribution with finite mean, while $\mathcal{H}$ excludes those with atoms at 0.

[31] Actually, this statement is more general, as it holds for any distribution with nonnegative support, even with infinite mean. Clearly, if the average waiting time is infinite, an agent's best reply involves reneging.

agents who are becoming pessimistic about their individual state find it optimal to restart the queue whenever the waiting-time distribution does not satisfy the NBUE property.

As a consequence, $H \in \mathcal{H}^N$ is a necessary condition for agents to have incentives not to renege;[32] further, for any $H \in \mathcal{H}^N$, when agents play a non-reneging strategy $\sigma \in \Sigma^{NR}$, one can find an NBUE distribution $H \in \mathcal{H}^N$ that is "payoff and service rate equivalent."

4.3 Finite-Dimensional Problem

It follows from Section 3.3 and Proposition 1 that, when focusing on non-reneging strategies, the payoff and service rate are determined by two summary statistics of the distribution H and by one threshold belief only.

In this section, I characterize the set of implementable pairs of statistics in two steps. First, I identify the admissible set, that is, the set of pairs corresponding to some distribution $H \in \mathcal{H}^N$. Then, in Section 4.3.2, I characterize the set of summary statistics that satisfy the capacity constraint. The proofs for this section are in Appendix B.2.

It is convenient to define, for any distribution H, the following summary statistics,

$$\mu^H := \int_0^\infty t \, \mathrm{d}H(t) \,, \qquad \delta^H := \int_0^\infty e^{-(\rho_0+\rho_1)t} \, \mathrm{d}H(t) \,.$$

4.3.1 Admissible Set

For which pairs (δ, μ) does there exist a distribution $H \in \mathcal{H}^N$ such that $(\delta, \mu) = \left(\delta^H, \mu^H\right)$? This is what I call admissibility. By definition, the two statistics are the mean and moment generating function evaluated at $-(\rho_0 + \rho_1)$ of some NBUE distribution. Lemma 2 solves for the set of pairs satisfying this restriction.

Lemma 2. *The following are equivalent:*

1. *There exists a distribution $H \in \mathcal{H}^N$ such that*

$$\mu^H = \mu, \qquad \delta^H = \delta;$$

2. *it holds that $(\delta, \mu) \in \Gamma^{NR}$, where*

$$\Gamma^{NR} := \left\{ (\delta, \mu) \in (0,1) \times (0, \infty) : e^{-(\rho_0+\rho_1)\mu} \le \delta \le \mathbb{E}\left[e^{-(\rho_0+\rho_1)\mathrm{Exp}(\mu)} \right] \right\},$$

and $\mathrm{Exp}[\mu]$ *is an exponential random variable with mean μ.*

[32] Whether $H \in \mathcal{H}^N$ is also sufficient to induce non-reneging, and what is the characterization of the optimal strategy within Σ given any $H \in \mathcal{H}$ remain open questions. Nevertheless, as shown in Section 5, it is without loss to restrict attention to non-reneging strategies; for this case, I characterize the best reply in closed form.

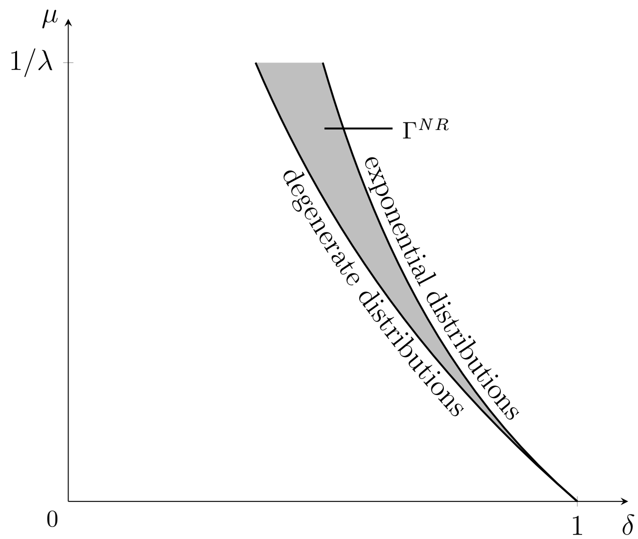


Figure 3: The set Γ^{NR} for $(\rho_0, \rho_1, \lambda) = (1/2, 1/2, 1)$.

The set Γ^{NR} is depicted in Figure 3. Note that, because the function $t \mapsto e^{-(\rho_0+\rho_1)t}$ is convex, the southwest boundary of the set corresponds to degenerate distributions, that is to distributions assigning probability 1 to some $\mu \in (0, \infty)$. The northeast boundary of Γ^{NR} corresponds to the set of exponential distributions; it is a well-known result in reliability theory that this family of distributions is "extreme" within the NBUE family in a sense made precise in the proof. From the perspective of incentives, when the waiting time is exponentially distributed, the non-reneging constraint is binding at all times: agents are served at a constant rate, independent of their arrival time in the queue.

4.3.2 Feasible Set

The feasibility constraint (C) can be restated as a restriction on the set of summary statistics. Before identifying the feasible set, I need to characterize, for any distribution H, the best reply of the agents, as it affects the actual throughput. Given Proposition 2, for $H \in \mathcal{H}^N$, the domain of the incentive constraint (IC) can be restricted to Σ^{NR}. In turn, I replace the strategy σ that appears in the feasibility constraint (C) with the unique solution to (IC) within the class of non-reneging strategies. By Proposition 1, when focusing on non-reneging strategies, it is without loss to restrict attention to cutoff strategies, that is, to strategies $\sigma \in \Sigma^{NR}$ such that $\overline{p}_\sigma = 1$. Let $\sigma_p \in \Sigma^{NR}$ be the cutoff strategy that sets $\underline{p}_\sigma = p$.

Define

$$p^*(\delta,\mu)=\begin{cases}0 & \text{if } \beta(\mu,\delta)\leq\alpha(\mu)-1<-1\,,\\ \dfrac{\rho_0}{\rho_0+\rho_1}\left(1-\dfrac{\beta(\delta,\mu)}{W_{-1}\left(e^{-1+\alpha(\mu)}\beta(\delta,\mu)\right)}\right) & \text{if } \beta(\delta,\mu)\in(\alpha(\mu)-1,0) \\ & \text{and } \alpha(\mu)<0,\\ \dfrac{\rho_0}{\rho_0+\rho_1} & \text{if } \beta(\delta,\mu)\geq 0\text{, or } \alpha(\mu)>0,\end{cases}$$

where W_{-1} is the (negative branch of the) Lambert function. Further, let

$$\alpha(\mu)=-(\rho_0+\rho_1)\frac{\theta_1-\theta_0}{\theta_1-c\,\mu}\mu,$$
$$\beta(\delta,\mu)=-\frac{\rho_0(\theta_1-c\,\mu)+(1-\delta)\rho_1(\theta_0-c\,\mu)}{\delta\rho_0(\theta_1-c\,\mu)}\,.$$

Lemma 3. *Fix a distribution $H\in\mathcal{H}$. There exists a unique best reply within Σ^{NR}. It is a cutoff strategy, with cutoff $p^*\left(\delta^H,\mu^H\right)$.*

Lemma 3 is a special case of Lemma 13, which is stated and proved in the Appendix. The optimal cutoff $p^*(\delta,\mu)$ is increasing in μ and increasing in δ if and only if

$$\frac{\rho_0}{\rho_0+\rho_1}\theta_1+\frac{\rho_1}{\rho_0+\rho_1}\theta_0-c\,\mu\geq 0.$$

I am now ready to introduce the set of feasible summary statistics.

Lemma 4. *The following are equivalent:*

1. *given $H\in\mathcal{H}$, and an optimal $\sigma\in\Sigma^{NR}$ given H, $\lim_{t\to\infty}\frac{1}{t}N_t\leq\lambda$;*
2. *it holds that $\left(\delta^H,\mu^H\right)\in\Gamma^\lambda$, where:*

$$\Gamma^\lambda:=\left\{(\delta,\mu)\in(0,1)\times\mathbb{R}_+:\mu\geq\frac{1}{\lambda}-\frac{(1-\delta)\frac{\rho_1}{\rho_0+\rho_1}t\left(p^*(\delta,\mu)\right)}{1-\delta+\delta p^*(\delta,\mu)}\right\}.$$

Figure 4 illustrates the set Γ^λ for two sets of parameters. The set Γ^λ is an upper contour set: for a fixed δ, the service rate (equal to the arrival rate at the queue in steady state) is decreasing in μ. Intuitively, an increase in the average waiting time μ has two effects: first, by having each agent wait longer before being served, it lowers the arrival rate; second, agents react by joining the queue at a higher belief, further reducing the arrival rate. Therefore, on the southern boundary of the set Γ^λ, denoted $\partial\Gamma^\lambda$, the constraint (C) is binding.

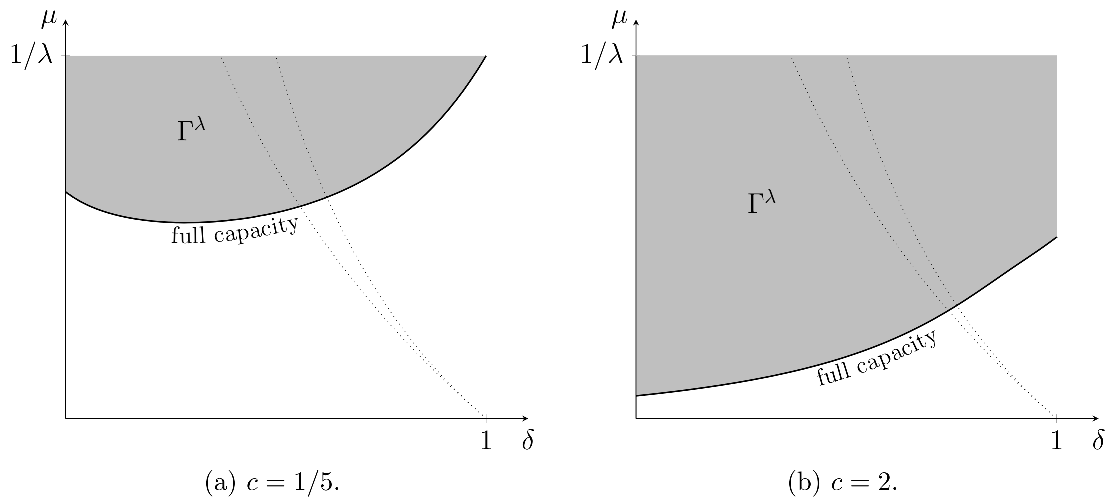


Figure 4: The set Γ^λ for $(\theta_1, \theta_0, \rho_0, \rho_1, \lambda) = (1, -3/4, 1/2, 1/2, 1)$. The dotted lines mark the boundary of the admissible set.

4.4 Optimal Queueing Discipline

Combining the previous results, I define the set of implementable summary statistics Γ as

$$\Gamma := \Gamma^\lambda \cap \Gamma^{NR}.$$

Having identified the admissible and feasible set of summary statistics, it remains to describe the preferences of the designer over this set. Before stating the main result, I revisit the designer's problem in terms of queueing disciplines and discuss her role in mitigating congestion.

The results are stated in terms of waiting-time distributions but have a natural interpretation as queueing disciplines. For ease of exposition, I state the relationship between queueing disciplines and waiting-time distributions in the following definitions.

Definition 6. *Agents are served according to a first-come first-served discipline if the waiting-time distribution H is a degenerate distribution that puts mass 1 on some $\mu > 0$.*

Definition 7. *Agents are served according to a service-in-random-order discipline if the waiting-time distribution H is an exponential distribution with support $[0, \infty)$.*

Definition 8. *Agents are served according to a service-in-random-order discipline with a minimum wait requirement if the waiting-time distribution has constant hazard rate and support $[t, \infty)$, for some $t > 0$.*

When the queueing discipline is first-come first-served, an agent joining the queue, in steady state, faces no uncertainty over the time of his next service. Under service-in-random-order, with or without a minimum wait requirement, queueing agents (with a time-in-queue greater than the minimum wait requirement t) are served pairwise independently at a constant rate.

A noteworthy consequence of Lemma 2 is that as far as payoffs are concerned, it suffices to consider only these three classes of queueing disciplines. In fact, the classes of corresponding waiting-time distributions span the set Γ^{NR}. If agents are served in order of arrival, the pair of summary statistics lies on the west boundary of Γ. If the waiting time is exponentially distributed, the pair of summary statistics lies on the east boundary of Γ. Finally, each point in the interior of Γ is achieved by a shifted exponential distribution that can be generated by serving agents in random order with a minimum wait requirement $t > 0$.

Because the designer maximizes the aggregate payoff and each agent achieves the same realized payoff in equilibrium, her preferences over the set Γ coincide with the preferences of each of the agents. However, there is scope for the intervention by a designer, because agents do not internalize the congestion effects of their actions. To shed light on the problem faced by the designer, I now present a payoff decomposition that highlights the source of the externality. (The formal derivation can be found in the Appendix.)

Fix a waiting-time distribution $H \in \mathcal{H}$. The payoff from a cutoff strategy $\sigma_p \in \Sigma^{NR}$ can be written as

$$\begin{aligned}\mathbb{E}_{\sigma_p,H}&\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t\right)\right]\\ &= \left(\lim_{t\to\infty}\frac{1}{t}\cdot N_t^i\right)\cdot\left(m\left(\delta^H,p\right)\theta_1 + \left(1-m\left(\delta^H,p\right)\right)\theta_0 - c\,\mu^H\right),\end{aligned} \tag{6}$$

where

$$m\left(\delta^H,p\right) := \frac{(1-\delta^H)\frac{\rho_0}{\rho_0+\rho_1} + \delta^H p}{(1-\delta^H)+\delta^H p},$$

and the service rate induced by σ and H equals

$$\lim_{t\to\infty}\frac{1}{t}\cdot N_t^i = \Lambda\left(\delta^H,\mu^H,p\right) := \frac{1}{\mu^H + (1-m\left(\delta^H,p\right))\,t(p)}.$$

As is apparent from (6), the payoff from the strategy σ_p is the product of the rate at which the agent is served and the expected total payoff collected between service times. The latter is a function of $m\left(\delta^H,p\right)$, the probability of being served when the state is θ_1. The relationship between $m\left(\delta^H,p\right)$ and $\Lambda\left(\delta^H,\mu^H,p\right)$ is easy to understand. From the elementary renewal theorem, the expected service rate equals the inverse of the average time between services. When joining the queue, the agent always expects to spend in line an amount of time μ^H. Moreover, after being served, he waits an amount of time $t(p)$ before joining the queue whenever the realized payoff is θ_0, which occurs a proportion $1-m\left(\delta^H,p\right)$ of the time.

An agent's cutoff choice affects the rate at which he is served $\Lambda(\delta^H,\mu^H,p)$. Because of scarcity, a "tragedy of the commons" arises. The designer must guarantee through an appropriate choice of a pair of summary statistics (δ, μ), that the service rate induced by the agents' best reply does not exceed the capacity λ.

As expected, the designer finds it optimal to allocate the full amount of the resource available. Moreover, any of the three queueing disciplines arises as optimal for some admissible set of parameters $(\theta_1, \theta_0, c, \rho_0, \rho_1, \lambda) \in \mathbb{R}_+ \times \mathbb{R}_- \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$.

Theorem 1.

(i) *For any solution* $(H^*, \sigma^*) \in \mathcal{H}^N \times \Sigma$ *to* [P], *the constraint* (C) *is binding, i.e.,* $(\delta^{H^*}, \mu^{H^*}) \in \partial\Gamma^\lambda$.

(ii) *For some* $t > 0$, *there exist admissible parameters* $(\theta_1, \theta_0, \rho_0, \rho_1, \lambda)$ *such that either first-come-first-served or service-in-random-order without a wait requirement or with a minimum wait requirement* t *is optimal for some* $c \geq 0$.

Proof. (i) I show that the designer's preferences over (δ, μ) are characterized by upward-sloping indifference curves. It then follows that $(\delta^H, \mu^H) \in \partial\Gamma^\lambda$, implying that (C) is binding. The designer's preferences coincide with the agents' preferences. Since agents are best-replying, the change in the threshold belief can be neglected (as follows from the envelope theorem). Then, given (6), because

$$\frac{\partial m(\delta, p)}{\partial \delta} = \frac{p}{((1-\delta)+\delta p)^2} \frac{\rho_0}{\rho_0+\rho_1} > 0,$$

$$\frac{\partial \Lambda(\delta, \mu, p)}{\partial \mu} = -(\Lambda(\delta, \mu, p))^2 < 0, \qquad \frac{\partial \Lambda(\delta, \mu, p)}{\partial \delta} = (\Lambda(\delta, \mu, p))^2 \frac{\partial m(\delta, p)}{\partial \delta} > 0,$$

the payoff is increasing in δ and decreasing in μ.

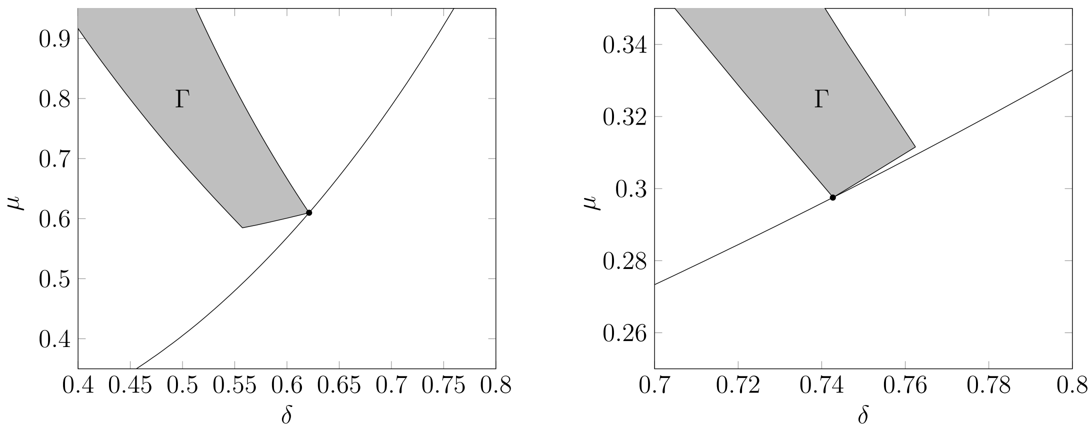
(ii) I construct an example (see Figure 5; details are presented in the Appendix) such that first-come first-served and service-in-random-order are uniquely optimal for two cost levels, c',c'', $c' > c''$, respectively. After replacing the cutoff p with the agents' best reply from Lemma 3, the designer's objective function is continuous in (δ, μ) and in c. Moreover, without loss, the domain of the program [P] can be restricted to some compact subset of Γ containing $\partial\Gamma^\lambda$. The fact that some other point along $\partial\Gamma^\lambda$ can arise as optimal for some $c \in (c', c'')$ follows from the maximum theorem. □

4.5 Discussion

4.5.1 Thickness vs. Congestion

The designer faces a tradeoff between minimizing the waiting cost, and maximizing the value of the service provided. The two-dimensional characterization of the implementable mechanisms allows me to distinguish the effect of these two forces.

By Theorem 2, the designer finds it optimal to allocate the full amount of the available resource. As a result, the first term in (6) equals λ, and the optimal pair $(\delta, \mu) \in \Gamma$



(a) Service-in-random-order is optimal for $c = 1/5$.

(b) First-come first-served is optimal for $c = 2$.

Figure 5: The optimal queueing discipline for $(\theta_1, \theta_0, \rho_0, \rho_1, \lambda) = (1, -3/4, 1/2, 1/2, 1)$.

maximizes the second term among all pairs in $\partial\Gamma^\lambda$. On the one hand, the statistic δ affects the probability of serving agents whose prevailing state is θ_1, $m\left(\delta, p^*(\delta, \mu)\right)$. On the other hand, the statistic μ determines the equilibrium length of the queue $\lambda\mu$, by Little's law.

Ideally, the designer would like to deter agents from joining the queue when their belief is low, and thereby allocate the capacity to agents with higher valuation. A higher δ helps achieve this goal. In fact, given $H \in \mathcal{H}$, an agent who joins the queue with a belief p, expects his valuation at the next service to be θ_1 with probability

$$\int_0^\infty \left(e^{-(\rho_0+\rho_1)s}p + \left(1 - e^{-(\rho_0+\rho_1)s}\right)\frac{\rho_0}{\rho_0+\rho_1}\right)\mathrm{d}H(s) = \delta^H p + (1-\delta^H)\frac{\rho_0}{\rho_0+\rho_1}.$$

Roughly, the higher δ^H is, the weaker the incentive to join at a low belief.

For a fixed μ, the exponential distribution $H(t) = 1 - e^{-t/\mu}$ maximizes δ^H among all NBUE distributions H. When the length of the queue has little impact on the aggregate payoff, this consideration resolves the tradeoff: sufficient conditions for service-in-random-order to be optimal are provided in the following lemma.[33] The proof is in the Appendix.

Lemma 5. *Assume that* $c = 0$ *and*

$$\frac{\rho_0}{\rho_0+\rho_1}\theta_1 + \frac{\rho_1}{\rho_0+\rho_1}\theta_0 \geq 0. \tag{7}$$

Then, service-in-random-order is optimal.

The left-hand side of (7) is a lower bound to the expected lump-sum payoff at each service: if an agent were to be served at random points in time, a lump-sum θ_1 would

[33] The assumption is used in Theorem 2.

accrue a fraction $\rho_0/(\rho_0+\rho_1)$ of the time. In the proof, the assumption guarantees that an increase in δ never induces agents to join at a lower belief. However, the scenario identified in Lemma 5 is extreme, as an agent can guarantee a strictly positive payoff without engaging in experimentation.

When $c > 0$, considerations about the queue length cannot be ignored: the average waiting time μ directly affects payoffs, and service-in-random-order is not always optimal, as formalized in Lemma 6.

Lemma 6. *Fix any admissible set of parameters $(\theta_1, \theta_0, c, \rho_0, \rho_1, \lambda)$. There exists a $c' > c$ for which service-in-random order is not optimal.*

When queueing is costly, serving agents whose state is likely to be θ_1 comes at the cost of a longer queue. Having a high proportion of agents who feed back in the line exacerbates congestion. When the waiting cost is high,[34] the designer finds it optimal to use a first-come first-served discipline, even if this implies forgoing the possibility to serve returning agents as soon as they rejoin. The next lemma formalizes the idea that the benefit from a first-come first-served queueing discipline comes from shortening the queue. It is an immediate consequence of the fact that indifference curves are upward sloping and the pair of statistics for degenerate distributions lies on the west boundary of the implementable set Γ.

Lemma 7. *If first-come first-served order is optimal, then it minimizes the average waiting time μ (equivalently, the queue length) among all the feasible disciplines.*

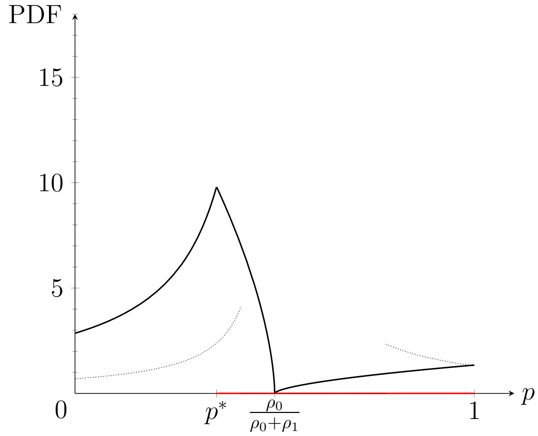
A failure of (7) is a sufficient condition for first-come first-served to minimize the queue length. When experimentation is needed to achieve positive payoffs, the feedback effect is unavoidable but not necessarily detrimental to welfare. Unfortunately, while it is easy to compute the optimal queueing discipline for any given set of parameters, tighter comparative statics appear elusive, because of the two-level maximization problem.

4.5.2 Steady-State Distributions

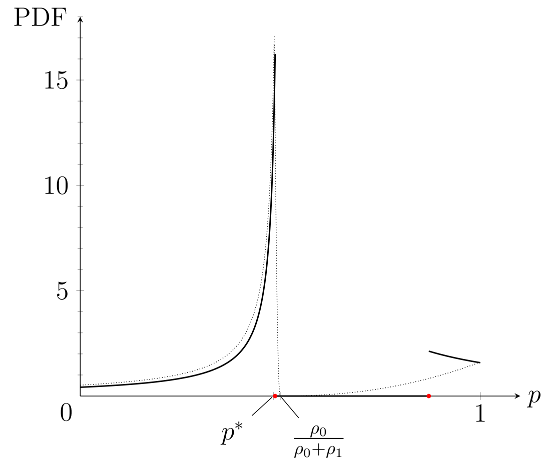
The characterization of equilibria in terms of a few statistics eliminates the need to compute the joint distribution of state variables (p, a, w) for the unit mass of agents. Nevertheless, the steady-state marginal distribution over posterior beliefs is easy to compute and provides additional insights into the trade-off highlighted in the previous section.

Figure 6 depicts the steady-state distribution of posterior beliefs generated by the optimal queueing disciplines, service-in-random-order and first-come-first-served respectively respectively, for two possible cost levels. When service is rendered in random order, the total capacity is allocated across all queueing agents. In this case, the density function peaks at the check-in beliefs 1 and p^*. Intuitively, service-in-random-order generates a thick pool of

[34] Because rescaling (θ_1, θ_0, c) amounts to rescaling payoffs but does not affect the implementable set Γ, increasing c is equivalent to decreasing the gain from targeting "high types" $\theta_1 - \theta_0$.



(a) Service-in-random-order discipline for $c = 1/5$.



(b) First-come first-served discipline for $c = 2$.

Figure 6: Density of the cross-sectional distribution of posterior beliefs for $(\theta_1, \theta_0, \rho_0, \rho_1, \lambda) = (1, -3/4, 1/2, 1/2, 1)$. In red, the range of beliefs at which agents are served. The dotted lines correspond to the distributions generated by the suboptimal disciplines.

potential customers, in the sense of a favorable distribution over the wide range of posterior beliefs entertained by served agents.[35]

When agents are served in order of arrival, there are only two beliefs at which they receive service. It is worth forgoing the possibility to serve agents as soon as they join when the distribution is likely to be unbalanced, that is, when there is a large proportion of agents with low beliefs joining the queue. Even if the distribution of beliefs is endogenous and depends on the queueing discipline, for a fixed discipline, a higher waiting cost induces agents to join at higher beliefs, giving rise to an unbalanced distribution. As a result, first-come first-served arises as optimal in these cases.

5 Allowing for Reneging

The solution to the designer's problem when she is not restricted to $\sigma \in \Sigma^{NR}$ is closely related to the design of an optimal menu. In this section, I employ an indirect approach to solve the unrestricted design problem. First, I extend the baseline model to the case in which the designer is allowed to offer a menu of queues. In Section 5.3, I characterize the optimal menu. Section 5.4 shows that the optimal menu can be "virtually" implemented with a single queueing discipline if reneging is allowed.

[35]Strictly speaking, the thickness of the market in the usual sense, that is, the number of transactions (in my model, the service rate), is not under the designer's control; here, instead, the idea of thickness relates to the fraction of agents of type θ_1 that are served.

5.1 Menus and Strategies

Because agents' posterior beliefs are private information, the designer faces a screening problem. Hence, efficiency can potentially be improved by the introduction of a menu. When offered a menu, an agent self-selects into one of multiple customer classes (i.e., he chooses to join one of several queues), with each class offering service according to a different waiting-time distribution.

Formally, a menu is a collection of queueing disciplines, each of which is indexed by some $Q \in \mathbf{Q}$, where the set $\mathbf{Q}$ is understood to be sufficiently large. As I show, it suffices to consider binary menus. For now, let a nonempty set $\mathbf{Q}$ be given.

5.1.1 Strategies

It is necessary to modify the definition of strategies, as compared to Section 2.1.1. In particular, at any action time τ_k, an agent now chooses both whether to queue or not, and, if he chooses to queue, which line to join.

Formally, a strategy for agent i, $i \in [0,1]$, is a sequence of random times and random variables,

$$\sigma = \{(\tau_k, a_k, Q_k)\}_{k=1}^{\infty},$$

where

(i) $0 \leq \tau_1 < \tau_2 < \ldots,$

(ii) for any $k \in \mathbb{N}$,

$$\tau_k = \{t > \tau_{k-1} : \mathrm{d}N_t > 0\} \wedge \tilde{\tau}_k,$$

where $\tilde{\tau}_k$ is a predictable stopping time adapted to the filtration $\left(\mathcal{F}_t^i\right)_{t\geq 0}$,[36]

(iii) for any $k \in \mathbb{N}$, $a_k \in \{0,1\}$ and $Q_k \in \mathbf{Q}$ are $\mathcal{F}_{\tau_k}^i$-measurable random variables,

(iv) for any $k \in \mathbb{N}$, if $a_k = 0$, then $a_{k+1} = 1$ a.s.

The new element in the definition of a strategy is the sequence $\{Q_k\}_{k\in\mathbb{N}}$. The behavior at τ_k is now described by a pair of random variables, a_k and Q_k. At any action time τ_k at which the agent arrives at the queue, i.e., $a_k = 1$, $Q_k \in \mathbf{Q}$ indicates which queue from the menu he joins.[37] Notice that the (piecewise-constant) action process defined by a strategy $\left(a_t^i, Q_t^i\right)_{t\geq 0}$ is now two-dimensional and takes values in $\{0,1\} \times \mathbf{Q}$.

[36] For any $t \geq 0$, $\mathcal{F}_t^i$ is now the σ-algebra generated by $(N_s^i)_{0\leq s\leq t}$, $\left\{\theta_s^i, s \leq t : \mathrm{d}N_t^i > 0\right\}$, and $(a_s^i, Q_s^i)_{0\leq s\leq t}$.

[37] The value of the random variable Q_k at those times at which $a_k = 0$ does not affect the outcome. Hereafter, when describing strategies, I forego specifying the values taken by Q_k at these times.

5.1.2 Menus

In line with Section 3.1, the designer observes the aggregate distribution of agents' actions, and, for each queueing agent, the time-in-queue; that is, for each $Q \in \mathbf{Q}$, she observes the mass of agents waiting in that queue and, for each of them, the time elapsed since the last time the agent joined that queue. As before, for any given agent, she does not observe past allocations, past realized payoffs or past actions. In this sense, both switching queues and restarting have the effect of resetting an agent's time-in-queue, and either of these actions makes the agent indistinguishable from a "newcomer."

The designer commits to a collection $\mathbf{Q}$ of queueing disciplines. In steady state, the designer's problem can be stated as a choice over collections of waiting-time distributions $\{H_Q : Q \in \mathbf{Q}\}$, $H_Q \in \mathcal{H}$, for all Q. As in the baseline model, the feasibility constraint requires that the aggregate service rate does not exceed λ. In this sense, a menu creates classes of customers, each of which is allocated part of the total capacity, but the maximal total capacity is unchanged.

5.1.3 Equilibrium

The definition of steady-state equilibrium requires minor changes. Following the same argument as before, without loss, attention can be restricted to strategies that are Markov in $\left(t, p_t^i, a_t^i, Q_t^i, w_t^i\right)$ (calendar time, posterior belief, current action, current queue, and time-in-queue).With abuse of notation, I denote by Σ the set of stationary Markov strategies, that is, strategies that do not condition on calendar time.Similarly, let $\mathcal{M}^+$ be the set of distributions over $[0,1] \times \{0,1\} \times \mathbf{Q} \times \mathbb{R}_+$. At any time t, the joint distribution of state variables (p, a, Q, w) is described by $M_t \in \mathcal{M}^+$, and the evolution of M_t is determined by the queueing discipline and the agents' strategies.

Definition 9. *A symmetric steady-state equilibrium is a tuple $(\sigma, \{H_Q : Q \in \mathbf{Q}\}, M) \in \Sigma \times \mathcal{H}^{\mathbf{Q}} \times \mathcal{M}^+$, such that*

1. *(Optimality) The strategy $\sigma = \sigma^i$ maximizes*

$$\mathbb{E}_{\sigma^i, \{H_Q : Q \in \mathbf{Q}\}} \left[\limsup_{T \to \infty} \frac{1}{T} \left(\int_0^T \theta_t^i \, \mathrm{d}N_t^i - \int_0^T c\, a_t^i \, \mathrm{d}t \right) \right], \tag{8}$$

2. *(Stationarity) The distribution function $M_t = M$ is independent of t.*

The strategy $\sigma^i \in \Sigma$ and the menu $\{H_Q : Q \in \mathbf{Q}\}$ define the distribution of the times the agent is served $(N_t^i)_{t \geq 0}$. Hence, the process Q_t induced by the strategy σ does affect expected payoffs, even if it does not explicitly appear in (8).

5.2 Binary Menus Suffice

Even if the actions available to the agents are now more complicated objects, the same techniques apply. The definition of the sufficient statistics (Section 3.2) and the definitions of non-reneging and non-abandoning strategies (Definitions 3 and 4) extend to this richer environment with no changes. Moreover, the sufficient statistics still determine the payoff and the service rate given any stationary Markov strategy. This section shows that attention can be restricted to binary menus. I begin by illustrating the reasoning and then state the result. Fix a menu $\{H_Q : Q \in \mathbf{Q}\}$.

First, the set of strategies can be partitioned into equivalence classes based on the three pairs of sufficient statistics identified in Lemma 1. Specifically, let any pair of strategies $\sigma, \sigma' \in \Sigma$ belong to the same equivalence class if $y_\sigma = y_{\sigma'}$, $d_\sigma = d_{\sigma'}$, and $q_\sigma = q_{\sigma'}$, so that strategies in the same class yield the same payoff and the same service rate.

Second, Proposition 1 remains valid: given the menu $\{H_Q : Q \in \mathbf{Q}\}$, any best reply is a non-abandoning strategy. (See the proof of Proposition 1 in Appendix B for the precise argument.)

Focusing on non-abandoning strategies, each equivalence class can be identified by a pair of induced waiting-time distributions. This follows from two observations. By Section 3.3, when an agent plays a non-abandoning strategy, the sufficient statistics are determined solely by the induced waiting-time distributions and the cutoff belief. In turn, for any pair of induced waiting-time distributions, there exists a unique optimal cutoff (see Lemma 13 in the Appendix). Consequently, it suffices to index equivalence classes with the pair of induced distributions.

Identifying each strategy $\sigma \in \Sigma$ with the induced distributions $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ is convenient because $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ are simple objects: they belong to $\mathcal{H}$, and they are convolutions of countably many truncated versions of the distributions H_Q, $Q \in \mathbf{Q}$.

Given these observations, a menu of queueing disciplines can be understood as a set of possible induced distributions. The designer solves what is essentially a static delegation problem in which she faces two types of agents. The two types ("high" and "low") differ in the realized payoff at the last service (θ_1 and θ_0) and, hence, in whether their belief is above or below the invariant probability of state θ_1, $\rho_0/(\rho_0 + \rho_1)$. Faced with a menu, each type of agent selects the preferred waiting-time distribution from those that can be "engineered" by repeatedly leaving and joining (potentially) different queues over time.

Combining these observations, it follows that it is without loss of generality to restrict attention to binary menus and non-reneging strategies. The next lemma, the proof of which is in the spirit of the revelation principle, states this formally.

Lemma 8. *If $\sigma \in \Sigma$ is optimal given $\{H_Q : Q \in \mathbf{Q}\} \in \mathcal{H}^{\mathbf{Q}}$, then there exists a binary menu $\{H_0, H_1\} \in \mathcal{H} \times \mathcal{H}$ and a strategy $\sigma' \in \Sigma^{NR}$ such that*

(i) *σ' is optimal within Σ given $\{H_0, H_1\}$;*

(ii) *under σ', for any $t \geq 0$, $Q_t = 1$ if and only if $p_t = 1$;*

(iii) *$(\sigma', \{H_0, H_1\})$ and $(\sigma, \{H_Q : Q \in \mathbf{Q}\})$ yield the same payoffs and the same service rate.*

Proof. Set $H_0 = \hat{H}_{\sigma,0}$, and $H_1 = \hat{H}_{\sigma,1}$, where $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ are defined in (3). If, when offered the menu $\{\hat{H}_{\sigma,0}, \hat{H}_{\sigma,1}\}$, the agent uses the best non-reneging strategy that prescribes choosing $\hat{H}_{\sigma,1}$ at check-in when the belief is 1 and joining $\hat{H}_{\sigma,0}$ otherwise, his payoff is unchanged. Moreover, any other strategy is suboptimal. In fact, when faced with the original menu, the agent is able to induce the waiting-time distributions $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ and, hence, also any possible convolution of truncations thereof. Yet, he finds it optimal to induce the distributions $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$. Hence, when best-replying to the binary menu $\{H_0, H_1\}$, he has a best reply $\sigma' \in \Sigma^{NR}$ satisfying (ii). □

Property (ii) is an incentive compatibility condition: when checking in, each agent finds it optimal to join the queue designed for his type. As attention is restricted to binary menus hereafter, I refer to strategies satisfying condition (ii) as incentive compatible strategies.

5.3 Optimal Menu

Given Lemma 8, the designer's problem can be stated as

$$\text{[M]} \qquad \sup \mathbb{E}_{\sigma,\{H_0,H_1\}} \left[\limsup_{T\to\infty} \frac{1}{T} \left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t \right) \right],$$

over binary menus $\{H_0, H_1\} \in \mathcal{H} \times \mathcal{H}$ and strategies $\sigma \in \Sigma^{NR}$, subject to

$$\sigma \in \arg\max_{\sigma^i \in \Sigma} \mathbb{E}_{\sigma^i,\{H_0,H_1\}} \left[\limsup_{T\to\infty} \frac{1}{T} \left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \,\mathrm{d}t \right) \right],$$

$$\lim_{t\to\infty} \frac{1}{t} N_t^i \leq \lambda,$$

$$(\hat{H}_{\sigma,0}, \hat{H}_{\sigma,1}) = (H_0, H_1). \tag{9}$$

It is worth noting that restricting the designer to a non-reneging strategy is without loss. In fact, the designer's problem could be stated as a choice among pairs of induced waiting-time distributions. The chosen pair of distributions would then correspond to an equivalence class of strategies: the constraint (9) is simply a criterion for selecting one strategy from this class. Next I characterize the optimal menu. The proof of Theorem 2 is contained in Appendix B.3. Recall that $H^* \in \mathcal{H}^N$ denotes the queueing discipline that solves the constrained problem [P].

Theorem 2. *There exists a solution $(\sigma, \{H_0^*, H_1^*\}) \in \Sigma^{NR} \times \mathcal{H}^N \times \mathcal{H}^N$. It is such that $\mu^{H_0^*} \geq \mu^{H_1^*}$, and either of the following holds:*

(i) *(pooling menu) $H_0^* = H_1^* = H^*$;*

(ii) *(separating menu) $H_0^* \neq H_1^*$, and*

  (a) *(FCFS/SIRO menu) H_0^* is degenerate at $\mu^{H_0^*}$ and H_1^* is exponential;*

  (b) *(low-type IC binds) any best reply to $\{H_1^*, H_1^*\}$ yields the same payoff as $(\sigma, \{H_0^*, H_1^*\})$.*

  (c) *(no balking) $\underline{p}_\sigma = 0$.*

The optimal menu involves only waiting-time distributions satisfying the NBUE property. The result is intuitive but not obvious: while for the high type, any waiting-time distribution that induces no reneging is necessarily NBUE, the same does not hold for the low type. However, because the payoff is decreasing in δ^{H_0} and increasing in δ^{H_1}, the optimal waiting-time distribution for the low type is indeed NBUE (see Figure 3).

The optimal menu can be of two types, pooling or separating. If it is pooling, it coincides with the optimal queueing discipline characterized in Theorem 1. If it is separating, the agents are offered a choice between two queues, one serving according to a first-come first-served queueing discipline, the other serving in random order.

The agents joining the queue with a high belief, that is, immediately after having received a positive lump-sum payoff, are served in random order. The agents joining the queue with a belief below the invariant probability are served according to a first-come first-served queueing discipline. Interestingly, in this case, agents always queue: the cutoff belief is 0, and agents join the queue even immediately after having received a negative payoff.

Finally, the incentive constraint for the agent joining with a low belief is binding. In other words, each agent is indifferent between being offered the menu $\{H_0^*, H_1^*\}$ and being offered the single queueing discipline H_1^*. Yet, the queueing discipline H_1^* is not implementable because it induces an aggregate service rate larger than the total capacity.

The solution sheds light on the previous results. Consistent with the tradeoff highlighted in Section 4.4, the first-come-first-served discipline is well-suited to agents joining with a low belief. Interestingly, the designer never pursues the objective of minimizing the queue length; on the contrary, the length of the queue is sometimes maximal.

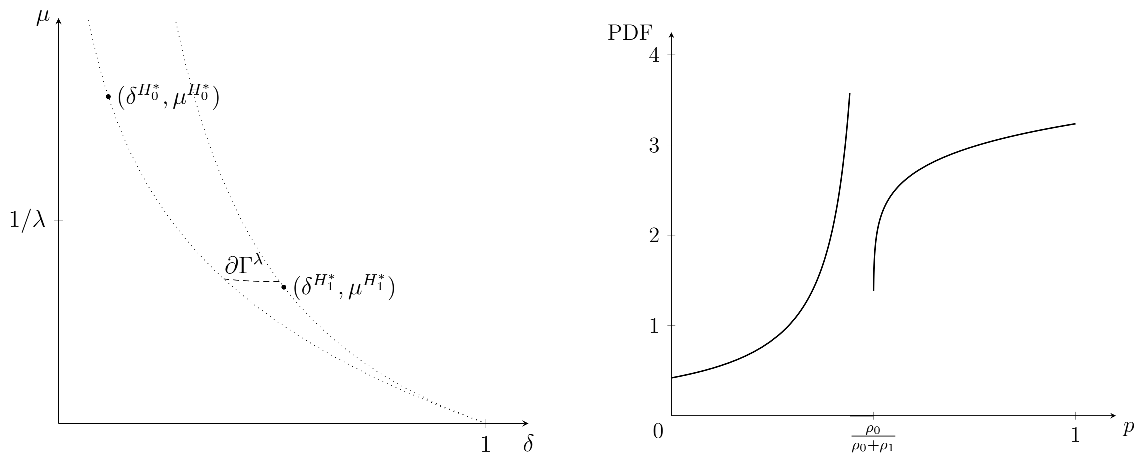
When the optimal menu is separating, the value of the information acquired at each service is maximized. Learning is so valuable that agents have incentives to queue at all times. From the perspective of the individual experimentation problem, this does not mean that "exploring" is valuable at any belief. Joining the queue has an option value: it guarantees the right to be served at some point in the future when the belief will be higher. The reason that an agent joins at a belief of 0 is that he is certain that he will not engage in "exploration" for some time. Figure 7 illustrates an example of an optimal separating menu. The left panel depicts the pair of summary statistics associated with the optimal menu. The right panel plots the steady-state distribution of posterior beliefs.

5.4 Optimal Queueing Discipline

As demonstrated below, the optimal menu can be "virtually" implemented, in the sense that it is possible to implement an outcome arbitrarily close to it with a single queueing discipline. This discipline takes advantage of reneging.

As discussed, the payoff and service rate induced by a non-abandoning strategy depend only on the pair of induced waiting-time distributions. Consequently, a single queue implements the outcome induced by some binary menu (together with an incentive-compatible strategy) if the pair of waiting-time distributions induced by the agents' best reply coincides with the pair of disciplines offered in the menu.

Now, I introduce the notion of virtual implementability used in Theorem 3.



(a) Summary statistics for the optimal pair $\{H_0^*, H_1^*\}$.

(b) Density of the cross-sectional distribution of posterior beliefs.

Figure 7: Optimal menu for $(\theta_1, \theta_0, c, \rho_0, \rho_1, \lambda) = (1, -3/4, 1/20, 2/3, 2/3, 1)$.

Definition 10. *A menu $\{H_0, H_1\} \in \mathcal{H} \times \mathcal{H}$ is virtually implementable if there exists a feasible and incentive-compatible sequence $(H_n, \sigma_n) \in \mathcal{H} \times \Sigma$ such that $H_{\sigma_n,0} \to H_0$ and $H_{\sigma_n,1} \to H_1$.*

Owing to the simple structure of the optimal menu, one can construct a queueing discipline $H \in \mathcal{H}$ such that the pair of waiting-time distributions induced by the agents' best reply is arbitrarily close to $\{H_0^*, H_1^*\}$.

Theorem 3. *The optimal menu $\{H_0^*, H_1^*\}$ is virtually implementable.*

In the proof, relegated to Appendix B.4, given an optimal separating menu $\{H_0^*, H_1^*\}$, $H_0^* \neq H_1^*$ (by Theorem 2 the result trivially holds otherwise), I consider a sequence of waiting-time distributions $\{H_n\}_{n\in\mathbb{N}} \in \mathcal{H}$, defined by

$$H_n(t) = \begin{cases} 1 - e^{-t/(\mu^{H_1^*}+1/n)} & t \leq 1/n, \\ 1 - e^{-1/(n\mu^{H_1^*}+1)} & 1/n < t < \mu^{H_0^*} - \varepsilon_n, \\ 1 & t \geq \mu^{H_0^*} - \varepsilon_n, \end{cases}$$

where $0 < \varepsilon_n \ll \mu^{H_0^*}$ is such that $\varepsilon_n \to 0$ as $n \to \infty$. The distributions H_n have the property that the hazard rate is constant in the interval $[0, 1/n]$. As a consequence, agents are able to "engineer" for themselves the exponential distribution H_1^* by restarting at intervals of times $1/n$. It can be shown that, for any n, a strategy that prescribes such behavior at beliefs higher than the invariant probability, and no-reneging otherwise, is indeed a best reply to H_n. Additionally, as $n \to \infty$, the cost of having random service for agents who just checked in becomes virtually costless for the agents with low belief.

The result has a striking implication. Not only is reneging desirable, in the sense that it allows to achieve larger aggregate payoffs relative to the constrained design problem, but it also permits to implement the optimal menu with a single queue.

6 Benchmarks

The analysis in the paper rests on three main assumptions: learning, anonymity, and no transfers. In this section, I provide benchmarks that allow to understand the role and the interaction of these three features. In short, if any of these three assumptions is lifted, the first best, discussed in Section 6.1, is restored.

6.1 First-Best

In this section, I consider the first-best problem in which the designer does not have to satisfy the incentive compatibility constraint. Formally, I consider the program

$$\text{[FB]} \qquad \sup \mathbb{E}_{\sigma,H}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \;\mathrm{d}t\right)\right],$$

over $H \in \mathcal{H}$, and $\sigma \in \Sigma$, subject to

$$\lim_{t\to\infty}\frac{1}{t}N_t^i \leq \lambda.$$

Proposition 3. *The value of the first-best program is equal to $\lambda\theta_1$. In particular, there exists a sequence $(p_n, H_n) \in (0, \rho_0/(\rho_0+\rho_1)) \times \mathcal{H}$ such that*

$$\lim_{n\to\infty}\mathbb{E}_{\sigma_{p_n},H_n}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i \,\mathrm{d}N_t^i - \int_0^T c\, a_t^i \;\mathrm{d}t\right)\right] = \lambda\theta_1,$$
$$\lim_{t\to\infty}\frac{1}{t}N_t^i \leq \lambda. \tag{10}$$

and $p_n \to \rho_0/(\rho_0+\rho_1)$, as $n \to \infty$.

Proof. For any n, set H_n equal to a degenerate distribution that puts mass 1 on μ_n, where μ_n is the unique positive solution of

$$\Lambda\left(e^{-(\rho_0+\rho_1)\mu_n}, \mu_n, p_n\right) = \lambda.$$

By definition, for any sequence $\{p_n\}_{n\geq 0}$ such that $p_n \to \rho_0/(\rho_0+\rho_1)$, $\mu_n \to 0$ as $t(p_n) \to \infty$. As a result, using the decomposition in (6), it is readily verified that the limit payoff is equal to $\lambda\theta_1$. □

The supremum is not achieved: if an agent were to queue only when his belief is (weakly) above the invariant probability $\rho_0/(\rho_0+\rho_1)$, his long-run average payoff would be zero.

Nevertheless, the designer can virtually achieve a total payoff of $\lambda\theta_1$ by having agents join the queue at a belief arbitrarily close to the invariant probability.

The intuition is easy to grasp. Roughly, the designer faces a restless bandit problem in which serving one agent corresponds to pulling one arm. Because the state changes independent of the designer's action, she would like to adopt an "exploiting" policy by serving an agent repeatedly and as frequently as possible, as long as the realized payoff is θ_1. However, with the double continuum of agents and time, translating this intuition into a well-defined policy is problematic. Specifically, an allocation is defined as a collection of counting processes, such that each agent is served finitely many times in any finite interval. As a result, over any interval of time, the designer needs to serve a positive measure of distinct agents. She could do so by having agents join the queue as long as their belief exceeds some $\underline{p} < \rho_0/(\rho_0 + \rho_1)$.

Following the reasoning of the proof, consider the case of the first-come first-served queueing discipline. If agents use the cutoff strategy with threshold $\underline{p}$, two types of agent are served at any point in time, one having a belief below the invariant probability and the other having one above. Both the proportion of agents of each of these two types, and the beliefs they entertain when they are served are endogenously determined in steady state, and thus the feasibility constraint (10) holds with equality.

The aggregate payoff is increasing in $\underline{p}$: as the cutoff increases, the proportion of served agents with the highest belief increases; moreover, the two beliefs at which agents are served also increase. The set of non-reneging strategies satisfying the feasibility constraint (10) with equality is not closed: for any sequence of cutoff beliefs $\{p_n\}_{n=1}^{\infty}$ such that $p_n \to \rho_0/(\rho_0+\rho_1)$, the service rate for the limit strategy is zero.[38]

The result confirms the earlier discussion. The designer cannot implement (allocations arbitrarily close to) the first-best in some equilibrium of the queueing game, because of the congestion externality. In particular, the designer needs to provide incentives to the agents to check in at a rate no larger than the total capacity λ. If the designer were to offer the sequence (δ_n, μ_n), agents would choose a best-reply strategy that violates the feasibility constraint.

6.2 Non-Anonymity

In this section, I show that if the designer is able to condition on past histories, the first-best can be "virtually" achieved. First, I informally introduce the definition of an allocation mechanism in this case and then describe the simple mechanism that approximates the first-best.

As in the baseline game, agents queue to be served; at all times, the designer chooses how to allocate the resource to queueing agents as a function of their individual histories

[38] The problem would have a maximum if, for example, an agent became idle after being served. In other words, assume that an agent who has been served at time t receives a payoff of 0 if served before $t + \Delta$, for some $\Delta > 0$, irrespective of his current state (the transition of which is the same as before). In this case, agents cannot be served at a belief higher than $e^{-\Delta(\rho_0+\rho_1)} + \left(1 - e^{-\Delta(\rho_0+\rho_1)}\right) \rho_0/(\rho_0 + \rho_1)$, and the designer has a bound on the higher cutoff strictly below $\rho_0/(\rho_0 + \rho_1)$.

of allocations and time-in-queue.[39] As in the model with anonymity, agents do not report their types. At all times, they only choose whether to queue.

In the following, I construct a sequence of mechanisms, in particular quota mechanisms, that achieve the first-best in the limit.

Assume that the designer commits to serving agents according to a first-come first-served discipline, but forbids some agents to be served at specific points in time; that is, if an agent reaches the head of the line at some time when he is forbidden to be allocated the good, he is served as soon as his no-service period ends. The no-service time intervals are chosen as a function of past allocations. In particular, if an agent is served at t, he cannot be served before $t + \Delta$, for some $\Delta > 0$ to be determined. Moreover, the designer divides the time horizon in blocks of length T such that $\lambda T \in \mathbb{N}$; agents are entitled to be served at most λT times per block.

Now, I claim that by taking T "sufficiently large," and a sequence of $\Delta > 0$ converging to zero, one can construct an equilibrium in the induced game that achieves an aggregate payoff arbitrarily close to the first-best. Informally, by taking T sufficiently large, I can compute a lower bound on agents' payoff by applying a renewal theorem. For a fixed T, taking a sequence of $\Delta > 0$ converging to zero guarantees that the payoff from this lower bound converges to $\lambda\theta_1$.

Specifically, fix $\Delta > 0$ and $T > 0$. Consider the strategy such that an agent queues whenever his belief is above some $\underline{p}^\Delta < \rho_0/(\rho_0 + \rho_1)$ to be determined, and whenever his belief is above $\rho_0/(\rho_0 + \rho_1)$ and Δ time has elapsed since being last served. Let $\underline{p}^\Delta$ be implicitly defined by[40]

$$\lambda = \frac{\underline{p}^\Delta + (1 - e^{-(\rho_0+\rho_1)\Delta})\frac{\rho_1}{\rho_0+\rho_1}}{\underline{p}^\Delta \Delta + (1 - e^{-(\rho_0+\rho_1)\Delta})\frac{\rho_1}{\rho_0+\rho_1} t(\underline{p}^\Delta)}. \tag{11}$$

By construction,[41] if an agent uses this strategy and the waiting time before being served is $\mu \geq 0$ (a small number),

$$\lim_{t\to\infty} \frac{1}{t} N_t^i \to \lambda + o(\mu) \leq \lambda, \text{ a.s.}$$

Hence, for any $\varepsilon > 0$, there exists T large enough that, within a block $[t, t + T]$,

$$\Pr\left[N_{t+T}^i - N_t^i \leq T\left(\lambda + o(\varepsilon)\right) + \varepsilon\right] \leq \varepsilon.$$

Moreover, because by construction, the (aggregate) service rate does not exceed λ, in steady state the queue must be empty, irrespective of the strategy used by the agents. Consequently,

[39] Notice that, at any t, the processes $(N_s^i)_{s\leq t}$ and $(w_s^i)_{s\leq t}$ suffice to summarize the past history of actions of agent i, $(a_s^i)_{s\leq t}$.

[40] The constant $\Delta > 0$ is understood to be sufficiently small that $t(\underline{p}^\Delta) > \Delta$.

[41] The right-hand side of (11) equals $\lim_{t\to\infty} N_t^i/t$ for this strategy if the waiting time before being served is zero. This can be checked by using (21) in the Appendix.

an agent joining the queue is served with no delay provided that an amount of time Δ elapsed since he was last served. As a result, an agent's expected payoff from this strategy equals

$$\lambda\left(\frac{\underline{p}^{\Delta}}{\underline{p}^{\Delta}+(1-e^{-(\rho_0+\rho_1)\Delta})\frac{\rho_1}{\rho_0+\rho_1}}\theta_1+\left(1-\frac{\underline{p}^{\Delta}}{\underline{p}^{\Delta}+(1-e^{-(\rho_0+\rho_1)\Delta})\frac{\rho_1}{\rho_0+\rho_1}}\right)\theta_0\right)+o(\varepsilon),$$

for some ε that can be taken arbitrarily small as $T\to\infty$. The result of Lemma 9 then follows.

Lemma 9. *Without anonymity, there exist mechanisms that yield an aggregate payoff arbitrarily close to $\lambda\theta_1$.*

6.3 Transfers

When monetary transfers are allowed, conceptually, the allocation problem falls into the framework analyzed by Bergemann and Välimäki (2010) and Athey and Segal (2013). Over time, agents are incentivized to truthfully report their types, i.e., their posterior beliefs, by a VCG-type mechanism.

To apply Bergemann and Välimäki (2010) and Athey and Segal (2013)'s result, the arguments need to be adapted to account for a few technical differences. While the absence of discounting is not crucial, as shown by Hörner et al. (2015) (see their Proposition 2), the continuum of dates and the continuum of agents pose a few challenges.

Instead, directly applying the standard externality argument is straightforward and makes it possible to demonstrate that with transfers the first-best can be (virtually) achieved, while balancing the budget. The idea is to levy an "entry fee" that causes each agent to internalize the congestion externality. Interestingly, in contrast to the scheduling example in Bergemann and Välimäki (2010), the continuum of agents simplifies the computation of the social externality cost of an agent's action.

Lemma 10. *For any $\varepsilon>0$, there exists a budget-balanced mechanism that achieves a payoff larger that $\lambda\theta_1-\varepsilon$.*

Proof. Consider the following sequence of mechanisms parameterized by $\Delta>0$. Let H^{Δ} be the cumulative distribution function of a degenerate distribution at Δ, i.e., $H^{\Delta}=\mathbf{1}_{[\Delta,\infty)}$. Agents are offered the waiting-time distribution H^{Δ}, and, at check-in they are charged a fee ξ^{Δ}, defined next. Let ξ^{Δ} be the Lagrange multiplier associated with the constraint (12) in the following program

$$[\mathrm{CE}^{\Delta}]\qquad \sup_{\sigma\in\Sigma}\mathbb{E}_{\sigma,H^{\Delta}}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T\theta_t^i\,\mathrm{d}N_t^i-\int_0^T c\,a_t^i\;\mathrm{d}t\right)\right],$$

subject to

$$\ell_{\sigma,H^{\Delta}}\leq\lambda, \tag{12}$$

where $\ell_{\sigma,H^\Delta} \in \mathbb{R}$ is the almost-sure limit of the random variable N_t^i/t, that is,

$$\ell_{\sigma,H^\Delta} = \lim_{t\to\infty} \frac{1}{t} N_t^i, \text{ a.s.}$$

When agents are offered the waiting-time distribution H^Δ with an entry fee ξ^Δ, the best-reply problem can be written as

$$\sup_{\sigma\in\Sigma} \mathbb{E}_{\sigma,H^\Delta} \left[\limsup_{T\to\infty} \frac{1}{T} \left(\int_0^T \theta_t^i \, \mathrm{d}N_t^i - \int_0^T c\, a_t^i \; \mathrm{d}t \right) - \left(\lim_{t\to\infty} \frac{1}{t} N_t^i \right) \xi^\Delta \right].$$

Because for any $\sigma \in \Sigma$, the service rate converges almost surely to ℓ_{σ,H^Δ}, an agent's best reply solves [CE$^\Delta$]. To balance the budget, the designer redistributes the proceeds from the entry fee as a flow payoff $\lambda\xi^\Delta$ to all agents. Taking $\Delta \to 0$, by Proposition 3, the payoff converges to $\lambda\theta_1$. □

6.4 Known-Own State

In the model, learning plays a crucial role. In particular, because an agent's payoff for a given strategy is not only determined by the average (steady-state) waiting time, the ranking of different queueing disciplines depends on the parameters. Conversely, when agents observe the evolution of their state, the ranking of queueing disciplines is unambiguous.

Lemma 11. *If each agent observes his state at all times, the service-in-random-order discipline dominates any other queueing discipline.*

Proof. First, notice that the payoff in the best equilibrium is bounded above by

$$\lambda\theta_1 - \frac{\rho_0}{\rho_0+\rho_1} c.$$

In any symmetric equilibrium, the designer would like to allocate all of the capacity to agents of type θ_1. By assumption, only queueing agents can be served, and the equilibrium is symmetric, and hence in equilibrium, at best only type θ_1 agents queue. By assumption, the initial individual state is θ_1 with probability $\rho_0/(\rho_0+\rho_1)$. It follows that the fraction of type θ_1 agents in the population equals $\rho_0/(\rho_0+\rho_1)$ at all times, explaining the upper bound.

I now prove that service-in-random-order achieves this upper bound.

In this case, because of the memorylessness of the waiting-time distribution and of the state transitions, to analyze the best reply, it suffices to compare two possible strategies: the strategy that prescribes queueing at all times and the strategy that prescribes queueing only when the type is θ_1. Suppose that the (exponential) waiting-time distribution has mean μ. Then, the payoff from queueing at all times equals

$$\frac{1}{\mu}\left(\frac{\rho_0}{\rho_0+\rho_1}\theta_1 + \frac{\rho_1}{\rho_0+\rho_1}\theta_0 \right) - c.$$

Instead, the payoff from queueing only when the state equals θ_1 is

$$\frac{\rho_0}{\rho_0+\rho_1}\left(\frac{1}{\mu}\theta_1 - c\right).$$

Because the designer chooses a queueing discipline such that the service rate equals λ, in the second case, the equilibrium achieves the upper bound. Moreover, for any μ, the strategy of queueing only when the state is θ_1 dominates the strategy that prescribes queueing at all times. This concludes the proof.

However, I additionally compute the payoff under these two strategies for the first-come first discipline. I show that in that case, (constrained) inefficient equilibria do exist. It is immediate that also for first-come first-served, it suffices to restrict attention to these two best-reply candidates. Under the first-come first-served discipline with average waiting time μ, the payoff from queueing at all times is

$$\frac{1}{\mu}\frac{\rho_0}{\rho_0+\rho_1}\theta_1 - c.$$

Instead, under the strategy that prescribes queueing only when the state is θ_1, the payoff is equal to

$$\frac{\rho_0}{\rho_0+\rho_1}\left(\rho_1\frac{e^{-\rho_1\mu}}{1-e^{-\rho_1\mu}}\theta_1 - c\right).$$

In this case, for a fixed μ, either strategy can be optimal. □

The intuition is straightforward. When served in order of arrival, an agent can find it optimal to queue even when his state is θ_0 because queueing has an option value. Not only does queueing give the right to be allocated the good at some point in the future, but because the service time is deterministic, an agent can ensure that he is not served when the state is θ_0. Clearly, having agents of type θ_0 in line is detrimental to welfare because it increases congestion.

Instead, if service occurs in random order, queueing has no option value, as all agents receive the same treatment independent of their time-in-queue. As a result, service-in-random-order yields an aggregate payoff (weakly) higher than first-come first-served.

Why does the tradeoff between thickness and congestion disappear in the absence of learning? The key distinction is that, in this case, the evolution of agents' types is not affected by the allocation. Each agent observes his state θ_t^i at all times, irrespective of his consumption experience. In the baseline model, the aggregate distribution of agents' types (their posterior beliefs) is determined endogenously. In turn, this distribution impacts the expected payoff from each service, giving rise to the tradeoff.

7 Concluding Remarks

In this paper, I study the optimal design of a queue when agents learn from past consumption experiences. In this setup, a menu to screen agents takes the form of multiple queues (or customer classes) with agents being served in a different order within each of them. The optimal menu is (at most) binary and has a simple structure. When it is optimal to offer two distinct queues, agents are served in a first-come first-served manner in one queue, and in random order in the other queue; if pooling is optimal, the single queue is either first-come first-served or service-in-random-order, possibly with a minimum wait requirement. Owing to the flexibility of the queueing setup, the optimal menu can always be implemented with a single queue by allowing agents to renege and rejoin at will.

The model is stylized in many respects. I believe that the techniques developed to characterize equilibrium payoffs extend to richer learning technologies, provided that the stationarity of the environment is preserved. However, the assumption of a continuum of agents is important because it allows one to formulate the best-reply problem as a simple Markov decision problem. On the one hand, a model with a finite number of agents would allow for a finer analysis of the strategic interaction among them, beyond the general-equilibrium effect captured by the current model. On the other hand, to the extent that the problem reduces to a "two-level" optimization problem, I believe that the main insights would not be overturned in a setting with a large, but finite, number of agents.

More broadly, because of the stationarity of the environment and the independence assumption, important aspects of queueing and learning via experimentation are missing. Allowing for variable capacity would be useful to study the welfare implication of congestion in an environment with fluctuations.[42] This is a natural extension. Correlation in agents' valuations would introduce the possibility of observational learning that is currently missing in my model.

[42] Interestingly, the peer-to-peer landing platform Zopa, a two-tier queueing mechanism to allocate lenders' funds to borrowing opportunities, prioritizes returning lenders. However, one may expect economic fluctuations to play a key role in this environment.

References

Asmussen, Søren (2008) *Applied probability and queues*, New York: Springer.

Athey, Susan and Ilya Segal (2013) "An Efficient Dynamic Mechanism," *Econometrica*, 81, 2463–2485.

Bassamboo, Achal and Ramandeep Singh Randhawa (2016) "Scheduling homogeneous impatient customers," *Management Science*, 62, 2129–2147.

Bergemann, Dirk and Alessandro Pavan (2015) "Introduction to Symposium on Dynamic Contracts and Mechanism Design," *Journal of Economic Theory*, 159, 679–701.

Bergemann, Dirk and Juuso Välimäki (2010) "The Dynamic Pivot Mechanism," *Econometrica*, 78, 771–789.

Bloch, Francis and David Cantala (2016) "Dynamic Assignment of Objects to Queuing Agents," *American Economic Journal: Microeconomics*, forthcoming.

Chen, Hong and David D. Yao (2013) *Fundamentals of queueing networks: Performance, asymptotics, and optimization*, New York: Springer.

Cripps, Martin W. and Caroline D. Thomas (2016) "Strategic Experimentation in Queues," Working Paper.

Dolan, Robert J (1978) "Incentive Mechanisms for Priority Queuing Problems," *The Bell Journal of Economics*, 9, 421.

Guess, Frank and Frank Proschan (1988) "Mean residual life: Theory and applications," in P. R. Krishnaiah and C.R. Rao eds. *Handbook of Statistics*, Volume 7: Elsevier, Ch. 12, 215–224.

Hassin, Refael (1985) "On the optimality of first come last served queues," *Econometrica*, 53, 201–02.

——— (2016) *Rational Queueing*: Chapman and Hall/CRC.

Hassin, Refael and Moshe Haviv (2003) *To queue or not to queue: Equilibrium behavior in queueing systems*, New York: Springer.

Haviv, Moshe (2013) *Queues: A Course in Queueing Theory*, New York: Springer-Verlag.

Hörner, Johannes, Satoru Takahashi, and Nicolas Vieille (2015) "Truthful Equilibria in Dynamic Bayesian Games," *Econometrica*, 83, 1795–1848.

Jovanovic, Boyan and Robert W. Rosenthal (1988) "Anonymous sequential games," *Journal of Mathematical Economics*, 17, 77–87.

Krähmer, Daniel and Roland Strausz (2015) "Dynamic Mechanism Design," in Tilman Borgers ed. *An introduction to the theory of mechanism design*, New York: Oxford University Press.

Leshno, Jacob D. (2015) "Dynamic matching in overloaded waiting lists," Working Paper.

Naor, Pinhas (1969) "The regulation of queue size by levying tolls," *Econometrica: journal of the Econometric Society*, 15–24.

Platz, Trine Tornøe and Lars Peter Østerdal (2016) "The curse of the first-in-first-out queue discipline," Working Paper.

Puterman, Martin L. (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, New York: Wiley.

Rosenberg, Dinah, Antoine Salomon, and Nicolas Vieille (2013) "On Games of Strategic Experimentation," *Games and Economic Behavior*, 82, 31–51.

Shaked, Moshe and J. George Shanthikumar (2007) *Stochastic orders*, New York: Springer-Verlag.

Stokey, Nancy L. (2008) *The Economics of Inaction: Stochastic Control models with fixed costs*, Princeton: Princeton University Press.

Sun, Yeneng (2006) "The exact law of large numbers via Fubini extension and characterization of insurable risks," *Journal of Economic Theory*, 126, 31–69.

Thomas, Caroline (2014) "Strategic Experimentation with Congestion," Working Paper.

Thomke, Stefan and Donald Reinertsen (2012) "Six myths of product development," *Harvard Business Review*, 90, 84–94.

Appendix A: General setup

A.1 Queueing Discipline

In this section, I provide the general definition of a queueing discipline. In order to do so, I first define the filtration $(\mathcal{G}_t)_{t\geq 0}$ describing the information available to the designer.

Let $A(t) \in [0, t]$ denote the cumulative mass of arrivals at the queue over the interval $[0, t]$. Formally, given the realization of agents' impulse controls $\{(\tau_k, a_k)\}_{k=1}^{\infty}$,

$$A(t) \coloneqq \int_0^t \int_0^1 \mathbf{1}_{\{\exists k \in \mathbb{N}\colon \tau_k^i = s,\, a_k^i = 1\}} \,\mathrm{d}i \,\mathrm{d}s,$$

where $\mathbf{1}_A$ is the indicator function of the event A. The stochastic process $A(t)$ has non-decreasing non-negative right-continuous sample paths.

Let $R(x, t) \in [0, 1]$ be the proportion of agents arriving at time x who renege before time $t \geq x$. Formally, whenever $\int_x^{x+\varepsilon} \mathrm{d}A(x) > 0, \forall \varepsilon > 0$ small enough, $R(x, t)$ is defined as

$$R(x,t) \coloneqq \lim_{\varepsilon \to 0} \frac{\displaystyle\int_x^{x+\varepsilon} \int_x^t \int_0^1 \mathbf{1}_{\{\exists k \in \mathbb{N}\colon \tau_k^i = y,\, a_k^i = 1, \tau_{k+1}^i = s,\, \mathrm{d}N_s^i = 0\}} \,\mathrm{d}i \,\mathrm{d}s \,\mathrm{d}y}{\displaystyle\int_x^{x+\varepsilon} \mathrm{d}A(x)}.$$

Otherwise, if $\int_x^{x+\varepsilon} \mathrm{d}A(x) = 0$ for any $\varepsilon < \varepsilon'$, for some $\varepsilon' > 0$, then $R(x, t) \coloneqq 0$. For each x, the stochastic process $R(x, \cdot)$ has non-decreasing non-negative right-continuous sample paths. By definition, for each t, the stochastic process $R(\cdot, t)$ has left-continuous sample paths.

The information available to the designer is modeled by the filtration $(\mathcal{G}_t)_{t\geq 0}$, where $\mathcal{G}_t$ is the σ-algebra generated by $(A(s))_{s\leq t}$, $(\{R(x, s) \mid x \leq t\})_{s\leq t}$, and $(\{G(x, s) \mid x \leq s < t\})_{s\leq t}$, which I define next.

Definition 11. *A queueing discipline is a process $G(x, t)$ taking values in $[0, 1]$, such that*

1. *for every x, $(G(x, t))_{t\geq 0}$ is a non-decreasing right-continuous process predictable with respect to the filtration $(\mathcal{G}_t)_{t\geq 0}$, and $G(x, x) = 0$;*
2. *for each t, $(G(x, t))_{x\geq 0}$ has left-continuous sample paths.*

I interpret $G(x, t)$ as the proportion of agents joining at time x served by time t. Condition 1 guarantees that this interpretation is meaningful, and enforces the informational restriction that the designer cannot condition the allocation at t on the action of agents at time t. Condition 2 is technical; it rules out, for example, the undesirable situation in which at some point all agents who have been in queue for strictly more than t have been served, while none of those who been waiting for exactly t are.

Definition 12. *A queueing discipline G is feasible if*

1. *(No-service of absents) For all $\underline{x}$, $\overline{x}$, and t, $\underline{x} < \overline{x} < t$,*

$$\int_{[\underline{x},\overline{x}]} G(x,t)\,\mathrm{d}A(x) \leq \int_{[\underline{x},\overline{x}]} (1 - R(x,t))\,\mathrm{d}A(x).$$

2. *(Capacity constraint) For any $t \geq 0$, $s \geq 0$, $s \leq t$,*

$$\int_{[0,t]} (G(x,s) - G(x,t))\,\mathrm{d}A(x) \leq \lambda\,(t-s)\,.$$

Condition 1 states that the designer cannot serve a proportion of agents larger than the remaining one. Condition 2 reflects the capacity constraint.

A.2 Agent's MDP

Given a queueing discipline, and an aggregate strategy profile for his opponents, agent i chooses his best reply to maximize

$$\mathbb{E}_{\sigma^i,\sigma^{-i},G}\left[\limsup_{T\to\infty}\frac{1}{T}\left(\int_0^T \theta_t^i\,\mathrm{d}N_t^i - \int_0^T c\,a_t\;\mathrm{d}t\right)\right]\,.$$

In this section, I argue that, restricting attention to (non-stationary) Markov strategies is without loss. In particular, an agent's best-reply problem can be stated as a Markov decision problem (MDP).

Recall that at any time t, the designer observes for each agent $i \in [0,1]$ only the current action and the time-in-queue. Specifically, the designer does not observe the private histories of posterior beliefs, nor does she observe agents' private histories of actions before $t - w_t^i$. Hence, the private history of posterior beliefs $\left(p_s^i\right)_{s<t}$, actions $\left(a_s^i\right)_{s<t}$ and time-in-queue $\left(w_s^i\right)_{s<t}$ are payoff irrelevant.

Fix an initial joint distribution over posterior beliefs, time-in-queue, and actions that satisfies two requirements. First, the average posterior belief equals the invariant probability of θ_1, $\rho_0/(\rho_0+\rho_1)$. Second, the joint distribution puts mass zero on triples such that the-time-in-queue is positive and the action is 0 (not queueing). Take the point of view of agent i. Because there is a continuum of agents, agent i has negligible influence on the other agents. Moreover, there is no aggregate uncertainty: given a (history-dependent) strategy profile, the aggregate distribution of actions and posterior beliefs follows a deterministic path. As a result, agent i faces an MDP with state space $(\mathbb{R}_+ \times [0,1] \times \{0,1\} \times \mathbb{R}_+)$, calendar time, posterior belief, current action, and time-in-queue.[43]

[43] It may appear odd to include the action as a state variable. However, this simplifies the description of the agents' MDP. Because strategies are defined as impulse controls, the problem faced by the agent can be described as a semi-Markov decision process (see Puterman, 1994, Chap. 11). At any action time ("decision epoch" using Puterman's terminology), the agent chooses an action and a predicable stopping time. The distribution of the next action time is determined by the chosen stopping time, and by the agent's action and time-in-queue to the extent that these two variables affect the probability of being served.

In particular, an agent's payoff depends on his opponents' actions only through their effect on the offered waiting-time distribution. Moreover, for a fixed queueing discipline, and given a symmetric strategy profile for his opponents, the offered waiting-time distribution is a deterministic function of time. Specifically, given a queueing discipline, a strategy profile, and a distribution $M_t : [0,1] \times \{0,1\} \times \mathbb{R}_+ \to [0,1]$, the processes $A(t)$ and $R(x,t)$ are a deterministic function of time. At each time t, the agent computes the offered waiting-time distributions H_t. That is, should he join the queue at time t, he would be served by time x with probability $H_t(x)$, where

$$H_t(x) = \int_t^x \frac{1}{1 - R(t,s)} G(t, \mathrm{d}s).$$

I conclude that given the offered waiting-time distribution, the agent faces an MDP, so that there is no loss of generality in restricting attention to Markov strategies.

A.2.1 Observability and Off-Path

Because of the lack of aggregate uncertainty, the aggregate distribution of actions, that is, the evolution of the queue, is perfectly predictable by an agent. Hence, the assumption that the agent does not observe the queue is without loss.

But should then the strategy of an agent define his behavior at any possible history of aggregate actions? As pointed out by Jovanovic and Rosenthal (1988, Sec. 4, Remark (a)), with a continuum of agents, it is unclear how to define equilibrium refinements. Nevertheless, since no individual agent can affect the aggregate distribution of actions, in order to characterize equilibrium outcomes, it suffices to restrict attention to Nash equilibrium in which each agent best replies to the aggregate distribution of other agents' actions. Therefore, I abstain from discussing the agent's strategy for histories off the equilibrium path.

Appendix B: Proofs

B.1 Proofs for Section 3

B.1.1 Proof of Lemma 1.

First, I compute the transition probabilities for the semi-Markov process π_t. Given $\tau \in \{\tau_0(n), n \geq 0\}$, so that $p_\tau = 0$, using (2),

$$\Pr_{\sigma,H}\left[p_{T(\tau)} = 1\right] = \mathbb{E}_{\sigma,H}\left[\left(1 - e^{-(\rho_0+\rho_1)(T(\tau)-\tau)}\right)\frac{\rho_0}{\rho_0+\rho_1}\right] = \frac{\rho_0}{\rho_0+\rho_1}\left(1 - d_\sigma(0)\right).$$

Similarly, given $\tau \in \{\tau_1(n), n \geq 1\}$, so that $p_\tau = 1$, using (2),

$$\begin{aligned}\Pr_{\sigma,H}\left[p_{T(\tau)} = 1\right] &= \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} + \frac{\rho_0}{\rho_0+\rho_1}\left(1 - e^{-(\rho_0+\rho_1)(T(\tau)-\tau)}\right)\right] \\ &= \frac{\rho_0}{\rho_0+\rho_1} + d_\sigma(1)\frac{\rho_1}{\rho_0+\rho_1}.\end{aligned}$$

Let

$$M_\sigma := \begin{pmatrix} 1 - \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0)) & \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0)) \\ (1-d_\sigma(1))\frac{\rho_1}{\rho_0+\rho_1} & \frac{\rho_0}{\rho_0+\rho_1} + d_\sigma(1)\frac{\rho_1}{\rho_0+\rho_1} \end{pmatrix},$$

denote the transition matrix, and $(1-m_\sigma, m_\sigma)$ be the unique stationary distribution of the positive recurrent irreducible chain with transition matrix M_σ. It is easy to check that

$$m_\sigma = \frac{(1-d_\sigma(0))\frac{\rho_0}{\rho_1+\rho_0}}{(1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0} + (1-d_\sigma(0))\frac{\rho_0}{\rho_1+\rho_0}}.$$

For $j \in \{0,1\}$, given $\tau \in \{\tau_j(n), n \geq 0\}$, the expected length of time before the next transition is, by definition, $y_\sigma(j)$. First, assume that $y_\sigma(j) < \infty$ for $j \in \{0,1\}$. Because $H \in \mathcal{H}$ (no atoms at 0), for $\tau \in \cup_{j\in\{0,1\}}\{\tau_j(n), n \geq 0\}$, there exist $\varepsilon > 0$ and $\varepsilon' > 0$, such that

$$\Pr\left[T(\tau) - \tau \leq \varepsilon\right] \leq 1 - \varepsilon'.$$

Moreover the transition matrix M_σ is unichain. As a result, by Th. 11.4.2, Ch. 11 in Puterman (1994), the evaluation equations can be used to compute the long-run average payoff. Hence, the long-run average payoff, denoted $\gamma(\sigma; H)$ equals

$$\begin{aligned} \gamma(\sigma; H) =& \frac{m_\sigma(\theta_1 - c\,q_\sigma(1)) + (1-m_\sigma)(\theta_0 - c\,q_\sigma(0))}{m_\sigma y_\sigma(1) + (1-m_\sigma)y_\sigma(0)} \\ =& \frac{\frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))(\theta_1 - c\,q_\sigma(1)) + (1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0}(\theta_0 - c\,q_\sigma(0))}{\frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))\,y_\sigma(1) + (1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0}y_\sigma(0)}. \end{aligned} \tag{13}$$

The time between two upward jumps in the belief, that is, the time between $\tau_1(n)$ and $\tau_1(n+1)$, for some $n \in \mathbb{N}$, is independent of n. Hence, the average number of upward jumps per unit of time converges almost surely to the inverse of the mean inter-arrival time (see for example Asmussen, 2008, Proposition 1.4). The same holds for downward jumps. As a result, the service rate converges almost surely to a constant. From direct inspection of (13), the rate at which the agent collects lump sums, equals

$$\begin{aligned} \lim_{t\to\infty}\frac{1}{t}N_t =& \frac{1}{m_\sigma y_\sigma(1) + (1-m_\sigma)y_\sigma(0)} \\ =& \frac{(1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0} + \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))}{(1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0}y_\sigma(0) + \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))\,y_\sigma(1)}, \text{a.s.} \end{aligned} \tag{14}$$

Because the game is symmetric and I analyze the steady-state, the long run fraction of time the agent is queueing coincides with the length of the queue in steady-state, and is equal to

$$\frac{m_\sigma q_\sigma(1) + (1-m_\sigma)q_\sigma(0)}{m_\sigma y_\sigma(1) + (1-m_\sigma)y_\sigma(0)} = \frac{(1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0}q_\sigma(0) + \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))\,q_\sigma(1)}{(1-d_\sigma(1))\frac{\rho_1}{\rho_1+\rho_0}y_\sigma(0) + \frac{\rho_0}{\rho_0+\rho_1}(1-d_\sigma(0))\,y_\sigma(1)}.$$

I now discuss the case in which $y_\sigma(j) = \infty$ for some $j \in \{0, 1\}$. In this case the strategy σ is absorbing. That is, given the agent's initial belief p_0, $\mathbb{E}_{\sigma,H}[\tau_{p_0}(0)] = \infty$. The expected (long-run) average payoff diverges to $-\infty$ or is equal to 0, depending on

$$\lim_{T\to\infty} \mathbb{E}_{\sigma,H}\left[\int_0^T a_t \,\mathrm{d}s\right] \gtreqqless 0$$

If the previous limit is 0, then the (long-run) average payoff is 0; it diverges to $-\infty$ otherwise. In both cases, the service rate equals zero. □

B.1.2 Proof of Proposition 1.

Notation and Preliminary Observations. As mentioned at the beginning of Section 3, it is assumed that $\gamma(\sigma; H) > 0$ (Lemma 14 shows that this assumption is without loss). Note that, since by assumption $\theta_0 < 0$ and $c \geq 0$, (13) readily implies that $\gamma(\sigma; H) > 0$ only if $\theta_1 - c\, q_\sigma(1) > 0$. This guarantees that $\gamma(\sigma; H) > 0$ is strictly increasing in $d_\sigma(1)$, and strictly decreasing in $d_\sigma(0)$.

As shown, a strategy $\sigma \in \Sigma$ can be described in terms of one state variable only, the posterior belief. I now introduce some notation to describe stationary Markov strategies in a way that exploits this recursivity. Fix a (pure) strategy $\sigma \in \Sigma$. For any $p \in [0, 1]$, let $\tau_\sigma : [0, 1] \to \mathbb{R}_+$ be such that, under σ, if at some t, $p_t = p$, and $\tau_k \leq t < \tau_{k+1}$, then

$$\tilde{\tau}_{k+1} = t + \tau_\sigma(p),$$

where, by the definition in Section 2.1.1, $\tilde{\tau}_{k+1}$ is the time at which an agent playing σ reneges conditional on not having been served prior that time (that is, conditional on $T(t) \geq \tilde{\tau}_{k+1}$). Similarly, let $a_\sigma : [0, 1] \to \{0, 1\}$ be such that, along the path induced by σ,

$$a_k = a_\sigma(p_{\tau_k}).$$

The maps τ_σ and a_σ completely characterize the strategy $\sigma \in \Sigma$: for any starting belief p and action (on the recurrent path induced by σ), the agent adjusts his action either after an interval of time $\tau_\sigma(p)$, or when his belief jumps, whichever occurs first.

For convenience, I also introduce additional notation to describe the evolution of beliefs. Recall that the belief evolves deterministically along a history with no service. I define $\varphi(p, t)$, $\varphi(p, t) : [0, 1] \times \mathbb{R} \to [0, 1]$ to be the Bayesian update of the belief p after an interval of time of length t along a path with no service. Specifically, given (2),

$$\varphi(p, t) = e^{-(\rho_0+\rho_1)t} p + \left(1 - e^{-(\rho_0+\rho_1)t}\right) \frac{\rho_0}{\rho_0 + \rho_1}.$$

Note that the function $\varphi(p, t)$ is defined also for $t < 0$: in this case, if an agent holds an initial belief $\varphi(p, t)$, after an interval of time $-t$, his revised belief, conditional on service, equals p. Throughout, let $t(p, q)$, $t : [0, 1] \times [0, 1] \to \mathbb{R}$ be the time it takes for the belief to go from p to q, in the absence of jumps. That is, $t(p, q)$ is implicitly defined by

$$\varphi\left(p, t(p, q)\right) = q.$$

Again, it is convenient to let the function t take both negative and positive values.

Overview of the proof. The proof proceeds by contradiction and is divided in two steps. Assume that $\sigma \in \Sigma$ is optimal given $H \in \mathcal{H}$. First, I show that $\overline{p}_\sigma = 1$ and that σ does not involve abandoning at beliefs above the invariant probability. Then, I show that abandoning is suboptimal also at lower beliefs.

A strategy $\sigma \in \Sigma$ affects the payoff only via the sufficient statistics in Lemma 1. However, for an arbitrary strategy σ, the computation of the sufficient statistics can be intricate. In the proof, for a given strategy $\sigma \in \Sigma$, I look for candidate strategies that improve upon it by considering modifications of σ for which the induced change in the sufficient statistics is easy to compute.

No abandoning at $p \geq \rho_0/(\rho_0+\rho_1)$ and $\overline{p}_\sigma = 1$. The proof relies on a simple idea. Along the history with no service, the belief evolves deterministically irrespective of whether the agent is queueing or not. When the agent abandons the queue at some belief above the invariant probability, the only payoff-relevant change that takes place before rejoining the queue next, is a decrease in his belief. Hence, the payoff can be potentially improved by anticipating the time at which he rejoins. Yet, one needs to determine the continuation strategy after rejoining. Next I show that this can be done in a straightforward way: the agent, upon rejoining, behaves as if he had rejoined according to the original strategy.

Formally, assume that the strategy $\sigma \in \Sigma$ either involves abandoning at $p > \rho_0/(\rho_0+\rho_1)$, or is such that $\overline{p}_\sigma < 1$, or both. Note that from (13), the payoff $\gamma(\sigma; H)$ is strictly decreasing in $q_\sigma(1)$ and $y_\sigma(1)$, and strictly increasing in $d_\sigma(1)$.

Let $p' := \sup\{p : a_\sigma(p) = 0\}$. If $\overline{p}_\sigma < 1$, $p' = 1$; otherwise, p' is the belief of the agent the first time he abandons the queue after joining it with a belief of 1. Also, let $p'' := \sup\{p < p' : a^\sigma(p) = 1\}$. If $\overline{p}_\sigma < 1$, $p'' = \overline{p}_\sigma$; otherwise, p'' is the belief at which the agent rejoins the queue after having abandoned it with a belief p'.

Define the strategy $\tilde{\sigma} \in \Sigma$ such that

$$\tau_{\tilde{\sigma}}(p) = \begin{cases} \tau_\sigma(p) & p \leq \rho_0/(\rho_0+\rho_1), \text{and } p > p', \\ \tau_\sigma\left(\varphi\left(p'', t\left(p', p\right)\right)\right) & p' \geq p \geq \rho_0/(\rho_0+\rho_1), \end{cases}$$

and

$$a_{\tilde{\sigma}}(p) = \begin{cases} a_\sigma(p) & p \leq \rho_0/(\rho_0+\rho_1), \text{and } p > p', \\ a_\sigma\left(\varphi\left(p'', t\left(p', p\right)\right)\right) & p' \geq p \geq \rho_0/(\rho_0+\rho_1). \end{cases}$$

According to the strategy $\tilde{\sigma}$, the agents rejoins the queue at p', and his behavior from then on coincides with the behavior of an agent who joined at p''.

I claim that $y_{\tilde{\sigma}}(1) < y_\sigma(1)$, $q_{\tilde{\sigma}}(1) = q_\sigma(1)$, and $d_{\tilde{\sigma}}(1) > d_\sigma(1)$, contradicting the optimality of σ. In fact, it is immediate that $y_{\tilde{\sigma}}(1) = y_\sigma(1) - t(p', p'')$. By the definition of $d_\sigma(1)$, given $\tau \in \{\tau_1(n), n \geq 1\}$, because σ involves abandoning, and the agent does not queue for beliefs in $[p'', p']$,

$$
\begin{aligned}
d_\sigma(1) = & \Pr_{\sigma,H}\left[T(\tau) \leq \tau + t(1,p')\right] \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau) \leq \tau + t(1,p')\right] \\
& + \Pr_{\sigma,H}\left[T(\tau) > \tau + t(1,p'')\right] \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-t)} \mid T(\tau) > \tau + t(1,p'')\right].
\end{aligned}
$$

Moreover, $\Pr_{\sigma,H}\left[T(\tau) > \tau + t(1,p'')\right] > 0$, for otherwise σ would be a non-abandoning strategy. It is easy to see that

$$
\begin{aligned}
d_{\tilde\sigma}(1) = & \Pr_{\sigma,H}\left[T(\tau) \leq \tau + t(1,p')\right] \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau) \leq \tau + t(1,p')\right] \\
& + \Pr_{\sigma,H}\left[T(\tau) > \tau + t(1,p'')\right] e^{(\rho_0+\rho_1)t(p',p'')} \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-t)} \mid T(\tau) > \tau + t(1,p'')\right] \\
& > d_\sigma(1),
\end{aligned}
$$

yielding the desired contradiction.

No abandoning at $p < \rho_0/(\rho_0+\rho_1)$. When the belief is below the invariant probability, the agent is growing optimistic over time. Hence, proving that waiting before rejoining is suboptimal is more subtle. In the proof I show that, starting from an abandoning strategy σ, payoffs can be improved by "closing the gaps," that is, by immediately rejoining upon reneging and, at the same time, adjusting the initial cutoff, p_σ.

From (13), the payoff $\gamma(\sigma; H)$ is strictly decreasing in $q_\sigma(0)$, $y_\sigma(0)$, and $d_\sigma(0)$. Assume by contradiction that σ involves abandoning at some belief below the invariant probability of state θ_1, $\rho_0/(\rho_0+\rho_1)$. I claim that there exists a strategy $\tilde\sigma$ such that $y_{\tilde\sigma}(0) < y_\sigma(0)$, $q_{\tilde\sigma}(0) = q_\sigma(0)$, and $d_{\tilde\sigma}(0) < d_\sigma(0)$, contradicting the optimality of σ.

Let $p' := \inf\{p > \underline{p}_\sigma : a^\sigma(p) = 0\}$ be the lowest belief at which the agent abandons the queue after joining at $\underline{p}_\sigma$, and $p'' := \sup\{p > \underline{p}^\sigma : a^\sigma(p) = 1\}$ be the belief at which the agent returns to the queue after p'. Note that such a belief is well-defined, and in particular $p'' < \rho_0/(\rho_0+\rho_1)$; for otherwise, the strategy σ would yield a payoff of 0, contradicting the maintained assumption.

Given the structure of the strategy σ, for $\tau \in \{\tau_0(n), n \geq 1\}$,

$$
\begin{aligned}
y_\sigma(0) = & \Pr_{\sigma,H}\left[T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] \cdot \mathbb{E}_{\sigma,H}\left[T(\tau) - \tau \mid T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] \\
& + \Pr_{\sigma,H}\left[T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right] \cdot \mathbb{E}_{\sigma,H}\left[T(\tau) - \tau \mid T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right],
\end{aligned}
\tag{15}
$$

and

$$
\begin{aligned}
d_\sigma(0) = & \Pr_{\sigma,H}\left[T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] \cdot \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] \\
& + \Pr_{\sigma,H}\left[T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right] \cdot \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right],
\end{aligned}
\tag{16}
$$

where again, by definition of abandoning strategy, $\Pr_{\sigma,H}\left[T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right] > 0$.

I now consider a class of strategies that includes σ, and show that σ is suboptimal within this class. Specifically, each strategy in this class is characterized by a pair of beliefs, $(\underline{\tilde{p}}, \tilde{p}'')$ such that (i) $\underline{\tilde{p}} < \tilde{p}'' < \rho_0/(\rho_0+\rho_1)$ (ii) $\tilde{p}'' \geq \varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p''))$.

For a pair of beliefs $\left(\underline{\tilde{p}}, \tilde{p}''\right)$ satisfying these restrictions, let $\sigma_{\underline{\tilde{p}},\tilde{p}''}$ be the strategy such that (see Figure 8),

$$\tau_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(p) = \begin{cases} t(p, \underline{\tilde{p}}) & p < \underline{\tilde{p}} \\ \tau_\sigma(\varphi(\underline{p}_\sigma, t(\underline{\tilde{p}}, p))) & \underline{\tilde{p}} \leq p \leq \varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p')), \\ t(p, \tilde{p}'') & \varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p')) < p < \tilde{p}'', \\ \tau_\sigma\left(\varphi\left(p'', t\left(\tilde{p}'', p\right)\right)\right) & \tilde{p}'' \leq p \leq \rho_0/(\rho_0+\rho_1), \\ \tau_\sigma(p) & p > \rho_0/(\rho_0+\rho_1), \end{cases}$$

and

$$a_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(p) = \begin{cases} 0 & p < \underline{\tilde{p}} \\ a_\sigma(\varphi(\underline{p}_\sigma, t(\underline{\tilde{p}}, p))) & \min\{\underline{p}_\sigma, \underline{\tilde{p}}\} \leq p \leq \varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p')), \\ 0 & \varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p')) < p < \tilde{p}'', \\ a_\sigma\left(\varphi\left(p'', t\left(\tilde{p}'', p\right)\right)\right) & \tilde{p}'' \leq p \leq \rho_0/(\rho_0+\rho_1), \tilde{p}'' \leq p \leq \rho_0/(\rho_0+\rho_1), \\ a_\sigma(p) & p > \rho_0/(\rho_0+\rho_1). \end{cases}$$

In words, according to the strategy $\sigma_{\underline{\tilde{p}},\tilde{p}''}$, starting from a belief of 0, the agent joins the queue as soon as the belief reaches $\underline{\tilde{p}}$. After joining, the agent behaves "as if" he had joined at $\underline{p}$: the Markov strategy is adjusted so that the induced paths of actions are identical. Because σ involves reneging at p', according to $\sigma_{\underline{\tilde{p}},\tilde{p}''}$ the agent reneges at $\varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p'))$. Note that this is not necessarily the first time the agent reneges after joining: σ may involve restarting the queue multiple times before abandoning, as so would do $\sigma_{\underline{\tilde{p}},\tilde{p}''}$.

The second parameter $\tilde{p}''$ is the belief at which the agent rejoins after $\varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p'))$. According to $\sigma_{\underline{\tilde{p}},\tilde{p}''}$ the agent rejoins at $\tilde{p}''$ and thereafter follows the path of actions prescribed by σ after p''. Note that if $\varphi(\underline{\tilde{p}}, t(\underline{p}_\sigma, p')) = \tilde{p}''$, then $\sigma_{\underline{\tilde{p}},\tilde{p}''}$ is a non-abandoning strategies.

By construction, for $\tau \in \{\tau_0(n), n \geq 1\}$, (cf. (15) and (16))

$$\begin{aligned} y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0) = & \Pr{}_{\sigma,H}\left[T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] \\ & \cdot \left(\mathbb{E}_{\sigma,H}\left[T(\tau) - \tau \mid T(\tau) - \tau \leq t(\underline{p}_\sigma, p')\right] + t(\underline{p}_\sigma, \underline{\tilde{p}})\right) \\ & + \Pr{}_{\sigma,H}\left[T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right]\left(\mathbb{E}_{\sigma,H}\left[T(\tau) - \tau \mid T(\tau) - \tau \geq t(\underline{p}_\sigma, p'')\right] + t(p'', \tilde{p}'')\right), \end{aligned} \tag{17}$$

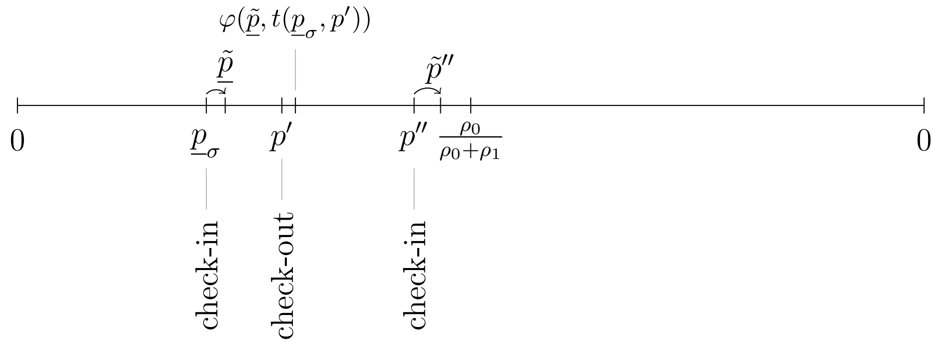


Figure 8: Strategy that involves abandoning at some $p' < \rho_0/(\rho_0+\rho_1)$

and

$$\begin{aligned} d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0) =& \Pr{}_{\sigma,H}\left[T(\tau)-\tau \le t(\underline{p}_\sigma, p')\right] \\ & \cdot e^{-(\rho_0+\rho_1)t(\underline{p}_\sigma,\underline{\tilde{p}})} \cdot \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau)-\tau \le t(\underline{p}_\sigma, p')\right] \\ & + \Pr{}_{\sigma,H}\left[T(\tau)-\tau \ge t(\underline{p}_\sigma, p'')\right] \\ & \cdot e^{-(\rho_0+\rho_1)t(p'',\tilde{p}'')}\mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau)-\tau \ge t(\underline{p}_\sigma, p'')\right]. \end{aligned} \tag{18}$$

The choice of the optimal strategy within the class of strategies $\sigma_{\underline{\tilde{p}},\tilde{p}''}$, can be formulated as a choice over pairs $(t(\underline{p}_\sigma,\underline{\tilde{p}}), t(p'',\tilde{p}''))$ subject to suitable feasibility constraints. Specifically, the constraint (i) and (ii) imply that the pair $(t(\underline{p}_\sigma,\underline{\tilde{p}}), t(p'',\tilde{p}''))$ must satisfy

$$\begin{cases} t(\underline{p}_\sigma,\underline{\tilde{p}}) \ge -t(0,\underline{p}_\sigma), \\ t(\underline{p}_\sigma,\underline{\tilde{p}}) + t(p'',\tilde{p}'') \ge -t(p',p''). \end{cases} \tag{19}$$

There is a one-to-one mapping between pairs $(t(\underline{p}_\sigma,\underline{\tilde{p}}), t(p'',\tilde{p}''))$ and strategies $\sigma_{\underline{\tilde{p}},\tilde{p}''}$. I conclude the proof by showing that the optimal pair must be lying at the boundary, that is, it must satisfy (19) with equality. Proceeding by contradiction, suppose that the optimal pair is interior. Then,

$$\frac{\partial\gamma^H(\sigma_{\underline{\tilde{p}},\tilde{p}''})}{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}\frac{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(\underline{p}_\sigma,\underline{\tilde{p}})} + \frac{\partial\gamma^H(\sigma_{\underline{\tilde{p}},\tilde{p}''})}{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}\frac{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(\underline{p}_\sigma,\underline{\tilde{p}})} = 0,$$

and

$$\frac{\partial\gamma^H(\sigma_{\underline{\tilde{p}},\tilde{p}''})}{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}\frac{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(p'',\tilde{p}'')} + \frac{\partial\gamma^H(\sigma_{\underline{\tilde{p}},\tilde{p}''})}{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}\frac{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(p'',\tilde{p}'')} = 0.$$

Or equivalently, using the fact that the payoff is strictly monotone in the sufficient statistics,

$$\frac{\frac{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(\underline{p}_\sigma,\underline{\tilde{p}})}}{\frac{\partial y_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(p'',\tilde{p}'')}} = \frac{\frac{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(\underline{p}_\sigma,\underline{\tilde{p}})}}{\frac{\partial d_{\sigma_{\underline{\tilde{p}},\tilde{p}''}}(0)}{\partial t(p'',\tilde{p}'')}}. \tag{20}$$

But from (17) and (18), this is impossible, because

$$\frac{e^{-(\rho_0+\rho_1)t(\underline{p}_\sigma,\underline{\tilde{p}})} \cdot \mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau)-\tau \leq t(\underline{p}_\sigma,p')\right]}{e^{-(\rho_0+\rho_1)t(p'',\tilde{p}'')}\mathbb{E}_{\sigma,H}\left[e^{-(\rho_0+\rho_1)(T(\tau)-\tau)} \mid T(\tau)-\tau \geq t(\underline{p}_\sigma,p'')\right]} > 1,$$

whereas (20) holds true only if this ratio equals 1. Because the constraint (19) is slack when evaluated at σ (as $p' < p''$), this concludes the proof that any optimal strategy must be non abandoning.

Additional remark on the proof. Note that the proof extends to the case of menus. In that case, the strategy must define not only the (Markovian) stopping time and entry/exit action, but also the chosen queueing discipline. Specifically, alongside the function a_σ : $[0,1] \to \{0,1\}$, one must define a function $Q_\sigma : [0,1] \to \mathbf{Q}$ such that, along the path induced by σ,

$$Q_k = Q_\sigma(p_{\tau_k}).$$

The same proof applies, with the function a_σ replaced by the pair of functions (a_σ, Q_σ), yielding that abandoning is suboptimal also in a more general setup with menus.

B.2 Proofs for Section 4

B.2.1 Preliminaries.

Fix an offered waiting-time distribution $H \in \mathcal{H}$. Given Proposition 1, it suffices to restrict attention to non-abandoning strategies $\sigma \in \Sigma^{NA}$ for which $\overline{p}_\sigma = 1$. As discussed in Section 3.3, the sufficient statistics for such a strategy$\sigma \in \Sigma^{NA}$ are functions of the threshold belief $\underline{p}_\sigma$, and of some summary statistics of the induced waiting-time distributions, $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$. Moreover, whenever $\gamma(\sigma; H) > 0$, $\hat{H}_{\sigma,0}$ and $\hat{H}_{\sigma,1}$ are convolutions of countably many truncated versions of H, and hence belong to $\mathcal{H}$. The next lemma identifies the set of pairs (δ, λ) for which there exists a distribution $H \in \mathcal{H}$ such that $(\delta^H, \lambda^H) = (\delta, \lambda)$.

Lemma 12. *The following are equivalent:*

1. *There exists a distribution $H \in \mathcal{H}$*

$$\mu^H = \mu, \qquad \delta^H = \delta;$$

2. it does hold that $(\delta,\mu)\in\Gamma^+$, *where*

$$\Gamma^+ := \left\{(\delta,\mu)\in(0,1)\times(0,\infty) : \delta \geq e^{-(\rho_0+\rho_1)\mu}\right\}.$$

Proof. Let $(\delta,\mu)\in\Gamma$. If $\delta = e^{-(\rho_0+\rho_1)\mu}$, the degenerate distribution that puts mass 1 on μ yields the desired statistics. For the case $\delta \neq e^{-(\rho_0+\rho_1)\mu}$, consider a distribution H that randomizes between $\{\varepsilon/q, (\mu-\varepsilon)/(1-q)\}$, with probability $(q, 1-q)$, where $0<\varepsilon\ll\mu$ and $q>0$ are chosen to satisfy

$$q\, e^{-(\rho_0+\rho_1)\varepsilon/q} + (1-q)\, e^{-(\rho_0+\rho_1)(\mu-\varepsilon)/(1-q)} = \delta.$$

If ε was 0, the previous equation would have a unique root $q\in(0,1)$. Because the left-hand side is continuous in ε, there exist an $\varepsilon>0$, and a $q\in(0,1)$ such that the equality is satisfied. Clearly $H\in\mathcal{H}$ and by construction $(\mu^H,\delta^H)=(\mu,\delta)$.

To prove the other direction, notice that $e^{-(\rho_0+\rho_1)x}$ is convex in x, so that for a given mean μ, the minimum value for the other statistics is achieved by the degenerate distribution that puts mass 1 on μ. This, together with the fact that $e^{-(\rho_0+\rho_1)x}<1$ for any $x>0$, concludes the proof that for any $H\in\mathcal{H}$, $(\delta,\mu)\in\Gamma^+$. $\square$

Substituting for y_σ, q_σ, and d_σ from Section 3.3 in (13), the payoff from a strategy $\sigma\in\Sigma^{NA}$ can be written as

$$\gamma(\sigma;H) = G\left(\delta^{\hat{H}_{\sigma,0}}, \mu^{\hat{H}_{\sigma,0}}, \delta^{\hat{H}_{\sigma,1}}, \mu^{\hat{H}_{\sigma,1}}, p_\sigma\right),$$

where

$$G(\delta_0,\mu_0,\delta_1,\mu_1,p) := \frac{\left((1-\delta_0)\frac{\rho_0}{\rho_0+\rho_1}+\delta_0 p\right)(\theta_1-\mu_1 c) + (1-\delta_1)\frac{\rho_1}{\rho_0+\rho_1}(\theta_0-\mu_0 c)}{\left((1-\delta_0)\frac{\rho_0}{\rho_0+\rho_1}+\delta_0 p\right)\mu_1 + (1-\delta_1)\frac{\rho_1}{\rho_0+\rho_1}\left(\mu_0+\frac{1}{\rho_0+\rho_1}\ln\left(\frac{\rho_0}{(1-p)\rho_0-p\rho_1}\right)\right)}.$$

Replacing for y_σ, q_σ, and d_σ in (14), the (a.s. limit of the) long-run service rate induced by σ is equal to

$$\lim_{t\to\infty}\frac{1}{t}N_t = \Lambda\left(\delta^{\hat{H}_{\sigma,0}}, \mu^{\hat{H}_{\sigma,0}}, \delta^{\hat{H}_{\sigma,1}}, \mu^{\hat{H}_{\sigma,1}}, p_\sigma\right),$$

where

$$\begin{aligned}\Lambda(\delta_0,\mu_0,\delta_1,\mu_1,p)\\ := \frac{\delta_0 p + (1-\delta_0)\frac{\rho_0}{\rho_0+\rho_1} + (1-\delta_1)\frac{\rho_1}{\rho_0+\rho_1}}{\left(\delta_0 p + (1-\delta_0)\frac{\rho_0}{\rho_0+\rho_1}\right)\mu_1 + (1-\delta_1)\frac{\rho_1}{\rho_0+\rho_1}\left(\mu_0+\frac{1}{\rho_0+\rho_1}\ln\left(\frac{\rho_0}{(1-p)\rho_0-p\rho_1}\right)\right)}.\end{aligned} \tag{21}$$

For later purposes, notice that the following identities hold. (The decomposition in Section 4.4 is a special case of these identities.)

$$\begin{aligned}G(\delta_0,\mu_0,\delta_1,\mu_1,p) =\\ \Lambda(\delta_0,\mu_0,\delta_1,\mu_1,p)\left(m(\delta_0,\delta_1,p)(\theta_1-\mu_1 c) + (1-m(\delta_0,\delta_1,p))(\theta_0-\mu_0 c)\right),\end{aligned} \tag{22}$$

and

$$\Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p) = \frac{1}{m(\delta_0, \delta_1, p)\, \mu_1 + (1 - m(\delta_0, \delta_1, p))\, (\mu_0 + t(p))}, \tag{23}$$

where

$$m(\delta_0, \delta_1, p) := \frac{(1 - \delta_0) \frac{\rho_0}{\rho_0 + \rho_1} + \delta_0 p}{(1 - \delta_0) \frac{\rho_0}{\rho_0 + \rho_1} + \delta_0 p + (1 - \delta_1) \frac{\rho_1}{\rho_1 + \rho_0}}. \tag{24}$$

The next lemma shows that for any tuple of summary statistics, $(\delta_0, \mu_0, \delta_1, \mu_1)$, there exists a unique optimal cutoff strategy within the class of non-reneging strategies. Let $p^* : \Gamma \times \Gamma \to [0, \rho_0/(\rho_0 + \rho_1)]$ be defined as

$$p^*(\delta_0, \mu_0, \delta_1, \mu_1) := \begin{cases} 0 & \text{if } \beta(\delta_0, \mu_0, \delta_1, \mu_1) \leq \alpha(\mu_0, \mu_1) - 1 < -1\,, \\ \frac{\rho_0}{\rho_0 + \rho_1} & \text{if } \beta(\delta_0, \mu_0, \delta_1, \mu_1) \geq 0, \text{ or } \alpha(\mu_0, \mu_1) > 0, \\ \frac{\rho_0}{\rho_0 + \rho_1} \left(1 - \frac{\beta(\delta_0, \mu_0, \delta_1, \mu_1)}{W_{-1}\left(e^{-1 + \alpha(\mu_0, \mu_1)} \beta(\delta_0, \mu_0, \delta_1, \mu_1)\right)}\right) & \text{otherwise,} \end{cases}$$

where W_{-1} is the (negative branch of the) Lambert function and

$$\begin{aligned} \alpha(\mu_0, \mu_1) &= -(\rho_0 + \rho_1) \frac{\theta_1 \mu_0 - \theta_0 \mu_1}{\theta_1 - c\mu_1}, \\ \beta(\delta_0, \mu_0, \delta_1, \mu_1) &= -\frac{\rho_0 (\theta_1 - c\mu_1) + (1 - \delta_1)\, \rho_1 (\theta_0 - c\mu_0)}{\delta_0 \rho_0 (\theta_1 - c\mu_1)}. \end{aligned} \tag{25}$$

Lemma 13. *Given $(\delta_0, \mu_0, \delta_1, \mu_1) \in \Gamma^+ \times \Gamma^+$, there exists a unique $p \in [0, \rho_0/(\rho_0 + \rho_1)]$ that solves*

$$\max_{p \in [0, \rho_0/(\rho_0 + \rho_1)]} G(\delta_0, \mu_0, \delta_1, \mu_1, p).$$

It equals $p^*(\delta_0, \mu_0, \delta_1, \mu_1)$.

Proof. Taking first-order conditions yields

$$\Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p) \Bigg(- G(\delta_0, \mu_0, \delta_1, \mu_1, p) \left(\delta_0 \mu_1 + \frac{\rho_1}{\rho_0 + \rho_1} \frac{1 - \delta_1}{(1 - p)\, \rho_0 - p\rho_1} \right) + \delta_0 (\theta_1 - c\mu_1) \Bigg) = 0, \tag{26}$$

where the left-hand side is the derivative of $G(\delta_0, \mu_0, \delta_1, \mu_1, p)$ with respect to p. Differentiating further, it is readily verified that the second-order conditions hold whehever the first-order conditions do and $G(\delta_0, \mu_0, \delta_1, \mu_1, p) > 0$.

Using the notation introduced in (25), and omitting the arguments of α and β, the first-order conditions simplify, after some algebra, to,

$$\frac{\rho_0}{\rho_0(1-p)-\rho_1 p} = \frac{W_{-1}\left(e^{-1+\alpha}\beta\right)}{\beta}. \tag{27}$$

If $\beta+1>\alpha$ and $\alpha<0$,

$$\frac{W_{-1}\left(e^{-1+\alpha}\beta\right)}{\beta} > \frac{W_{-1}\left(e^{-1+\alpha}\left(-1+\alpha\right)\right)}{-1+\alpha} = 1\,,$$

and (27) has a unique root in $(0, \rho_0/(\rho_0+\rho_1))$.

Inspection of the derivative of $G(\delta_0,\mu_0,\delta_1,\mu_1,p)$ with respect to p reveals that the function, whenever non-negative, is decreasing only if it is strictly positive. Hence, the optimal p is equal to zero if only if $G(\delta_0,\mu_0,\delta_1,\mu_1,p)$ is decreasing at $p=0$. This yields that $p^*(\delta_0,\mu_0,\delta_1,\mu_1)=0$ if and only if

$$(1-\delta_0)\,\frac{\rho_0}{\rho_0+\rho_1}\,(\theta_1 - c\,\mu_1) + (1-\delta_1)\,\frac{\rho_1}{\rho_0+\rho_1}\,(\theta_0 - c\mu_0) - \rho_0\delta_0\,(\theta_1\mu_0 - \theta_0\mu_1) \geq 0. \tag{28}$$

On the other hand, the payoff is decreasing in p at $\rho_0/(\rho_0+\rho_1)$ if and only if

$$\lim_{p\uparrow\frac{\rho_0}{\rho_0+\rho_1}} G(\delta_0,\mu_0,\delta_1,\mu_1,p) = 0^+.$$

Computing this limit, one obtains that $p^*(\delta_0,\mu_0,\delta_1,\mu_1)=\rho_0/(\rho_0+\rho_1)$ if and only if

$$\frac{\rho_0}{\rho_0+\rho_1}\,(\theta_1 - c\,\mu_1) + (1-\delta_1)\,\frac{\rho_1}{\rho_0+\rho_1}\,(\theta_0 - c\,\mu_0) \leq 0. \tag{29}$$

□

It is convenient to define

$$G^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right) \coloneqq G\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right). \tag{30}$$

In the sequel, I denote with $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)}G^*$ the gradient of the function G^*. It is easy to see that the function G^* is Gateaux differentiable in $(\delta_0,\mu_0,\delta_1,\mu_1)$ whenever $G^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)$ is strictly positive. A similar remark holds for $p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)$ whenever interior, because it is a composition of Gateaux differentiable functions. Additionally, I denote with $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)}\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)$ and $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)}m\left(\delta_0,\delta_1,\mu,p^*\left(\delta,\mu,\delta,\mu\right)\right)$ the gradient of these composite functions viewed as functions on $(0,1)\times(0,\infty)\times(0,1)\times(0,\infty)$.

Note that Lemma 3 is a special case of Lemma 13, obtained by taking $\delta_0=\delta_1=\delta$, and $\mu_0=\mu_1=\mu$. When no confusion arises, I denote $G\left(\delta,\mu,\delta,\mu,p\right)$ with $G(\delta,\mu,p)$, $\Lambda\left(\delta,\mu,\delta,\mu,p\right)$ with $\Lambda(\delta,\mu,p)$, $m\left(\delta,\delta,p\right)$ with $m\left(\delta,p\right)$, $p^*\left(\delta,\mu\right)$ with $p^*\left(\delta,\mu,\delta,\mu\right)$, and $G^*\left(\delta,\mu\right)$ with $G\left(\delta,\mu,\delta,\mu,p^*\left(\delta,\mu,\delta,\mu\right)\right)$.

The following fact assembles some technical results to be used in the sequel.

Fact 1. *Let* $(\delta_0, \mu_0, \delta_1, \mu_1) \in \Gamma^+ \times \Gamma^+$ *and* $p \in [0, \rho_0/(\rho_0+\rho_1))$ *be such that* $G(\delta_0, \mu_0, \delta_1, \mu_1) > 0$.

1. $G(\delta_0, \mu_0, \delta_1, \mu_1, p)$ *and* $G^*(\delta_0, \mu_0, \delta_1, \mu_1)$ *are strictly decreasing in* μ_0, μ_1 *and* δ_0, *and strictly increasing in* δ_1.

2. $p^*(\delta_0, \mu_0, \delta_1, \mu_1)$ *is increasing in* μ_0 *and* μ_1, *increasing in* δ_0 *(strictly if interior), and decreasing in* δ_1.

3. $p^*(\delta, \mu)$ *is increasing in* μ. *It is increasing in* δ *if and only if*

$$\left(\frac{\rho_1}{\rho_0+\rho_1}\theta_0 + \frac{\rho_0}{\rho_0+\rho_1}\theta_1 - c\mu\right)(\theta_1 - c\mu)^{-1} \geq 0. \tag{31}$$

Lemma 13 has an immediate important consequence: in the baseline setting with a single queue, there always exists an equilibrium that yields strictly positive payoffs. As a result, the designer can always guarantee that the aggregate payoff is strictly positive.

Lemma 14. *For any set of admissible parameters* $(\theta_1, \theta_0, c, \rho_0, \rho_1, \lambda) \in \mathbb{R}_{>0} \times \mathbb{R}_{<0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$, *there exists an equilibrium* $(\sigma, H, M) \in \Sigma \times \mathcal{H} \times \mathcal{M}$ *that yields a strictly positive payoff.*

Proof. Let $H^u(t) = 1 - e^{t/u}$, for some $u > 0$ to be determined, so that $\delta^{H^u} = 1/(1+(\rho_0+\rho_1)u)$ and $\mu^{H^u} = u$. When the offered waiting-time distribution is equal to H^u, agents are served-in-random-order, so any best reply involves no reneging. By Lemma 13 the unique best reply strategy is a cutoff strategy with cutoff $p^*(\delta^{H^u}, \mu^{H^u})$. Replacing δ_0 and δ_1 with δ^{H^μ}, and μ_0 and μ_1 with μ^{H^u} in the left-hand side of (28), one obtains

$$\frac{(\rho_0+\rho_1)(\theta_0 - cu)u}{1 + u(\rho_0+\rho_1)} < 0. \tag{32}$$

Hence, for any u, the optimal cutoff $p^*(\delta^{H^u}, \mu^{H^u})$ is strictly positive.

I now show that for any $\lambda > 0$ there exists a $u > 0$ such that

$$\Lambda(\delta^{H^u}, \mu^{H^u}, p^*(\delta^{H^u}, \mu^{H^u})) = \lambda, \tag{33}$$

where

$$\Lambda(\delta, \mu, p) = \frac{1}{\mu + (1 - m(\delta, p))\, t(p)}, \qquad m(\delta, p) = \frac{(1-\delta)\frac{\rho_0}{\rho_0+\rho_1} + \delta p}{(1-\delta) + \delta p}.$$

(These equations are obtained by replacing δ_0 and δ_1 with δ, and μ_0 and μ_1 with μ in (23) and (24).) The function $\Lambda(\delta^{H^u}, \mu^{H^u}, p^*(\delta^{H^u}, \mu^{H^u}))$ is continuous in u. By (29), $p^*(\delta^{H^u}, \mu^{H^u}) < \rho_0/(\rho_0+\rho_1)$ for some u small enough, and $\Lambda(\delta^{H^u}, \mu^{H^u}, p^*(\delta^{H^u}, \mu^{H^u}))$ diverges to infinity as $u \to 0$. Moreover, there exists a $\overline{u}$ such that $\Lambda(\delta^{H^{\overline{u}}}, \mu^{H^{\overline{u}}}, p^*(\delta^{H^{\overline{u}}}, \mu^{H^{\overline{u}}})) = 0$. Hence, by the intermediate value theorem, there exists some $0 < u' < \overline{u}$ such that (33) is satisfied, so that

H^u and $\sigma_{p^*(\delta^{H^u},\mu^{H^u})}$ are part of an equilibrium.[44] Because $0 < p^*(\delta^{H^u}, \mu^{H^u}) < \rho_0/(\rho_0+\rho_1)$, the first-order conditions (26) hold. Note that in any equilibrium, the payoff is non-negative because each agent can guarantee a payoff of 0 by never queueing. Hence, $\theta_1 - cu > 0$. This, together with (26), implies that the payoff is strictly positive. □

B.2.2 Proof of Proposition 2.

(1.) I show that, if $H \notin \mathcal{H}^N$, then an agent finds it optimal to renege (and immediately rejoin) at some $p > \rho_0/(\rho_0 + \rho_1)$.

Because payoffs are determined by the summary statistics $\delta^{\hat{H}_{\sigma,0}}$, $\mu^{\hat{H}_{\sigma,0}}$, $\delta^{\hat{H}_{\sigma,1}}$, $\mu^{\hat{H}_{\sigma,0}}$, *a priori*, studying the best-reply problem requires characterizing the set of statistics that an agent can generate by restarting the queue potentially arbitrarily many times. In the proof, I compute the set of statistics for a restricted class of non-abandoning strategies. I refer to this class as constant restart strategies, and denote the set of such strategies with $\Sigma^K \supset \Sigma^{NR}$. I show that, given $H \notin \mathcal{H}^N$, any strategy $\sigma \in \Sigma^{NR}$ is suboptimal within this class. A strategy $\sigma \in \Sigma^K$ is characterized by the threshold belief $\underline{p}_\sigma$ and a constant $k_\sigma \leq \operatorname{supp}\sup H$. It involves no reneging at beliefs below the invariant probability. At beliefs above the invariant probability, it prescribes the agent to renege and rejoin the queue after constant intervals of times of length $k_\sigma > 0$. Formally,

$$\tau_\sigma(p) = \begin{cases} t(p, \underline{p}_\sigma) & p \leq \underline{p}_\sigma, \\ \infty & \underline{p}_\sigma < p < \rho_0/(\rho_0 + \rho_1), \\ k_\sigma \lceil t(1,p)/k_\sigma \rceil - t(1,p) & p \geq \rho_0/(\rho_0 + \rho_1), \end{cases}$$

and

$$a_\sigma(p) = \begin{cases} 0 & p \leq \underline{p}_\sigma, \\ 1 & \text{otherwise.} \end{cases}$$

Note that any non-reneging strategy $\sigma \in \Sigma^{NR}$ is a constant restart strategy such that $H(k_\sigma) = 1,\ k_\sigma \in (0, \infty]$. The following lemma (the proof of which is immediate and omitted) characterizes the distributions $\hat{H}_{0,\sigma}$, and $\hat{H}_{\sigma,1}$, for any strategy $\sigma \in \Sigma^K$.

Lemma 15. *Fix $\sigma \in \Sigma^K$, and let $H \in \mathcal{H}$ be the offered waiting-time distribution. Then, $\hat{H}_{0,\sigma} = H$, and $\hat{H}_{\sigma,1}$ satisfies*

$$1 - \hat{H}_{\sigma,1}(t) = (1 - H(k_\sigma))^{\lfloor t/k_\sigma \rfloor} \left(1 - H\left(t - \lfloor t/k_\sigma \rfloor k_\sigma\right)\right).$$

[44] It can be shown that (33) has a unique positive root (details available upon request). Consequently, there exists a unique equilibrium in which service is rendered in random order.

Moreover, $\mu^{\hat{H}_{\sigma,1}} = M^H(k_\sigma)$*, and* $\delta^{\hat{H}_{\sigma,1}} = D^H(k_\sigma)$*, where*

$$M^H(k) \coloneqq \frac{1}{H(k)} \int_0^k (1 - H(t))\, \mathrm{d}t,$$

$$D^H(k) \coloneqq \frac{(\rho_0 + \rho_1) \displaystyle\int_0^k e^{-(\rho_0+\rho_1)t} H(t)\, \mathrm{d}t + H(k) e^{-(\rho_0+\rho_1)k}}{1 - (1 - H(k))\, e^{-(\rho_0+\rho_1)k}}.$$

The remainder of the proof relies on one step of the proof of Lemma 2. In particular, there I show that if $H \in \mathcal{H}^N$, then $\left(\delta^H, \mu^H\right) \in \Gamma^{NR}$. Here, I use the fact that if $H \notin \mathcal{H}^N$, then $\left(\delta^H, \mu^H\right) \notin \Gamma^{NR}$.

I distinguish two cases (Lemma 16 and Lemma 17), depending on whether $H \in \mathcal{H}$ has a bounded or unbounded support. In both cases, I show that there exist a k and a strategy $\sigma \in \Sigma^K$ with $k_\sigma = k$ such that $\mu^{\hat{H}_{\sigma,1}} < \mu^H$, and $\delta^{\hat{H}_{\sigma,1}} > \delta^H$. By Fact 1, the result follows. The proofs of Lemma 16 and Lemma 17 rely on some well-known statistical results that are collected, together with some auxiliaries lemmas, in Section B.2.3.

Lemma 16. *Let* $H \in \mathcal{H}$ *have bounded support with higher endpoint* $\overline{w}$*. If* $\left(\delta^H, \mu^H\right) \notin \Gamma^{NR}$*, there exists a* $k \in \mathbb{R}$*,* $0 < k < \overline{w}$*, such that*

$$M^H(k) < \mu^H, \qquad D^H(k) > \delta^H. \tag{34}$$

Proof. The first step consists in relating the statistics $M^H(k)$ and $D^H(k)$ to statistics of the residual life. In Section B.2.3 I show that for any random variable with distribution $H \in \mathcal{H}$,

$$\mu^H = (1 - H(k))\, m^H(k) + H(k) M^H(k), \tag{35a}$$

$$\delta^H = \left(1 - (1 - H(k))\, e^{-(\rho_0+\rho_1)k}\right) D(k) + (1 - H(k))\, e^{-(\rho_0+\rho_1)k} d(k), \tag{35b}$$

where $m^H(k)$, $m^H : [0, \overline{w}] \to \mathbb{R}$ is the mean residual life function,[45] and $d^H(k)$, $d^H : [0, \overline{w}] \to [0, 1]$ is defined below. Formally, let W be a random variable with distribution H. Then,

$$m^H(w) \coloneqq \begin{cases} \mathbb{E}\left[W - w \mid W > w\right], & \text{for } w \le \overline{w}, \\ 0, & \text{otherwise.} \end{cases} \tag{36}$$

Similarly, define

$$d^H(w) \coloneqq \begin{cases} \mathbb{E}\left[e^{-(\rho_0+\rho_1)(W-w)} \mid W > w\right], & \text{for } w \le \overline{w}, \\ 0, & \text{otherwise.} \end{cases} \tag{37}$$

[45] The concept of mean residual life of a random variable is borrowed from reliability theory (see Shaked and Shanthikumar, 2007). The statistics $M(t)$ appears in that literature in relation with age-replacement policies (see for example Sec. 3.6, Ch. 4 in Unnikrishnan Nair N., P.G. Sankaran, and N. *Quantile-Based Reliability Analysis,Quantile-Based Reliability Analysis* [Birkhäuser Basel, 2013]).

The equations (35) allow me to focus on the statistics of the residual life, which has been studied in the reliability literature. In the remainder of the proof I show, in three claims, that if $\left(\delta^H, \mu^H\right) \notin \Gamma^{NR}$, then there exists a k such that

$$m^H(k) > \mu^H, \qquad d^H(k) < \delta^H,$$

which, by (35), implies (34).

Claim 1. *There exists $w^\ddagger < \overline{w}$ such that $\left(m^H(w), d^H(w)\right) \in \Gamma^{NR}$ for $w \geq w^\ddagger$.*

Proof. By Theorem 4, $m^H(w)$ is right-continuous and strictly positive, and crosses zero at most once. Moreover, it only admits upward jumps because for any $\varepsilon > 0$,

$$m^H(w) - m^H(w-\varepsilon) = \frac{H(w) - H(w-\varepsilon)}{1 - H(w-\varepsilon)} m(w) - \int_{w-\varepsilon}^{w} \frac{1 - H(t)}{1 - H(w-\varepsilon)} \,\mathrm{d}t. \tag{38}$$

Hence, because $m^H(\overline{w}) = 0$, $m^H(w)$ must be decreasing in $\left(w^\ddagger, \overline{w}\right)$, for some $t^\ddagger$ sufficiently close to $\overline{w}$. By Lemma 18, $\left(d^H(w^\ddagger), m^H(w^\ddagger)\right) \in \Gamma^{NR}$. ☐

Let $w^\dagger := \inf\left\{w : \left(d^H(t), m^H(t)\right) \notin \Gamma^{NR}, t < w\right\}$. By Claim 1, $w^\dagger < \infty$.

Claim 2. $\lim_{\varepsilon \downarrow 0} \left(m^H(w^\dagger) - m^H(w^\dagger - \varepsilon)\right) \geq 0$.

Proof. If m^H is discontinuous at $w^\dagger$, by (38), the function has an upward jump so that the statement is verified. If $m(w)$ is continuous, then it is increasing at $w^\dagger$, for otherwise, by Lemma 18, $m^H(w^\dagger - \varepsilon) \in \Gamma^{NR}$, for some $\varepsilon > 0$, a contradiction to the definition of $w^\dagger$. ☐

Claim 3. $m^H(w^\dagger) > \mu^H$.

Proof. Assume by contradiction that $m^H(w^\dagger) \leq \mu^H$. By Lemma 19, it must be that $m^H(w) > \mu^H$ for some $w < w^\dagger$, for otherwise $\left(\delta^H, \mu^H\right) \in \Gamma^{NR}$. Moreover, $m^H(0) = \mu^H$. Because m^H only admits upward jumps, there exists at $0 < w^* < w^\dagger$ such that $m^H(w^*) = \mu^H$, and $m^H(w) \leq \mu^H$ for $w \in [w^*, w^\dagger]$. In this case, by Lemma 19,

$$\left(d^H(w^*), m^H(w^*)\right) \in \Gamma^{NR},$$

a contradiction to the definition of $w^\dagger$. It follows that $m^H(w^\dagger) > \mu^H$. Hence,

$$d^H(w^\dagger) \leq \frac{1}{1 + (\rho_0 + \rho_1) m^H(w^\dagger)} \leq \frac{1}{1 + (\rho_0 + \rho_1)\mu^H} \leq \delta^H,$$

where the first follows from $(d^H(t^\dagger), m^H(t^\dagger)) \in \Gamma^{NR}$, and the last from $\left(\delta^H, \mu^H\right) \notin \Gamma^{NR}$. ☐

☐

Lemma 17. *If $H \in \mathcal{H}$ has unbounded support, and there exists a $\sigma \in \Sigma^{NR}$ that is optimal given H, then $H \in \mathcal{H}^N$.*

Proof. The proof follows the one of Lemma 16. Claim 2 and Claim 3 apply with no change, while the argument in Claim 1 must be adapted as follows. By assumption, at any $w \geq 0$, and at any belief p on the recurrent path induced by $\sigma \in \Sigma^{NR}$ and H, the following incentive constraint is satisfied,

$$\begin{aligned}
&\int_0^\infty \Bigg(\left(e^{-(\rho_0+\rho_1)t} p + \left(1 - e^{-(\rho_0+\rho_1)t}\right) \frac{\rho_0}{\rho_0+\rho_1} \right) \theta_1 \\
&\quad + \left(1 - e^{-(\rho_0+\rho_1)t} p - \left(1 - e^{-(\rho_0+\rho_1)t}\right) \frac{\rho_0}{\rho_0+\rho_1} \right) \theta_0 \\
&\quad - (c + \gamma(\sigma; H)) \Bigg) \mathrm{d} \left(\frac{H(t+w) - H(w)}{1 - H(w)} \right) \\
\geq &\int_0^\infty \Bigg(\left(e^{-(\rho_0+\rho_1)t} p + \left(1 - e^{-(\rho_0+\rho_1)t}\right) \frac{\rho_0}{\rho_0+\rho_1} \right) \theta_1 \\
&\quad + \left(1 - e^{-(\rho_0+\rho_1)t} p - \left(1 - e^{-(\rho_0+\rho_1)t}\right) \frac{\rho_0}{\rho_0+\rho_1} \right) \theta_0 - (c + \gamma(\sigma; H)) \Bigg) \mathrm{d}H(t).
\end{aligned}$$

The constraint can be interpreted as a "one-step deviation" principle. It compares the gain from remaining in the queue (left-hand side) and the gain from restarting the queue (right-hand side).

Because, along the history with no service, the belief converges to the invariant probability $\rho_0/(\rho_0 + \rho_1)$, the inequality holds in the limit as $w \to \infty$ only if

$$\lim \sup_{w \to \infty} m(w) \leq \mu.$$

This implies that $w^\dagger := \inf\{t : m(t) \leq \mu, t \geq w\}$ is finite. By Lemma 18,

$$\left(m^H(w^\dagger), d^H(w^\dagger)\right) \in \Gamma^{NR}.$$

The result then follows from Claim 2 and Claim 3 in Lemma 16. □

(2.) If $H \in \mathcal{H}^N$, by Lemma 2, $(\delta^H, \mu^H) \in \Gamma^{NR}$. By Lemma 3, the optimal strategy σ within σ^{NR} is a cutoff strategy with cutoff $p^*(\delta^H, \mu^H)$. As shown in the proof of Lemma 2, for any pair $(\delta, \mu) \in \Gamma^{NR}$ there exists a distribution H' that is a convolution of an exponential distribution and a degenerate distribution, and such that $(\delta^{H'}, \mu^{H'}) = (\delta, \mu)$. In other words, H' is either degenerate, or a shifted (to the right) exponential. Clearly, in the first case agents have no incentives to renege, as they are served in order of arrival. For the second case, assume that H' is a shifted exponential with lower endpoint s. Reneging and rejoining before s is dominated (strictly if $c > 0$) by postponing the check-in time. Restarting after s is dominated for the same reason abandoning is suboptimal when the agent faces an exponential distribution, because it introduces "gaps" in the induced waiting-time distribution. As a result, the optimal strategy within Σ given H' is a non-reneging strategy. Because the optimal non-reneging strategy given H' is only a function of $(\delta^{H'}, \mu^{H'})$, and $(\delta^{H'}, \mu^{H'}) = (\delta^H, \mu^H)$, σ is optimal within Σ^{NR} given H'. Last, by the results in Section 3 the service rate is unchanged because it is only a function of $(\delta^{H'}, \mu^{H'})$.

□

B.2.3 Auxiliary Results

Derivation of (35). First, it is standard to show that (by Fubini's theorem)

$$m(w) = \frac{1}{1 - H(w)} \int_w^\infty 1 - H(t)\,\mathrm{d}t,$$

when $1 - H(w) > 0$. If $H(w) = 0$, then $m(w) = \mu^H$. Equation (35a) immediately follows.

By Lemma 20, $d(w)$ can be written as

$$d(w) = (\rho_0 + \rho_1) \int_w^\infty e^{-(\rho_0+\rho_1)(t-w)} \frac{H(t) - H(w)}{1 - H(w)}\,\mathrm{d}t\,.$$

Equation (35b) immediately follows.

Lemma 18. *Let W be a random variable with distribution $H \in \mathcal{H}$. If there exists a $w^\dagger$ such that $m^H(w) \le m^H(w^\dagger)$ for any $w \ge w^\dagger$, then $\left(d^H(w^\dagger), m^H(w^\dagger)\right) \in \Gamma^{NR}$.*

Proof. The random variable $\left[W - w^\dagger \mid W \ge w^\dagger\right]$ is an NBUE random variable. In fact, by definition, $\mathbb{E}\left[W - w^\dagger \mid W \ge w^\dagger\right] = m^H(w^\dagger)$, and its mean residual life function at $w \ge 0$ equals $m^H(w^\dagger + w)$. By Lemma 2,

$$\left(d^H(w^\dagger), m^H(w^\dagger)\right) = \left(\mathbb{E}\left[W - w^\dagger \mid W \ge w^\dagger\right], \mathbb{E}\left[e^{-(\rho_0+\rho_1)\left(W-w^\dagger\right)} \mid W \ge w^\dagger\right]\right) \in \Gamma^{NR}.$$

□

Lemma 19. *Fix $H \in \mathcal{H}$. If there exist w' and $w'' > w'$ such that $m^H(w') \ge m^H(w)$ for all $w \in [w', w'']$, and $\left(d^H(w''), m^H(w'')\right) \in \Gamma^{NR}$, then $\left(d^H(w'), m^H(w')\right) \in \Gamma^{NR}$.*

Proof. For any $w \in [w', w'']$, the difference $m^H(w'') - m^H(w)$ (respectively $d^H(w'') - d^H(w)$) can be written as functions of (i) the values taken by the conditional distribution $\frac{H(\cdot) - H(w')}{1 - H(w')}$ in $[w', w'']$, (ii) the difference $w'' - w$, and (iii) $m^H(w'')$ (respectively $d^H(w'')$). Hence, one can construct a non-negative random variable with distribution $\tilde{H}$ such that

$$\begin{aligned} m^{\tilde{H}}(t) &= m^H\left(w' + t\right) && \text{for } t \le w'' - w', \\ d^{\tilde{H}}(t) &= d^H\left(w' + t\right) && \text{for } t \le w'' - w', \end{aligned}$$

and arbitrary $m^{\tilde{H}}(t)$ for $t > w'' - w'$, as long as $m : \mathbb{R}_+ \to \mathbb{R}$ satisfies the properties in Theorem 4. In particular, let

$$\tilde{H}(s) = \begin{cases} \dfrac{H\left(w' + s\right) - H\left(w'\right)}{1 - H\left(w'\right)} & \text{for } s \le w'' - w', \\ \dfrac{1 - H\left(w''\right)}{1 - H\left(w'\right)} F(s - (w'' - w')) & \text{for } s \ge w'' - w', \end{cases}$$

where $F : \mathbb{R} \to [0,1]$ is the cdf of a random variable with decreasing mean residual life such that $(\delta^F, \mu^F) = (m^H(w''), d^H(w''))$. The existence of F is guaranteed by the proof of Lemma 2. As a result, $\tilde{H}$ has decreasing mean residual life and hence is an NBUE distribution, i.e., $\tilde{H} \in \mathcal{H}^N$. By Lemma 2, $(\delta^{\tilde{H}}, \mu^{\tilde{H}}) \in \Gamma^{NR}$, and by construction $(\delta^{\tilde{H}}, \mu^{\tilde{H}}) = \big(d^H(w'), m^H(w')\big)$. □

The following Lemma seems to be standard but, as I could not find a proof, I give one.

Lemma 20. *Let X be a non-negative random variable with distribution function F_X. Then*

$$\int_{\mathbb{R}^+} e^{-(\rho_0+\rho_1)x} F_X(\mathrm{d}x) = (\rho_0 + \rho_1) \int_0^\infty e^{-(\rho_0+\rho_1)x} F(x)\,\mathrm{d}x\,.$$

Proof. Define the random variable $Z := e^{-(\rho_0+\rho_1)X}$. Z is a non-negative random variable and, for any $x \geq 0$ its distribution function is $F_Z(x) = 1 - F_X\left(-\ln(x)/(\rho_0+\rho_1)\right)$. Hence

$$\begin{aligned}\mathbb{E}\left[Z\right] &= \int_0^\infty (1 - F_Z(x))\,\mathrm{d}x = \int_0^\infty \left(F_X\left(-\ln(x)/(\rho_0+\rho_1)\right)\right)\mathrm{d}x \\ &= (\rho_0+\rho_1)\int_0^\infty e^{-(\rho_0+\rho_1)x} F(x)\,\mathrm{d}x\,.\end{aligned}$$

□

For convenience, I report a standard result from statistics. (See (36) for the definition of the mean residual function.)

Theorem 4 (Th. 2.1 in Guess and Proschan (1988))**.** *Consider the following conditions*

(i) $m : [0,\infty) \to [0,\infty)$.

(ii) $m(0) > 0$.

(iii) *m is right continuous (not necessarily continuous).*

(iv) *$m(t) + t$ is increasing on $[0,\infty)$.*

(v) *when there exists t_0 such that $m(t_0^-) := \lim_{t\to t_0^-} m(t) = 0$, then $m(t) = 0$ holds for $t \in [t_0,\infty)$. Otherwise, when there does not exist such a t_0 with $m(t_0^-) = 0$, then $\int_0^\infty 1/m(u)\,\mathrm{d}u = \infty$ holds.*

A function m satisfies (i)–(v) *if and only if m is the mean residual life function of a non-degenerate at* 0 *life distribution.*

B.2.4 Proof of Lemma 2.

Let $H \in \mathcal{H}^N$. First, $\left(\delta^H, \mu^H\right) \in \Gamma^+$ (see Lemma 12). Second, any NBUE random variable with mean μ is smaller than $\text{Exp}[\mu]$ in the convex stochastic order[46], where $\text{Exp}[\mu]$ is the exponential random variable with the mean μ (see Shaked and Shanthikumar, 2007, Chap. 3, Th. A.55). Because the function $x \mapsto e^{-(\rho_0+\rho_1)x}$ is convex, this implies that any $H \in \mathcal{H}^N$ satisfies

$$\delta^H \leq \mathbb{E}\left[e^{-(\rho_0+\rho_1)\text{Exp}(\mu)}\right] = \frac{1}{1+(\rho_0+\rho_1)\,\mu}.$$

This proves that for any $H \in \mathcal{H}^N$, $\left(\delta^H, \mu^H\right) \in \Gamma^{NR}$. It remains to show the converse. Let $(\delta, \mu) \in \Gamma^{NR}$. Note that any degenerate distribution belongs to $\mathcal{H}^N$. For $\mu > 0$, let $\text{D}(\mu)$ denote the random variable degenerate at μ. Then, by the properties of the moment generating function,

$$\mathbb{E}\left[e^{-(\rho_0+\rho_1)(\alpha \text{D}(\mu)+(1-\alpha)\text{Exp}(\mu))}\right] = e^{-(\rho_0+\rho_1)\alpha\mu}\frac{1}{1+(\rho_0+\rho_1)\,(1-\alpha)\,\mu}.$$

Let α be such that

$$e^{-(\rho_0+\rho_1)\alpha\mu}\frac{1}{1+(\rho_0+\rho_1)\,(1-\alpha)\,\mu} = \delta. \tag{39}$$

Because $\delta \in \left[e^{-(\rho_0+\rho_1)\mu}, \frac{1}{1+(\rho_0+\rho_1)\mu}\right]$, and the left-hand side is decreasing in α, for $\alpha \in [0,1]$, there exists a unique root α to (39) in $[0,1]$. This shows that one can find an H which is a convolution of a degenerate distribution and an exponential distribution, and such that $(\delta, \mu) = \left(\delta^H, \mu^H\right)$. Because convolutions of IHR distributions are IHR, the random variable $\alpha\text{D}(\mu) + (1-\alpha)\text{Exp}(\mu)$ is IHR, and hence it is an NBUE random variable. Because $\mu < \infty$, and neither $\text{D}(\mu)$, nor $0 < \text{Exp}(\mu)$ have atoms at 0, $H \in \mathcal{H}^N$. □

B.2.5 Proof of Lemma 4

First, by the previous results, for any $H \in \mathcal{H}$ such that $\delta^H = \delta$ and $\mu^H = \mu$, the almost-sure limit of N_t/t as $t \to \infty$ induced by H together with the best non-reneging strategy $\sigma_{p^*(\delta,\mu)}$ equals $\Lambda\left(\delta, \mu, p^*(\delta,\mu)\right)$. In the proof, I show that $\Lambda\left(\delta, \mu, p^*(\delta,\mu)\right)$ is decreasing in μ. The result of the lemma follows.

I consider two cases. First, if $p^*(\delta,\mu) = 0$, by differentiating $\Lambda\left(\delta, \mu, p\right)$ with respect to μ (see (23)), the result is immediate. Instead, if $p^*(\delta,\mu) > 0$, using the payoff decomposition,

[46] The random variable X is said to be smaller than Y in the convex order if

$$\mathbb{E}\left[\phi(X)\right] \leq \mathbb{E}\left[\phi(Y)\right] \text{ for all convex functions } \phi : \mathbb{R} \to \mathbb{R},$$

provided the expectation exists.

one can write the first order conditions as

$$\frac{\partial\Lambda(\delta,\mu,p)}{\partial p}\cdot\Big(m(\delta,p)\,\theta_1+(1-m(\delta,p))\,\theta_0-c\,\mu\Big)\\ +\Lambda(\delta,\mu,p)\,\frac{\partial m(\delta,p)}{\partial p}(\theta_1-\theta_0)=0. \tag{40}$$

Because $\partial m(\delta,p)/\partial p>0$ the first-order conditions hold only if

$$\frac{\partial\Lambda(\delta,\mu,p)}{\partial p}<0.$$

Last, the optimal cutoff $p^*(\delta,\mu)$ is increasing in μ. Combining this with the fact that $\partial\Lambda(\delta,\mu,p)/\partial\mu<0$, the result follows.

□

B.2.6 Calculations for Theorem 1

For a given set of parameters, the two pairs of summary statistics (δ,μ) that solve

$$\Lambda(\delta,\mu,p^*(\delta,\mu))=\lambda,$$

and lie on the west and east boundaries of the admissible set are easy to compute numerically. As explained in Lemma 2, and given the definitions 6 and 7, they correspond respectively to the first-come first-served and service-in-random-order disciplines. Let $(\theta_1,\theta_0,\rho_0,\rho_1,\lambda)=(1,-3/4,1/2,1/2,1)$.

When $c=1/5$, the summary statistics associated to the first-come first-served discipline are $(\delta_{\text{FCFS}},\mu_{\text{FCFS}})\simeq(0.557,0.585)$, and this discipline yields a payoff $G^*(\delta_{\text{FCFS}},\mu_{\text{FCFS}})=0.275$. The statistics associated to the service-in-random-order discipline are $(\delta_{\text{SIRO}},\mu_{\text{SIRO}})\simeq(0.621,0.610)$, and the payoff is equal to $G^*(\delta_{\text{SIRO}},\mu_{\text{SIRO}})\simeq0.325$.

It can be numerically verified that the slope of the indifference curve at $(\delta_{\text{SIRO}},\mu_{\text{SIRO}})$ is larger than the slope of the locus of points satisfying $\Lambda(\delta,\mu,p^*(\delta,\mu))=\lambda$, and the opposite is true at $(\delta_{\text{FCFS}},\mu_{\text{FCFS}})$ (see Figure 5). To conclude the proof, it remains to show that any other pair of statistics lying on $\partial\Gamma^\lambda\cap\Gamma$ yields a payoff no larger than $G^*(\delta_{\text{SIRO}},\mu_{\text{SIRO}})$. However, because (31) holds true at $(\delta_{\text{SIRO}},\mu_{\text{SIRO}})$, by the same argument as Lemma 5, no pair (δ,μ) such that $\mu<\mu_{\text{SIRO}}$ can be optimal. The result then follows from the fact that $\delta\leq\delta_{\text{SIRO}}$ for any $(\delta,\mu)\in\partial\Gamma^\lambda\cap\Gamma$.

When $c=2$, the summary statistics associated to the first-come first-served discipline are $(\delta_{\text{FCFS}},\mu_{\text{FCFS}})\simeq(0.743,0.300)$, and this discipline yields a payoff $G^*(\delta_{\text{FCFS}},\mu_{\text{FCFS}})\simeq0.039$. The statistics associated to the service-in-random-order discipline are $(\delta_{\text{SIRO}},\mu_{\text{SIRO}})\simeq(0.762,0.312)$, and the payoff is equal to $G^*(\delta_{\text{SIRO}},\mu_{\text{SIRO}})\simeq0.035$. In this case, $p^*(\delta,\mu)$ is decreasing in δ at $(\delta_{\text{FCFS}},\mu_{\text{FCFS}})$. It follows that $\Lambda(\delta,\mu,p^*(\delta,\mu))$ is increasing in δ and decreasing in μ, so that the lower boundary of the set Γ is increasing (see Figure 5). To show that serving in random order is optimal, it suffices to show that the slope of the indifference curve at $(\delta_{\text{FCFS}},\mu_{\text{FCFS}})$ is smaller than the slope of the locus of points satisfying

$\Lambda(\delta, \mu, p^*(\delta, \mu)) = \lambda$, and that the two curves intersect at most once, which can be easily verified numerically (see Figure 5).

B.2.7 Proof of Lemma 5

Let $(\delta, \mu) \in \Gamma^{NR}$ and $h \in \mathbb{R}^2_+$ be a direction such that

$$\nabla_{(\delta,\mu)} G^*(\delta, \mu) \cdot h = 0. \tag{41}$$

The proof uses the decomposition in (6). Given the assumption, the optimal cutoff $p^*(\delta, \mu)$ is increasing in μ and δ. (See Fact 1.) Hence,

$$\nabla_{(\delta,\mu)} m(\delta, p^*(\delta, \mu)) \cdot h > 0.$$

By (6), (41) holds only if

$$\nabla_{(\delta,\mu)} \Lambda(\delta, \mu, p^*(\delta, \mu)) \cdot h < 0.$$

If follows that the optimal pair $(\delta, \mu) \in \Gamma^{NR}$ must lie at the east boundary of the set Γ^{NR}. □

B.2.8 Proof of Lemma 6

The proof proceeds by contradiction. Let c_0 denote the initial cost. Let $(\delta^*_{c_0}, \mu^*_{c_0})$ be the optimal choice of the designer given the initial parameters. By assumption, $\delta^*_{c_0} = 1/(\mu^*_{c_0}(\rho_0 + \rho_1)+1)$. Suppose that for any $c > c_0$, the service-in-random-order discipline remains optimal.

Claim 4.

$$(\delta^*_{c_0}, \mu^*_{c_0}) \in \Gamma^{NR}_c$$

Proof. If (δ, μ) are the statistics of an exponential distribution, then $p^*(\delta_{c_0}, \mu_{c_0}) > 0$ (see proof of Lemma 14). For a fixed (δ, μ), the function $p^*(\delta, \mu)$ is increasing in c. Moreover, the function $\Lambda(\delta, \mu, p)$ is decreasing in p when evaluated at $p^*(\delta, \mu)$ (see (40)). Because Γ^{NR} is an upper contour set, the result follows. □

Hence, if service-in-random-order remains optimal, then $\delta^*_c > \delta^*_{c_0}$, and $\mu^*_c < \mu^*_{c_0}$, since they both lie on the curve of descriptive statistics associated to the exponential family. Because $\Lambda(\delta, \mu, p)$ is decreasing in μ, and increasing in δ, $p^*(\delta_c, \mu_c) > p^*(\delta_{c_0}, \mu_{c_0})$. Thus, taking a sequence of $c > c_0$ such that $c \to \infty$, $p^*(\delta^*_c, \mu^*_c) \to \rho_0/(\rho_0 + \rho_1)$.

For any $c > c_0$, let $h = (h^c_\delta, h^c_\mu) \in \mathbb{R}^2_+$ be a direction such that

$$\nabla_{(\delta,\mu)} G^*(\delta^*_c, \mu^*_c) \cdot h^c = 0,$$

where $(\delta^*_c, \mu^*_c) \in \Gamma^{NR}$ is the optimal pair given the cost parameter c. For convenience, set without loss $h^c_\mu = 1$. Using the payoff decomposition in (6), and the fact that the capacity constraint is binding to simplify the expressions,

$$\begin{aligned}
&\nabla_{(\delta,\mu)}\Lambda\left(\delta_c^*,\mu_c^*,p^*(\delta_c^*,\mu_c^*)\right)\cdot h^c \\
&\qquad = \lambda^2\Bigg(-h_\mu^c + t(p^*(\delta_c^*,\mu_c^*))\left(\nabla_{(\delta,\mu)}m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)\cdot h^c\right) \\
&\qquad\quad - \frac{1-m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)}{\rho_0(1-p^*(\delta_c^*,\mu_c^*))-\rho_1 p^*(\delta_c^*,\mu_c^*)}\nabla_{(\delta,\mu)}p^*(\delta_c^*,\mu_c^*)\cdot h^c\Bigg),
\end{aligned} \tag{42}$$

whereas the variation in the second term in (6) is

$$\left(\nabla_{(\delta,\mu)}m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)\cdot h^c\right)(\theta_0-\theta_1) - c\,h_\mu^c. \tag{43}$$

Because the optimal pair (δ_c^*,μ_c^*) is assumed to lie on the east boundary,

$$\nabla_{(\delta,\mu)}\Lambda\left(\delta_c^*,\mu_c^*,p^*(\delta_c^*,\mu_c^*)\right)\cdot h^c \le 0.$$

Hence, by definition of h, the expression in (43) must be nonnegative.

Because $(\theta_1-\theta_0)/c \to 0$ while $t(p^*(\delta_c^*,\mu_c^*)) \to \infty$, for c sufficiently large,

$$\nabla_{(\delta,\mu)}m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)\cdot h^c \ge \frac{c\,h_\mu^c}{\theta_0-\theta_1} > \frac{2}{t(p^*(\delta_c^*,\mu_c^*))}.$$

I obtain a contradiction by showing that, in the limit, (42) is positive.

First, note that along the sequence as $c\to\infty$, the capacity constraint is satisfied, by (6), it must hold that

$$\lim_{c\to\infty}\left(1-m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)\right)t(p^*(\delta_c^*,\mu_c^*)) \to 1/\lambda.$$

On the other hand, by definition of $t(p)$, this implies that

$$\frac{1-m^*\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)}{\rho_0(1-p^*(\delta_c^*,\mu_c^*))-\rho_1 p^*(\delta_c^*,\mu_c^*)} \to \infty. \tag{44}$$

Second, by the first-order conditions, and the definition of h, because $h_\mu^c > 0$ (from (46)),

$$-\frac{h_\delta^c}{\delta^2} + \frac{\rho_0+\rho_1}{\rho_0(1-p^*(\delta_c^*,\mu_c^*))-\rho_1 p^*(\delta_c^*,\mu_c^*)}\nabla_{(\delta,\mu)}p^*(\delta_c^*,\mu_c^*)\cdot h^c < 0.$$

For this to hold in the limit as $c\to\infty$, either the second term must converge to a constant or $\nabla_{(\delta,\mu)}p^*(\delta_c^*,\mu_c^*)\cdot h^c < 0$. In the first case, the fact that (42) is eventually positive follows from $m\left(\delta_c^*,p^*(\delta_c^*,\mu_c^*)\right)\to 1$. In the second, it is implied by (44). This concludes the proof.

B.3 Proofs for Section 5

B.3.1 Proof of Theorem 2

Preliminaries. Before solving the problem [M], I introduce some notation and state a few preliminary results. Fix a binary menu $(H_0, H_1) \in \mathcal{H} \times \mathcal{H}$. Strategies that satisfy the incentive compatibility constraint (9) are cutoff strategies, and Lemma 13 characterizes the optimal cutoff. Note that the first-oder conditions in Lemma 13 can be restated as

$$G(\delta_0, \mu_0, \delta_1, \mu_1, p) = \frac{\theta_1 - c\mu_1}{\mu_1 + \frac{\rho_1}{\rho_0 + \rho_1}\frac{1-\delta_1}{\delta_0}\frac{1}{(1-p)\rho_0 - p\rho_1}}. \tag{45}$$

For convenience, let

$$\kappa(\delta_0, \mu_0, \delta_1, \mu_1) := \frac{1-\delta_1}{\delta_0}\frac{1}{(1-p^*(\delta_0, \mu_0, \delta_1, \mu_1))\rho_0 - p^*(\delta_0, \mu_0, \delta_1, \mu_1)\rho_1}, \tag{46}$$

so that, when the optimal cutoff is interior, the following identity holds:

$$m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)) = \frac{\frac{1}{1-\delta_1}\frac{\rho_0}{\rho_0+\rho_1} - \frac{1}{(\rho_0+\rho_1)\kappa(\delta_0, \mu_0, \delta_1, \mu_1)}}{\frac{1}{1-\delta_1}\frac{\rho_0}{\rho_0+\rho_1} + \frac{\rho_1}{\rho_1+\rho_0} - \frac{1}{(\rho_0+\rho_1)\kappa(\delta_0, \mu_0, \delta_1, \mu_1)}}. \tag{47}$$

I shall refer to these three equalities several times in the remainder of the proof, as well as to the payoff decomposition in (22) and to Fact 1.

Recall that from (30),

$$G^*(\delta_0, \mu_0, \delta_1, \mu_1) := G(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)).$$

Similarly, define:

$$\Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) := \Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)).$$

Relaxed problem. I start by solving the relaxed program

[RP] $$\max G^*(\delta_0, \mu_0, \delta_1, \mu_1)$$

over $(\delta_0, \mu_0) \in \Gamma^+$ and $(\delta_1, \mu_1) \in \Gamma^{NR}$ subject to

$$G^*(\delta_0, \mu_0, \delta_1, \mu_1) \geq G^*(\delta_1, \mu_1, \delta_1, \mu_1), \tag{IC-0}$$
$$G^*(\delta_0, \mu_0, \delta_1, \mu_1) \geq G^*(\delta_0, \mu_0, \delta_0, \mu_0), \tag{IC-1}$$
$$\Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \leq \lambda. \tag{C}$$

The problem is relaxed for a few reasons. First, the non-reneging constraint for the low type is dropped. Second, requiring that $(\delta_1, \mu_1) \in \Gamma^{NR}$ only guarantees that the high type does not have an incentive to renege and rejoin the queue H_1. However the incentive problem is more involved as the high type could renege and rejoin the other queue. I first state the solution of the program [RP], then conclude the proof of Theorem 2, and last present the proof of the maximization.

Lemma 21. *There exists a solution to* [RP]. *It is such that* (C) *binds and either of the following holds:*

(i) $\delta_0 = \delta_1$ *and* $\mu_0 = \mu_1$*;*

(ii) $\delta_0 = e^{-(\rho_0+\rho_1)\mu_0}$, $\mu_0 \geq \mu_1$, *and* $\delta_1 = 1/(1+(\rho_0+\rho_1)\mu_1)$, *and* (IC-0) *is binding.*

Conclusion of the proof of Theorem 2. It remains to prove that, in the case of a separating menu, the agents have no incentives to renege. Because H_0 is degenerate, (IC-0) suffices to guarantee that the low type does not have incentives to renege. Because $\mu_1 \leq \mu_0$ and $\delta_1 > \delta_0$, and the waiting-time distribution H_1 is memoryless, the high type, even along an arbitrarily long history with no service, does not find it optimal to leave his queue and join the queue H_0. Because H_0 is degenerate, there is no need to check for additional deviations, and this concludes the proof.

Proof of Lemma 21 It is convenient to state [RP] as a problem over $(\delta_0, \mu_0, \delta_1, \mu_1) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$, by formulating the domain restrictions $(\delta_0, \mu_0) \in \Gamma^+$ and $(\delta_1, \mu_1) \in \Gamma^{NR}$ as explicit constraints

$$e^{-(\rho_0+\rho_1)\mu_1} - \delta_1 \leq 0, \tag{WB-1}$$

$$e^{-(\rho_0+\rho_1)\mu_0} - \delta_0 \leq 0, \tag{WB-0}$$

$$\delta_1 - \frac{1}{1+(\rho_0+\rho_1)\,\mu_1} \leq 0. \tag{EB-1}$$

Define the Lagrangian function as

$$\begin{aligned} L(\delta_0, \mu_0, \delta_1, \mu_1, \boldsymbol{\eta}) = {} & G^*(\delta_0, \mu_0, \delta_1, \mu_1) + \\ & + \eta_1 (\lambda - \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1)) \\ & + \eta_2 (G^*(\delta_0, \mu_0, \delta_1, \mu_1) - G^*(\delta_1, \mu_1, \delta_1, \mu_1)) \\ & + \eta_3 (G^*(\delta_0, \mu_0, \delta_1, \mu_1) - G^*(\delta_0, \mu_0, \delta_0, \mu_0)) \\ & + \eta_4 \left(\delta_0 - e^{-(\rho_0+\rho_1)\mu_0}\right) + \eta_5 \left(\frac{1}{1+(\rho_0+\rho_1)\,\mu_1} - \delta_1\right) + \eta_6 \left(\delta_1 - e^{-(\rho_0+\rho_1)\mu_1}\right), \end{aligned}$$

where $\boldsymbol{\eta} \in \mathbb{R}^6_+$ is a vector of multiplier.

If $(\delta_0^*, \mu_0^*, \delta_1^*, \mu_1^*) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$ and $\boldsymbol{\eta}^* \geq \mathbf{0}$, $\boldsymbol{\eta}^* \neq \mathbf{0}$ are such that

(a) the constraints (IC-1), (IC-0),(C), (WB-1), (WB-0), (EB-1) and the complementary slackness conditions are satisfied;

(b) $L(\delta_0^*, \mu_0^*, \delta_1^*, \mu_1^*, \boldsymbol{\eta}^*) \geq L(\delta_0, \mu_0, \delta_1, \mu_1, \boldsymbol{\eta}^*)$, for any $(\delta_0, \mu_0, \delta_1, \mu_1) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$,

then $(\delta_0^*, \mu_0^*, \delta_1^*, \mu_1^*)$ is optimal.

In the following, I first derive necessary conditions for a $(\delta_0, \mu_0, \delta_1, \mu_1) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$ and $\boldsymbol{\eta} \geq \mathbf{0}$, $\boldsymbol{\eta} \neq \mathbf{0}$ to satisfy (a) and (b). It is then easy to show that for any set of parameters, such a pair exists and the optimal $(\delta_0, \mu_0, \delta_1, \mu_1)$ must satisfy the conditions in the statement of Lemma 21.

First, let assume, throughout the claims 5–10, that $(\delta_0, \mu_0, \delta_1, \mu_1) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$ and $\boldsymbol{\eta} \geq \mathbf{0}$, $\boldsymbol{\eta} \neq \mathbf{0}$ satisfy (a) and (b).

Claim 5. *It holds that* $(\delta_0, \mu_0) \in \Gamma^{NR}$.

Proof. The proof proceeds by contradiction. If $(\delta_0, \mu_0) \notin \Gamma^{NR}$, $\mu_0 > \mu_1$. Otherwise, if $\mu_0 \leq \mu_1$, $\delta_0 > 1/(1 + (\rho_0 + \rho_1)\mu_0) \geq 1/(1 + (\rho_0 + \rho_1)\mu_1) \geq \delta_1$, where the first inequality follows from $(\delta_0, \mu_0) \notin \Gamma^{NR}$ and the last from $(\delta_1, \mu_1) \in \Gamma^{NR}$. This, by Fact 1, would violate the constraint (IC-1). The fact that $(\delta_0, \mu_0) \notin \Gamma^{NR}$ implies that $\eta_4 = 0$. Also by Fact 1, if $\mu_0 > \mu_1$, (IC-0) is satisfied only if $\delta_0 < \delta_1$. This in turns implies that (IC-1) is slack, so that $\eta_2 = 0$.

By definition of Γ^{NR}, if $(\delta_0, \mu_0) \notin \Gamma^{NR}$, then $\delta_0 > 1/(1 + \mu_0(\rho_0 + \rho_1))$. Replacing into (28), it is easily verified that $p^*(\delta_0, \mu_0, \delta_1, \mu_1) > 0$, whenever $\delta_0 > 1/(1 + \mu_0(\rho_0 + \rho_1))$. On the other hand, because the payoff is bounded away from zero, $p^*(\delta_0, \mu_0, \delta_1, \mu_1) < \rho_0/(\rho_0 + \rho_1)$. Hence, the first-order conditions hold.

Because the function $G^*(\delta_0, \mu_0, \delta_1, \mu_1)$ is decreasing in μ_0 and in δ_0, there exists a direction $h = (h_{\delta_0}, h_{\mu_0}, 0, 0)$, $h_{\delta_0} > 0$, and $h_{\mu_0} < 0$, such that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0.$$

By (45), then $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \kappa(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0$, and by (47)

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)) \cdot h = 0.$$

As a consequence, by the decomposition in (22), because $h_{\mu_0} < 0$, it must hold that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0.$$

Because along the direction h, the incentive constraint (IC-1) is relaxed, for (b) to hold, it must be that $\eta_1 = 0$. But then there exists a direction $h' = \left(h'_{\delta_0}, h'_{\mu_0}, 0, 0\right)$, $h'_{\delta_0} > 0$, $h'_{\mu_0} < 0$, such that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h' = 0,$$

and such that along the direction h', the incentives constraint (IC-1) is relaxed and all other constraints are satisfied. As a result, the Lagrangian is increasing along that direction. This violates (b) and provides the desired contradiction. □

Claim 6. *If* $\mu_0 < \mu_1$*, then* $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$.

Proof. First, as before, if $\mu_0 < \mu_1$, (IC-1) implies that $\delta_1 > \delta_0$ so that both (IC-0) and (WB-1) are slack, and $\eta_3 = \eta_6 = 0$. Assume that $p^*(\delta_0, \mu_0, \delta_1, \mu_1) > 0$, so that the first-order conditions hold.

Consider a change along a direction $h = (0, h_{\mu_0}, 0, h_{\mu_1})$, where $h_{\mu_0} > 0$, and $h_{\mu_1} < 0$ are such that

$$\frac{h_{\mu_0}}{h_{\mu_1}} = -\frac{m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{1 - m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))},$$

so that the "total queueing cost" is unchanged, see (22).

From (45), if the total change in the payoff is negative,

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \kappa(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0.$$

Thus, $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)) \cdot h > 0$. Consequently, by the decomposition in (22), if the total change in the payoff is negative, it must be that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0.$$

Because the constraints (IC-1), (EB-1) and (WB-1) are relaxed along the direction h, it must be that $\eta_1 = 0$. But in this case there exists some $h'_{mu_1} < h_{mu_1}$ such that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot (0, h_{\mu_0}, 0, h'_{\mu_1}) > 0.$$

Again, the constraints (IC-1), (EB-1) and (WB-1) are relaxed along this direction, and hence (b) is violated, yielding the desired contradiction. □

Claim 7. *If* (WB-1) *is slack, then* (IC-0) *is binding.*

Proof. By assumption $\eta_6 = 0$. Suppose by contradiction that (IC-0) is slack, so that $\eta_3 = 0$. Consider a change along a direction $h = (h_{\delta_0}, h_{\mu_0}, 0, h_{\mu_1}, 0)$, $h_{\delta_0} < 0$, $h_{\mu_0} > 0$, and $h_{\mu_1} < 0$ such that

$$\frac{h_{\mu_0}}{h_{\mu_1}} = -\frac{m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{1 - m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))},$$

and

$$h_{\delta_0} + (\rho_0 + \rho_1)e^{-(\rho_0+\rho_1)\mu_0} h_{\mu_0} = 0.$$

I now distinguish two cases depending on whether the first-order conditions hold or not.

Case 1: $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$**.** Since $\partial m(\delta_0, \delta_1, p)/\partial\delta_0 < 0$, by (22),

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0 \Rightarrow \nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0.$$

Because $\partial\Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p)/\partial\mu_1 < 0$, one can find $h'_{\mu_1} < h_{\mu_1}$ such that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot \left(h_{\delta_0}, h_{\mu_0}, 0, h'_{\mu_1}\right) = 0,$$

all other constraints are relaxed or unchanged, and

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot \left(h_{\delta_0}, h_{\mu_0}, 0, h'_{\mu_1}\right) > 0.$$

contradicting the optimality of $(\delta_0, \mu_0, \delta_1, \mu_1)$.

Case 2: $p^*(\delta_0, \mu_0, \delta_1, \mu_1) > 0$**.** In this case the first-order conditions holds. Hence, from (45), for $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h \leq 0$ to hold,

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \kappa(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0.$$

Thus, $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1)) \cdot h > 0$. Consequently, if $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0$, then

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0.$$

Because all active constraints are either unchanged or relaxed along h, for (b) to hold, it must be that $\eta_1 = 0$. But then, given Lemma 22 there exist an $h'_{\mu_1} < h_{\mu_1}$ such that the following inequalities holds:

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot \left(h_{\delta_0}, h_{\mu_0}, 0, h'_{\mu_1}\right) \leq 0,$$
$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot \left(h_{\delta_0}, h_{\mu_0}, 0, h'_{\mu_1}\right) > 0.$$

This contradicts the optimality of $(\delta_0, \mu_0, \delta_1, \mu_1)$. □

Claim 7 has in important consequence: $\mu_1 \leq \mu_0$. In fact, $\mu_1 > \mu_0$ is incompatible with binding (IC-0). (See the proof of Claim 6.)

Claim 8. *If $\mu_0 > \mu_1$, then* (WB-0) *and* (EB-1) *are binding.*

Proof. I consider two cases, depending on whether the first-order conditions hold or not.

Case 1: $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$**.** Suppose first that (WB-0) is slack, so that $\eta_4 = 0$. In this case,

$$\frac{\partial\Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{\partial\delta_0}$$
$$= -\frac{(\mu_0 - \mu_1)\,\rho_0}{(1-\delta_0)\,\rho_0 + (1-\delta_1)\,\rho_1} \frac{\partial\Lambda(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{\partial\mu_0},$$

and

$$\frac{\partial G^*\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\delta_0}$$
$$=-\frac{\left(\mu_0\theta_1-\mu_1\theta_0\right)\rho_0}{\left(1-\delta_0\right)\rho_0\theta_1+\left(1-\delta_1\right)\rho_1\theta_0}\frac{\partial G^*\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\mu_0}.$$

Because $\theta_0<0$,

$$-\frac{\left(\mu_0\theta_1-\mu_1\theta_0\right)\rho_0}{\left(1-\delta_0\right)\rho_0\theta_1+\left(1-\delta_1\right)\rho_1\theta_0}<-\frac{\left(\mu_0-\mu_1\right)\rho_0}{\left(1-\delta_0\right)\rho_0+\left(1-\delta_1\right)\rho_1}.$$

As a result, there exists a direction $h=\left(h_{\delta_0},h_{\mu_0},0,0\right)$, $h_{\delta_0}<0$, $h_{\mu_0}>0$, along which

$$\nabla_{\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}L\left(\delta_0,\mu_0,\delta_1,\mu_1,\boldsymbol{\eta}\right)\cdot h>0,$$

a contradiction. Hence (WB-1) must bind. Assume next that (EB-1) is slack. In this case,

$$\frac{\partial\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\delta_1}$$
$$=\frac{\left(\mu_0-\mu_1\right)\rho_1}{\left(1-\delta_0\right)\rho_0+\left(1-\delta_1\right)\rho_1}\frac{\partial\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\mu_1},$$

and

$$\frac{\partial G^*\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\delta_1}$$
$$=\frac{\left(\mu_0\theta_1-\mu_1\theta_0\right)\rho_1}{\left(1-\delta_0\right)\rho_0\theta_1+\left(1-\delta_1\right)\rho_1\theta_0}\frac{\partial G^*\left(\delta_0,\mu_0,\delta_1,\mu_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)}{\partial\mu_1}.$$

Again,

$$\frac{\left(\mu_0\theta_1-\mu_1\theta_0\right)\rho_1}{\left(1-\delta_0\right)\rho_0\theta_1+\left(1-\delta_1\right)\rho_1\theta_0}>\frac{\left(\mu_0-\mu_1\right)\rho_1}{\left(1-\delta_0\right)\rho_0+\left(1-\delta_1\right)\rho_1}.$$

Because along a direction $h=(0,0,h_{\delta_1},h_{\mu_1},)$, $h_{\delta_1}>0$ and h_{μ_1}, equation IC-0 is relaxed, the constraint (EB-1) binds must bind for (b) to hold.

Case 2: $p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)>0$**.** Suppose that (EB-1) is slack. Consider a change along a direction $h=\left(h_{\delta_0},h_{\mu_0},0,0\right)$, $h_{\delta_0}>0$, $h_{\mu_0}<0$ such that $\nabla_{\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}G^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\cdot h=0$. From the first-order conditions (45), $\nabla_{\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}\kappa\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\cdot h=0$, so that

$$\nabla_{\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}m\left(\delta_0,\delta_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)\cdot h=0.$$

By the decomposition in (22), if the total change in the payoff is zero, it must be that

$$\nabla_{\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}\Lambda^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\cdot h<0.$$

By the usual argument, because all other constraints are relaxed or left unchanged, this is incompatible with (b).

Also, it is possible to do so while preserving incentives, as the constraint (IC-0) was slack. This contradicts the optimality of $(\delta_0, \mu_0, \delta_1, \mu_1)$, and implies that (EB-1) binds.

Assume now that (WB-1) is slack. Consider a change along a direction $h = (h_{\delta_0}, h_{\mu_0}, h_{\delta_1}, 0)$, $h_{\delta_0} > 0$, $h_{\mu_0} < 0$, and $h_{\delta_1} > 0$ such that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^* (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0,$$

and

$$h_{\delta_1} - (\rho_0 + \rho_1) \frac{1}{1 + (\rho_0 + \rho_1)\mu_1} h_{\mu_1} = 0.$$

From the first-order conditions (45), $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \kappa (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0$, and hence $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} m (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0$. By the decomposition in (22), it must be that

$$\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \Lambda^* (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h < 0.$$

This is incompatible with (b) as all other constraints are either relaxed or unchanged along this direction. □

Claim 9. *If* $(\mu_0, \delta_0) \neq (\mu_1, \delta_1)$, *and* $\Lambda^* (\delta_1, \mu_1, \delta_1, \mu_1) > \Lambda^* (\delta_0, \mu_0, \delta_1, \mu_1)$ *then* $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$.

Proof. Proceeding by contradiction, assume that $p^* (\delta_0, \mu_0, \delta_1, \mu_1) > 0$. Since $\mu_1 \leq \mu_0$, (IC-0) and $(\mu_0, \delta_0) \neq (\mu_1, \delta_1)$ imply that $\delta_1 > \delta_0$.

If $p^* (\delta_1, \mu_1, \delta_1, \mu_1) > 0$, by the first-order conditions (45), the following equalities hold

$$\begin{aligned} &\kappa (\delta_1, \mu_1, \delta_1, \mu_1) = \kappa (\delta_0, \mu_0, \delta_1, \mu_1) \\ &m (\delta_1, \delta_1, p^* (\delta_1, \mu_1, \delta_1, \mu_1, p)) = m (\delta_0, \delta_1, p^* (\delta_0, \mu_0, \delta_1, \mu_1)) . \end{aligned}$$

Since, the constraint (IC-0) is binding, and $\mu_0 \geq \mu_1$

$$\Lambda^* (\delta_1, \mu_1, \delta_1, \mu_1) \leq \Lambda^* (\delta_0, \mu_0, \delta_1, \mu_1) .$$

I conclude the proof by showing that if $p^* (\delta_0, \mu_0, \delta_1, \mu_1) > 0$, then $p^* (\delta_1, \mu_1, \delta_1, \mu_1) = 0$ is impossible. I show that for a given μ_1 and δ_1 when moving south east along the indifference curve of $G^* (\delta_0, \mu_0, \delta_1, \mu_1)$ in the (μ_0, δ_0) space, the threshold belief $p^* (\delta_0, \mu_0, \delta_1, \mu_1)$ always increases. Since $(\delta_1, \mu_1, \delta_1, \mu_1)$ and $(\delta_0, \mu_0, \delta_1, \mu_1)$ lie on the same indifference curve and (δ_1, μ_1) is at the south east of (δ_0, μ_0), the result follows.

Let $h = (h_{\delta_0}, h_{\mu_0}, 0, 0)$, $h_{\delta_0} > 0$, and $h_{\mu_0} < 0$ be such that $\partial G^* (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0$. Then, from (45), $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} \kappa (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0$. Since $h_{\delta_0} > 0$, it follows from the definition of the function $\kappa (\delta_0, \mu_0, \delta_1, \mu_1)$ that $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} p^* (\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0$, concluding the proof. □

From the proof of Claim 9, it follows that if $(\mu_0, \delta_0) \neq (\mu_1, \delta_1)$ and $(\delta_0, \mu_0, \delta_1, \mu_1)$ solves [RP], either also $(\delta_1, \mu_1, \delta_1, \mu_1)$ solves [RP], or $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$.

In the statement of Lemma 21, the case (ii) refers to the situation in which $G^*(\delta_0, \mu_0, \delta_1, \mu_1) > G^*(\delta_1, \mu_1, \delta_1, \mu_1)$.

Claim 10. *The constraint* (C) *binds in* $(\delta_0, \mu_0, \delta_1, \mu_1)$.

Proof. If $(\delta_0, \mu_0) = (\delta_1, \mu_1)$ (case (i) in Theorem 2), the result follows from Theorem 1.

If the designer needs to offer a menu to maximize efficiency (case (ii) in Theorem 2), by the previous claim, $p^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h = 0$. In this case,

$$\frac{\partial G^*(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{\partial \delta_0} = \frac{(\mu_0\theta_1 - \mu_1\theta_0)\,\rho_0}{(1-\delta_0)\,\rho_0\theta_1 + (1-\delta_1)\,\rho_1\theta_0} \frac{\partial G^*(\delta_0, \mu_0, \delta_1, \mu_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))}{\partial \mu_0}.$$

By Claim 9, the first-order conditions do not hold and $p^*(\delta_0, \mu_0, \delta_1, \mu_1) = 0$; hence (28) holds strictly. Consequentely,

$$\frac{(\mu_0\theta_1 - \mu_1\theta_0)\,\rho_0}{(1-\delta_0)\,\rho_0\theta_1 + (1-\delta_1)\,\rho_1\theta_0} < \frac{1}{(\rho_0+\rho_1)\delta_0}.$$

Hence, there exist a direction $h = (h_{\delta_0}, h_{\mu_0}, 0, 0)$, $h_{\delta_0} \in \mathbb{R}_+$, and $h_{\mu_0} \in \mathbb{R}_-$ such that the constraint (WB-0) is not violated, and $\nabla_{(\delta_0,\mu_0,\delta_1,\mu_1)} G^*(\delta_0, \mu_0, \delta_1, \mu_1) \cdot h > 0$. If (C) is slack, this contradicts (b). It follows that (C) must be binding.

□

Last, it remains to show that there always exist a $\boldsymbol{\eta} \in \mathbb{R}^6_+$ and a $(\delta_0, \mu_0, \delta_1, \mu_1) \in (0,1) \times (0,\infty) \times (0,1) \times (0,\infty)$ that satisfies (a) and (b). As clear from the proof, for any $\boldsymbol{\eta}$ such that $\eta_1 > 0$, $\eta_2 > 0$, $\eta_3 = 0$, $\eta_4 > 0$, $\eta_5 > 0$, and $\eta_6 = 0$ either of the two candidates satisfies (a)–(b). □

B.3.2 Auxiliary results.

Lemma 22. *The service rate induced by the cutoff strategy* $p^*(\delta_0, \mu_0, \delta_1, \mu_1)$*, that is* $\Lambda(\delta_1, \mu_1, \delta_0, \mu_0, p^*(\delta_0, \mu_0, \delta_1, \mu_1))$*, is decreasing in* δ_0 *whenever the cutoff is interior.*

Proof. The payoff $G^*(\delta_0, \mu_0, \delta_1, \mu_1)$ is decreasing in δ_0. I show that $m(\delta_0, \delta_1, p^*(\delta_0, \mu_0, \delta_1, \mu_1))$ is increasing in δ_0. It follows from the decomposition in (22) that $\Lambda(\delta_1, \mu_1, \delta_0, \mu_0, p^*(\delta_0, \mu_0, \delta_1, \mu_1))$ must be decreasing in δ_0.

Note that

$$\frac{\partial m(\delta_0, \delta_1, p)}{\partial p} = -\left(\frac{\delta_0}{\frac{\rho_0}{\rho_0+\rho_1} - p}\right) \frac{\partial m(\delta_0, \delta_1, p)}{\partial \delta_0},$$

and

$$\frac{\partial m\left(\delta_0,\delta_1,p\right)}{\partial\delta_0}=-\frac{\left(1-\delta_1\right)\frac{\rho_1}{\rho_0+\rho_1}}{\left(\left(1-\delta_0\right)\frac{\rho_0}{\rho_0+\rho_1}+\delta_0 p+\left(1-\delta_1\right)\frac{\rho_1}{\rho_1+\rho_0}\right)^2}\left(\frac{\rho_0}{\rho_0+\rho_1}-p\right)<0.$$

Moreover

$$\frac{\partial p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)}{\partial\delta_0}=\frac{\rho_0}{\rho_0+\rho_1}\cdot\frac{\beta}{\delta_0}\left(\frac{1}{1+W_{-1}\left(e^{-1+\alpha}\beta\right)}\right)=\frac{\frac{\rho_0}{\rho_0+\rho_1}-p}{\delta_0}\left(\frac{W_{-1}\left(e^{-1+\alpha}\beta\right)}{1+W_{-1}\left(e^{-1+\alpha}\beta\right)}\right)>0\,.$$

As a result, the total change in $m\left(\delta_0,\delta_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)\right)$ equals

$$\frac{\partial m\left(\delta_0,\delta_1,p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right),\right)}{\partial\delta_0}=\frac{\partial m\left(\delta_0,\delta_1,p\right)}{\partial\delta_0}\left(\frac{1}{1+W_{-1}\left(e^{-1+\alpha}\beta\right)}\right)>0$$

where the last inequality follows from $e^{-1+\alpha}\beta\in(-1/e,0)$.

□

Lemma 23. *The service rate induced by the cutoff strategy* $p^*\left(\delta_0,\mu_0,\delta_1,\mu_1\right)$*, that is* $\Lambda^*\left(\delta_1,\mu_1,\delta_0,\mu_0\right)$*, is decreasing in* μ_0 *and* μ_1*.*

Proof. I consider two cases. First, assume $p^*(\delta_0,\mu_0,\delta_1,\mu_1)=0$. By differentiating (21), it is immediate to verify that

$$\frac{\partial\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p\right)}{\partial\mu_i}<0,\qquad i=0,1,\tag{48}$$

for any $p\in[0,\rho/(\rho_0+\rho_1))$. Moreover, the optimal cutoff is weakly decreasing in μ_i, $i=0,1$. The result follows.

If instead $p^*(\delta_0,\mu_0,\delta_1,\mu_1)>0$, differentiating (22) with respect to p, the first-order conditions can be written as

$$\begin{aligned}&\frac{\partial\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p\right)}{\partial p}\cdot\\&\cdot\left(m\left(\delta_0,\delta_1,p\right)\left(\theta_1-c\mu_1\right)+\left(1-m\left(\delta_0,\delta_1,p\right)\right)\left(\theta_0-c\mu_0\right)\right)+\\&\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p\right)\cdot\frac{\partial m\left(\delta_0,\delta_1,p\right)}{\partial p}\left(\left(\theta_1-c\mu_1\right)-\left(\theta_0-c\mu_0\right)\right)=0.\end{aligned}$$

It is easy to verify that

$$\frac{\partial m\left(\delta_0,\delta_1,p\right)}{\partial p}>0\,.$$

Whenever the payoff is strictly positive, $(\theta_1-c\mu_1)-(\theta_0-c\mu_0)>0$, so that the first-order conditions hold only if

$$\frac{\partial\Lambda\left(\delta_0,\mu_0,\delta_1,\mu_1,p\right)}{\partial p}<0.$$

Because the optimal cutoff is increasing in μ_i, and given (48),

$$\frac{\partial \Lambda\left(\delta_0, \mu_0, \delta_1, \mu_1, p\right)}{\partial \mu_i} < 0, \qquad i = 0, 1.$$

□

B.4 Proof of Theorem 3

Clearly, it suffices to prove the result for the case in which the optimal menu is separating. Define the sequence of distributions $\{H_n\}_{n\in\mathbb{N}} \in \mathcal{H}$ by:

$$H_n(t) = \begin{cases} 1 - e^{-t/(\mu^{H_1^*}+1/n)} & t < 1/n, \\ 1 - e^{-1/(n\mu^{H_1^*}+1)} & 1/n \leq t < \mu^{H_0^*} - \varepsilon_n, \\ 1 & t \geq \mu^{H_0^*} - \varepsilon_n. \end{cases}$$

for a sequence $\varepsilon_n \to 0$ as $n \to \infty$ to be specified. I show that, for n large enough, the following strategy $\sigma_n \in \Sigma$ is a best reply. According to σ_n, the agent does not renege at beliefs below the invariant probability of state θ_1, $\rho_0/(\rho_0 + \rho_1)$. The strategy σ_n prescribes reneging and rejoining the queue at interval of times $1/n$ for beliefs above $\rho_0/(\rho_0+\rho_1)$. Also, $\overline{p}_{\sigma_n} = 1$ and $\underline{p}_{\sigma_n}$ is the optimal cutoff, as defined in Lemma 13, given the pair of induced waiting-time distributions $\hat{H}_{\sigma_n,0}$ and $\hat{H}_{\sigma_n,1}$ which are easily verified to be

$$\hat{H}_{\sigma_n,0}(t) = H_n(t), \qquad \hat{H}_{\sigma_n,1}(t) = 1 - e^{-t/(\mu^{H_1^*}+1/n)}.$$

Note that H_n is not an NBUE distribution. I now show that σ_n dominates any other non-abandoning strategy. Because $H : \mathbb{R}_+ \to [0, 1]$ is constant in $[1/n, \mu^{H_0^*} - \varepsilon_n]$, reneging when the time-in-queue belongs to this interval is dominated by either reneging at $1/n$ or not reneging. Hence, it remains to show that inducing the pair $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ is preferred to inducing either $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,0})$ or $(\hat{H}_{\sigma_n,1}, \hat{H}_{\sigma_n,1})$. By Theorem 2, (H_0^*, H_1^*) is strictly preferred to (H_0^*, H_0^*); therefore, $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ is preferred to $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,0})$ for n large enough. Moreover, one can find a sequence $\varepsilon_n \to 0$, $\varepsilon_n \geq 0$ such that the agent is indifferent between $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ and (H_0^*, H_1^*), so that $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ is strictly preferred to $(\hat{H}_{\sigma_n,1}, \hat{H}_{\sigma_n,1})$.

Last, it remains to show that the feasibility constraint is satisfied along the sequence. Note that $p^*(\delta^{\hat{H}_{\sigma_n,0}}, \mu^{\hat{H}_{\sigma_n,0}}, \delta^{\hat{H}_{\sigma_n,1}}, \mu^{\hat{H}_{\sigma_n,0}}) = 0$ for n large enough. In fact, as shown in Claim 9, whenever the optimal menu is separating, the first-order conditions do not hold. Because the agent is indifferent between $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ and (H_0^*, H_1^*), and $\delta^{\hat{H}_{\sigma_n,0}} > \delta^{H_0^*}$ and $\delta^{\hat{H}_{\sigma_n,1}} > \delta^{H_1^*}$, it follows that the service rate induced by the best reply to $(\hat{H}_{\sigma_n,0}, \hat{H}_{\sigma_n,1})$ is smaller than (H_0^*, H_1^*), and this concludes the proof. □