The Effects of Information on Abandonment and Congestion in Non-Stationary Priority Queues

Philipp Afèche
Rotman School of Management, University of Toronto, philipp.afeche@rotman.utoronto.ca

Junqi Hu
Strome College of Business, Old Dominion University, jhu@odu.edu

Rouba Ibrahim
School of Management, University College London, rouba.ibrahim@ucl.ac.uk

Vahid Sarhangian
Department of Mechanical and Industrial Engineering, University of Toronto, sarhangian@mie.utoronto.ca

In observable priority service systems, e.g., the emergency department of a hospital, customers are typically left to infer their own information about their queue positions in the waiting room. Because such inferences are likely inaccurate, the system manager may choose to provide accurate information. This paper examines the impact of providing (queue position) information on key operational performance measures. We consider a two-class, multi-server priority queue with time-varying arrivals and state- and information-dependent abandonment rates. We compare two information levels: no information (abandonment rates depend on customers' perceived queue positions) and full information (abandonment rates depend on customers' true queue positions). For each level, we establish fluid approximations of the queue length process, and the existence, uniqueness, and asymptotic stability of a periodic fluid model equilibrium. We leverage these approximations for analytical comparison results of these information levels in terms of the abandonment rate, queue-length, and total cost of each class. These results yield the following guidelines: (1) Providing queue position information does not uniformly improve performance, but instead involves two trade-offs: (i) within each priority class, between queue length and abandonment; (ii) across priority classes, i.e., information may benefit one class while harming the other. These effects are most pronounced when the high-priority class alternates between under- and overloaded. (2) Therefore, the information policy that minimizes the total cost depends on the load, customer perceptions, and the relative costs of delay versus abandonment. (3) Effective information design may require class-specific disclosure or operational levers, e.g. staffing, to eliminate trade-offs.

Key words: State-dependent abandonment; information-dependent abandonment; priority queues; time-varying arrivals; service systems

---

1. Introduction

Customer abandonment is prevalent in service systems and can have a significant impact on system performance. In most service settings, some form of information on the current "state" of the system, or waiting time estimates, is available. This information may be communicated (and, hence, fully controlled) by the provider in unobservable (virtual) queues. In observable service settings, customers are left to infer their own information. Because such inferences are noisy and possibly inaccurate, the system manager may choose to provide accurate information through technologyenabled solutions, e.g., Smartphone Apps. In this paper, we examine the impact of providing (queue position) information on key operational performance measures of the system.

Our primary motivation comes from the Emergency Department (ED) of a hospital, where patients with different medical urgency (and, hence, service priority) share a common waiting room and typically do not receive real-time information regarding their anticipated delays, position, or priority level. Previous (observational) empirical studies (e.g., Batt and Terwiesch 2015, Bolandifar et al. 2019) have found that observable (time-varying) characteristics of the waiting room (number waiting, service speed) and those of the patients (severity levels) impact their abandonment decisions. This raises the question of whether abandonment rate and other system-level performance measures of EDs can be improved by providing information, in form of queue positions or waiting time predictions, to the patients. Examining the effects of providing information about queue positions or delays through experimental or quasi-experimental studies is particularly challenging in the complex setting of an ED. In addition to ethical and quality of care concerns, abandonment rates are also typically small, hence requiring long experiments to obtain sufficiently large samples. A rare example is Westphal et al. (2022) which evaluates the impact of providing operational (current and next stage of care) and wait time information (expected remaining time in ED) through a field-study implementation of a web-based phone application called MyED (Westphal et al. 2020).

In this paper we propose and study a novel model to identify the performance impact of providing (queue position) information in observable service settings, such as the ED. Our goal is to provide insights into the related key trade-offs and how these trade-offs depend on two key system features, priority service and time-varying arrivals, which are prevalent in observable service settings.

We consider a Markovian multi-server two-class queueing system with time-varying Poisson arrivals, operating under a strict priority discipline. Customers in the same priority class are scheduled First-Come, First-Served (FCFS). We model abandonment rates as state-dependent through an increasing and concave function that maps customers' *perceived* queue positions to their abandonment rates. The perceived queue positions depend on the information available to the customers. Our analysis compares the following two levels of information. (1) *No information*: Customers only observe the total queue length, but neither the priority classes of others, nor the FCFS dynamics, nor the positions of individual customers in the queue. In this case, we assume that each customer perceives their position to be a fraction of the total queue length, where this fraction may depend on the customer's class and elapsed waiting time. (2) *Full information*: customers are aware of the priority classes and customer positions, i.e., observe their true queue position.

Mapping to the ED, the high- and low-priority classes in our model correspond to emergency severity index (ESI) levels 2 and 3, and ESI levels 4 and 5, respectively, as the most urgent patients in ESI level 1 do not abandon, cf. Batt and Terwiesch (2015). Accordingly, the capacity in our model corresponds to the residual capacity available to patients in ESI levels 2 or higher. We note that, although our primary motivation comes from observable service settings, our models and results are relevant to virtual queues where information is communicated through delay announcements. In virtual queues, no and full information correspond to communicating, respectively, only the up-to-date total number of customers in the system, and the exact customer positions in the queue.

We compare the long-run average performance of these two information levels for each class in terms of (i) the abandonment rate, (ii) the queue-length, and (iii) the *total cost*, i.e., the cost-weighted sum of (i) and (ii). To this end, we study a many-server fluid approximation. The fluid approximation captures the first-order impact of the information level and yields analytical performance comparison results under the fluid model equilibria. Our analysis yields new insights on the interactions between information and key system features, namely priority service and time-varying arrivals. These insights lead to important implications regarding the provision of information in observable priority service systems with time-varying arrivals such as the ED.

Contributions and Summary of Main Results

1. We study a *novel model of a two-class priority queue with time-varying arrivals and abandonment rates that are state- and information-dependent.* Under full information, each customer knows their exact queue position. Under no information, each customer perceives their queue position as a class- and waiting-time-dependent fraction of the total queue length. We show that this intractable model is accurately approximated by a tractable parsimonious model with a class-dependent constant position fraction. We develop our analytical results for this parsimonious model.

2. *Fluid limits and equilibrium analysis.* We derive fluid approximations of the transient dynamics of the system for each information level and justify them through a Strong Functional Law of Large Number (FSLLN) for the fluid-scale queue length processes. We further leverage an extended Lyapunov method for time-varying systems to establish the convergence of the fluid limits to asymptotic periodic equilibria under periodic arrival rates. The approach is general and can be used to establish the equilibrium behaviour of other time-varying queueing systems. Further, the results lead to new observations on the impact of information on the equilibrium behavior of the system. In particular, the period of the equilibrium queue length may not be the same as that of the arrival rate, depending on the information level.

3. *Performance comparisons of full versus no position information.* We show that information provision does not uniformly improve performance but involves two distinct and interacting trade-offs (see §6.1 for a detailed summary). (i) *Within-class trade-off*: For each priority class, full information may reduce congestion but increase abandonment, or vice versa, depending on system load and how customers perceive their queue positions. (ii) *Cross-class trade-off (HP vs. LP)*: Information may benefit one priority class while harming the other. This trade-off holds for each

metric (queue length, abandonment rate, and total cost). Trade-offs (i) and (ii) both arise in a parameter regime relevant to the ED, where the HP class alternates between under- and overloaded.

Therefore, the information policy that minimizes the total cost, at the class or system level, depends on the load, customer perceptions, and the relative costs of delay versus abandonment. Proposition 2 formalizes this result for systems with alternating HP under-/overload. To isolate the effects of time-varying arrivals and priority service on these trade-offs, we build up to this result via a sequence of results for systems with increasing complexity (Propositions 3, 4, 5, 6, 7 and 8).

These results show that one-size-fits-all transparency policies may be suboptimal, and that effective information design for the ED may require class-specific information disclosure or complementary operational levers, such as staffing, to eliminate adverse trade-offs. For example, the trade-off between priority classes may be eliminated by ensuring that the HP class remains underloaded.

Organization of the Paper In §2 we review the related literature. §3 presents the queueing model. In §4 and §5 we present the fluid limits of the queueing model under the different information levels and establish their convergence to periodic equilibria. In §6, we compare the equilibrium system performance under full versus no information. Finally, in §7 we conclude with a discussion of managerial implications and directions for future work.

2. Related Literature

Our work relates to three bodies of literature that we briefly discuss below: Theoretical studies on delay announcement or state information in service systems; queueing models with customer abandonment; and empirical studies of customer abandonment in service systems.

1. *Theoretical studies on delay announcements or state information in service systems.* This body of (theoretical) studies focuses on aspects of system information, e.g., its accuracy, equilibrium analysis for a given information scheme, and performance comparison across different schemes.

(i) There is a large literature on communicating wait-time information in service systems; see Ibrahim (2018) for a recent survey. A primary focus of this body of literature has been to provide a single announcement at the arrival epoch of the customer; see, e.g., Ibrahim et al. (2017), Bassamboo and Ibrahim (2021). A related body of literature focuses on the impact of providing lagged delay information, e.g., Nirenberg et al. (2018), Lakrad et al. (2022) on the performance of the system. Armony et al. (2009) study the impact of fixed and Last-to-Enter-Service (LES) delay announcements on customer behavior in queues with abandonment, by analyzing a fluid approximation and characterizing the resulting system equilibria. In contrast, *we focus on sequentially updated delay information and evaluate the impact of providing queue position information instead.*

(ii) A stream of papers study and compare the effects of different information provision strategies on the *balking* behavior of rational, utility-maximizing customers, and the resulting performance in terms of throughput, welfare and profit (e.g., Hassin 1986, Chen and Frank 2004, Burnetas and Economou 2007, Guo and Zipkin 2007, 2008, Hassin and Roet-Green 2020). Their general takeaway is that more information may or may not be beneficial. However, *contrary to our model, these papers ignore abandonment and limit attention to stationary single-server FIFO systems.*

(iii) A related stream of work focuses on ticket queues (e.g., Xu et al. 2007, Jennings and Pender 2016, and Kuzu et al. 2019). In ticket queues, customers wait in a virtual queue but are informed of their queue position through a numbered ticket upon arrival. Ticket queues also lead to partial queue information, as some customers may abandon without notifying anyone, and hence render the queue position information on the tickets inaccurate. The literature on ticket queues focuses on understanding the impact of this inaccurate information on system performance. *In contrast, our focus is on comparing performance under different levels of information for priority queues.*

(iv) We model the impact of information on abandonment behaviour through state-dependent (individual) abandonment rates. Earlier studies have used state-dependent abandonment rates to model and compare different delay-announcement strategies. Whitt (1999) and Jouini et al. (2009, 2011) assume that customers react to the delay announcement by balking, but may also subsequently renege if they decide to join. The announcement impacts the queueing performance through the transition rates of the corresponding birth-and-death processes. In our model, the available information impacts the abandonment rates through a general function of the actual state. *In contrast to previous work, we consider a setting where the information (and, hence, the abandonment rates) dynamically changes over time. In addition, we consider time-varying arrivals that cannot be captured through simple birth-and-death processes.*

2. *Queueing models with customer abandonment.* In this body, we group (theoretical) studies that focus on system analysis and optimization, taking the information setting as given.

(i) Our work is closer to the large literature on queueing models that capture abandonment through exogenous rate functions, see, e.g., Bassamboo and Randhawa (2016), Dong and Ibrahim (2021), and Pender (2017) and the references therein. Closely related to our work are studies that consider state-dependent (individual) abandonment rates. For instance, Whitt (2005a,b) propose and study Markovian queueing models with state-dependent abandonment rates to approximate performance in queueing models with general abandonment distributions. *However, they consider a single-class with stationary arrivals and do not examine different information levels, as we do.*

Our study relies on a fluid approximation of the stochastic queueing system. Fluid models are commonly used to approximate performance in queueing models with abandonment (e.g., Whitt 2006, Liu and Whitt 2011b, Dong et al. 2015, Yu et al. 2021a). We derive fluid approximations for the transient dynamics of the system under time- and state-dependent rates using the strong approximation framework of Mandelbaum et al. (1998). We further establish the existence and study the periodic equilibrium of the fluid models under periodic arrivals. Fluid approximations of the time-dependent equilibrium behaviour of queueing systems are also proposed in Heyman and Whitt (1984) and Liu and Whitt (2011a). Perry and Whitt (2016) and Dong and Perry (2020) also establish the existence of a periodic equilibrium for many-server queues but without abandonment and using ad-hoc approaches. In contrast, we leverage general methods from the literature on nonlinear dynamical systems (e.g., Khalil 2002) to establish the results.

(ii) Studies of rational abandonment assume either that customers cannot observe the queue (e.g., Haviv and Ritov 2001, Shimkin and Mandelbaum 2004, Ata et al. 2017, Ata and Peng 2017), or that customers have full information about the queue state (e.g. Hassin 1985, Assaf and Haviv 1990, Afèche and Sarhangian 2015, Cui et al. 2022).

3. *Empirical studies of customer abandonment in service systems.* Several studies have empirically investigated customer abandonment behaviour in both unobservable and observable queues.

(i) In the unobservable (virtual) setting, Akşin et al. (2017) and Yu et al. (2017) use structural estimation models to estimate and understand the mechanism through which delay announcements impact callers' abandonment decisions. Yu et al. (2021b) conduct a field experiment in a call center to examine how delay information impact reference-dependent behaviour of customers. Yu et al. (2022) conduct a randomized field experiment in a ride sharing setting to examine the impact of wait time information and its progress on abandonment in virtual queues.

(ii) Closer to our work are empirical studies concerned with observable or "semi-observable" settings. This body of literature has focused on understanding the impact of the visible aspects of the queue, i.e., queue length, position, and service speed on the abandonment behaviour of customers; see Aksin et al. (2022) and the references therein. The majority of the studies are concerned with single class queues, with the exception of Batt and Terwiesch (2015) and Bolandifar et al. (2019) that are concerned with abandonment behaviour of patients in the multiclass setting of EDs. Queue length has been found to be a primary measure with an increasing effect on customer abandonment, even after controlling for wait. This has been observed in retail (deli) queue (Lu et al. 2013) as well as EDs (Batt and Terwiesch 2015, Bolandifar et al. 2019). Buell (2021) finds evidence from grocery queues that queue position - relative to the total length of the queue - also matters. Previous studies, e.g., Janakiraman et al. (2011), have suggested that customer utility is the combination of disutility from remaining wait, and utility from making progress. In addition to queue length, the service speed has also been observed to impact customer behaviour. Aksin et al. (2022) find, using lab experiments, that the sequence of observed service times impact the abandonment behaviour. Batt and Terwiesch (2015) find evidence from an ED that, in addition to queue length and service speed, patients also respond differently to arrivals of patients with higher severity (and hence priority). Bolandifar et al. (2019) finds that patients of different severity levels have heterogeneous abandonment responses to observable queue features.

In this work, we do not directly model individual customer behaviours. Instead, we model abandonment through a general state-dependent rate function that maps each customer's perception of her position to an abandonment rate. The resulting abandonment behaviour is, however, consistent with the empirical findings on customer abandonment discussed above. In particular, we assume the abandonment rate function is concave and increasing in the queue length. Hence, customers abandon faster from longer queues and making progress at the end of the queue results in a smaller reduction in abandonment probability than for customers who are closer to the head of the queue. Also, observing faster service times helps customers progress towards lower states faster, and hence reduces their abandonment probabilities. Finally, in the no information setting, we assume that customers are not aware of the priority levels of other customers, but different priority classes may have heterogeneous beliefs about their queue positions. *Our modeling framework allows us to derive new insights on the impact of system dynamics - in particular priority discipline and non-stationary arrivals - on abandonment and congestion. These features are prevalent in observable service systems such as EDs. As such, our results are relevant for design of information provision technology and design of future field studies such as Westphal et al. (2022).*

3. Model

We study a queueing system with s identical servers and two customer classes indexed by $k = 1, 2$. Class 1 customers have preemptive priority over class 2 customers; we refer to class 1 as HP (high-priority) and class 2 as LP (low-priority). Customers in the same class are scheduled FCFS.

Arrivals to class k follow a Poisson process with rate $\lambda_k(t)$, where $\lambda_k(t)$ is assumed to be bounded and continuous. We assume that the arrival rate is a periodic function, i.e., $\lambda_k(t) = \lambda_k(t + d_k)$, for $t \geq 0$, where $d_k > 0, d_k \in \mathbb{Q}$ is the fundamental (i.e., smallest) period of the arrival rate function for class k. We also examine the special case with stationary arrivals, i.e., with $\lambda_k(t) = \lambda_k$ for all t (in this case the fundamental period d_k does not exist since $\lambda_k(t) = \lambda_k(t + d_k) \quad \forall\, d_k > 0$). Service times are exponentially distributed with class-dependent rates μ_k.

Customers waiting for service abandon the system once their patience expires. A key feature of our model is that a customer's patience varies with the system state and depends on the information design. Under information design I, the patience time (time to abandonment) of a waiting class k customer in position l of her class (i.e., with the l-th earliest arrival time among class k customers in queue) is exponentially distributed with state-dependent rate $\theta(q^I_{kl}(t)/s)$, where $q^I_{kl}(t)$ is the customer's *perceived queue position* at time t, and $\theta(\cdot)$ is a function that maps her perceived position to her abandonment rate. Note that the abandonment rate is determined by applying $\theta(\cdot)$ to the *scaled* perceived queue position, i.e., after dividing it by the number of servers s. Intuitively, scaling the queue position by s captures the effect of system size on customers' abandonment behaviour: as the number of servers (and the system size) increases, we expect the impact of perceiving larger queues to decrease proportionally. We make the following assumptions on $\theta(\cdot)$.

ASSUMPTION 1. *The abandonment rate function $\theta(\cdot)$ is continuous, strictly increasing, concave, and bounded, i.e., $\theta(x) \leq M$ for all $x \geq 0$ and some finite constant $M > 0$.*

It is natural to assume that $\theta(\cdot)$ increases in a customer's perceived queue position. The assumption that $\theta(\cdot)$ is concave reflects that the marginal increase in a customer's abandonment rate decreases as the *relative* increase in her perceived queue position gets smaller. For example, a customer's abandonment rate increases more if her perceived queue position increases from 10 to 11 (10% relative increase) than from 100 to 101 (1% relative increase).

We assume the same $\theta(\cdot)$ function for all information levels for analytical tractability and to cleanly isolate the interactions between information granularity and system characteristics, e.g., non-stationary arrival rates and priority classes; see also the discussion in Section 7.3.

REMARK 1. This model does not explicitly capture balking (consistent with much of the queueing control literature), but it does imply that customers with large perceived queue positions abandon soon after joining the system.

Customers' perceived queue position depends on the information they receive. Under full information (F), they are informed about their real-time queue position. Under no information (N), customers are only informed about the real-time total number of customers in the system. For each information design $I \in \{F, N\}$, we denote by $\{X_k^I(t) : t \geq 0\}$, $\{Q_k^I(t); t \geq 0\}$, and $\{R_k^I(t); t \geq 0\}$, respectively the processes that keep track of the number of class $k \in \{1, 2\}$ customers in the system, in the queue, and total number abandoned.

Full information. Customers observe the priority classes of all customers as well as their queue positions. They have knowledge of the scheduling policy in the system and, in particular, know that customers are served FCFS within each class. Each waiting customer's perceived queue position matches her exact position, i.e., $q_{kl}^F(t) = (\sum_{i=1}^{k-1} X_i^F(t) + l - s)^+$ for $l > (s - \sum_{i=1}^{k-1} X_i^F(t))^+$.

No information. Customers only observe the total queue length of the system. Customers neither know nor are informed about the scheduling policy. This assumption is reasonable in shared waiting rooms such as the ED or government services, where customers rarely experience the service and it may be difficult for them to infer the scheduling policy based on observation alone. A class k customer's perceived queue position depends on the current queue length and a class- and waiting-time-dependent relative position fraction $\beta_k(w_{kl}(t)) \in (0, 1]$, where $w_{kl}(t)$ is the elapsed waiting time of the class k customer in position l of her class by time t. Specifically, a class k waiting customer's perceived queueing position is given by: $q^N_{kl}(t) = \beta_k(w_{kl}(t)) \cdot (X^N_1(t) + X^N_2(t) - s)^+$ for $l > (s - \sum_{i=1}^{k-1} X^N_i(t))^+$. We note that assuming a class-dependent relative position function does not necessarily mean that customers are aware of their own priority class.

We assume that $\beta_k(\cdot)$ is a non-increasing function with $\lim\limits_{\|x\|\to\infty} \beta_k(x) = 0$, for $k = 1, 2$. This assumption reflects that, as their waiting time $w_{kl}(t)$ accumulates, a customer anticipates progressing forward in the queue, resulting in a decrease in $\beta_k(w_{kl}(t))$. This does not imply that their perceived "absolute" position, i.e., the product of $\beta_k(w_{kl}(t))$ by the current total queue length, $(X^N_1(t) + X^N_2(t) - s)^+$, decreases in their waiting time. Since customers are not informed of their exact queue positions nor the FCFS service discipline, they rely on the total queue length observed in the waiting room to infer their position. Therefore, although $\beta_k(w_{kl}(t))$ decreases over time, their perceived absolute position may change non-monotonically with respect to the waiting time.

Intuitively, $\beta_k(w_{kl}(t))$ reflects the current belief of class k customers about their queue positions under no position information. Let $\boldsymbol{w}(t)$ denote the vector of waiting times $w_{kl}(t)$ for all customers in queue at time t. We refer to this no-information model as the $\beta_k(\boldsymbol{w}(t))$ *model*.

Our goal is to compare the performance under full vs. no information. We focus on the long-run average queue-length, $\lim_{T\to\infty}(1/T)\int_0^T Q^I_k(t)dt$, and average abandonment rate, $\lim_{T\to\infty} R^I_k(T)/T$, for each class $k = 1, 2$. We also consider the weighted sum of these performance measures. Let c_k and V_k denote the class-k unit delay cost and abandonment cost, respectively. The total long-run average class k cost is given by $\lim_{T\to\infty}(1/T)\left(\int_0^T c_k Q^I_k(t)dt + V_k R^I_k(T)\right)$.
The $\beta_k(\boldsymbol{w}(t))$ model is analytically intractable, because the waiting time vector $\boldsymbol{w}(t)$ has time-varying, state-dependent and continuous components $w_{kl}(t)$ that depend on the functions $\theta()$ and $\beta_k()$. Hence, our analytical results focus on a simpler model with static position fractions $\beta_k \in (0, 1]$, where the constant β_k reflects the average belief of class-k customers about their queue positions during their wait. A class k customer's average perceived queue position is $q^N_{kl}(t) = \beta_k \cdot (X^N_1(t) + X^N_2(t) - s)^+$. We call this no information model with static position fractions the β_k *model*.

Importantly, we show via simulation (see Appendix A) that this tractable β_k model accurately approximates both, the equilibrium performance of the $\beta_k(\boldsymbol{w}(t))$ model, and its performance impact compared to full information. Specifically, this simulation study establishes two key results:
(i) The long-run average performance measures under the dynamic $\beta_k(\boldsymbol{w}(t))$ model (with wait-time dependent position fractions) are accurately approximated by a static β_k model (with appropriately chosen constant position fractions): Compared to the dynamic model, the best-fitting β_k model yields small relative errors in the long-run average numbers-in-system (below 0.4%) and average abandonment rates (below 0.3% under high load and below 2% under moderate load). Given the lack of empirical evidence on customers' perceived queue positions, we consider a family of exponential functions $\beta_k(x)$ for two variants of the *No Information* setting: (1) Customers are unaware

of priority classes. For this case we assume a common function, $\beta_1(x) = \beta_2(x)$. (2) Customers know their own priority class but not others' status. We model this case with class-dependent functions that reflect priority awareness, whereby HP customers perceive a more favorable initial position ($\beta_1(0) \le \beta_2(0)$) and a faster advancement rate ($\beta_1'(x) < \beta_2'(x)$) than LP customers.

(ii) The comparisons of No vs. Full information yield consistent results for the $\beta_k(\boldsymbol{w}(t))$ model and the corresponding best-fitting β_k model. (See Appendix A, Figure 10 for an illustrative example.)

In sum, (i)-(ii) support our focus on the β_k model to study the impact of No vs. Full information.

4. Fluid Approximation

In this section, we obtain a fluid approximation for the number-in-system process with or without queue position information. To this end, we consider a sequence of queueing systems described in Section 3, indexed by n. The arrival rate and number of servers scale up uniformly in n whereas service rates and the function $\theta(\cdot)$ remain unscaled.

Let s^n and $\lambda_k^n(t)$ denote, respectively, the number of servers and arrival rates in the n-th system. Denote by $\{X^{F,n}(t) := (X_1^{F,n}(t), X_2^{F,n}(t)) : t \ge 0\}$, and $\{X^{N,n}(t,\boldsymbol{\beta}) := (X_1^{N,n}(t,\boldsymbol{\beta}), X_2^{N,n}(t,\boldsymbol{\beta})) : t \ge 0\}$ the processes that keep track of the number of customers of both classes in system under full and no information, respectively, in the nth system and with $\boldsymbol{\beta} = (\beta_1, \beta_2)$. Similarly, denote by $\{Q^{F,n}(t) := (Q_1^{F,n}(t), Q_2^{F,n}(t)) : t \ge 0\}$, and $\{Q^{N,n}(t,\boldsymbol{\beta}) := (Q_1^{N,n}(t,\boldsymbol{\beta}), Q_2^{N,n}(t,\boldsymbol{\beta})) : t \ge 0\}$ the processes that keep track of the number of customers in queue, and by $\{R^{F,n}(t) := (R_1^{F,n}(t), R_2^{F,n}(t)) : t \ge 0\}$ and $\{R^{N,n}(t,\boldsymbol{\beta}) := (R_1^{N,n}(t,\boldsymbol{\beta}), R_2^{N,n}(t,\boldsymbol{\beta})) : t \ge 0\}$ processes that keep track of the number of customers abandoned under the two information levels. Let $A_k \equiv \{A_k(t) : t \ge 0\}$, $S_k \equiv \{S_k(t) : t \ge 0\}$, and $N_k \equiv \{N_k(t) : t \ge 0\}$ be independent unit-rate Poisson processes corresponding to the arrival, service, and abandonment processes, respectively. Then, under full information, the sample path of $X^{F,n}(t)$ is uniquely determined by the initial state $X^{F,n}(0)$ and the following equations:

$$\begin{aligned} X_1^{F,n}(t) &= X_1^{F,n}(0) + A_1\left(\int_0^t \lambda_1^n(u)du\right) - S_1\left(\mu_1 \int_0^t (X_1^{F,n}(u) \wedge s^n)du\right) \\ &\quad - N_1\left(\int_0^t \mathcal{A}_1^{F,n}(X^{F,n}(u))du\right), \end{aligned} \tag{1}$$

$$\begin{aligned} X_2^{F,n}(t) &= X_2^{F,n}(0) + A_2\left(\int_0^t \lambda_2^n(u)du\right) - S_2\left(\mu_2 \int_0^t (X_2^{F,n}(u) \wedge (s^n - X_1^{F,n}(u))^+)du\right) \\ &\quad - N_2\left(\int_0^t \mathcal{A}_2^{F,n}(X^{F,n}(u))du\right), \end{aligned} \tag{2}$$

where $\mathcal{A}_k^{F,n}(X^{F,n}(u))$ denotes the aggregate class k abandonment rate at time u under full information, defined as,

$$\mathcal{A}_1^{F,n}(X^{F,n}(u)) := \sum_{i=1}^{(X_1^{F,n}(u)-s^n)^+} \theta\left(\frac{i}{s^n}\right), \tag{3}$$

$$\mathcal{A}_2^{F,n}(X^{F,n}(u)) := \sum_{i=(X_1^{F,n}(u)-s^n)^+ +1}^{(X_1^{F,n}(u)+X_2^{F,n}(u)-s^n)^+} \theta\left(\frac{i}{s^n}\right). \tag{4}$$

Note that the queue length processes satisfy $Q_1^{F,n}(u) = X_1^{F,n}(u) \wedge s^n$, $Q_2^{F,n}(u) = X_2^{F,n}(u) \wedge (s^n - X_1^{F,n}(u))^+$ and the abandonment processes satisfy $R_1^F(t) = N_1\left(\int_0^t \mathcal{A}_1^{F,n}(X^{F,n}(u))du\right)$ and $R_2^F(t) = N_2\left(\int_0^t \mathcal{A}_2^{F,n}(X^{F,n}(u))du\right)$. Similarly, we obtain the sample path of $X^{N,n}(t,\boldsymbol{\beta})$ with initial state $X^{N,n}(0,\boldsymbol{\beta})$ and equations (1)–(2) with $X^{F,n}(t)$ and $\mathcal{A}_k^{F,n}(X^{F,n}(u)$ replaced by $X^{N,n}(t,\boldsymbol{\beta})$ and $\mathcal{A}_k^{N,n}(X^{N,n}(u,\boldsymbol{\beta}))$, where,

$$\mathcal{A}_1^{N,n}(X^{N,n}(u,\boldsymbol{\beta}),\boldsymbol{\beta}) := \theta\left(\frac{\beta_1(X_1^{N,n}(u,\boldsymbol{\beta}) + X_2^{N,n}(u,\boldsymbol{\beta}) - s^n)^+}{s^n}\right)(X_1^{N,n}(u,\boldsymbol{\beta}) - s^n)^+, \tag{5}$$

$$\mathcal{A}_2^{N,n}(X^{N,n}(u,\boldsymbol{\beta}),\boldsymbol{\beta}) := \theta\left(\frac{\beta_2(X_1^{N,n}(u,\boldsymbol{\beta}) + X_2^{N,n}(u,\boldsymbol{\beta}) - s^n)^+}{s^n}\right)\left(X_2^{N,n}(u,\boldsymbol{\beta}) - (s^n - X_1^{N,n}(u,\boldsymbol{\beta}))^+\right)^+. \tag{6}$$

For simplicity, we suppress $\boldsymbol{\beta}$ in the expressions of $\mathcal{A}_k^{N,n}(X^{N,n}(t,\boldsymbol{\beta}),\boldsymbol{\beta})$ and $X^{N,n}(t,\boldsymbol{\beta})$ in the remainder of §4 and in §5. By (1) and (2), in the n-th system, the number of class k customers at time t, $X_k^{I,n}(t)$, is equal to the initial number-in-system, plus the cumulative number of arrivals, minus the cumulative number of service completions and abandonments until time t. In turn, the aggregate class k service rate at time t equals a server's class k service rate, multiplied by the number of servers that are available for and working on class k customers.

By equations (3)–(6), the aggregate class k abandonment rate at time t under information level I, $\mathcal{A}_k^{I,n}(X^{I,n}(t))$, equals the sum of the individual abandonment rates of class k customers in queue at time t. That is, $\mathcal{A}_k^{I,n}(X^{I,n}(u)) = \sum_{l=1}^{X_k^{I,n}(t)} \theta(\frac{q_{kl}^I(t)}{s^n})$. Importantly, note that different information designs imply different aggregate abandonment rates for the same system state, $X^{I,n}(u)$.

The following result establishes that the fluid-scaled process $X^{I,n}(t)/n$ (for both information designs) converges to a unique deterministic fluid limit as $n \to \infty$. That is, the fluid limits are good approximations of the corresponding stochastic sample paths when the system is sufficiently large.

THEOREM 1. *Assume that as $n \to \infty$, s^n/n is increasing and $s^n/n \to s$, $\lambda_k^n(t)/n \to \lambda_k(t) < \infty$ uniformly on compact subsets of $\mathbb{R}_+$, and $X_k^{I,n}(0)/n \to x_k^I(0)$ almost surely for $k = 1, 2$, $I \in \{F, N\}$. Under information level I, as $n \to \infty$ the scaled process $\{X^{I,n}(t)/n : t \geq 0\}$ converges almost surely to $\{x^I(t) := (x_1^I(t), x_2^I(t)) : t \geq 0\}$ uniformly on compact subsets of $\mathbb{R}_+$, where $\{x^I(t) : t \geq 0\}$ is the unique solution of the following system of ordinary differential equations starting from initial condition $x^I(0)$:*

$$\dot{x}_1^I(t) = \lambda_1(t) - \mu_1(x_1^I(t) \wedge s) - A_1^I(x^I(t)), \tag{7}$$

$$\dot{x}_2^I(t) = \lambda_2(t) - \mu_2\left((s - x_1^I(t))^+ \wedge x_2^I(t)\right) - A_2^I(x^I(t)), \tag{8}$$

where,

$$A_1^I(x(t)) = \begin{cases} \int_0^{(x_1(t)-s)^+} \theta\left(u/s\right) du, & \textit{if } I = F, \\ \theta\left(\frac{\beta_1 (x_1(t)+x_2(t)-s)^+}{s}\right)(x_1(t)-s)^+, & \textit{if } I = N; \end{cases} \tag{9}$$

$$A_2^I(x(t)) = \begin{cases} \int_{(x_1(t)-s)^+}^{(x_1(t)+x_2(t)-s)^+} \theta\left(u/s\right) du, & \textit{if } I = F, \\ \theta\left(\frac{\beta_2 (x_1(t)+x_2(t)-s)^+}{s}\right)(x_2(t)-(s-x_1(t))^+)^+, & \textit{if } I = N. \end{cases} \tag{10}$$

For simplicity we suppress the argument $\boldsymbol{\beta}$ in $x_1^N(t,\boldsymbol{\beta})$ and $A_k^N(x(t),\boldsymbol{\beta})$ in §4 and §5. The proof of Theorem 1 is in Appendix B. The proof is based on verifying the conditions of Theorem 2.2 of Mandelbaum et al. (1998) which establishes a FSLLN for a general family of queueing processes. In our model, the main difficulty is to establish the Lipschitz continuity of the abandonment rate functions. Furthermore, in the case of full information, the abandonment rates are different for individual customers in the same priority class. This introduces additional technical difficulties when showing the convergence of the summations for the system abandonment rate in (3)–(6) to the integrals in (9) and (10). The assumption that s^n/n is increasing facilitates this proof step.

5. Equilibrium Analysis

In this section, we study the long-run behavior of the fluid models developed in Section 4.

Definition 1. A solution $\tilde{x}^I := \{\tilde{x}^I(t) : t \geq 0\}$ to the system of ODEs (7)–(8) under a fixed information level I is a **periodic equilibrium** if there exists a vector $(p_1,p_2) \in \mathbb{R}_+^2$ such that $\tilde{x}_k^I(t+p_k) = \tilde{x}_k^I(t)$ for all $t \geq 0$ and $k = 1,2$. The smallest (p_1,p_2) pair (if it exists) is referred to as the **fundamental period** of the equilibrium.

Remark 2. Note that, when $\tilde{x}_k^I(t+p_k) = \tilde{x}_k^I(t)$ holds for arbitrarily small p_k, i.e., $\tilde{x}_k^I(t) = \tilde{x}_k^I$ is a constant for some $k \in \{1,2\}$, the fundamental period of the equilibrium does not exist. When $\tilde{x}_k^I(t)$ is a constant for both $k = 1,2$, then the periodic equilibrium reduces to an equilibrium point.

Our first result establishes the existence of a unique periodic equilibrium under the assumption that the service rates for the two classes are the same. Recall that $\lambda_k(t)$ has fundamental period d_k and denote by $d := lcm(d_1,d_2)$ the least common multiple of d_1,d_2 (such d must exist since $d_k \in \mathbb{Q}$). Then, d is a period of the total arrival rate of the system, i.e., $\Lambda(t) := \lambda_1(t)+\lambda_2(t)$, although it may not be its fundamental period. In particular, the fundamental period of $\Lambda(t)$ does not exist when $\Lambda(t)$ is static (see case (ii) of Example 1), and can be either equal to or smaller than d.

Proposition 1. *Assume that $\mu_1 = \mu_2 = \mu$ and $\beta_1 = \beta_2 = \beta$. Under information level $I \in \{F,N\}$, the system (7)-(8) has a unique periodic equilibrium $\tilde{x}^I$ with period (p_1^I,p_2^I), where $(p_1^F,p_2^F) = (d_1,d)$ and $(p_1^N,p_2^N) = (d,d)$.*

We prove Proposition 1 by establishing the existence of a fixed point of the Poincaré map (see Definition 4) with respect to the system of two-dimensional ODEs (7)–(8). The approach is general and can be adapted to establish the existence and uniqueness of a periodic equilibrium for other time-varying queueing models. Since the monotonicity of Poincaré map (see Proposition C.1 in Appendix C.1) only applies to one-dimensional ODEs, we assume equal service rates to convert our two-dimensional ODE to a one-dimensional one in terms of the total queue length $x_1 + x_2$ for $x = (x_1, x_2) \in \mathbb{R}^2$. The assumption of equal service rates and class-independent β is however not necessary for the existence of the equilibrium. Numerical experiments suggest that Proposition 1 continues to hold if $\mu_1 \neq \mu_2$ and $\beta_1 \neq \beta_2$ (see Example 1). Further, we note that under stationary arrivals, the periodic equilibrium reduces to a single equilibrium point $\tilde{x}^I \in \mathbb{R}^2_+$. In this case, under each information level I, Proposition 1 implies that there exists a unique equilibrium point $\tilde{x}^I$.

By (7)–(10), the HP number-in-system process under full information solely depends on the HP arrival rate, $\lambda_1(t)$, whereas the HP and LP number-in-system processes under no information and the LP number-in-system process under full information depend on the arrival rates of both classes. As a result, p_1^F is determined by the period of the HP arrival rate, d_1, while p_1^N, p_2^I, for $I \in \{F, N\}$, are determined by the period of the total arrival rate, d.

Note that (p_1^I, p_2^I) is a period, but not necessarily the fundamental period of the periodic equilibrium $\tilde{x}^I$. However, numerically we observe instances where (p_1^I, p_2^I) is the fundamental period, which implies that the period of the equilibrium could indeed depend on the information level. That is, information could have a first-order impact on the dynamics of the system by changing the period of its equilibrium. This is in contrast with other examples of fluid models in the literature, where the equilibria of the trajectories coincide with that of the arrival rate (e.g., Heyman and Whitt 1984, Dong and Perry 2020). We illustrate this numerically in the following example.

Example 1. Consider the system with $s = 20$, $\mu_1 = 1, \mu_2 = 2$, $\beta_1 = 0.8, \beta_2 = 0.9$, $\theta(x) = 4.2 - 4e^{-x}$. In this case, $\mu_1 < \mu_2$, $\beta_1 < \beta_2$, and $\theta(0) = 0.2 < \mu_2$. Consider the following two sets of sinusoidal arrival rates: (i) $\lambda_1(t) = 20(1 - 0.8\sin(\pi t/5))$, $\lambda_2(t) = 20(1 - 0.8\sin(\pi t/3))$; and (ii) $\lambda_1(t) = 20(1 - 0.8\sin(\pi t/12))$, $\lambda_2(t) = 20(1 + 0.8\sin(\pi t/12))$. In each case, the number-in-system trajectory converges o a periodic equilibrium $(\tilde{x}_1^I, \tilde{x}_2^I)$ for each information level I. In particular,

(i) If $\lambda_1(t) = 20(1 - 0.8\sin(\pi t/5))$ and $\lambda_2(t) = 20(1 - 0.8\sin(\pi t/3))$. Then $(d_1, d_2) = (10, 6)$ and $d = 30$, and the periodic equilibrium $(\tilde{x}_1^I, \tilde{x}_2^I)$ has a fundamental period $(p_1^I, p_2^I) = (30, 30)$, for $I = N$, and $(p_1^I, p_2^I) = (10, 30)$, for $I = F$. In particular, as shown in Figure 1, the fundamental period for the HP class is 30 under No Information and 10 under Full Information. This illustrates that the fundamental period of the equilibrium can depend on the information design.

(ii) If $\lambda_1(t) = 20(1 - 0.8\sin(\pi t/12))$ and $\lambda_2(t) = 20(1 + 0.8\sin(\pi t/12))$. Then $d_1 = d_2 = d = 24$, and the periodic equilibrium $(\tilde{x}_1^I, \tilde{x}_2^I)$ has a fundamental period $(p_1^I, p_2^I) = (24, 24)$ for $I \in \{F, N\}$,

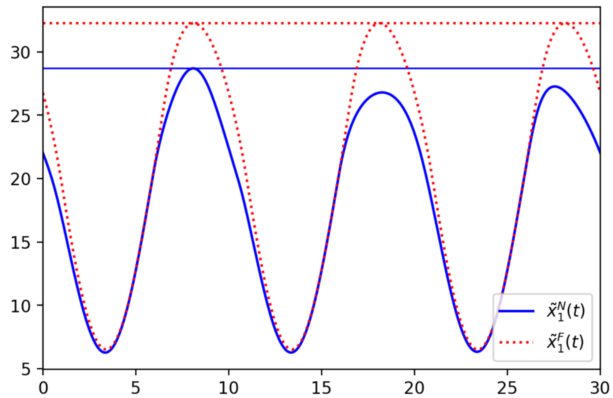


Figure 1 **Trajectories of HP number-in-system processes in periodic equilibrium (**$\mu_1 = 1, \mu_2 = 2,$ $\beta_1 = 0.8, \beta_2 = 0.9,$ $s = 20, \theta(x) = 4.2 - 4e^{-x}, \lambda_1(t) = 20(1 - 0.8\sin(\pi t/5)), \lambda_2(t) = 20(1 - 0.8\sin(\pi t/3))$**).**

(a) HP number-in-system, $p_1^I = 24$.

(b) LP number-in-system, $p_2^I = 24$.

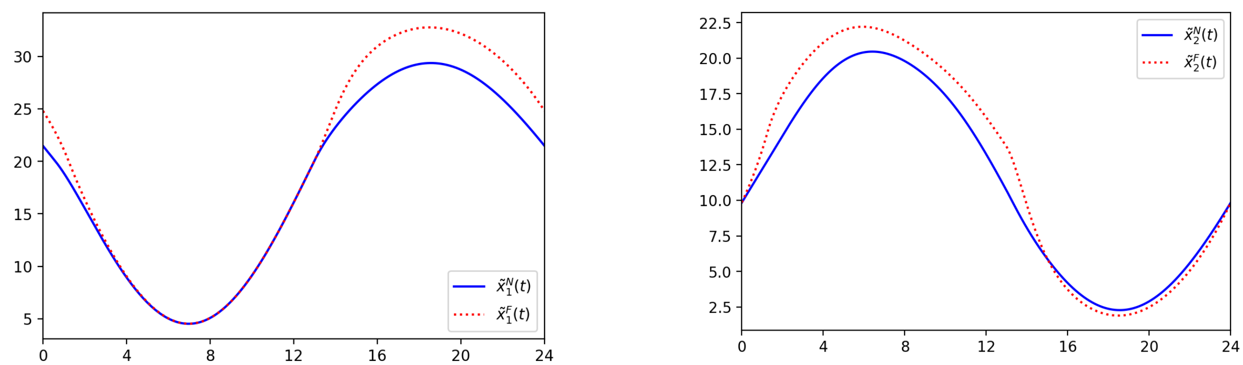


Figure 2 **Trajectories of number-in-system processes in periodic equilibrium (**$\mu_1 = 1, \mu_2 = 2, s = 20, \beta_1 = 0.8, \beta_2 = 0.9,$ $\theta(x) = 4.2 - 4e^{-x}, \lambda_1(t) = 20(1 + 0.8\sin(\pi t/12)), \lambda_2(t) = 20(1 - 0.8\sin(\pi t/12))$**).**

see Figure 2. In this case, the total arrival rate $\Lambda(t) = 40$ is constant, which illustrates that the fundamental period of the equilibrium need not agree with the period of the total arrival process.

Next, we investigate whether the fluid model is *asymptotically periodic*, i.e., whether starting with any initial condition the trajectories converge to the periodic equilibrium $\tilde{x}^I$ as $t \to \infty$. To this end, we examine the stability of the periodic equilibrium. Let $x^I(t)$ be the unique solution to the system of ODEs (7)–(8) under information level I with initial condition $x^I(0)$. Denote by $f_1^I(t, x^I(t))$ and $f_2^I(t, x^I(t))$ the RHS of equations (7) and (8). That is, $f_k^I(t, x^I(t))$ denotes the net flow rate of class k customers under information level I, at time t and state $x^I(t)$. Let $y_k^I(t) := x_k^I(t) - \tilde{x}_k^I(t)$ denotes the deviation of the trajectory of the number-in-system process $x_k^I(t)$ from the periodic equilibrium $\tilde{x}_k^I(t)$, for $k = 1, 2$, and $I \in \{F, N\}$. Consider the following system:

$$\dot{y}_1^I(t) = f_1^I(t, y + \tilde{x}^I) - f_1^I(t, \tilde{x}^I) =: \tilde{g}_1^I(t, y), \tag{11}$$

$$\dot{y}_2^I(t) = f_2^I(t, y + \tilde{x}^I) - f_2^I(t, \tilde{x}^I) =: \tilde{g}_2^I(t, y). \tag{12}$$

Note that although $\lambda_k(t)$ in $f_k^I(t, y + \tilde{x}^I)$ and $f_k^I(t, \tilde{x}^I)$ cancel out, the system (11)–(12) is not time-invariant since $\tilde{g}_k^I$ depends on the the time-varying trajectory $\tilde{x}^I(t)$. Moreover, the solution of (7)–(8) depends on the initial state via $\lambda(0)$.

Definition 2. $(0,0)$ is an **equilibrium point** of $\dot{y} = g(t, y)$ if $g(t, 0) = 0$, for $t \geq 0$.

Observe that $y = (0,0)$ is an equilibrium point for system (11)–(12) since $\tilde{g}_1^I(t, 0) = 0$ and $\tilde{g}_2^I(t, 0) = 0$. The following definition formalizes the notion of stability for time-varying trajectories.

Definition 3. Let $g(t, y)$ be a Lipschitz function defined on $\mathbb{R}_+ \times \mathbb{R}^2$, and $g(t, 0) = 0$. The equilibrium point $y = (0,0)$ of $\dot{y} = g(t, y)$ is **globally uniformly asymptotically stable** if for any initial condition $y(0)$, $\lim_{t\to\infty} |y(t)| = 0$, where $|\cdot|$ is the standard Euclidean norm.

That is, if the origin is a globally asymptotically stable equilibrium point of a system, then a trajectory starting from an arbitrary point converges to the origin as t tends to infinity.

Theorem 2. *Assume $\theta(0) > 0$, $\mu_1 = \mu_2 = \mu$, and $\beta_1 = \beta_2 = \beta$. Then $(y_1, y_2) = (0,0)$ is a globally uniformly asymptotically stable equilibrium for (11)–(12) under information level $I \in \{F, N\}$.*

Theorem 2 implies that, under any of the information levels, the corresponding fluid model converges to its periodic equilibrium as $t \to \infty$. In the case of stationary arrivals, since the periodic equilibrium reduces to a single equilibrium point $\tilde{x}^I \in \mathbb{R}_+^2$, the theorem implies that starting from any initial condition, the trajectories converge to $\tilde{x}^I$ as $t \to \infty$.

The proof of Theorem 2 relies on an extended Lyapunov method (see Theorem 6 in Appendix C.3). Lyapunov methods are commonly used to prove the asymptotic stability of stationary systems, where one needs to find a positive definite Lyapunov function $V(y)$ for trajectory $y \in \mathbb{R}^n$ with a negative definite derivative $\dot{V}(y)$. For time-varying systems, the stability of the equilibrium point, in general, depends on (the initial) time. Therefore, one needs to find Lyapunov function candidates $V(t, y)$ on $\mathbb{R}^+ \times \mathbb{R}^n$. This requires satisfying more strict conditions for the positive definiteness of $V(t, y)$ to hold and the natural candidates typically used for stationary systems in the literature fail to satisfy the conditions. Therefore, to facilitate the proof, we assume positive abandonment rates (to bound $\dot{V}(t, y)$), equal service rates and class-independent β (to bound or cancel out the cross-product terms in $\dot{V}(t, y)$). Nevertheless, we find numerically that these assumptions are not necessary for the stability of the equilibrium (see Example 1).

Remark 3. To prove Proposition 1 and Theorem 2 we introduce and apply general results from the literature on nonlinear dynamical systems. As such, the methods we use here can be adapted to examine the long-run behaviour of other time-varying queueing models as well.

6. Performance Comparisons

In this section, we study how information impacts the system's equilibrium performance measures. Our overarching objective is to understand which information design minimizes the aggregate costs due to delay and abandonment (or, equivalently, throughput loss).

To this end, we begin by formalizing the notation for the performance metrics in the fluid equilibrium. Recall that $\tilde{x}^I$ denotes the periodic equilibrium number-in-system process under information level I. For information design $I \in \{F, N\}$, let $\bar{x}_k^I$ denote the time-average number-in-system for class k, and $\bar{x}^I := (\bar{x}_1^I, \bar{x}_2^I)$. Let $\bar{Q}_k^I$ denote the time-average queue length and $\bar{A}_k^I$ the time-average system abandonment rate for class k. Under stationary arrivals, the process $\tilde{x}^I$ is constant over time, so that $\bar{x}_k^I = \tilde{x}_k^I$, $\bar{Q}_1^I = (\bar{x}_1^I - s)^+$, $\bar{Q}_2^I = (\bar{x}_1^I - (s - \bar{x}_2^I)^+)^+$ and $\bar{A}_k^I = A_k^I(\bar{x}^I)$. Under non-stationary arrivals, Proposition 1 and the definition of d imply that d is a period of $\tilde{x}_k^I$, for $k = 1, 2$, so that $\bar{x}_k^I = \frac{1}{d}\int_0^d \tilde{x}_k^I(t)dt$ and $\bar{A}_k^I = \frac{1}{d}\int_0^d A_k^I(\tilde{x}^I(t))dt$. The time-average queue lengths are defined as follows:

$$\bar{Q}_1^I := \frac{1}{d}\int_0^d (\tilde{x}_1^I(t) - s)^+ dt \quad \text{and} \quad \bar{Q}_2^I := \frac{1}{d}\int_0^d (\tilde{x}_2^I(t) - (s - \tilde{x}_1^I(t))^+)^+ dt.$$

To develop the results, we express the time-average system abandonment rate, $\bar{A}_k^I$, in two ways. First, we write $\bar{A}_k^I$ as the difference between the time averages of the system arrival rate and the system service rate during an interval of length d. This holds because by Proposition 1, the in- and outflows are balanced during such an interval, so (7) and (8) yield the following equations:

$$\bar{A}_1^I = \frac{1}{d}\int_0^d \lambda_1(t)dt - \frac{\mu_1}{d}\int_0^d (\tilde{x}_1^I(t) \wedge s)dt \text{ for } I \in \{F, N\}, \tag{13}$$

$$\bar{A}_2^I = \frac{1}{d}\int_0^d \lambda_2(t)dt - \frac{\mu_2}{d}\int_0^d \left((s - \tilde{x}_1^I(t))^+ \wedge \tilde{x}_2^I(t)\right) dt \text{ for } I \in \{F, N\}. \tag{14}$$

Second, we write $\bar{A}_k^I$ by averaging the sum of individual customers' abandonment rates, given by (9) and (10), over a time interval of length d. These individual customer abandonment rates depend on the system's information level. This approach yields the following equations:

$$\bar{A}_1^N := \frac{1}{d}\int_0^d \theta\left(\beta_1 \frac{(\tilde{x}_1^N(t) + \tilde{x}_2^N(t) - s)^+}{s}\right)(\tilde{x}_1^N(t) - s)^+ dt, \tag{15}$$

$$\bar{A}_1^F := \frac{1}{d}\int_0^d \int_0^{(\tilde{x}_1^F(t)-s)^+} \theta(u/s)\, du dt, \tag{16}$$

$$\bar{A}_2^N := \frac{1}{d}\int_0^d \theta\left(\beta_2 \frac{(\tilde{x}_1^N(t) + \tilde{x}_2^N(t) - s)^+}{s}\right)(\tilde{x}_2^N(t) - (s - \tilde{x}_1^N(t))^+)^+ dt, \tag{17}$$

$$\bar{A}_2^F := \frac{1}{d}\int_0^d \int_{(\tilde{x}_1^F(t)-s)^+}^{(\tilde{x}_1^F(t)+\tilde{x}_2^F(t)-s)^+} \theta(u/s)\, du dt. \tag{18}$$

We refer to equations (13)-(18) to help discuss some results and the underlying intuition.

Since the performance metrics under the No information regime depend on the parameter $\boldsymbol{\beta}$, in this section we express this dependence explicitly by writing $\tilde{x}_k^N(t, \boldsymbol{\beta})$, $\bar{x}_k^N(\boldsymbol{\beta})$, $\bar{Q}_k^N(\boldsymbol{\beta})$ and $\bar{A}_k^N(\boldsymbol{\beta})$.

Let $\rho_k(t) := \lambda_k(t)/s\mu_k$ be the class k load at time t. Let $\rho_k := \frac{1}{d}\int_0^d \rho_k(t)dt$, $\underline{\rho}_k := \min_{t\geq 0}\rho_k(t)$, and $\bar{\rho}_k := \max_{t\geq 0}\rho_k(t)$ be the average, minimum, and maximum class k load, respectively.

As noted in Section 3, we integrate the performance effects of queueing and abandonment for each class in the following total cost metric. Define TC_k^I, the time-average **total cost** for class k customers under information design I, as the weighted sum of delay and abandonment costs:

$$TC_k^I = c_k \cdot \bar{Q}_k^I + V_k \cdot \bar{A}_k^I, \quad k = 1, 2,\ I \in \{N, F\}, \tag{19}$$

where c_k is the delay cost per unit time per class-k customer waiting in queue, and V_k is the cost per class-k abandonment. This total cost metric is common in the literature on queueing control with abandonment, e.g., Atar et al. (2010), and is suitable for guiding managerial recommendations as it quantifies the trade-off between queueing and abandonment. Under state-independent abandonment rates, minimizing this total cost is equivalent to minimizing holding cost under a modified holding cost rate, because in this case the abandonment rate is proportional to the queue length. However, this proportional relationship does not hold in the case of state-dependent abandonment rates, as in our model, so the total cost depends on both abandonment and queue lengths.

We make the impact of the information design on this total cost metric precise by defining the following total cost difference for class k under No vs. Full Information:

$$TC_k^{N-F}(\boldsymbol{\beta}) := TC_k^N(\boldsymbol{\beta}) - TC_k^F = c_k\left(\bar{Q}_k^N(\boldsymbol{\beta}) - \bar{Q}_k^F + \frac{V_k}{c_k}(\bar{A}_k^N(\boldsymbol{\beta}) - \bar{A}_k^F)\right). \tag{20}$$

In §6.1 we present our main results on how the information design affects this total cost difference for the HP and LP class, identify related trade-offs, and discuss the underlying drivers and intuition in terms of how information effects queue lengths and abandonment rates, then conclude by outlining the analytical roadmap for our comparison results for these individual performance metrics. In §6.2-6.5 we present the latter comparison results for a model sequence with increasing complexity, starting with stationary single-class and ending with non-stationary two-class systems.

Remark 4. Whereas our results compare performance metrics in the fluid model equilibrium that arises in the large-system regime, these results are also robust for a moderate number of servers. Appendix E.1 illustrates the accuracy of the fluid approximations for such systems.

Remark 5. Our results measure congestion in terms of the average queue length. However, our results are robust if we instead consider the average waiting time, as we discuss in Appendix E.2.

6.1. Main Results and Analysis Roadmap

We highlight our main results on how the information design affects the total cost differences (20) and link these results to the underlying queue length and abandonment effects. Let $\bar{Q}_k^{N-F}(\boldsymbol{\beta}) :=$

$\bar{Q}_k^N(\boldsymbol{\beta}) - \bar{Q}_k^F$ and $\bar{A}_k^{N-F}(\boldsymbol{\beta}) := \bar{A}_k^N(\boldsymbol{\beta}) - \bar{A}_k^F$ denote, respectively, the class-k average queue length difference and abandonment rate difference under No vs. Full information, so (20) is equivalent to

$$TC_k^{N-F}(\boldsymbol{\beta}) = c_k \left(\bar{Q}^{N-F}(\boldsymbol{\beta}) + \frac{V_k}{c_k} \bar{A}^{N-F}(\boldsymbol{\beta}) \right). \tag{21}$$

By (21), if both the queue length difference $\bar{Q}_k^{N-F}(\boldsymbol{\beta})$ and the abandonment rate difference $\bar{A}_k^{N-F}(\boldsymbol{\beta})$ have the same signs, the sign of the total cost difference $TC_k^{N-F}(\boldsymbol{\beta})$ aligns with the rankings of these performance measures, independent of the value of $\frac{V_k}{c_k}$. However, if $\bar{Q}_k^{N-F}(\boldsymbol{\beta})$ and $\bar{A}_k^{N-F}(\boldsymbol{\beta})$ have opposite signs, there exists a threshold value of $\frac{V_k}{c_k}$, such that the sign of the total cost difference $TC_k^{N-F}(\boldsymbol{\beta})$ aligns with the sign of the queue length difference $\bar{Q}_k^{N-F}(\boldsymbol{\beta})$ for $\frac{V_k}{c_k}$ below this threshold, and with the sign of the abandonment rate difference $\bar{A}^{N-F}(\boldsymbol{\beta})$ for $\frac{V_k}{c_k}$ above this threshold.

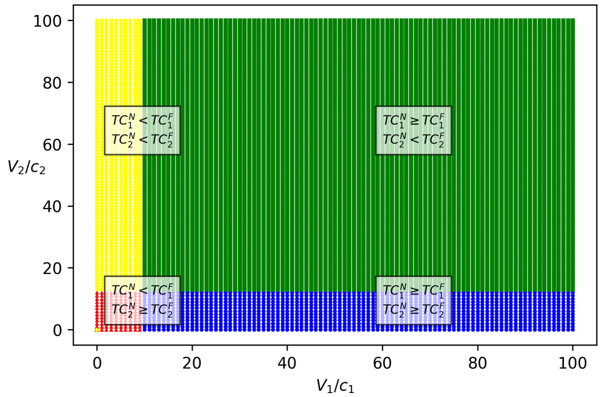
Our key findings focus on the total cost differences (20) for non-stationary two-priority systems with alternating under-/overload. We consider the same sets of model parameters that we used to demonstrate the robustness of using for No information the best-fitting β_k model to approximate the $\beta_k(\boldsymbol{w}(t))$ model (see Appendix A, Table 1). For each parameter set, we evaluate the total cost differences (20) as follows: For No information we evaluate the fluid equilibrium metrics $\bar{Q}_k^N(\boldsymbol{\beta}^*)$ and $\bar{A}_k^N(\boldsymbol{\beta}^*)$ for the β_k model with the constants $\boldsymbol{\beta}^*$ that yield the best fit for the chosen $\beta_k()$ functions (cf. Appendix A, Tables 2 and 3). For full information we evaluate the fluid equilibrium metrics $\bar{Q}_k^F$ and $\bar{A}_k^F$. We then evaluate the sign of $TC_k^{N-F}(\boldsymbol{\beta}^*)$ for $(\frac{V_1}{c_1}, \frac{V_2}{c_2}) \in \mathbb{R}_+^2$.

All our numerical results show a consistent structure. We illustrate this structure with representative examples for class-independent (Figure 3) and class-dependent (Figure 4) $\beta_k()$ functions. Each figure summarizes the total cost rankings (left) and the performance measures (right).

In the class-independent case (Figure 3), the best-fitting β_k model yields a larger average perceived position fraction for HP vs. LP customers, i.e., $\boldsymbol{\beta}^* = (0.71, 0.24)$: HP customers are relatively pessimistic as they fail to account for their faster advancement and lower wait in queue. The mean waitimg times and corresponding fractions are $W_1 = 0.148$, $\beta_1(W_1) = 0.87$ for HP, and $W_2 = 1.239$, $\beta_2(W_2) = 0.37$ for LP. By contrast, in the class-dependent case (Figure 4), the best-fitting β_k model yields a smaller average perceived position fraction for HP vs. LP customers, i.e., $\boldsymbol{\beta}^* = (0.19, 0.36)$: HP customers are relatively optimistic as they do account for their faster advancement and shorter wait. The mean waitimg times and corresponding fractions are $W_1 = 0.185$, $\beta_1(W_1) = 0.25$ for HP, and $W_2 = 1.148$, $\beta_2(W_2) = 0.59$ for LP customers.

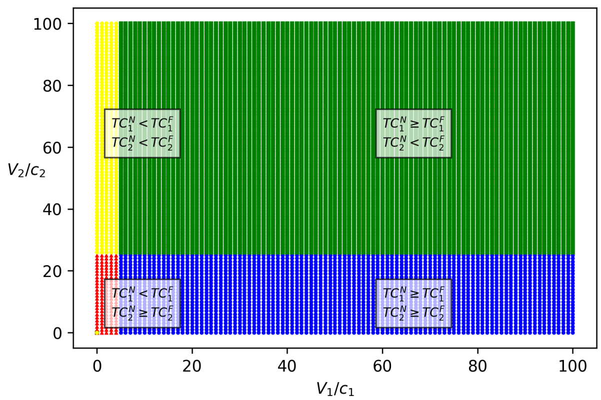
The following discussion focuses on the case of class-independent $\beta_k(x)$ shown in Figure 3, but the same result structure and reasoning apply to the case of class-dependent $\beta_k(x)$.

The left panel of Figure 3 highlights the key finding, namely, the information design may have consistent or opposite effects on the total cost of the HP and LP classes:



k	$\bar{Q}_k^N(\boldsymbol{\beta}^*)$	$\bar{Q}_k^F$	$\bar{Q}_k^{N-F}(\boldsymbol{\beta}^*)$
1	4.64	6.27	-1.63
2	18.00	14.94	3.06
k	$\bar{A}_k^N(\boldsymbol{\beta}^*)$	$\bar{A}_k^F$	$\bar{A}_k^{N-F}(\boldsymbol{\beta}^*)$
1	3.94	3.78	0.17
2	11.06	11.31	-0.25

Figure 3 **Ranking of total cost under No information Best-fitting β^* model and Full information model as functions of V_1/c_1 and V_2/c_2 ($\boldsymbol{\beta}(x) = (e^{-2x}, e^{-2x}), \boldsymbol{\beta}^* = (0.71, 0.24)$, $\mu_1 = \mu_2 = 1$, $s = 100, \theta(x) = 1.5 - e^{-x}, \rho_1 = 0.8,\ \rho_2 = 0.2,\ \ \lambda_k(t) = s\rho_k(1 - 0.5\sin(\pi t/12))$.**



k	$\bar{Q}_k^N(\boldsymbol{\beta}^*)$	$\bar{Q}_k^F$	$\bar{Q}_k^{N-F}(\boldsymbol{\beta}^*)$
1	6.20	6.27	-0.07
2	16.82	14.94	1.87
k	$\bar{A}_k^N(\boldsymbol{\beta}^*)$	$\bar{A}_k^F$	$\bar{A}_k^{N-F}(\boldsymbol{\beta}^*)$
1	3.79	3.78	0.02
2	11.24	11.31	-0.07

Figure 4 **Ranking of total cost under No information Best-fitting β^* model and Full information model as functions of V_1/c_1 and V_2/c_2 ($\boldsymbol{\beta}(x) = (0.3e^{-2x}, e^{-x}), \boldsymbol{\beta}^* = (0.19, 0.36)$, $\mu_1 = \mu_2 = 1$, $s = 100, \theta(x) = 1.5 - e^{-x}, \rho_1 = 0.8,\ \rho_2 = 0.2,\ \ \lambda_k(t) = s\rho_k(1 - 0.5\sin(\pi t/12))$.**

1. For HP customers no (N) information yields a larger total cost than full (F) information if the ratio of abandonment cost to waiting cost $V_1/c_1 > 9.59$, and the opposite holds if $V_1/c_1 < 9.9$.
2. For LP customers, no (N) information yields a smaller total cost than full (F) information if the ratio of abandonment cost to waiting cost $V_2/c_2 > 12.24$, and the opposite holds if $V_2/c_2 < 12.24$.

In sum, the information design has the same effect on both classes for $(\frac{V_1}{c_1}, \frac{V_2}{c_2})$ in the blue and yellow areas, but involves a trade-off between the two classes for $(\frac{V_1}{c_1}, \frac{V_2}{c_2})$ in the red and green areas.

Two types of trade-offs drive these results, (1) a class-level trade-off between the two key performance metrics - queue length and abandonment rate, and (2) a HP vs. LP class trade-off. The table in the right panel of Figure 3 highlights these trade-offs: Compared to full (F) information, no (N) information yields a smaller queue length but a higher abandonment rate for the HP class, with opposite effects on the queue length and abandonment rate of the LP class. That is, the information design improves one metric but harms the other (the class-level trade-off), yet these performance effects go in opposite directions for the two classes (the HP vs. LP trade-off).

The following analytical results complement the numerical results discussed above. Proposition 2 pertains to the case where the HP class alternates between under- and overloaded. [1]

Proposition 2. *For two-priority systems with non-stationary periodic arrivals and at least occasional HP overload ($\bar{\rho}_1 > 1$), if $\beta_1 \geq 0.5$ and $\max_{t\geq 0} \tilde{x}_1^F(t) > s > \min_{t\geq 0} \tilde{x}_1^N(t, \boldsymbol{\beta})$, the equilibrium average total costs under no (N) and full (F) information compare as follows:*

1. *For HP: there is a threshold $\frac{V_1}{c_1}^*(\boldsymbol{\beta})$ such that $TC_1^N(\boldsymbol{\beta}) = TC_1^F$ at $\frac{V_1}{c_1} = \frac{V_1}{c_1}^*(\boldsymbol{\beta})$ and*
   (a) *If $\frac{V_1}{c_1} < \frac{V_1}{c_1}^*(\boldsymbol{\beta})$, then $TC_1^N(\boldsymbol{\beta}) < TC_1^F$.*
   (b) *If $\frac{V_1}{c_1} > \frac{V_1}{c_1}^*(\boldsymbol{\beta})$, then $TC_1^N(\boldsymbol{\beta}) > TC_1^F$.*
2. *For LP: if $\mu_2 \leq \theta(0)$, then for a given β_1, there are thresholds $\beta_2^*(\beta_1), \frac{V_2}{c_2}^*(\beta_1, \beta_2^*)$ such that $TC_2^N((\beta_1, \beta_2^*)) = TC_2^F$ at $\frac{V_2}{c_2} = \frac{V_2}{c_2}^*(\beta_1, \beta_2^*)$, and when $\beta_2 < \beta_2^*(\beta_1)$,*
   (a) *If $\frac{V_2}{c_2} < \frac{V_2}{c_2}^*(\beta_1, \beta_2^*)$, then $TC_2^N(\boldsymbol{\beta}) > TC_2^F$.*
   (b) *If $\frac{V_2}{c_2} > \frac{V_2}{c_2}^*(\beta_1, \beta_2^*)$, then $TC_2^N(\boldsymbol{\beta}) < TC_2^F$.*

Proposition 2 is consistent with the results shown in Figures 3 and 4, even though some of the parameters for the examples in these figures are beyond the assumptions stated in Proposition 2: e.g., in Figures 3 and 4 we have $\beta_1 < 0.5$ and $\mu_2 = 1 > 0.5 = \theta(0)$, whereas Proposition 2 requires $\beta_1 \geq 0.5$ for the HP ranking and $\mu_2 \leq \theta(0)$ for the LP ranking. These examples therefore suggest that Proposition 2 is robust and applicable to broader range of parameters regimes.

In sum, the cost difference $TC_k^{N-F}(\boldsymbol{\beta})$ depends on the $\frac{V_k}{c_k}$ ratio, the queue length difference $\bar{Q}_k^{N-F}(\boldsymbol{\beta})$ and the abandonment rate difference $\bar{A}_k^{N-F}(\boldsymbol{\beta})$. In general, these differences in performance metrics may have opposite signs, both across metrics for the same class, and across classes for the same metric (see Figures 3 and 4). As a result, the information design may involve cost trade-offs between the HP and LP classes, depending on the values of the ratios V_1/c_1 and V_2/c_2.

In general, the impact of the information design on the total cost of each class depends on the strengths of the identified trade-offs, specifically, (i) on the β functions that describe customers' perceived queue positions under No information, (ii) on the corresponding best-fitting β constants, and (iii) on the class- and metric-specific dependent β thresholds at which both information design yield the same performance. Note that (i) is ultimately an empirical question, whereas (ii) and (iii) are highly sensitive to the operational characteristics, i.e., the number and processing rate of servers, the arrival patterns and loads of the two classes, and the abandonment rate function $\theta(\cdot)$.

However, whereas the results for different problem instances will differ in some respects from those discussed above, they will share the key structural features highlighted above: (1) The effect

[1] Proposition 10 in Appendix D.5.2 covers the case where the HP class is uniformly underloaded.

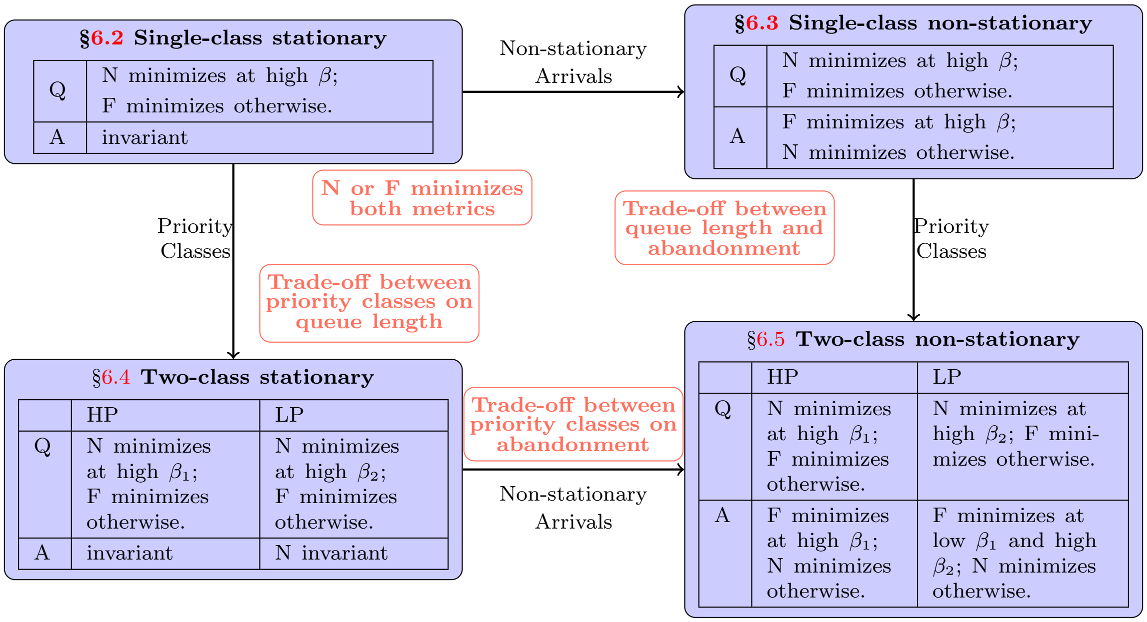


Figure 5 **Analytical roadmap for Sections 6.2-6.5 and summary of performance comparisons under no (N) versus full (F) information (Q = average number-in-system, A = average abandonment rate).**

of the information design on any class-level performance metric depends on the queue position perception parameter β, (2) Information design may involve a class-level trade-off between the two key performance metrics - queue length and abandonment rate, and (3) a trade-off between the HP and LP class. In Sections 6.2-6.5 we develop the building blocks for the results discussed above for a "full range" of β and load. Specifically, to fully understand the impact of information under various load patterns and levels, we study its impact on the individual system performance measures: numbers-in-system and abandonment rates.

Deriving and explaining these comparison results for a two-class system with non-stationary arrival rates poses significant challenges, particularly in the most practically relevant case when the HP load alternates between over- and under-loaded. Therefore, to highlight the individual and collective impact of prioritization and time-varying arrivals on the effects of information provision, we develop and present these results in the following sequence: In Section 6.2, we start with the single-class model with stationary arrivals. In Section 6.3, we study the impact of non-stationary arrivals for a single customer class. In Section 6.4, we study the impact of priority service by considering two priority classes with stationary arrival rates. In Section 6.5, we consider both time-varying arrivals and two priority classes. Figure 5 shows this analytical roadmap for Sections 6.2-6.5 and the results and trade-offs we identify along the way.

In a stationary, overloaded, single-class system, the queue-length ranking between N and F is driven by β. Specifically, small values of β reflect customer optimism, leading to longer queues under N, while large β values reflect customer pessimism, leading to shorter queues. In the non-stationary case (where HP is occasionally overloaded), N and F exhibit a fundamental queue–abandonment tradeoff governed by two thresholds. In a two-class priority system, the comparison between N and F is further complicated by cross-class trade-offs. In a stationary environment, the relative performance of each class is determined by its specific β_k parameter, where N may lead to smaller queues in one class and longer queues in the other class. When non-stationarity is introduced, these cross-class effects result in a trade-off between HP and LP abandonment rates. These dual trade-offs (between metrics of the same class, and across classes) demonstrate that the cost-minimizing information policy is driven not only by β and the system load, but also by the relative weights assigned to each class within the objective function.

Remark 6. Our results focus on systems with a fixed number of servers. Our results however generalize to cases where the number of servers are non-stationary and periodic; see Appendix E.3.

6.2. Single Class with Stationary Arrivals

In this section, we consider a single customer class with stationary arrivals. We omit the class subscript for simplicity. Under stationary arrivals, (13), (15) and (16) reduce, respectively, to:

$$\bar{A}^I = \lambda - \mu(\bar{x}^I \wedge s), \text{ for } I \in \{F, N\}, \tag{22}$$

$$\bar{A}^N := \theta\left(\beta\frac{(\bar{x}^N - s)^+}{s}\right)(\bar{x}^N - s)^+, \tag{23}$$

$$\bar{A}^F := \int_0^{(\bar{x}^F - s)^+} \theta\,(u/s)\,du. \tag{24}$$

These equations imply the following rankings of performance measures (we omit a formal proof).

Proposition 3. *For single-class systems with stationary arrivals, the equilibrium average number-in-system and abandonment rate under no (N) and full (F) information compare as follows:*

1. *If $\rho \le 1$, then $\bar{x}^N(\beta) = \bar{x}^F = s\rho$ and $\bar{A}^N(\beta) = \bar{A}^F = 0$.*
2. *If $\rho > 1$, then $s < \bar{x}^N(\beta), s < \bar{x}^F$ and $\bar{A}^N(\beta) = \bar{A}^F = \lambda - s\mu$. There is a threshold $\beta^* \in (0,1)$ such that $\bar{x}^N(\beta^*) = \bar{x}^F$ and:*
   (a) *If $\beta \in [0, \beta^*)$, then $\bar{x}^F < \bar{x}^N(\beta)$.*
   (b) *If $\beta \in (\beta^*, 1]$, then $\bar{x}^N(\beta) < \bar{x}^F$.*

Proposition 3.2 highlights how the effect of information in overloaded systems ($\rho > 1$) depends on the position fraction β. (Information has no effect in underloaded systems, as there is no queue.) There exists a (load-dependent) threshold β^* such that no information (N) yields a longer queue than full information (F) if $\beta < \beta^*$, and vice versa if $\beta > \beta^*$. Intuitively, this holds as a result of the following three factors: (i) The equilibrium abandonment rate is invariant to the information under stationary arrivals, see (22); (ii) the system abandonment rate under both information regimes increases in the queue length; and (iii) for any *fixed* queue length, the parameter β in the no-information regime has the following effects on abandonment: the individual and system abandonment rates are continuously increasing in β, lower compared to full information (F) for $\beta = 0$, and higher compared to F for $\beta = 1$; see (23) and (24). Therefore, for β below (above) the threshold β^*, no information yields a longer (shorter) queue of customers who abandon at a slower (faster) average rate, compared to customers under full information.

6.3. Single Class with Non-Stationary Periodic Arrivals

We turn to single-class systems with non-stationary arrivals. Compared to the stationary case, time-varying arrivals give rise to one more load regime, i.e., alternating between over- and underloaded ($\underline{\rho} < 1 < \bar{\rho}$). This regime is prevalent in practice and yields a key difference compared to stationary settings: a trade-off between queueing (Proposition 4) and abandonment (Proposition 5).

Proposition 4. *For single-class systems with non-stationary periodic arrivals, the equilibrium average number-in-system under no (N) and full (F) information compare as follows:*

1. *If* $\bar{\rho} \leq 1$, *then* $\bar{x}^N = \bar{x}^F$.
2. *If* $\bar{\rho} > 1$ *and* $\max\limits_{t\geq 0} \tilde{x}^N(t,0) > s$, *there is a threshold* $\beta_q^* \in (0, 0.5)$ *such that* $\bar{x}^N(\beta_q^*) = \bar{x}^F$ *and:*
   (a) *If* $\beta \in [0, \beta_q^*)$, *then* $\bar{x}^F < \bar{x}^N(\beta)$.
   (b) *If* $\beta \in (\beta_q^*, 1]$, *then* $\bar{x}^N(\beta) < \bar{x}^F$.

Proposition 4 builds on stronger results that rank the equilibrium number-in-system processes $\tilde{x}^N(t,\beta)$ and $\tilde{x}^F(t)$ for *all* t (see Lemma 4 in Appendix D.3). Proposition 4 shows that the corresponding ranking results for stationary systems (Proposition 3) are robust and naturally generalize to the non-stationary case, specifically in the important case where the system is overloaded at least some of the time ($\bar{\rho} > 1$, Part 2): Compared to full information, no information increases the *average* equilibrium number-in-system for small β, and weakly reduces this metric for large β.

Proposition 4.2 implies a further insight: The information design matters even if the system is *underloaded on average* (i.e., $\rho < 1$), so long as a queue forms *some of the time* under no information with $\beta = 0$ (i.e., if $\max_{t\geq 0} \tilde{x}^N(t,0) > s$).

Proposition 5. *For single-class systems with non-stationary periodic arrivals, the equilibrium average abandonment rates under no (N) and full (F) information compare as follows:*

1. *If* $\bar{\rho} \leq 1$, *or* $\underline{\rho} \geq 1$, *or more generally* $\min\limits_{t\geq 0} \tilde{x}^N(t,1) \geq s$, *then* $\bar{A}^N(\beta) = \bar{A}^F$.
2. *If* $\underline{\rho} < 1 < \bar{\rho}$ *and* $\max\limits_{t\geq 0} \tilde{x}^N(t,0) > s > \min\limits_{t\geq 0} \tilde{x}^N(t,1)$, $\exists\, \beta_a^* \in (0, 0.5)$ *such that* $\bar{A}^N(\beta_a^*) = \bar{A}^F$ *and:*
   (a) *If* $\beta \in [0, \beta_a^*)$, *then* $\bar{A}^N(\beta) < \bar{A}^F$.
   (b) *If* $\beta \in (\beta_a^*, 1]$, *then* $\bar{A}^N(\beta) > \bar{A}^F$.

Comparing Propositions 4.2 and 5.2 with Proposition 3 shows that non-stationary arrivals yield one key difference compared to stationary settings: A trade-off between queueing and abandonment in the practically most relevant regime where the system alternates between under- and overloaded ($\underline{\rho} < 1 < \bar{\rho}$):[2] For $\beta < \min(\beta_a^*, \beta_q^*)$ the no-information downside of higher congestion (a longer queue length) is offset by throughput gain, i.e., a lower abandonment rate; for $\beta > \max(\beta_a^*, \beta_q^*)$ this trade-off is reversed, as full information yields more congestion and higher throughput.

Intuitively, this trade-off follows because the information regime with the larger number-in-system also experiences a higher average server utilization, implying that more customers are served and fewer abandon, compared to the information regime with the smaller number-in-system.

In sum, information design in non-stationary systems that alternate between under- and overloaded is subject to a trade-off between two conflicting key objectives, minimizing congestion (queueing) and maximizing throughput (minimizing abandonment). The key implication is that the information design must carefully consider and balance the impact on *both* performance measures, queue length and abandonment, as well as the resulting costs and benefits. For example, with relatively more pessimistic customers (i.e., larger β), no information is preferable to full information if the resulting throughput loss and/or the cost of abandonment per customer are relatively small.

6.4. Two-Class Priority System with Stationary Arrivals

We turn to stationary systems with two priority classes. The interplay between two priority classes yields a key difference, compared to the single-class setting: The information design may have *opposite* effects on the queue lengths of the two classes. Building on (13)-(18), Proposition 6 makes these effects precise. Information does not affect the abandonment rates in stationary systems.

PROPOSITION 6. *For two-priority systems with stationary arrivals, the equilibrium average numbers-in-system and abandonment rates under no (N) and full (F) information compare as follows:*

1. *If $\rho_1 \le 1$, then $\bar{x}_1^N(\boldsymbol{\beta}) = \bar{x}_1^F = s\rho_1$ and $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = 0$. For LP customers:*
   - (a) *If $\rho_1 + \rho_2 \le 1$, then $\bar{x}_2^N(\boldsymbol{\beta}) = \bar{x}_2^F = s\rho_2$ and $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = 0$.*
   - (b) *If $\rho_1 + \rho_2 > 1$, then $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = \lambda_2 - s(1-\rho_1)\mu_2$, and $\exists\, \beta_2^* \in (0,1)$ such that $\bar{x}_2^N(\boldsymbol{\beta}) = \bar{x}_2^F$ and: (i) If $\beta_2 \in [0, \beta_2^*)$ then $\bar{x}_2^F < \bar{x}_2^N(\boldsymbol{\beta})$. (ii) If $\beta_2 \in (\beta_2^*, 1]$, then $\bar{x}_2^N(\boldsymbol{\beta}) < \bar{x}_2^F$.*
2. *If $\rho_1 > 1$, then $\bar{x}_1^N(\boldsymbol{\beta}) > s$, $\bar{x}_1^F > s$, $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = \lambda_1 - s\mu_1$, and $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = \lambda_2$.*
   There exist two increasing threshold functions $\beta_1^*(\beta_2) \in (0, 0.5)$ and $\beta_2^*(\beta_1) \in (0,1]$, which are decreasing in the LP load ρ_2, and partition the parameter space of $\boldsymbol{\beta}$ into four regions:
   - (a) *If $\beta_1 \in [0, \beta_1^*(\beta_2))$ and $\beta_2 \in [0, \beta_2^*(\beta_1))$, then $\bar{x}_1^N(\boldsymbol{\beta}) > \bar{x}_1^F$ and $\bar{x}_2^N(\boldsymbol{\beta}) > \bar{x}_2^F$.*
   - (b) *If $\beta_1 \in [0, \beta_1^*(\beta_2))$ and $\beta_2 \in (\beta_2^*(\beta_1), 1]$, then $\bar{x}_1^N(\boldsymbol{\beta}) > \bar{x}_1^F$ and $\bar{x}_2^N(\boldsymbol{\beta}) < \bar{x}_2^F$.*
   - (c) *If $\beta_1 \in (\beta_1^*(\beta_2), 1]$ and $\beta_2 \in [0, \beta_2^*(\beta_1))$, then $\bar{x}_1^N(\boldsymbol{\beta}) < \bar{x}_1^F$ and $\bar{x}_2^N(\boldsymbol{\beta}) > \bar{x}_2^F$.*
   - (d) *If $\beta_1 \in (\beta_1^*(\beta_2), 1]$ and $\beta_2 \in (\beta_2^*(\beta_1), 1]$, then $\bar{x}_1^N(\boldsymbol{\beta}) < \bar{x}_1^F$ and $\bar{x}_2^N(\boldsymbol{\beta}) < \bar{x}_2^F$.*

[2] The thresholds β_q^* and β_a^* are close but do not necessarily coincide; either threshold may be larger.

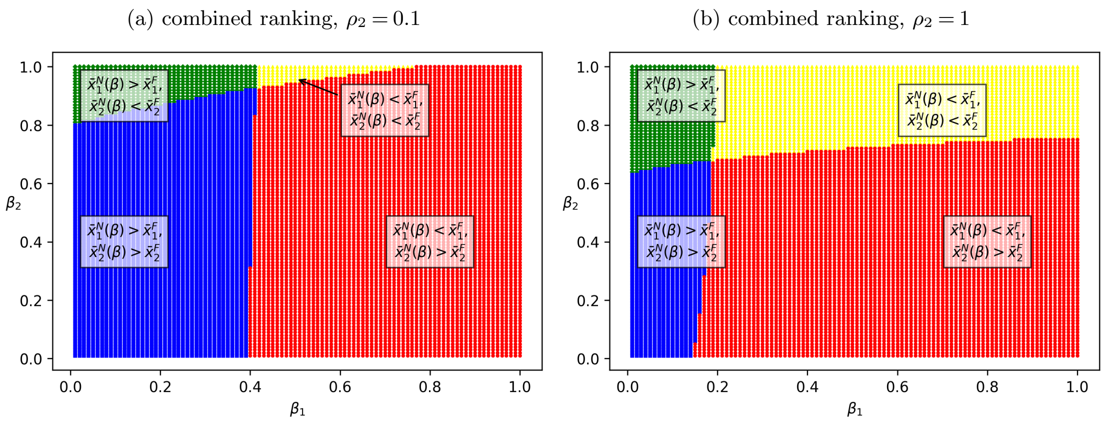


Figure 6 Stationary two-priority system: Ranking of equilibrium numbers-in-system under information regimes N and F, as function of $\boldsymbol{\beta}$ ($\mu_1 = \mu_2 = 1,\ s = 100, \theta(x) = 2 - e^{-x}, \rho_1 = 1.5$).

Underloaded HP class (Proposition 6.1): HP customers do not queue and utilize $s\rho_1$ servers. Therefore, LP customers experience a single-class stationary system with $s(1-\rho_1)$ servers, and Proposition 6.1 is consistent with Proposition 3: Compared to full information (F), no information (N) yields a larger (smaller) LP number for β_2 below (above) some threshold.

Overloaded HP class (Proposition 6.2): In this case, more information may have the *opposite* effect on the HP and LP queue lengths. Before elaborating on this result, we note that, whereas no LP customers are getting served in this regime with HP overload, the results of Proposition 6.2 are relevant as they also hold in two cases where some LP customers *are* getting served: (1) in the fluid limit for the practically important case with non-stationary arrivals and alternating HP under/overload (§6.5, Proposition 7), and (2) in stochastic systems with heavy HP load but $\rho_1 < 1$.

Proposition 6.2 delivers two main insights on the queue length effects of information:

(1) For each priority class in isolation, the effects of information and the underlying driver, are fully consistent with the single-class results of Proposition 3: For class k No information (N) yields a longer (shorter) queue than full information (F), if β_k is below (above) a threshold.

(2) Considering both priority classes *jointly*, the information design may have *opposite* effects on their queue lengths: Compared to full information, no information increases the queue of the relatively optimistic class (β below the class threshold) and decreases the queue of the relatively pessimistic class (β above the class threshold); see Parts 2(b) and 2(c).

Example 2. To illustrate Proposition 6.2, Figure 6 shows how the ranking of the equilibrium numbers-in-system under no and full information depends on $\boldsymbol{\beta}$ at low ($\rho_2 = 0.1$) and high ($\rho_2 = 1$) LP loads. Each point in Figure 6 corresponds to a (β_1, β_2) combination. Figure 6 shows:

1. *Information effect on HP queue length mainly depends on β_1:* Regardless of the LP load, the HP queue is shorter under full information for low β_1, and under no information for high β_1.

2. *Information effect on LP queue length also depends on LP load ρ_2:* Figure 6(a) for low LP load ($\rho_2 = 0.1$) shows the threshold $\beta_2^*(\beta_1) = 1$ for $\beta_1 > 0.75$; this means no information yields a *longer* LP queue ($\bar{x}_2^N(\boldsymbol{\beta}) > \bar{x}_2^F$), *even if* LP customers are maximally pessimistic, i.e., $\beta_2 = 1$. In contrast, Figure 6 (b) shows for high LP load ($\rho_2 = 1$) that $\beta_2^*(\beta_1) < 1$ for all β_1, so no information yields a *shorter* LP queue ($\bar{x}_2^N(\boldsymbol{\beta}) < \bar{x}_2^F$) if LP customers are sufficiently pessimistic, i.e., $\beta_2 \in (\beta_2^*(\beta_1), 1]$.

3. *Consistent vs. opposite information effects on HP and LP queue lengths:* The blue and yellow areas correspond to cases where information design has consistent effects on the queue lengths of both classes: The blue area corresponds to Part 2(a), i.e., full information minimizes the queue lengths of *both* classes. The yellow area corresponds to Part 2(d), i.e., no information minimizes both queue lengths. In contrast, the green area (Part 2(b)) and the red area (Part 2(c)) correspond to cases with an information design trade-off between high- and low-priority queue lengths, i.e., one design minimizes the HP queue length but the other minimizes the LP queue length.

6.5. Two-Class Priority System with Non-Stationary Periodic Arrivals

The case with two priority classes and non-stationary arrivals has three possible HP load regimes: (i) Uniformly underloaded HP class ($\bar{\rho}_1 \leq 1$). In this load regime, information only affects the LP class, and the results are consistent with those for single-class non-stationary systems (Propositions 4 and 5). We provide analytical results for this regime in Appendix D.5.2.
(ii) HP class alternating between under- and overloaded ($\underline{\rho}_1 < 1 < \bar{\rho}_1$). This practically prevalent load regime shows that the results for simpler systems in §6.2-6.4 are robust, and also gives rise to an important additional trade-off, between HP and LP abandonment, that does not arise in simpler systems. This is technically the most challenging regime. We present a combination of analytical results (Propositions 7 and 8) and numerical results (Example 3 and Figure 7).
(iii) Uniformly overloaded HP class ($\bar{\rho}_1 \geq 1$). In this load regime, information only affects the number-in-system, and the respective ranking results are consistent with those for two-class stationary systems (Proposition 6.2). For the sake of brevity, we omit these results and related discussion.

HP Class Alternating Between Under- And Overloaded ($\underline{\rho}_1 < 1 < \bar{\rho}_1$): This practically prevalent case is also technically the most challenging. We present a combination of analytical (Propositions 7 and 8, building blocks for Proposition 2) and numerical results (Example 3 and Figure 7). For one, we show that our results for simpler systems are robust, specifically, the two performance trade-offs: (i) between number-in-system and abandonment of the same class under non-stationary arrivals (Propositions 4 and 5 for single-class systems, Proposition 9 for LP class in two-class systems with uniform HP underload), and (ii) between the HP and LP numbers-in-system (Proposition 6.2 for stationary arrivals). Furthermore, we identify an *important additional trade-off, between HP and LP abandonment*, which is unique to this load regime.

Proposition 7. *For two-priority systems with non-stationary periodic arrivals and occasional HP overload ($\bar{\rho}_1 > 1$), information has the following effects on the equilibrium numbers-in-system:*

1. *If $\beta_1 = 0$, then for HP: $\tilde{x}_1^N(t,\boldsymbol{\beta}) \geq \tilde{x}_1^F(t)\ \forall t$, and $\bar{x}_1^N(\boldsymbol{\beta}) > \bar{x}_1^F$ if $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$. For LP:*
   (a) *$\beta_2 = 0$: Then $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,\boldsymbol{\beta})\ \forall t$, if $\min_{t\geq 0} \tilde{x}_1^F(t) \geq s$.*
   (b) *$\beta_2 = 1$: Then $\tilde{x}_2^F(t) > \tilde{x}_2^N(t,\boldsymbol{\beta})\ \forall t$, if $\max_{t\geq 0} \tilde{x}_1^F(t) > s$ and $\mu_2 \leq \theta(0)$.*
2. *If $\beta_1 \geq 0.5$ then for HP: $\tilde{x}_1^N(t,\boldsymbol{\beta}) \leq \tilde{x}_1^F(t)\ \forall t$, and $\bar{x}_1^N(\boldsymbol{\beta}) < \bar{x}_1^F$ if $\max_{t\geq 0} \tilde{x}_1^F(t) > s$. For LP:*
   (a) *$\beta_2 = 0$: Then $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,\boldsymbol{\beta})\ \forall t$, if $\max_{t\geq 0} \tilde{x}_1^F(t) > s$ and $\mu_2 \leq \theta(0)$.*
   (b) *$\beta_2 = 1$: Then there are LP load thresholds $\tilde{\rho}_2^1 < \tilde{\rho}_2^2$ such that*
      i. *$\tilde{x}_2^F(t) < \tilde{x}_2^N(t,\boldsymbol{\beta})\ \forall t$, if $\bar{\rho}_2 < \tilde{\rho}_2^1$ and $\min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$.*
      ii. *$\tilde{x}_2^F(t) > \tilde{x}_2^N(t,\boldsymbol{\beta})\ \forall t$, if $\underline{\rho}_2 > \tilde{\rho}_2^2$ and $\min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$, or $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$ and $\mu_2 \geq \theta(\infty)$.*

Remark 7. Proposition 7 specifies the *uniform ranking* of the equilibrium number-in-system processes at *all* times t, for four combinations of low and high β_1 and β_2, i.e., $\beta_1 \in \{0, [0.5, 1]\}$ and $\beta_2 \in \{0, 1\}$. These results imply the same rankings for the time *averages* of these processes, and also highlight how the ranking of these time averages depends more generally on the $\boldsymbol{\beta}$ parameters: The ranking of class-k average queue lengths under F vs. N reverses as β_k increases from low to high. Our numerical results illustrate this structure for the entire $\boldsymbol{\beta}$ range (Example 3, Figure 7).

The key insights from Proposition 7 on the queue length effects of information are consistent with those of Proposition 6.2 for the stationary two-class case with HP overload:

(1) For each priority class, compared to full information (F), no information (N) generally yields a longer queue if the class position parameter β is low, and a longer queue if β is high.

(2) The information design may have *opposite* effects on the queue lengths of the two classes. Specifically, compared to full information (F), no information (N) generally increases the queue of the optimistic class ($\beta = 0$) and decreases the queue of the pessimistic class ($\beta = 1$). Furthermore, information may also have opposite queue length effects if both classes are relatively pessimistic and the LP load is below a threshold: Specifically, by Part 2(b)i. of Proposition 7, for $\beta_1 \geq 0.5$ and $\beta_2 = 1$, full information increases the HP queue length, but reduces the LP queue length if the LP load is below a threshold ($\bar{\rho}_2 < \tilde{\rho}_2^1$). This is consistent with the stationary case; see Figure 6(a).

Proposition 8. *For two-priority systems with non-stationary periodic arrivals and HP class that alternates between under- and overloaded ($\underline{\rho}_1 < 1 < \bar{\rho}_1$), information has the following effects on the average abandonment rates:*

1. *If $\beta_1 = 0$, then for HP: $\bar{A}_1^N(\boldsymbol{\beta}) \leq \bar{A}_1^F$, with strict inequality iff $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s > \min_{t\geq 0} \tilde{x}_1^F(t)$. For LP customers:*
   (a) *$\beta_2 = 0$: Then $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F$ if $\min_{t\geq 0} \tilde{x}_1^F(t) \geq s$.*
   (b) *$\beta_2 = 1$: Then $\bar{A}_2^N(\boldsymbol{\beta}) > \bar{A}_2^F$, if $\max_{t\geq 0} \tilde{x}_1^F(t) > s > \min_{t\geq 0} \tilde{x}_1^F(t)$ and $\mu_2 \leq \theta(0)$.*

2. *If* $\beta_1 \geq 0.5$ *then for HP:* $\bar{A}_1^N(\boldsymbol{\beta}) \geq \bar{A}_1^F$, *with strict inequality iff* $\max_{t\geq 0} \tilde{x}_1^F(t) > s > \min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta})$. *For LP customers:*
   (a) $\beta_2 = 0$*: Then* $\bar{A}_2^N(\boldsymbol{\beta}) < \bar{A}_2^F$, *if* $\max_{t\geq 0} \tilde{x}_1^F(t) > s > \min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta})$ *and* $\mu_2 \leq \theta(0)$.
   (b) $\beta_2 = 1$*: Then* $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F$ *if* $\min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$.

Propositions 7 and 8 show that information design involves two abandonment-related trade-offs in non-stationary two-priority systems with alternating HP over-/underload:

(1) Between abandonment and number-in-system for each class. This trade-off is consistent with settings where a single class experiences queueing (Propositions 4.2, 5.2, and 9 in Appendix D.5.2).

(2) Between the HP and LP abandonment rates. This trade-off arises only in non-stationary systems with two classes that experience queueing. Specifically, if $\beta_1 = 0$ and $\beta_2 = 1$, no information (N) minimizes the HP abandonment rate whereas full information (F) minimizes the LP abandonment rate (see Part 1(b)), and vice versa if $\beta_1 \geq 0.5$ and $\beta_2 = 0$ (see Part 2(a)).

Numerical Study. We conclude the analysis of non-stationary two-class priority systems with a numerical study that shows how the theoretical ranking results of Propositions 7 and 8 for specific low/high (β_1, β_2) pairs extend to the entire $\boldsymbol{\beta}$ range. We focus on the time-averages of the equilibrium numbers-in-system and abandonment rates, because (i) these averages are of first-order importance, and (ii) the time-varying processes need not obey a uniform ranking at all times.

Example 3. We consider the same supply parameters ($\mu_1 = \mu_2 = 1$, $s = 100$) and abandonment rate function ($\theta(x) = 2 - e^{-x}$) as in Example 2, but the following more moderate and time-varying arrival rates: $\lambda_k(t) = s\rho_k(1 - 0.5\sin(\pi t/12))$, for $k = 1, 2$, where $\rho_1 = 0.9$ and $\rho_2 = 0.5$ are the average HP and LP loads, respectively. Note that $\theta(0) = \mu_2 < \theta(\infty)$, $\underline{\rho}_1 = 0.45 < 1$, and $\bar{\rho}_1 = 1.35 > 1$. These parameters satisfy the conditions in Parts 1(b) and 2(a) of Propositions 7 and 8. Figure 7 shows the rankings of the equilibrium average numbers-in-system (plot (a)) and average abandonment rates (plot (b)) under no (N) vs. full (F) information, as functions of $(\beta_1, \beta_2) \in [0,1] \times [0,1]$.

Figure 7 shows the following information design effects, consistent with Propositions 7 and 8:

1. *Trade-off between HP number-in-system and abandonment rate:* If HP customers are sufficiently optimistic (e.g. $\beta_1 < 0.1$), then full information (F) reduces the number-in-system but increases the abandonment rate, compared to no information (N); if HP customers are sufficiently pessimistic (e.g., $\beta_1 > 0.2$) then full information has the opposite effects.

2. *Trade-off or alignment between LP number-in-system and abandonment rate:* For LP customers, information design involves a similar trade-off as for HP, with the following qualification: For $\beta_1 \leq 0.4$, this trade-off arises for sufficiently low or high β_2. However, for $\beta_1 > 0.4$, this trade-off arises only if LP customers are sufficiently optimistic ($\beta_2 \lessapprox 0.6$), whereas no information minimizes *both*, number-in-system and abandonment, if LP are more pessimistic ($\beta_2 \gtrapprox 0.6$).

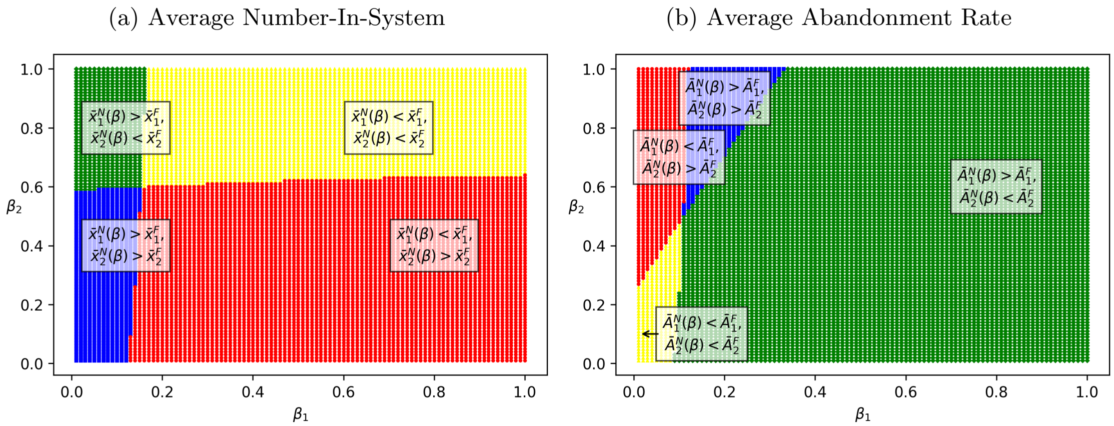


Figure 7 **Non-stationary two-priority system: Ranking of equilibrium performance metrics under information regimes N and F, as function of β**

$\mu_1 = \mu_2 = 1,\ s = 100,\ \rho_1 = 0.9,\ \rho_2 = 0.5,\ \ \lambda_k(t) = s\rho_k(1 - 0.5\sin(\pi t/12)),\ \theta(x) = 2 - e^{-x}.$

3. *Consistent vs. opposite information effects on HP and LP numbers-in-system:* Figure 7(a) shows that, consistent with the stationary case (Proposition 6.2, Figure 6(b)) information may have consistent (blue and yellow areas) or opposite effects (red and green areas) on the two classes.

4. *Consistent vs. opposite information effects on HP and LP abandonment rates:* Figure 7(b) similarly shows that the information design may have consistent effects (blue and yellow areas) or opposite effects (red and green areas) on the abandonment rates of the two classes, depending on customers' queue position perceptions in the absence of full information.

7. Conclusions

7.1. Summary of Main Results

This paper contributes theoretical models and insights on the effects of providing information in observable service systems with abandonment. We consider a Markovian queueing system with two priority classes (HP and LP), time-varying arrival rates and abandonment, a setting that is practically relevant but has hardly been studied in the information design literature. We propose a fairly flexible model of how information impacts customer abandonment, which captures key empirical findings in the literature pertaining to customer abandonment from observable queues.

Our results characterize the effects of information on key performance metrics of abandonment and number-in-system (or waiting time. In particular, our results provide insights on how these effects depend on the interplay between: (i) Customers' perceived queue position under no information; (ii) class-specific system load; (iii) temporal variability of arrival rates, and (iv) priority service. In the presence of time-varying arrivals and with two priority classes, we observe the following key effects and trade-offs:

- *Conflicting effects of information on different classes:* Using a total cost metric that aggregates delay and abandonment, we show that information provision can either uniformly improve or differentially impact system performance across priority classes. Alignment arises when the relative costs of abandonment and delay differ sufficiently across classes, so that information improves total cost for each class despite heterogeneous operational effects. Outside these regimes, information induces cross-class trade-offs, highlighting the need for class-dependent information design.
- *Number-in-system (and waiting time): HP-LP trade-off.* Providing information has *opposite* effects on the queue lengths of the HP and LP classes, if their customers have sufficiently different queue position perceptions. In such cases, compared to full information, no information increases congestion of the relatively optimistic class (β below a threshold) and decreases the queue of the relatively pessimistic class (β above a threshold). For example, for sufficiently low β_1 and high β_2, no information increases the HP queue and reduces the LP queue, compared to full information. (See Proposition 6 for overloaded stationary systems, and Proposition 7 and Figure 7(a) for systems with alternating HP under-/overload.)
- *Abandonment: HP-LP trade-off in systems with non-stationary HP under/overload.* Two-priority non-stationary systems give rise to an additional trade-off, between HP and LP abandonment, if the HP class alternates between under- and overloaded. In such cases, compared to full information, no information reduces the abandonment of the relatively optimistic class (β below a threshold) and increases the abandonment of the relatively pessimistic class (β above a threshold). For example, for sufficiently low β_1 and high β_2, no information reduces the HP abandonment and increases the LP abandonment, compared to full information. (See Proposition 8 and Figure 7(b).)

7.2. Managerial Implications

Our results imply that effective information provision requires (i) identifying for a particular system, given the load conditions and customer perceptions (β parameters), whether the above trade-offs indeed exist; and (ii) In cases where such trade-offs do exist, information design must carefully balance the queueing and abandonment costs of both classes, and/or consider hybrid designs (e.g., giving full information to one class but no information to the other class). Operational measures, such as the load, should be readily measurable, whereas customer perceptions, may be estimated using observational data or customer surveys. For example, the National Health Service (NHS) in the United Kingdom routinely conducts patient surveys to gather feedback on various aspects of their healthcare experience, including waiting times (see https://nhssurveys.org/).

As a concrete example, consider the following parameter regime relevant to an ED: HP class alternates between under- and overloaded regimes; HP customers are sufficiently optimistic (i.e., they perceive a low queue position) and LP customers are sufficiently pessimistic (i.e., they perceive a high queue position). In this case, providing accurate information minimizes LP abandonment but has the appositive impact on HP abandonment. Interestingly, the field study of Westphal et al. (2022) found that providing both operational and time information improved the sense of making progress in the ED for patients, but information only reduced abandonment when only operational information (and not waiting time information) was provided. Our results suggest that the impact of information on abandonment depend on both system load and customer perceptions. As such, these factors should be accounted for in design and analysis of future field studies.

Our results also provide insights on how other operational decisions can impact the effects of information. More specifically, the system manager can (partially) control the system load through staffing (or capacity allocation) decisions. Our results indicate that in scenarios with minimal time-variation in arrival patterns, managing the system load becomes the primary factor determining whether information provision impacts a specific class. Conversely, if arrival patterns do vary over time, the impact of system load management becomes more nuanced, as the effects of information are intertwined with the aforementioned trade-offs. In both scenarios, the system manager can eliminate the information-related performance trade-offs between the two priority classes by using staffing levels such that the HP class remains underloaded. Specifically, the potentially significant trade-off between the HP and LP abandonment rates in non-stationary settings under alternating HP under-/overload (Proposition 8) vanishes under uniform HP underload (Proposition 9).

7.3. Future Research

Our study motivates several future directions. In the following, we briefly discuss a few.

Our results motivate the design of more sophisticated state-dependent and/or hybrid information provision systems. One way to address the queue length and abandonment trade-offs is to explore time/state-dependent information designs. For systems with a priority service, a specific information design may have opposite effects on customers of different classes. Therefore, it would be interesting to study schemes that provide different information levels for different priority classes.

Our characterizations of the effects of information depend on customers' perceptions of their queue positions in the absence of accurate information. These results also suggest that "correcting" customer perceptions, instead of providing accurate queue information about their queue positions, may be sufficient for reducing abandonment. Vague delay announcements have have been previously studied in the context of unobserved (virtual) call center queues, e.g., Allon et al. (2011), Allon and Bassamboo (2011). Investigating the impact of vague announcements in observable settings, aimed at influencing customer perceptions, can be an interesting area of future research.

Our results assume that customers in both classes use the same θ function (that maps their perceived queue positions to individual abandonment rates). As noted in §3, this assumption is for analytical tractability, but also allows us to isolate the interactions between information granularity and system characteristics, i.e., non-stationary arrival rates and priority service. To accommodate heterogeneous abandonment behaviors across different classes, one can relax this assumption. Whether customers in different classes respond differently to their perceived queue positions, and if so, how, is ultimately an empirical question, and is certainly worthy of future research (see below). With respect to the implications of a model with class-dependent θ functions, it seems intuitively clear that our results would continue to hold if these functions are either "not too different", or if they differ in a way that reinforces the effects that our analysis identifies; on the other hand, some of our results may be reversed if the θ functions differ in a way that counters these effects.

Finally, another interesting direction is to investigate the estimation of the models introduced in this paper from data. Our modeling framework assumes that customers' abandonment rate is determined through a function (namely, θ) that maps the perceived position of the customer to an abandonment rate. Estimation of this function using data, which can be facilitated by imposing additional structure on the function, can be considered in future work.

References

Afèche P, Sarhangian V (2015) Rational abandonment from priority queues: equilibrium strategy and pricing implications. *Columbia Business School Research Paper* (15-93).

Aksin OZ, Gencer B, Gunes ED (2022) How observed queue length and service times drive queue behavior in the lab. *Available at SSRN 3387077* .

Akşin Z, Ata B, Emadi SM, Su CL (2017) Impact of delay announcements in call centers: An empirical approach. *Operations Research* 65(1):242–265.

Allon G, Bassamboo A (2011) The impact of delaying the delay announcements. *Operations research* 59(5):1198–1210.

Allon G, Bassamboo A, Gurvich I (2011) "we will be right with you": Managing customer expectations with vague promises and cheap talk. *Operations research* 59(6):1382–1394.

Armony M, Shimkin N, Whitt W (2009) The impact of delay announcements in many-server queues with abandonment. *Operations Research* 57(1):66–81.

Assaf D, Haviv M (1990) Reneging from processor sharing systems and random queues. *Mathematics of Operations Research* 15(1):129–138.

Ata B, Glynn P, Peng X (2017) An equilibrium analysis of a discrete-time markovian queue with endogenous abandonments. *Queueing Systems* 86(1-2):141–212.

Ata B, Peng X (2017) An equilibrium analysis of a multiclass queue with endogenous abandonments in heavy traffic. *Operations Research* 66(1):163–183.

Atar R, Giat C, Shimkin N (2010) The cμ/θ rule for many-server queues with abandonment. *Operations Research* 58(5):1427–1439.

Bagagiolo F (2012) Lecture notes: Ordinary differential equations. *Available at http://www.science.unitn.it/~bagagiol/noteODE.pdf* .

Bassamboo A, Ibrahim R (2021) A general framework to compare announcement accuracy: Static vs. les-based announcement. *Management Science* 67(7):4191–4208.

Bassamboo A, Randhawa RS (2016) Scheduling homogeneous impatient customers. *Management Science* 62(7):2129–2147.

Batt RJ, Terwiesch C (2015) Waiting patiently: An empirical study of queue abandonment in an emergency department. *Management Science* 61(1):39–59.

Bolandifar E, DeHoratius N, Olsen T, Wiler J (2019) An empirical study of the behavior of patients who leave the emergency department without being seen. *Journal of Operations Management* 65(5):430–446.

Buell RW (2021) Last-place aversion in queues. *Management Science* 67(3):1430–1452.

Burnetas A, Economou A (2007) Equilibrium customer strategies in a single server markovian queue with setup times. *Queueing Systems* 56(3):213–228.

Chen H, Frank M (2004) Monopoly pricing when customers queue. *IIE Transactions* 36(6):569–581.

Cui S, Veeraraghavan SK, Wang J, Zhang Y (2022) Observational reneging. *Available at SSRN 3290868* .

Dong J, Feldman P, Yom-Tov GB (2015) Service systems with slowdowns: Potential failures and proposed solutions. *Operations Research* 63(2):305–324.

Dong J, Ibrahim R (2021) Srpt scheduling discipline in many-server queues with impatient customers. *Management Science* 67(12):7708–7718.

Dong J, Perry O (2020) Queueing models for patient-flow dynamics in inpatient wards. *Operations Research* 68(1):250–275.

Guo P, Zipkin P (2007) Analysis and comparison of queues with different levels of delay information. *Management Science* 53(6):962–970.

Guo P, Zipkin P (2008) The effects of information on a queue with balking and phase-type service times. *Naval Research Logistics (NRL)* 55(5):406–411.

Hassin R (1985) On the optimality of first come last served queues. *Econometrica* 53(1):201–202.

Hassin R (1986) Consumer information in markets with random product quality: The case of queues and balking. *Econometrica: Journal of the Econometric Society* 1185–1195.

Hassin R, Roet-Green R (2020) On queue-length information when customers travel to a queue. *Manufacturing & Service Operations Management* .

Haviv M, Ritov Y (2001) Homogeneous customers renege from invisible queues at random times under deteriorating waiting conditions. *Queueing Systems* 38:495–508.

Heyman DP, Whitt W (1984) The asymptotic behavior of queues with time-varying arrival rates. *Journal of Applied Probability* 21(1):143–156.

Ibrahim R (2018) Sharing delay information in service systems: a literature survey. *Queueing Systems* 89(1-2):49–79.

Ibrahim R, Armony M, Bassamboo A (2017) Does the past predict the future? the case of delay announcements in service systems. *Management Science* 63(6):1762–1780.

Janakiraman N, Meyer RJ, Hoch SJ (2011) The psychology of decisions to abandon waits for service. *Journal of Marketing Research* 48(6):970–984.

Jennings OB, Pender J (2016) Comparisons of ticket and standard queues. *Queueing Systems* 84(1):145–202.

John D (2011) Little. little's law as viewed on its 50th anniversary. *Operations Research* 59(3):536–549.

Jouini O, Akşin Z, Dallery Y (2011) Call centers with delay information: Models and insights. *Manufacturing & Service Operations Management* 13(4):534–548.

Jouini O, Dallery Y, Akşin Z (2009) Queueing models for full-flexible multi-class call centers with real-time anticipated delays. *International Journal of Production Economics* 120(2):389–399.

Khalil H (2002) *Nonlinear Systems* (Prentice Hall).

Khalil HK (2014) *Nonlinear control* (Pearson Higher Ed).

Kuzu K, Gao L, Xu SH (2019) To wait or not to wait: The theory and practice of ticket queues. *Manufacturing & Service Operations Management* 21(4):853–874.

Lakrad F, Pender J, Rand R (2022) Queues with delayed information: A dynamical systems perspective. *SIAM Journal on Applied Dynamical Systems* 21(1):676–713.

Lebovitz NR (1999) *Ordinary Differential Equations*. URL http://people.cs.uchicago.edu/~lebovitz/odes.html.

Liu Y, Whitt W (2011a) Nearly periodic behavior in the overloaded $G/D/s+GI$ queue. *Stochastic Systems* 1(2):340–410.

Liu Y, Whitt W (2011b) A network of time-varying many-server fluid queues with customer abandonment. *Operations research* 59(4):835–846.

Lu Y, Musalem A, Olivares M, Schilkrut A (2013) Measuring the effect of queues on customer purchases. *Management Science* 59(8):1743–1763.

Mandelbaum A, Massey WA, Reiman MI (1998) Strong approximations for markovian service networks. *Queueing Systems* 30(1):149–201.

McNabb A (1986) Comparison theorems for differential equations. *Journal of mathematical analysis and applications* 119(1-2):417–428.

Nirenberg S, Daw A, Pender J (2018) The impact of queue length rounding and delayed app information on disney world queues. *2018 Winter Simulation Conference (WSC)*, 3849–3860 (IEEE).

Pender J (2017) Sampling the functional kolmogorov forward equations for nonstationary queueing networks. *INFORMS Journal on Computing* 29(1):1–17.

Perry O, Whitt W (2016) Chattering and congestion collapse in an overload switching control. *Stochastic Systems* 6(1):132–210.

Shimkin N, Mandelbaum A (2004) Rational abandonment from tele-queues: Nonlinear waiting costs with heterogeneous preferences. *Queueing Systems: Theory Appl.* 47(1-2):117–146.

Westphal M, Yom-Tov G, Parush A, Rafaeli A (2022) Reducing abandonment and improving attitudes in emergency departments: Integrating delay announcements into operational transparency to signal service quality. *Available at SSRN 4120485* .

Westphal M, Yom-Tov GB, Parush A, Carmeli N, Shaulov A, Shapira C, Rafaeli A, et al. (2020) A patient-centered information system (myed) for emergency care journeys: design, development, and initial adoption. *JMIR Formative Research* 4(2):e16410.

Whitt W (1999) Improving service by informing customers about anticipated delays. *Management science* 45(2):192–207.

Whitt W (2005a) Engineering solution of a basic call-center model. *Management science* 51(2):221–235.

Whitt W (2005b) Two fluid approximations for multi-server queues with abandonments. *Operations Research Letters* 33(4):363–372.

Whitt W (2006) Fluid models for multiserver queues with abandonments. *Operations research* 54(1):37–54.

Whitt W (2015) Notes on little's law: $L = \lambda W$. IEOR 4615: Service Engineering Lecture Notes, January 27, 2015, URL http://www.columbia.edu/~ww2040/.

Xu SH, Gao L, Ou J (2007) Service performance analysis and improvement for a ticket queue with balking customers. *Management science* 53(6):971–990.

Yu L, Iravani S, Perry O (2021a) A fluid-diffusion-hybrid limiting approximation for priority systems with fast and slow customers. *Operations Research* .

Yu Q, Allon G, Bassamboo A (2017) How do delay announcements shape customer behavior? an empirical study. *Management Science* 63(1):1–20.

Yu Q, Allon G, Bassamboo A (2021b) The reference effect of delay announcements: A field experiment. *Management Science* 67(12):7417–7437.

Yu Q, Zhang Y, Zhou YP (2022) Delay information in virtual queues: A large-scale field experiment on a major ride-sharing platform. *Management Science* 68(8):5745–5757.

(b_1, b_2)	(h_1, h_2)	$\theta(x)$	(ρ_1, ρ_2)
$(1, 1)$	$(1, 1)$	$2 - e^{-x}$	(0.5, 0.5)
$(0.5, 0.5)$	$(2, 2)$	$1.5 - e^{-x}$	(0.8, 0.2)
$(0.6, 1)$	$(2, 1)$		(0.2, 0.8)
$(0.3, 1)$			(1,1)
			(1.6,0.4)
			(0.4,1.6)

Table 1 **Parameter sets explored for model validation in Section A for $\beta_k(x) = b_k e^{-h_k x}$.**

Appendix A: Details of the Numerical Study of Section 3

In this appendix, we present the numerical study summarized in Section 3. In particular, we numerically show that the $\beta_k(\boldsymbol{w}(t))$ model can be accurately approximated by a best-fitting β_k model, and that the performance comparisons are preserved under the β_k model.

Parameter Choices. Given the absence of empirical evidence on the form of $\beta_k(x)$ in practice, we explore a family of exponential functions $\beta_k(x) = b_k e^{-h_k x}$ with varies parameters b_k, h_k for both cases. Specifically, b_k stands for the initial belief of the position of new arrivals, i.e., the value of $\beta_k(0)$. A smaller b_k indicates a smaller initial perceived position. For example, when $b_k = 1$, new arrivals perceive themselves as joining the end of the line; when $b_k = 0.5$, they consider themselves positioned in the middle of the queue upon arrival.

We consider two cases, each capturing one variant of the No Information setting:

Case 1: We assume that customers are not informed of any priority classes, including their own. In this case, it is reasonable to assume a common position function for both classes. That is, $\beta_1(x) = \beta_2(x)$.

Case 2: We assume that customers are informed of their own priority class (but not others). In this case, HP and LP customers may have different perceptions of their positions upon arrival and during their waits. Specifically, assume that $\beta_1(0) \leq \beta_2(0)$ (HP customers perceive themselves to be closer to the head of the queue than LP customers upon arrival) and $\beta_1'(x) < \beta_2'(x) < 0$ (LP customers are aware of their lower priority and therefore perceive their queue positions as improving at a slower rate).

The full set of parameter choices b_k, h_k used in the queue position fraction functions $\beta_k(x) = b_k e^{-h_k x}$ are detailed in the first two columns of Table 1 and plotted in Figure 8.

In addition to different choices of $\beta(x)$ functions, we assume $\mu_1 = 1, \mu_2 = 1$, $s = 100$, and consider two abandonment rate functions, $\theta(x)$, whose ranges are either comparable to or greater than the service rates; see the third column of Table 1. We investigate non-stationary sinusoidal arrival rates $\lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$ with different sets of average system loads, as shown in the last column of Table 1. The period of these arrival rates is 24. Specifically, we consider a moderately loaded system (where the sum of average loads equals 1) and a heavily loaded system (where the sum of average loads equals 2). For each case, we further examine three load configurations: balanced loads (HP:LP = 50:50), higher HP load (HP:LP = 80:20), and higher LP load (HP:LP = 20:80).

For each combination of parameters and specific form of $\beta_k(\cdot)$ in Table 1, we first generate the sample path of the number-in-system process, $X^N(t, \boldsymbol{\beta}(\boldsymbol{w}(t)))$, by simulating the corresponding arrival, service, and abandonment processes under the waiting-time-dependent queue position function $\beta_k(\boldsymbol{w}(t))$, where $\boldsymbol{\beta}(\boldsymbol{w}(t)) := (\beta_1(\boldsymbol{w}(t)), \beta_2(\boldsymbol{w}(t)))$.

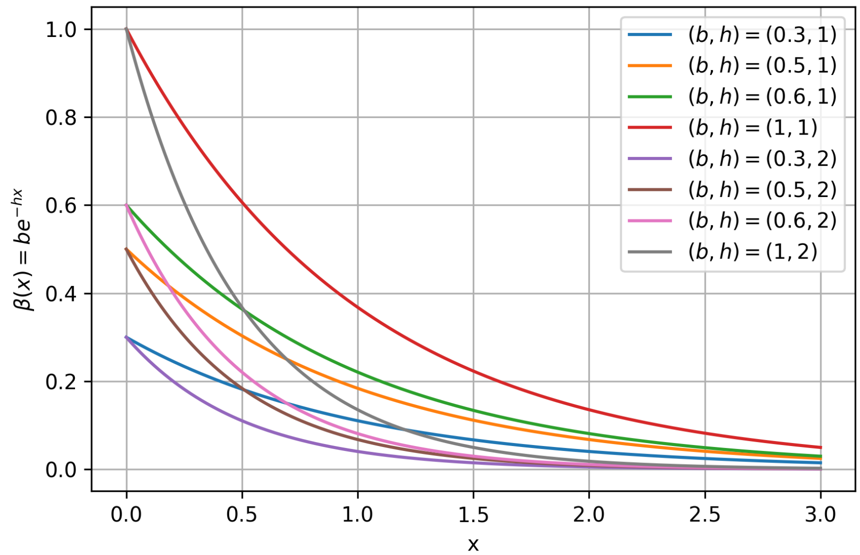


Figure 8 **Exponential $\beta_k(x) = b_k e^{-h_k x}$ with various pair of (b, h).**

(a) HP customers, $\beta_1(x) = \beta_2(x) = e^{-x}$ (b) LP customers, $\beta_1(x) = \beta_2(x) = e^{-x}$

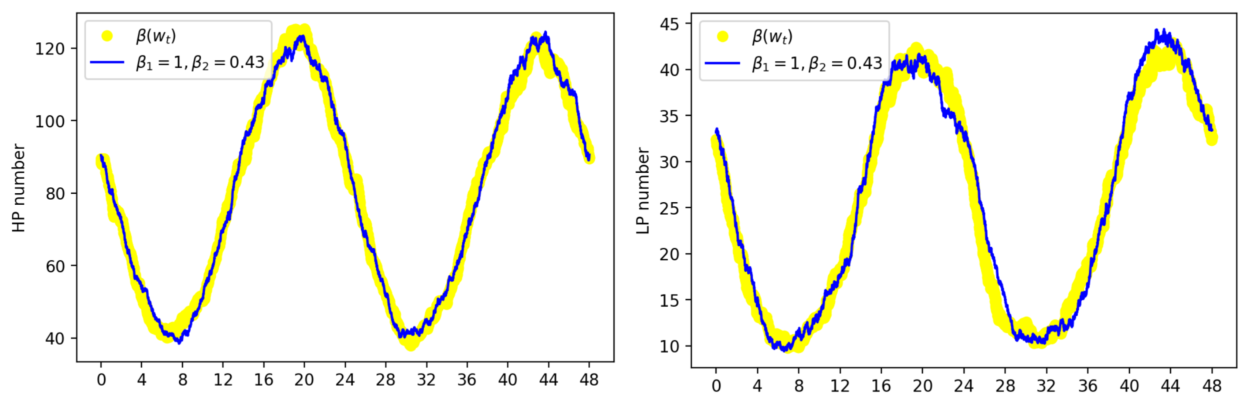


Figure 9 **Trajectories of the average sample path of the numbers-in-system processes under the best-fitting β_k models, compared with the average sample path of the $\beta_k(\boldsymbol{w}(t))$ model ($\mu_1 = 1, \mu_2 = 1$, $s = 100, \theta(x) = 1.5 - e^{-x}, \lambda_1(t) = 80(1 - 0.5\sin(\pi t/12)), \lambda_2(t) = 20(1 - 0.5\sin(\pi t/12))$).**

Simulation Settings. We generate the sample path of the $\beta_k(\boldsymbol{w}(t))$ model starting from an empty system. We initiate the sample path by generating arrival clocks for HP and LP customers. The following describes how the system is updated when different events occur:

1. When a new HP customer arrives, the HP number-in-system is incremented by 1. If all servers are busy with HP customers, the customer joins the HP queue and an abandonment clock is created; if all servers are busy but some are serving LP customers, one LP in service is randomly preempted and returned to the head of the LP queue, the new HP enters service with a service clock, and an updated abandonment clock is set for the preempted LP based on her accumulated waiting time and queue length; if there are available servers, the HP customer immediately enters service with a service clock.

2. When a new LP customer arrives, the LP number-in-queue is incremented by 1. If all servers are busy, the customer joins the LP queue and an abandonment clock is created; if there are available servers, the customer immediately enters service with a service clock.

3. When an HP customer completes service and leaves the system, the HP number-in-system is decreased by 1. If there are HP customers in the queue, the one with the earliest arrival time enters service; the customer's abandonment clock is removed, and a service clock is created. If there are no HP customers in the queue but some LP customers are waiting, the one with the earliest arrival time enters service; the customer's abandonment clock is removed, and a service clock is created.

4. When an LP customer completes service and leaves the system, the LP number-in-system is decreased by 1. If there are LP customers in the queue, the one with the earliest arrival time enters service; the customer's abandonment clock is removed, and a service clock is created.

5. When an HP or LP customer abandon the system, the corresponding HP or LP number-in-system is decreased by 1.

6. Ideally, we should continuously update the abandonment rates for customers waiting in queue since $\boldsymbol{w}(t)$ evolves continuously. However, due to the complexity of simulating a nonhomogeneous Poisson process with an endogenously changing rate function, we instead update the individual abandonment rates for all customers in queue whenever a new event occurs. Given the large scale of our simulated system (with $s = 100$), the system state updates approximately every 0.005 units of time for moderately loaded system (with an average total load of 1), and even more frequently under heavier loads. This ensures that the differences between our simulated abandonment processes and the actual abandonment processes remain small.

7. The next event is the one whose associated clock (arrival, service, or abandonment) expires first.

We use the initial period $t = 200$ as a warm-up and allow the simulation to run for a sufficient duration, up to $t = 1160$ (40 periods). Consequently, based on the system state updates after the warm-up period, we compute the performance measures for this system, including the long-run average numbers-in-system $(\bar{x}^N_{1,w}, \bar{x}^N_{2,w})$, the long-run average system abandonment rates $(\bar{A}^N_{1,w}, \bar{A}^N_{2,w})$ and the time-varying average numbers-in-system over a period $(\tilde{x}^N_{1,w}(t), \tilde{x}^N_{2,w}(t))$. We observe that, after the initial warm-up period, the trajectories of the sample paths appears to be periodic. We divide each sample path into intervals of one period in length and compute the performance measures by taking the average numbers-in-system and abandonment rates over these periods after the warm-up period.

Best-fitting β_k model. We identify the best-fitting β_k model by matching its long-run average numbers-in-system with those obtained under the $\beta_k(\boldsymbol{w}(t))$ model. In the non-stationary and periodic setting, the equilibrium of the system under β_k model is time-varying and periodic. For a given $\boldsymbol{\beta} = (\beta_1, \beta_2)$, we obtain the long-run average number-in-system $(\bar{x}^N_1(\boldsymbol{\beta}), \bar{x}^N_2(\boldsymbol{\beta}))$ from the simulated stochastic system. Then, for each pair of (β_1, β_2) with $\beta_k \in \{0, 0.01, \ldots, 1\}, k \in \{1, 2\}$, we search for the β_k model with $\boldsymbol{\beta}$ such that the average relative gap between $(\bar{x}^N_1(\boldsymbol{\beta}), \bar{x}^N_2(\boldsymbol{\beta}))$ and $(\bar{x}^N_{1,w}, \bar{x}^N_{2,w})$, defined as $MAPE_x(\boldsymbol{\beta}) := (|\bar{x}^N_1(\boldsymbol{\beta}) - \bar{x}^N_{1,w}|/\bar{x}^N_1(\boldsymbol{\beta}) + |\bar{x}^N_2(\boldsymbol{\beta}) - \bar{x}^N_{2,w}|/\bar{x}^N_{2,w})/2$, is minimized. The identified value of $\boldsymbol{\beta}$ is denoted as $\boldsymbol{\beta}^* = (\beta^*_1, \beta^*_2)$.

After obtaining the best-fitting β_k model that approximates the $\beta_k(\boldsymbol{w}(t))$ model in terms of the average number-in-system, we further examine the differences in the average system abandonment rates between the two models. We do so by evaluating the average relative gap of the long-run time-average system abandonment rates, defined as $MAPE_a(\boldsymbol{\beta}) := (|\bar{A}_1^N(\boldsymbol{\beta}) - \bar{A}_{1,w}^N|/\bar{A}_1^N(\boldsymbol{\beta}) + |\bar{A}_2^N(\boldsymbol{\beta}) - \bar{A}_{2,w}^N|/\bar{A}_{2,w}^N)/2$, between the $\beta_k(\boldsymbol{w}(t))$ model and the best-fitting β_k model. Note that for cases where the HP load is (almost) uniformly underloaded (i.e., $\rho_1 \leq 0.5$), the measure reduces to $MAPE_a(\boldsymbol{\beta}) := |\bar{A}_2^N(\boldsymbol{\beta}) - \bar{A}_{2,w}^N|/\bar{A}_{2,w}^N$, since the HP abandonment rate is either negligible (around 0.001) or zero.

Tables 2 and 3 present the average relative gaps between the $\beta_k(\boldsymbol{w}(t))$ model and the best-fitting β_k model to demonstrate that the $\beta_k(\boldsymbol{w}(t))$ model can be well approximated by an appropriately chosen β_k model.

Key finding. For a wide range of $\beta_k(x)$ functions, the long-run average numbers-in-system and abandonment rates of the $\beta_k(\boldsymbol{w}(t))$ model can be accurately approximated by appropriately choosing the (β_1, β_2) pair for our β_k model, under non-stationary two-class priority systems.

As illustrated in Tables 2 and 3, the best-fitting β_k model yields average numbers-in-system very close to those of the $\beta_k(\boldsymbol{w}(t))$ model, with average relative gaps ($MAPE_x(\boldsymbol{\beta}^*)$) of less than 0.4%. The best-fitting β_k model also maintains a small gap in the time-average system abandonment rates (less than 2%).

This finding supports the value of our parsimonious β_k model. It is quite intuitive that the β_k model can be "tuned" to match the performance under the more general $\beta_k(\boldsymbol{w}(t))$ model: Whereas the $\beta_k(\boldsymbol{w}(t))$ model tracks the waiting-time-dependent, and therefore heterogeneous, queue position beliefs of all present customers at all times, the appropriately chosen β_k parameters simply reflect the class-level averages of these beliefs over time. Therefore, the values of the matching (β_1, β_2) depend on both, the $\beta_k(x)$ function(s) and the waiting times experienced in each class.

We also perform the numerical studies for non-stationary sinusoidal arrival rates under different system loads (moderate or high, and balanced or unbalanced between priority classes) and other parameter choices for the abandonment rate function. And the results are aligned with our observations in this section.

Interpretation and Relevant Regimes of $\boldsymbol{\beta}^*$. Tables 2 and 3 also show how the best-fitting position fraction $\boldsymbol{\beta}^* = (\beta_1^*, \beta_2^*)$ varies with different parameter choices. In general, $\boldsymbol{\beta}^*$ can span nearly the entire range from 0 to 1 depending on the parameters.

For Case 1 with class-independent queue position functions, i.e., $\beta_1(x) = \beta_2(x)$, we consistently observe $\beta_1^* > \beta_2^*$, as HP customers have shorter waiting times than LP customers; e.g., the first four cases in Table 2.

For Case 2 with *class-dependent* $\beta_k(\boldsymbol{w}(t))$ function, since $\beta_1(0) \leq \beta_2(0)$ and $\beta_1'(x) < \beta_2'(x) < 0$, the values of the matching β_k model would reflect on average more optimistic HP customers and thus we may observe $\beta_1^* < \beta_2^*$; e.g., the last two cases with $\rho_1 = \rho_2 = 1$ and $\rho_1 = 1.6, \rho_2 = 0.4$ in Table 3.

Note that the impact of different $\beta_k(x)$ functions on the long-run average system abandonment rates is limited. This is because the variation in the form of $\beta_k(x)$ function is inherently bounded by 1 (by definition) and further constrained by the range of the $\theta(x)$ function. Additionally, in highly overloaded systems (as in Table 2), servers are almost always busy regardless of the $\beta_k(x)$ function used, resulting in a similar level of throughput across all systems. Since the difference in system abandonment rates corresponds to the difference in system throughputs, these differences should be small. On the other hand, when system loads are low, abandonment rates are low, and $\beta_k(x)$ has low impact on the system through abandonments.

Performance comparisons with full information model. To show that the identified best-fitting β_k model preserves the performance comparisons between the no information $\beta_k(\boldsymbol{w}(t))$ model and the full

Average Loads		$\beta_k(\boldsymbol{w}(t))$ model								Best-fitting β_k model						Average Relative Gap	
ρ_1	ρ_2	b_1	b_2	h_1	h_2	$\bar{x}^N_{1,w}$	$\bar{x}^N_{2,w}$	$\bar{A}^N_{1,w}$	$\bar{A}^N_{2,w}$	$\boldsymbol{\beta}^*_1$	$\boldsymbol{\beta}^*_2$	$\bar{x}^N_1(\boldsymbol{\beta}^*)$	$\bar{x}^N_2(\boldsymbol{\beta}^*)$	$\bar{A}^N_1(\boldsymbol{\beta}^*)$	$\bar{A}^N_2(\boldsymbol{\beta}^*)$	$MAPE_x$	$MAPE_a$
0.5	0.5	1	1	1	1	50.35	54.69	0	15.92	-	0.66	50.35	54.77	0.001	15.76	0.08%	1.05%
0.5	0.5	1	1	2	2	50.35	57.11	0.002	15.71	-	0.37	50.35	57.17	0.001	15.43	0.05%	1.77%
0.5	0.5	0.5	0.5	1	1	50.35	57.99	0.002	15.47	-	0.32	50.35	58.00	0.001	15.52	0.02%	0.36%
0.5	0.5	0.5	0.5	2	2	50.35	59.69	0.001	15.25	-	0.19	50.35	59.74	0.002	15.35	0.04%	0.62%
0.5	0.5	1	1	2	1	50.35	54.48	0.002	15.65	-	0.68	50.35	54.48	0.001	15.55	0.00%	0.60%
0.5	0.5	0.5	0.5	2	1	50.35	58.05	0.001	15.59	-	0.32	50.35	58.00	0.001	15.52	0.04%	0.45%
0.5	0.5	0.6	1	1	1	50.35	54.76	0.001	15.51	-	0.63	50.35	54.81	0.001	15.54	0.04%	0.24%
0.5	0.5	0.6	1	2	1	50.35	54.62	0.002	15.54	-	0.67	50.35	54.59	0.001	15.74	0.03%	1.32%
0.5	0.5	0.3	1	1	1	50.35	54.76	0.002	15.74	-	0.66	50.35	54.77	0.001	15.76	0.02%	0.13%
0.5	0.5	0.3	1	2	1	50.35	54.80	0.001	15.52	-	0.63	50.35	54.81	0.001	15.54	0.00%	0.17%
0.8	0.2	1	1	1	1	79.78	25.10	4.14	10.64	1	0.43	79.92	25.23	4.74	10.69	0.33%	0.43%
0.8	0.2	1	1	2	2	80.69	27.08	4.80	10.86	0.71	0.24	80.80	27.06	4.78	10.86	0.10%	0.07%
0.8	0.2	0.5	0.5	1	1	81.61	27.40	4.71	10.85	0.36	0.19	81.91	27.40	4.63	11.00	0.19%	1.31%
0.8	0.2	0.5	0.5	2	2	81.65	28.49	4.70	10.96	0.33	0.11	81.88	28.53	4.64	10.89	0.22%	0.63%
0.8	0.2	1	1	2	1	80.70	25.23	4.79	10.93	0.76	0.42	80.86	25.24	4.78	10.98	0.11%	0.44%
0.8	0.2	0.5	0.5	2	1	81.75	27.45	4.68	10.98	0.36	0.2	81.85	27.47	4.65	10.99	0.10%	0.14%
0.8	0.2	0.6	1	1	1	81.37	25.27	4.70	11.03	0.47	0.4	81.61	25.24	4.68	11.05	0.22%	0.19%
0.8	0.2	0.6	1	2	1	81.68	25.20	4.68	11.08	0.47	0.41	81.72	25.24	4.68	11.11	0.10%	0.34%
0.8	0.2	0.3	1	1	1	82.46	25.19	4.56	11.15	0.37	0.42	82.04	25.21	4.61	11.13	0.28%	0.22%
0.8	0.2	0.3	1	2	1	82.63	25.47	4.51	11.25	0.19	0.36	82.93	25.40	4.49	11.26	0.33%	0.13%
0.2	0.8	1	1	1	1	20.40	83.90	0	15.97	-	0.79	20.40	83.83	0.00	16.02	0.04%	0.31%
0.2	0.8	1	1	2	2	20.40	85.11	0.00	16.07	-	0.6	20.40	84.89	0.00	15.87	0.13%	1.25%
0.2	0.8	0.5	0.5	1	1	20.40	87.30	0.00	15.94	-	0.36	20.40	87.59	0.00	15.89	0.17%	0.31%
0.2	0.8	0.5	0.5	2	2	20.40	88.39	0.00	15.86	-	0.28	20.40	89.02	0.00	15.95	0.36%	0.56%
0.2	0.8	1	1	2	1	20.40	83.68	0.00	15.96	-	0.83	20.40	83.65	0.00	15.94	0.01%	0.12%
0.2	0.8	0.5	0.5	2	1	20.40	86.91	0.00	15.75	-	0.4	20.40	86.84	0.00	15.90	0.04%	0.95%
0.2	0.8	0.6	1	1	1	20.40	83.41	0.00	15.92	-	0.88	20.40	83.33	0.00	15.98	0.05%	0.39%
0.2	0.8	0.6	1	2	1	20.40	83.88	0.00	16.01	-	0.79	20.40	83.83	0.00	16.02	0.02%	0.05%
0.2	0.8	0.3	1	1	1	20.40	83.69	0.00	15.92	-	0.83	20.40	83.65	0.00	15.94	0.02%	0.09%
0.2	0.8	0.3	1	2	1	20.40	83.68	0.00	15.98	-	0.83	20.40	83.65	0.00	15.94	0.01%	0.25%

Table 2 Performance comparisons between $\beta_k(\boldsymbol{w}(t))$ model and best-fitting β_k model for moderate loads, and various $\beta_k(x)$ functions. ($s = 100, \theta(x) = 1.5 - e^{-x}$, $\mu_k = 1$, $\lambda_k(t) = \rho_k s(1 - 0.5\sin(\pi t/12)), \beta_k(x) = b_k e^{-h_k x}$, for $k = 1, 2$.)

information model, we compute the long-run average performance measures under the full information model. In the non-stationary and periodic setting, using the parameter sets in Table 1, we compute the long-run average number-in-system and time-average system abandonment rates of the stochastic systems under full information model, and compare these performance measures with those we obtained from the no information $\beta_k(\boldsymbol{w}(t))$ model and the best-fitting β_k model.

We observe that the ranking results of the long-run average number-in-system and abandonment rates are well preserved in all cases considered. An illustrative example is presented in Figure 10. Each point in this figure represents one of the ten parameter choice of the $\beta_k(\cdot)$ functions, as listed in the first 10 rows of Table 2.

Average Loads		$\beta_k(\boldsymbol{w}(t))$ model								Best-fitting β_k model						Average Relative Gap	
ρ_1	ρ_2	b_1	b_2	h_1	h_2	$\bar{x}^N_{1,w}$	$\bar{x}^N_{2,w}$	$\bar{A}^N_{1,w}$	$\bar{A}^N_{2,w}$	$\boldsymbol{\beta}^*_1$	$\boldsymbol{\beta}^*_2$	$\bar{x}^N_1(\boldsymbol{\beta}^*)$	$\bar{x}^N_2(\boldsymbol{\beta}^*)$	$\bar{A}^N_1(\boldsymbol{\beta}^*)$	$\bar{A}^N_2(\boldsymbol{\beta}^*)$	$MAPE_x$	$MAPE_a$
1	1	1	1	1	1	96.70	107.60	15.90	84.32	0.87	0.41	96.72	107.67	15.90	84.32	0.04%	0.02%
1	1	1	1	2	2	96.60	124.15	15.90	84.46	0.79	0.22	96.54	123.82	15.94	84.47	0.17%	0.13%
1	1	0.5	0.5	1	1	98.83	125.37	15.82	84.54	0.44	0.20	98.81	125.12	15.80	84.44	0.11%	0.14%
1	1	0.5	0.5	2	2	99.16	139.97	15.82	84.26	0.36	0.11	99.31	140.19	15.78	84.52	0.16%	0.29%
1	1	1	1	2	1	97.07	107.92	15.90	84.40	0.79	0.41	97.01	107.57	15.87	84.42	0.19%	0.12%
1	1	0.5	0.5	2	1	99.68	125.05	15.78	84.47	0.37	0.20	99.79	125.54	15.76	84.50	0.25%	0.11%
1	1	0.6	1	1	1	98.90	107.30	15.84	84.31	0.51	0.41	98.91	107.42	15.78	84.39	0.06%	0.22%
1	1	0.6	1	2	1	99.70	106.87	15.81	84.37	0.43	0.42	99.75	106.62	15.77	84.36	0.14%	0.15%
1	1	0.3	1	1	1	102.53	106.46	15.68	84.51	0.25	0.42	102.64	106.28	15.67	84.52	0.14%	0.05%
1	1	0.3	1	2	1	103.70	106.56	15.62	84.75	0.20	0.41	103.88	106.56	15.59	84.69	0.09%	0.12%
1.6	0.4	1	1	1	1	152.06	46.62	63.49	37.24	0.77	0.38	152.19	46.57	63.44	37.30	0.10%	0.12%
1.6	0.4	1	1	2	2	155.75	54.32	63.44	37.35	0.58	0.19	155.30	54.36	63.37	37.31	0.18%	0.10%
1.6	0.4	0.5	0.5	1	1	163.97	53.89	63.28	37.49	0.36	0.19	164.02	54.12	63.28	37.48	0.22%	0.01%
1.6	0.4	0.5	0.5	2	2	169.43	61.03	63.29	37.44	0.25	0.1	170.08	60.81	63.22	37.49	0.37%	0.11%
1.6	0.4	1	1	2	1	157.55	45.68	63.42	37.41	0.54	0.4	157.91	45.62	63.33	37.47	0.18%	0.16%
1.6	0.4	0.5	0.5	2	1	170.67	53.42	63.28	37.49	0.26	0.19	170.42	53.37	63.19	37.51	0.12%	0.10%
1.6	0.4	0.6	1	1	1	161.52	46.09	63.34	37.53	0.44	0.37	161.41	46.24	63.29	37.44	0.19%	0.15%
1.6	0.4	0.6	1	2	1	167.67	45.13	63.30	37.44	0.33	0.38	167.02	45.16	63.24	37.56	0.23%	0.20%
1.6	0.4	0.3	1	1	1	176.44	44.56	63.14	37.65	0.21	0.37	176.23	44.64	63.10	37.70	0.15%	0.08%
1.6	0.4	0.3	1	2	1	183.85	43.93	63.06	37.66	0.14	0.37	184.09	44.02	63.05	37.73	0.17%	0.10%
0.4	1.6	1	1	1	1	40.50	156.66	0	100.80	-	0.62	40.50	156.78	0	101.05	0.04%	0.25%
0.4	1.6	1	1	2	2	40.50	171.91	0	101.10	-	0.34	40.50	172.27	0	100.81	0.11%	0.29%
0.4	1.6	0.5	0.5	1	1	40.50	177.73	0	100.95	-	0.28	40.50	177.50	0	100.87	0.06%	0.08%
0.4	1.6	0.5	0.5	2	2	40.50	194.51	0	101.15	-	0.16	40.50	194.20	0	101.07	0.08%	0.08%
0.4	1.6	1	1	2	1	40.50	156.95	0	100.88	-	0.61	40.50	157.27	0	100.99	0.10%	0.11%
0.4	1.6	0.5	0.5	2	1	40.50	177.48	0	101.02	-	0.28	40.50	177.50	0	100.87	0.01%	0.14%
0.4	1.6	0.6	1	1	1	40.50	156.80	0	100.89	-	0.61	40.50	157.27	0	100.99	0.15%	0.09%
0.4	1.6	0.6	1	2	1	40.50	157.28	0	100.85	-	0.6	40.50	157.40	0	100.98	0.04%	0.13%
0.4	1.6	0.3	1	1	1	40.50	156.75	0	100.86	-	0.61	40.50	157.27	0	100.99	0.16%	0.12%
0.4	1.6	0.3	1	2	1	40.50	157.06	0	100.89	-	0.61	40.50	157.27	0	100.99	0.07%	0.09%

Table 3 **Performance comparisons between $\beta_k(\boldsymbol{w}(t))$ model and best-fitting β_k model for heavy loads, and various $\beta_k(x)$ functions. ($s=100, \theta(x)=1.5-e^{-x}$, $\mu_k=1$, $\lambda_k(t)=\rho_k s(1-0.5\sin(\pi t/12)), \beta_k(x)=b_k e^{-h_k x}$, for $k=1,2$.)**

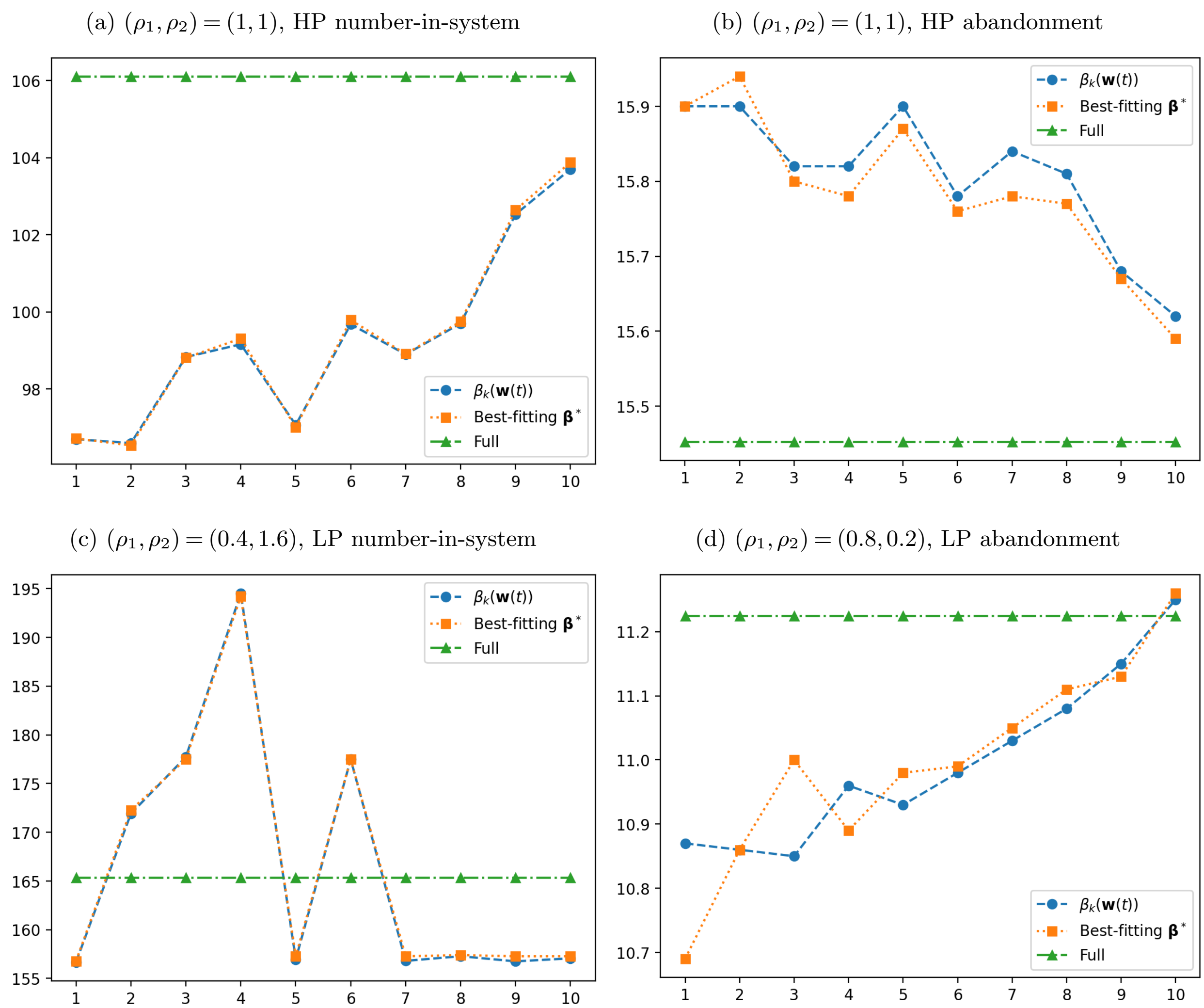


Figure 10 Comparisons of performance measures under No information $\beta_k(\boldsymbol{w}(t))$, No information Best-fitting β_k model with $\boldsymbol{\beta} = \boldsymbol{\beta}^*$, and Full information model ($\mu_1 = \mu_2 = 1$, $s = 100, \theta(x) = 1.5 - e^{-x}, \lambda_k(t) = s\rho_k(1 - 0.5\sin(\pi t/12))$.

Appendix B: Proof of Section 4: Theorem 1

In this section, we provide the proof of Theorem 1. We first introduce a useful Functional Strong Law of Large Numbers from Mandelbaum et al. (1998) in Appendix B.1, and then provide the proof of Theorem 1 for no and full information systems in Appendices B.2 and B.3, respectively.

B.1. Functional Strong Law of Large Numbers

Consider a sequence of stochastic processes $\mathbf{Q}^n := \{\mathbf{Q}^n(t)|t \geq 0\}$ for $n > 0$. The sample paths of $\mathbf{Q}^n$ are uniquely determined by $\mathbf{Q}^n(0)$ and the functional equations

$$\mathbf{Q}^n(t) = \mathbf{Q}^n(0) + \sum_{i \in I} K_i \left(\int_0^t \alpha_s^n(\frac{1}{n}\mathbf{Q}^n(s), i) ds \right) \mathbf{v}_i,$$

where $\{K_i(\cdot)|i \in I\}$ is a collection of mutually independent standard Poisson processes indexed by a countable or countably infinite set I, and are independent of $\mathbf{Q}^n(0)$, $\{\mathbf{v}_i \in \mathbb{V}|i \in I\}$ for a separable Banach space $\mathbb{V}$ with norm $|\cdot|$ is a collection of vectors such that

$$\sum_{i \in I} |\mathbf{v}_i| < \infty, \tag{25}$$

and $\{\alpha_t^n(\cdot, i)|t \geq 0, i \in I\}$ is a collection of real-valued non-negative Lipschitz function on $\mathbb{V}$ that jointly satisfy

$$\|\alpha_t^n(\cdot, i)\| \leq n\beta_t \gamma^{(i)} \tag{26}$$

for some locally integrable function β_t and $\gamma^{(i)} \in \mathbb{R}$, $i \in I$. Note that $\|\cdot\|$ is the **Lipschitz norm** for real-valued functions on $\mathbb{V}$, i.e.,

$$\|f\| := \sup_{x,y \in \mathbb{V}, x \neq y} \frac{|f(x) - f(y)|}{|x - y|} \vee |f(0)|.$$

THEOREM 3. ***Theorem 2.2 in Mandelbaum et al. (1998)*** *Assume that (25) and (26) hold. Moreover, assume that*

$$\lim_{n \to \infty} \sum_{i \in I} \int_0^t \left\| \frac{\alpha_s^n(\cdot, i)}{n} - \alpha_s(\cdot, i) \right\| ds = 0, \tag{27}$$

for all $t \geq 0$, where $\{\alpha_s(\cdot, i)|t \geq 0, i \in I\}$ is a collection of Lipschitz function. If $\{\mathbf{Q}^n(0)|n > 0\}$ is any family of random initial state vectors in $\mathbb{V}$, then

$$\lim_{n \to \infty} \frac{\mathbf{Q}^n(0)}{n} = \mathbf{Q}(0) \text{ a.s. implies } \lim_{n \to \infty} \frac{\mathbf{Q}^n(t)}{n} = \mathbf{Q}(t) \text{ a.s.},$$

where the convergence is uniform on compact sets in t, and $\mathbf{Q}$ is the unique deterministic process $\{\mathbf{Q}(t)|t \geq 0\}$ that solves the integral equation

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \int_0^t \sum_{i \in I} \alpha_s(\mathbf{Q}(s), i) \mathbf{v}_i ds, t \geq 0.$$

B.2. Proof of Theorem 1: No Information

We justify the fluid approximation by applying Theorem 3 to the scaled process under no information: $\{X^{N,n}(t)/n : t \geq 0\}$. Let $\mathbb{V} = \mathbb{R}^2$, $I = 1, \dots, 6$, $K_k = A_k, K_{k+2} = S_k, K_{k+4} = N_k$ for $k = 1, 2$, $\mathbf{v}_1 = (1, 0)', \mathbf{v}_2 = (0, 1)', \mathbf{v}_3 = \mathbf{v}_5 = (-1, 0)', \mathbf{v}_4 = \mathbf{v}_6 = (0, -1)'$. Then, (25) holds obviously. Moreover, for $x = (x_1, x_2) \in \mathbb{R}^2$, let

$$\alpha_t(x,1) = \lambda_1(t), \qquad \alpha_t^n(x,1) = \lambda_1^n(t);$$

$$\alpha_t(x,2) = \lambda_2(t), \qquad \alpha_t^n(x,2) = \lambda_2^n(t);$$

$$\alpha_t(x,3) = \mu_1(x_1 \wedge s), \qquad \alpha_t^n(x,3) = \mu_1(nx_1 \wedge s^n);$$

$$\alpha_t(x,4) = \mu_2\left((s - x_1)^+ \wedge x_2\right), \qquad \alpha_t^n(x,4) = \mu_2\left((s^n - nx_1)^+ \wedge nx_2\right);$$

$$\alpha_t(x,5) = \theta\left(\frac{\beta_1(x_1 + x_2 - s)^+}{s}\right)(x_1 - s)^+, \qquad \alpha_t^n(x,5) = \theta\left(\frac{\beta_1(nx_1 + nx_2 - s^n)^+}{s^n}\right)(nx_1 - s^n)^+;$$

$$\alpha_t(x,6) = \theta\left(\frac{\beta_2(x_1 + x_2 - s)^+}{s}\right)(x_2 - (s - x_1)^+)^+, \quad \alpha_t^n(x,6) = \theta\left(\frac{\beta_2(nx_1 + nx_2 - s^n)^+}{s^n}\right)(nx_2 - (s^n - nx_1)^+)^+.$$

Then, we need to verify the assumptions of 3, i.e., (26)–(27), and $\alpha_t(x,i)$ being Lipschitz for $i = 1, \dots, 6$.

First, we show that $\alpha_t(x,i)$ is Lipschitz, i.e., $\|\alpha_t(x,i)\| < \infty$, for $i = 1, \dots, 6$. Since $\alpha_t(x,1)$ and $\alpha_t(x,2)$ are independent of x, the proof is trivial. Recall that $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^2$, then

$$\|\alpha_t(x,3)\| = \sup_{x,y\in\mathbb{R}^2} \frac{|\mu_1(y_1 \wedge s) - \mu_1(x_1 \wedge s)|}{|y - x|} \leq \sup_{x,y\in\mathbb{R}^2} \frac{\mu_1|y_1 - x_1|}{|y - x|} \leq \sup_{x,y\in\mathbb{R}^2} \frac{\mu_1|y_1 - x_1|}{|y_1 - x_1|} = \mu_1.$$

$$\|\alpha_t(x,4)\| = \sup_{x,y\in\mathbb{R}^2} \frac{\mu_2|(s - y_1)^+ \wedge y_2 - (s - x_1)^+ \wedge x_2|}{|y - x|} =: \sup_{x,y\in\mathbb{R}^2} L_1(x,y).$$

1. If $x_1, y_1 > s$, $L_1(x,y) = 0$.
2. If $x_1 \leq s, y_1 > s$, then
$$L_1(x,y) = \frac{\mu_2|(s - x_1) \wedge x_2|}{|y - x|} \leq \frac{\mu_2|(y_1 - x_1) \wedge x_2|}{|y - x|} \leq \frac{\mu_2|(y_1 - x_1)|}{|y_1 - x_1|} = \mu_2.$$
3. If $x_1 > s, y_1 \leq s$, similar to the previous case, we can obtain that $L_1(x,y) \leq \mu_2$.
4. If $x_1, y_1 \leq s$, then
$$\begin{aligned} L_1(x,y) &= \frac{\mu_2|(s - y_1) \wedge y_2 - (s - x_1) \wedge x_2|}{|y - x|} \\ &= \begin{cases} \frac{\mu_2|(s-y_1)\wedge y_2 - x_2|}{|y-x|} \leq \frac{\mu_2|y_2 - x_2|}{|y_2 - x_2|} = \mu_2, & \text{if } x_1 + x_2 \leq s; \\ \frac{\mu_2|(s-y_1)\wedge y_2 - (s-x_1)|}{|y-x|} \leq \frac{\mu_2|(s-y_1)-(s-x_1)|}{|y_1 - x_1|} = \mu_2, & \text{if } x_1 + x_2 > s. \end{cases} \end{aligned}$$

Thus, $\|\alpha_t(x,4)\| \leq \mu_2 < \infty$.

$$\|\alpha_t(x,5)\| = \sup_{x,y\in\mathbb{R}^2} \frac{\left|\theta\left(\frac{\beta_1(y_1+y_2-s)^+}{s}\right)(y_1 - s)^+ - \theta\left(\frac{\beta_1(x_1+x_2-s)^+}{s}\right)(x_1 - s)^+\right|}{|y - x|} =: \sup_{x,y\in\mathbb{R}^2} L_2(x,y).$$

Note that, $\theta'(x)x \leq \theta(x)$, for $x > 0$, since θ is increasing and concave. Therefore, $\theta'(x)x \leq M$. For $x, y \in \mathbb{R}^2$, without loss of generality, assume that $x_1 + x_2 \leq y_1 + y_2$, then

1. If $x_1, y_1 > s$,
$$\begin{aligned} L_2(x,y) &= \frac{\left|\theta\left(\beta_1\frac{y_1+y_2-s}{s}\right)(y_1 - s) - \theta\left(\beta_1\frac{x_1+x_2-s}{s}\right)(x_1 - s)\right|}{|y - x|} \\ &\leq \frac{\left|\theta\left(\beta_1\frac{y_1+y_2-s}{s}\right)(y_1 - x_1)\right|}{|y - x|} + \frac{\left|\theta\left(\beta_1\frac{y_1+y_2-s}{s}\right) - \theta\left(\beta_1\frac{x_1+x_2-s}{s}\right)\right|(x_1 - s)}{|y - x|} \\ &\leq \frac{|M(y_1 - x_1)|}{|y_1 - x_1|} + \frac{\left|\theta'\left(\beta_1\frac{x_1+x_2-s}{s}\right)\frac{y_1+y_2-x_1-x_2}{s/\beta_1}\right|(x_1 + x_2 - s)}{|y - x|} \\ &\leq M + M\frac{|y_1 - x_1| + |y_2 - x_2|}{|y - x|} \leq M + \sqrt{2}M. \end{aligned}$$

2. If $x_1 \le s, y_1 > s$,
$$L_2(x,y) = \frac{\left|\theta\left(\beta_1 \frac{y_1+y_2-s}{s}\right)(y_1-s)\right|}{|y-x|} \le \frac{M(y_1-x_1)}{|y-x|} \le M.$$
3. If $x_1 > s, y_1 \le s$, then similarly we can obtain that $L_2(x,y) \le M$.
4. If $x_1, y_1 \le s$, then $L_2(x,y) = 0$.

Therefore, $\|\alpha_t(x,5)\| \le (1+\sqrt{2})M < \infty$.
$$\begin{aligned}\|\alpha_t(x,6)\| &= \sup_{x,y\in\mathbb{R}^2} \frac{\left|\theta\left(\frac{\beta_2(y_1+y_2-s)^+}{s}\right)(y_2-(s-y_1)^+)^+ - \theta\left(\frac{\beta_2(x_1+x_2-s)^+}{s}\right)(x_2-(s-x_1)^+)^+\right|}{|y-x|}\\ &=: \sup_{x,y\in\mathbb{R}^2} L_3(x,y).\end{aligned}$$

For $x, y \in \mathbb{R}^2$, without loss of generality, assume that $x_1+x_2 \le y_1+y_2$, then

1. If $x_1, y_1 > s$,
$$\begin{aligned} L_3(x,y) &= \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)y_2 - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)x_2\right|}{|y-x|}\\ &\le \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_2-x_2)\right|}{|y-x|} + \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)-\theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)\right|x_2}{|y-x|}\\ &\le M + \frac{\left|\theta'\left(\beta_2\frac{x_1+x_2-s}{s}\right)\frac{y_1+y_2-x_1-x_2}{s/\beta_2}\right|(x_1+x_2-s)}{|y-x|}\\ &\le M + M\frac{|y_1-x_1|+|y_2-x_2|}{|y-x|} \le M + \sqrt{2}M.\end{aligned}$$
2. If $x_1 \le s, y_1 > s$,
   (a) If $x_1+x_2 \le s$, then
$$L_3(x,y) = \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)y_2\right|}{|y-x|} \le \frac{M|y_1+y_2-x_1-x_2|}{|y-x|} \le \sqrt{2}M.$$
   (b) If $x_1+x_2 > s$, then
$$\begin{aligned} L_3(x,y) &= \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)y_2 - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)(x_1+x_2-s)\right|}{|y-x|}\\ &\le \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_1+y_2-s) - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)(x_1+x_2-s)\right|}{|y-x|}\\ &\le \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_1+y_2-x_1-x_2)\right|}{|y-x|} + \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)-\theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)\right|(x_1+x_2-s)}{|y-x|}\\ &\le 2M\frac{|y_1-x_1|+|y_2-x_2|}{|y-x|} \le 2\sqrt{2}M.\end{aligned}$$
3. If $x_1 > s, y_1 \le s$, then $y_1+y_2 > s$, and
$$\begin{aligned} L_3(x,y) &= \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_1+y_2-s) - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)x_2\right|}{|y-x|}\\ &\le \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)y_2 - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)x_2\right|}{|y-x|} \le \sqrt{2}M.\end{aligned}$$
4. If $x_1, y_1 \le s$,
   (a) If $x_1+x_2 > s$, then
$$L_3(x,y) = \frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_1+y_2-s) - \theta\left(\beta_2\frac{x_1+x_2-s}{s}\right)(x_1+x_2-s)\right|}{|y-x|} \le 2\sqrt{2}M.$$

(b) If $x_1+x_2\le s, y_1+y_2>s$, then

$$L_3(x,y)=\frac{\left|\theta\left(\beta_2\frac{y_1+y_2-s}{s}\right)(y_1+y_2-s)\right|}{|y-x|}\le M\frac{|y_1+y_2-x_1-x_2|}{|y-x|}\le\sqrt{2}M.$$

(c) If $y_1+y_2\le s$, then $L_3(x,y)=0$.

Thus, $\|\alpha_t(x,6)\|\le 2\sqrt{2}M<\infty$.

Using a similar analysis, we can obtain that $\frac{\alpha_t^n(x,i)}{n}$ is Lipschitz (i.e., $\|\alpha_t^n(x,i)\|\le nM^*$ for some constant $M^*<\infty$) for $i=1,\dots,6$.

Lastly, we prove equation (27). For $i=1,2$, since $\lambda_i^n(t)/n\to\lambda_i(t)$ is bounded, by the dominated convergence theorem we have

$$\lim_{n\to\infty}\int_0^t\left\|\frac{\alpha_u^n(x,i)}{n}-\alpha(x,i)\right\|du=\lim_{n\to\infty}\int_0^t\left\|\frac{\lambda_i^n(t)}{n}-\lambda_i(t)\right\|du=0.$$

For $i=4,\dots,6$, $\alpha_t^n(\cdot,i)$ are independent of t, thus it is sufficient to prove that $\lim_{n\to\infty}\left\|\frac{\alpha_t^n(x,i)}{n}-\alpha(x,i)\right\|=0$. Since $s^n/n\to s$, we have $\lim_{n\to\infty}\left\|\frac{s^n}{n}-s\right\|=0$. Moreover, $\alpha_t^n(x,i)$ and $\alpha(x,i)$ are continuous functions of x with $\lim_{n\to\infty}\frac{\alpha_t^n(x,i)}{n}=\alpha(x,i)$ for $i=4,\dots,6$. Therefore, $\lim_{n\to\infty}\left\|\frac{\alpha_t^n(x,i)}{n}-\alpha(x,i)\right\|=0$, for $i=4,\dots,6$.

Apply Theorem 3 we can obtain the desired results.

B.3. Proof of Theorem 1: Full Information

For full information system, the proof is similar to the no information case except that now

$$\alpha_t(x,5)=\int_0^{(x_1-s)^+}\theta\left(\frac{u}{s}\right)du,\qquad \alpha_t^n(x,5)=\sum_{i=1}^{\lfloor nx_1-s^n\rfloor^+}\theta\left(\frac{i}{s^n}\right);$$
$$\alpha_t(x,6)=\int_{(x_1-s)^+}^{(x_1+x_2-s)^+}\theta\left(\frac{u}{s}\right)du,\quad \alpha_t^n(x,6)=\sum_{i=\lfloor nx_1-s^n\rfloor^++1}^{\lfloor nx_1+nx_2-s^n\rfloor^+}\theta\left(\frac{i}{s^n}\right),$$

where $\lfloor x\rfloor$ is the floor of x. Therefore, to apply Theorem 3, we just need to verify (26)–(27), and $\alpha_t(x,i)$ being Lipschitz for $i=5,6$.

First, we prove that $\alpha_t(x,i)$ is Lipschitz for $i=5,6$. Note that,

$$\begin{aligned}\|\alpha_t(x,5)\|&=\sup_{x,y\in\mathbb{R}^2}\frac{\left|\int_0^{(y_1-s)^+}\theta\left(\frac{u}{s}\right)du-\int_0^{(x_1-s)^+}\theta\left(\frac{u}{s}\right)du\right|}{|y-x|}\\&=\sup_{x,y\in\mathbb{R}^2}\frac{\left|\int_{(x_1-s)^+}^{(y_1-s)^+}\theta\left(\frac{u}{s}\right)du\right|}{|y-x|}\le\sup_{x,y\in\mathbb{R}^2}\frac{M|y_1-x_1|}{|y-x|}\le M<\infty.\end{aligned}$$

$$\begin{aligned}\|\alpha_t(x,6)\|&=\sup_{x,y\in\mathbb{R}^2}\frac{\left|\int_{(y_1-s)^+}^{(y_1+y_2-s)^+}\theta\left(\frac{u}{s}\right)du-\int_{(x_1-s)^+}^{(x_1+x_2-s)^+}\theta\left(\frac{u}{s}\right)du\right|}{|y-x|}\\&=\sup_{x,y\in\mathbb{R}^2}\frac{\left|\int_{(x_1+x_2-s)^+}^{(y_1+y_2-s)^+}\theta\left(\frac{u}{s}\right)du-\int_{(x_1-s)^+}^{(y_1-s)^+}\theta\left(\frac{u}{s}\right)du\right|}{|y-x|}\\&\le\sup_{x,y\in\mathbb{R}^2}\frac{M|y_1+y_2-x_1-x_2|}{|y-x|}+\frac{M|y_1-x_1|}{|y-x|}\le(1+\sqrt{2})M<\infty.\end{aligned}$$

Next, we show that (26) hold, i.e., $\|\alpha_t^n(x,i)\|$ is Lipschitz, for $i=5,6$. Note that for $a\ge 0$, $na-s^n-1\le\lfloor na-s^n\rfloor\le na-s^n$, which implies that $\frac{1}{n}\lfloor na-s^n\rfloor\to a-s$ as $n\to\infty$. Thus, for $x,y\in\mathbb{R}$,

$$\frac{\lfloor ny-s^n\rfloor-\lfloor nx-s^n\rfloor}{n(y-x)}\to 1\text{ as }n\to\infty.$$

Then there exists a constant n_1 such that for $n \geq n_1$,

$$\frac{\lfloor ny - s^n \rfloor - \lfloor nx - s^n \rfloor}{n(y-x)} \leq 2.$$

Similarly, we can find a constant n_2 such that for $n \geq n_2$,

$$\frac{\lfloor ny - nx \rfloor}{n(y-x)} \leq 2.$$

For $y > x \geq 0$, $n \geq n^* := \max\{n_1, n_2\}$,

$$\left(\lfloor ny - s^n \rfloor^+ - \lfloor nx - s^n \rfloor^+\right)^+ = \begin{cases} \lfloor ny - s^n \rfloor - \lfloor nx - s^n \rfloor \leq 2n(y-x) & \text{if } y > x > \frac{s^n}{n}, \\ \lfloor ny - s^n \rfloor \leq \lfloor ny - nx \rfloor \leq 2n(y-x) & \text{if } y > \frac{s^n}{n} \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

That is, for $x, y \in \mathbb{R}, n \geq n^*$, we have

$$\left(\lfloor ny - s^n \rfloor^+ - \lfloor nx - s^n \rfloor^+\right)^+ \leq 2n|y-x|.$$

Therefore,

$$\begin{aligned} \|\alpha_t^n(x,5)\| &= \sup_{x,y\in\mathbb{R}^2} \frac{\left|\sum_{i=1}^{\lfloor ny_1 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right) - \sum_{i=1}^{\lfloor nx_1 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right)\right|}{|y-x|} \\ &\leq \sup_{x,y\in\mathbb{R}^2} \frac{M\left(\lfloor ny_1 - s^n\rfloor^+ - \lfloor nx_1 - s^n\rfloor^+\right)^+}{|y-x|} \leq \sup_{x,y\in\mathbb{R}^2} \frac{nM|y_1 - x_1|}{|y-x|} \leq nM. \end{aligned}$$

$$\begin{aligned} \|\alpha_t^n(x,6)\| &= \sup_{x,y\in\mathbb{R}^2} \frac{\left|\sum_{i=\lfloor ny_1 - s^n\rfloor^+ + 1}^{\lfloor ny_1 + ny_2 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right) - \sum_{i=\lfloor nx_1 - s^n\rfloor^+ + 1}^{\lfloor nx_1 + nx_2 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right)\right|}{|y-x|} \\ &= \sup_{x,y\in\mathbb{R}^2} \frac{\left|\sum_{i=\lfloor nx_1 + nx_2 - s^n\rfloor^+ + 1}^{\lfloor ny_1 + ny_2 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right) - \sum_{i=\lfloor nx_1 - s^n\rfloor^+ + 1}^{\lfloor ny_1 - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right)\right|}{|y-x|} \\ &\leq \sup_{x,y\in\mathbb{R}^2} \frac{M\left(\lfloor ny_1 + ny_2 - s^n\rfloor^+ - \lfloor nx_1 + nx_2 - s^n\rfloor^+\right)^+}{|y-x|} + \frac{M\left(\lfloor ny_1 - s^n\rfloor^+ - \lfloor nx_1 - s^n\rfloor^+\right)^+}{|y-x|} \\ &\leq \sup_{x,y\in\mathbb{R}^2} \frac{nM|y_1 + y_2 - x_1 - x_2|}{|y-x|} + \frac{nM|y_1 - x_1|}{|y-x|} \leq n(1+\sqrt{2})M. \end{aligned}$$

Lastly, we show (27), i.e., the convergence of $\frac{\alpha_t^n(x,i)}{n}$ to $\alpha_t(x,i)$, for $i = 5, 6$. To do so, we first prove the following useful equation:

$$\frac{1}{n}\sum_{i=1}^{\lfloor nx - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right) \to \int_0^{(x-s)^+} \theta\left(\frac{u}{s}\right) du, \forall x \in \mathbb{R}. \tag{28}$$

Indeed, since $\frac{s^n}{n} \leq s$, we have $\lfloor nx - s^n\rfloor^+ \geq \lfloor s^n\left(\frac{x-s}{s}\right)\rfloor^+$ and

$$\frac{1}{n}\sum_{i=1}^{\lfloor nx - s^n\rfloor^+} \theta\left(\frac{i}{s^n}\right) = \frac{s^n}{n}\sum_{x\in\mathcal{P}_n} \frac{1}{s^n}\theta(x) + \frac{1}{n}\sum_{x\in\mathcal{Q}_n} \theta(x),$$

where

$$\begin{aligned} \mathcal{P}_n &= \{\frac{i}{s^n} | i \in \mathbb{Z}, 0 \leq i < s^n\left(\frac{x-s}{s}\right)^+\} = \{\frac{i}{s^n} | i \in \mathbb{Z}, 0 \leq i \leq \left\lfloor s^n\left(\frac{x-s}{s}\right)\right\rfloor^+\}, \\ \mathcal{Q}_n &= \{\frac{i}{s^n} | i \in \mathbb{Z}, s^n\left(\frac{x-s}{s}\right)^+ \leq i \leq \lfloor nx - s^n\rfloor^+\} = \{\frac{i}{s^n} | i \in \mathbb{Z}, \left\lceil s^n\left(\frac{x-s}{s}\right)\right\rceil^+ \leq i \leq \lfloor nx - s^n\rfloor^+\}. \end{aligned}$$

Note that,

$$\lim_{n\to\infty}\frac{|\mathcal{Q}_n|}{n}=\lim_{n\to\infty}\frac{\lfloor nx-s^n\rfloor^+-s^n\left(\frac{x-s}{s}\right)^+}{n}\le\lim_{n\to\infty}\frac{(nx-s^n)^+-s^n\left(\frac{x-s}{s}\right)^+}{n}$$
$$=\lim_{n\to\infty}\frac{s^n}{n}\left(\frac{x}{s^n/n}-1\right)^+-\lim_{n\to\infty}\frac{s^n}{n}\left(\frac{x-s}{s}\right)^+=0.$$

Therefore,

$$\frac{1}{n}\sum_{x\in\mathcal{Q}_n}\theta(x)\le\frac{M|\mathcal{Q}_n|}{n}\to 0.$$

Moreover, by the convergence of the Riemann sum, we can obtain that:

$$\frac{s^n}{n}\sum_{x\in\mathcal{P}_n}\frac{1}{s^n}\theta(x)+\theta\left(\left(\frac{x-s}{s}\right)^+\right)\left(\left(\frac{x-s}{s}\right)^+-\left(\frac{1}{s^n}\left\lfloor s^n\left(\frac{x-s}{s}\right)\right\rfloor^+\right)\right)\to s\int_0^{\frac{(x-s)^+}{s}}\theta(u)du. \tag{29}$$

Noting that the second term on the left of (29) converges to 0 as $n\to\infty$. Thus,

$$\frac{1}{n}\sum_{i=1}^{\lfloor nx-s^n\rfloor^+}\theta\left(\frac{i}{s^n}\right)\to s\int_0^{\frac{(x-s)^+}{s}}\theta(u)du=\int_0^{(x-s)^+}\theta\left(\frac{u}{s}\right)du.$$

By equation (28), we can obtain that $\frac{\alpha_t^n(x,5)}{n}\to\alpha_t(x,5)$, and $\frac{\alpha_t^n(x,5)}{n}+\frac{\alpha_t^n(x,6)}{n}\to\alpha_t(x,5)+\alpha_t(x,6)$, which further implies that $\frac{\alpha_t^n(x,6)}{n}\to\alpha_t(x,6)$. The proof is complete. ☐

Appendix C: Stability Theorems and Proofs for Section 5

In this appendix, we present the proofs of results in Section 5. Specifically, we introduce the concepts and theorems we use for the proofs of Proposition 1 and Theorem 2 in Appendices C.1 and C.3; and provide the proofs of these two propositions in Appendices C.2 and C.4, respectively.

C.1. Pre-requisites for the Proof of Proposition 1

In this section, we introduce concepts and theorems we use for the existence proof of periodic equilibrium in Proposition 1. We first obtain the definition and property of the Poincaré map.

Definition 4. Consider a single differential equation $\dot{x}=f(t,x)$ and assume that $f(t,x)$ is periodic in t with period T, for $x\in\mathbb{R}$. The **Poincaré map** associated with $\dot{x}=f(t,x)$ is the map $\phi(x_0)=x_1$, where $x(t)$ is the solution of the ODE with $x(0)=x_0, x_1=x(T)$.

The Poincaré map is monotone as shown in the following proposition.

Proposition C.1. Let $\phi:J\to\mathbb{R}$ be the Poincaré map for $\dot{x}=f(t,x)$, where J is an interval. Then, for $a,b\in J$, we have $a<b\Rightarrow\phi(a)<\phi(b)$.

The next theorem provides conditions for the existence of fixed point of a function.

Theorem 4 **(One-dimensional Brouwer fixed-point theorem)**. *Every continuous function from a closed interval into itself has a fixed point.*

The following theorem gives a set of conditions under which an initial value problem has a unique solution.

Theorem 5 **(Picard's existence theorem)**. *Consider the initial value problem $\dot{x}=f(t,x)$, $x(t_0)=x_0$. Suppose $f(t,x)$ is uniformly Lipschitz continuous in x and continuous in t, then for some value $\epsilon>0$, there exists a unique solution x to the initial value problem on the interval $[t_0-\epsilon,t_0+\epsilon]$.*

C.2. Proof of Proposition 1

We use notation $\uparrow$ for increasing and convergence, i.e., $s^n/n \uparrow s$. We start from the no information case, i.e., $I = N$. We begin with a single-class version of the original theorem. That is, we show that if there exist a solution $x^N(t) \in \mathbb{R}$ to the following one-dimensional ordinary differential equation such that $\limsup_{t\to\infty} x^N(t) < \infty$, then there exist a periodic solution with the same period as $\lambda(t)$:

$$\dot{x}(t) = \lambda(t) - \mu(x(t) \wedge s) - \theta\left(\frac{\beta(x(t)-s)^+}{s}\right)(x(t)-s)^+ = f_N(t,x), \tag{30}$$

where $\lambda(t+d) = \lambda(t)$ for some $d > 0$. Note that, $f_N(t,x)$ is periodic in t with period d, i.e., for a given x, $f_N(t,x) = f_N(t+d,x)$.

First, we show that any solution to (30) starting from a finite initial condition is bounded. Note that, when $x(t) > s$, $f_N(t,x) = \lambda(t) - \mu s - \theta\left(\beta\frac{x(t)-s}{s}\right)(x(t)-s)$. Since $\lambda(t)$ is bounded and $\theta\left(\beta\frac{x-s}{s}\right)(x-s) \uparrow \infty$ as $x \to \infty$, then there exist a $\hat{x}$ such that $\theta\left(\beta\frac{\hat{x}-s}{s}\right)(\hat{x}-s) = \max_{t\geq 0}\lambda(t) - \mu s$, and $f_N(t,x) \leq 0$ for $\forall x \geq \hat{x}$. Therefore, $x(t) \leq \max\{\hat{x}, s\} < \infty$.

Next, let $x(t,\xi)$ be the solution to (30) with $x(0) = \xi \geq 0$. Then, (30) has a periodic solution with period d if, for every t, $x(t+d,\xi) = x(t,\xi)$. Note that $x(t+d,\xi) = x(t, x(d,\xi))$. Thus, it suffices to show that $\xi = x(d,\xi)$. We define the Poincaré map associated with the ODE as follows:

$$\phi(\xi) = x(d,\xi).$$

Then, showing that $\xi = x(d,\xi)$ amounts to showing that ξ is a fixed point of $\phi(\cdot)$. First, note that $\phi(\cdot)$ is a continuous function. We can obtain this continuous dependence on the initial conditions by applying Theorem 6.3.1 in Lebovitz (1999) since $f_N(t,x)$ is Lipschitz.

Thus, by Theorem 4, to show that $\phi(\cdot)$ has a fixed point, it suffices to show that there exists a finite closed interval $[\xi_1, \xi_2]$ such that $\phi([\xi_1,\xi_2]) \subset (\xi_1,\xi_2)$.

Define

$$\xi_1 := \inf\{x(kd), k \in \mathbb{N}\}$$

and

$$\xi_2 := \sup\{x(kd), k \in \mathbb{N}\}.$$

Then, by our previous analysis, we have that $\xi_1, \xi_2 < \infty$. If $\xi_1 = \xi_2$, then $x(t)$ is a constant, and so it is itself a periodic solution. If $\xi_1 < \xi_2$, then for $\xi \in (\xi_1,\xi_2)$, there exist k_1, k_2 such that

$$x(k_1 d) < \xi < x(k_2 d).$$

Thus, by the monotonicity of the Poincaré map, i.e., Proposition C.1:

$$\phi(\xi) = x(d,\xi) < x(d, x(k_2 d)) = x(d + k_2 d, x(0)) = x((k_2+1)d) \leq \xi_2.$$

We can also show that $\phi(\xi) \geq \xi_1$ as follows:

$$\phi(\xi) = x(d,\xi) > x(d, x(k_1 d)) = x(d + k_1 d, x(0)) = x((k_1+1)d) \geq \xi_1.$$

Thus, $\xi_1 \le \phi(\xi) \le \xi_2$. By the continuity of $\phi(\cdot)$, it follows that

$$\phi([\xi_1, \xi_2]) \subset (\xi_1, \xi_2),$$

which completes the proof. That is, there exist a periodic solution with period d to (30).

We note that $f_N(t,x)$ is uniformly (in t) Lipschitz continuous in x, so that the periodic solution to (30), which we refer to as $\tilde{x}(t)$, must be unique by Theorem 5, i.e., the Picard's existence theorem.

Now, we go back to the two-dimensional system of ordinary differential equations (7) and (8). For any solution $(x_1^N(t),\, x_2^N(t))$ of (7) and (8) such that $x_1^N(0)+x_2^N(0)=\tilde{x}_0(0)$, let $\lambda(t)=\lambda_1(t)+\lambda_2(t)$, then $x_0^N(t)=x_1^N(t)+x_2^N(t)$ is a solution to (30). Moreover, since the unique solution to (30) with initial condition $\tilde{x}_0(0)$ is $\tilde{x}_0(t)$, we must have $x_0^N(t)=\tilde{x}_0(t)$. That is, any solution $x(t)$ of (7) and (8) with initial conditions sum up to $\tilde{x}_0(0)$ must satisfy $x_1(t)+x_2(t)=\tilde{x}_0(t)$.

Now, consider (7) with $\tilde{x}_0(t)$ plugged in, i.e.,

$$\dot{x}_1(t) = \lambda_1(t) - \mu(x_1(t) \wedge s) - \theta \left(\frac{\beta(\tilde{x}_0(t) - s)^+}{s} \right) (x_1(t) - s)^+. \tag{31}$$

Recall that $\lambda(t)$ is periodic with period d, thus $\tilde{x}_0(t)$ is also periodic with period d, regard $\tilde{x}_0(t)$ as given and using a similar analysis to (31) as we find $\tilde{x}_0(t)$ in the first step, it must be that there exists a solution $\tilde{x}_1(t)$ to (31) which is periodic with period d. Let $\tilde{x}_2(t)$ be a solution to (7) and (8) with initial condition $\tilde{x}_2(0)=\tilde{x}_0(0)-\tilde{x}_1(0)$, then we must have $\tilde{x}_2(t)=\tilde{x}_0(t)-\tilde{x}_1(t)$. Since both $\tilde{x}_0(t)$ and $\tilde{x}_1(t)$ are periodic in d, $\tilde{x}_2(t)$ is also periodic in d, and the result follows.

For the full information case, by (7), (9), and the one-dimensional proof above, we can obtain the unique periodic solution with period d_1. Denote the solution as $\tilde{x}_1^I$, and plug into (8), for $I=F$. Then, the result follows from a similar argument as the last part of the proof for $I=N$. $\square$

C.3. Stability Theorem for Time-Varying System

THEOREM 6. ***Theorem 4.3 in Khalil (2014)*** *Let $y=(0,0)$ be an equilibrium point for (11)–(12), and $V:[0,\infty)\times\mathbb{R}^2\to\mathbb{R}$ be a continuous differentiable function such that*

$$W_1(y) \le V(t,y) \le W_2(y), \tag{32}$$

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial y} g(t,y) \le -W_3(y), \ \textit{for } \forall t \ge 0, \forall y \in \mathbb{R}^2, \tag{33}$$

where $W_i(y)$ are continuous positive definite functions on $\mathbb{R}^2$ for $i=1,2,3$. If $W_1(y)$ is radially unbounded, then $y=(0,0)$ is globally uniformly asymptotically stable.

Note that, a function $f(y):\mathbb{R}^2\to\mathbb{R}$ is **radially unbounded** if $|y|\to\infty \Rightarrow f(y)\to\infty$.

C.4. Proof of Theorem 2

We prove this theorem using Theorem 6. We show the proof for $I=N$, and the case for $I=F$ is similar to this case, so we omit it for brevity.

To apply Theorem 6, we need to find a continuous differentiable function $V(t,y)$ and continuous positive definite functions $W_i(y_1,y_2)$ for $i=1,2,3$, such that equations (32) and (33) are satisfied for $t\ge 0, \tilde{x}^N(t)\in\mathbb{R}_+^2, y\in\mathbb{R}^2$ such that $\tilde{x}^N + y \in \mathbb{R}_+^2$.

Let

$$V(t,y)=W_1(y)=W_2(y)=\frac{1}{2}y_1^2+\frac{1}{2}(y_1+y_2)^2.$$

Then, (32) always holds, and it is clear that $W_1(y)$ is radially unbounded. Let

$$W_3(y)=\min\{\theta(0),\mu\}y_1^2+\min\{\theta(0),\mu\}(y_1+y_2)^2.$$

The only thing left is to show that (33) holds, i.e.,

$$\dot V(y)=y_1\tilde g_1^N(y)+(y_1+y_2)(\tilde g_1^N(y)+\tilde g_2^N(y))\le -W_3(y).$$

By system of equations (7) and (8) we can obtain that

$$\begin{aligned}
\tilde g_1^N(y)&=\mu_1(\tilde x_1^N(t)\wedge s)-\mu_1((\tilde x_1^N(t)+y_1)\wedge s)\\
&\quad+\theta\left(\frac{\beta(\tilde x_1^N(t)+\tilde x_2^N(t)-s)^+}{s}\right)(\tilde x_1^N(t)-s)^+\\
&\quad-\theta\left(\frac{\beta(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s)^+}{s}\right)(\tilde x_1^N(t)+y_1-s)^+,\\
\tilde g_2^N(y)&=\mu_2((s-\tilde x_1^N(t))^+\wedge\tilde x_2^N(t))-\mu_2((s-\tilde x_1^N(t)-y_1)^+\wedge(\tilde x_2^N(t)+y_2))\\
&\quad+\theta\left(\frac{\beta(\tilde x_1^N(t)+\tilde x_2^N(t)-s)^+}{s}\right)(\tilde x_2^N(t)-(s-\tilde x_1^N(t))^+)^+\\
&\quad-\theta\left(\frac{\beta(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s)^+}{s}\right)(\tilde x_2^N(t)+y_2-(s-\tilde x_1^N(t)-y_1)^+)^+.
\end{aligned}$$

For $(y_1,y_2)\ne(0,0)$, we show (33) by cases as follows:

1. When $\tilde x_1^N(t)+\tilde x_2^N(t)\le s,\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2\le s,\tilde x_1^N(t)+y_1\le s$, then

$$\tilde g_1^N(y)=-\mu y_1,\quad \tilde g_2^N(y)=-\mu y_2.$$

Thus,

$$\dot V(y)=-\mu y_1^2-\mu(y_1+y_2)^2\le -W_3(y).$$

2. When $\tilde x_1^N(t)+\tilde x_2^N(t)\le s,\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2> s,\tilde x_1^N(t)+y_1\le s$, then $y_1+y_2>0$, and

$$\begin{aligned}
\tilde g_1^N(y)&=-\mu y_1,\\
\tilde g_2^N(y)&=\mu(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1-s)\\
&\quad-\theta\left(\beta\frac{\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s}{s}\right)(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s).
\end{aligned}$$

Then,

$$\begin{aligned}
&(y_1+y_2)(\tilde g_1^N(y)+\tilde g_2^N(y))\\
&=(y_1+y_2)(\mu(\tilde x_1^N(t)+\tilde x_2^N(t)-s)-\theta\left(\beta\frac{\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s}{s}\right)(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s))\\
&\le(y_1+y_2)(\mu(\tilde x_1^N(t)+\tilde x_2^N(t)-s)-\theta(0)(\tilde x_1^N(t)+\tilde x_2^N(t)+y_1+y_2-s))\\
&=(y_1+y_2)(\mu-\theta(0))(\tilde x_1^N(t)+\tilde x_2^N(t)-s)-\theta(0)(y_1+y_2)^2\\
&\le\begin{cases}-(y_1+y_2)^2(\mu-\theta(0))-\theta(0)(y_1+y_2)^2=-\mu(y_1+y_2)^2 & \text{if } \theta(0)\ge\mu\\ -\theta(0)(y_1+y_2)^2 & \text{if } \theta(0)<\mu\end{cases}\\
&=-\min\{\theta(0),\mu\}(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot V(y)\le -\mu y_1^2-\min\{\theta(0),\mu\}(y_1+y_2)^2\le -W_3(y).$$

3. When $\tilde{x}_1^N(t)+\tilde{x}_2^N(t)\le s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2>s, \tilde{x}_1^N(t)+y_1>s$, then $y_1>0$, and

$$\tilde{g}_1^N(y)=\mu(\tilde{x}_1^N(t)-s)-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+y_1-s),$$
$$\tilde{g}_2^N(y)=\mu\tilde{x}_2^N(t)-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_2^N(t)+y_2).$$

Note that, by a similar analysis as the previous cases, we can obtain that

$$\begin{aligned}
y_1\tilde{g}_1^N(y) &\le (s-\tilde{x}_1^N(t))(\mu-\theta(0))y_1-\theta(0)y_1^2\\
&\le -\min\{\theta(0),\mu\}y_1^2.\\
(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) &\le (y_1+y_2)(\mu-\theta(0))(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)-\theta(0)(y_1+y_2)^2\\
&\le -\min\{\theta(0),\mu\}(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot{V}(y)\le -\min\{\theta(0),\mu\}y_1^2-\min\{\theta(0),\mu\}(y_1+y_2)^2=-W_3(y).$$

4. When $\tilde{x}_1^N(t)+\tilde{x}_2^N(t)>s, \tilde{x}_1^N(t)\le s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2\le s, \tilde{x}_1^N(t)+y_1\le s$, then $y_1+y_2<0$, and

$$\begin{aligned}
\tilde{g}_1^N(y)&=-\mu y_1,\\
\tilde{g}_2^N(y)&=-\mu(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_2-s)+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s).
\end{aligned}$$

Then,

$$\begin{aligned}
(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) &\le -\mu(y_1+y_2)^2+(y_1+y_2)(\theta(0)-\mu)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\\
&\le -\min\{\theta(0),\mu\}(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot{V}(y)\le -\mu y_1^2-\min\{\theta(0),\mu\}(y_1+y_2)^2\le -W_3(y).$$

5. When $\tilde{x}_1^N(t)+\tilde{x}_2^N(t)>s, \tilde{x}_1^N(t)\le s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2>s, \tilde{x}_1^N(t)+y_1\le s$, then

$$\begin{aligned}
\tilde{g}_1^N(y)&=-\mu y_1,\\
\tilde{g}_2^N(y)&=\mu y_1+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\\
&\quad-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s).
\end{aligned}$$

Then,

$$\begin{aligned}
&(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y))\\
&\le (y_1+y_2)\left(\left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\right)-\theta(0)(y_1+y_2)^2\\
&\le -\theta(0)(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot{V}(y)\le -\mu y_1^2-\theta(0)(y_1+y_2)^2\le -W_3(y).$$

6. When $\tilde{x}_1^N(t)+\tilde{x}_2^N(t)>s, \tilde{x}_1^N(t)\le s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2>s, \tilde{x}_1^N(t)+y_1>s$, then $y_1>0$, and

$$\begin{aligned}
\tilde{g}_1^N(y) &= \mu(\tilde{x}_1^N(t)-s)-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+y_1-s),\\
\tilde{g}_2^N(y) &= \mu(s-\tilde{x}_1^N(t))+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\\
&\quad-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_2^N(t)+y_2)^+.
\end{aligned}$$

Then,

$$y_1\tilde{g}_1^N(y)\le(\mu-\theta(0))(\tilde{x}_1^N(t)-s)y_1-\theta(0)y_1^2\le-\min\{\mu,\theta(0)\}y_1^2.$$

$$\begin{aligned}
(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) &= (y_1+y_2)\Bigg(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\\
&\quad-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s)\Bigg)\\
&\le-\theta(0)(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot{V}(y)\le-\min\{\mu,\theta(0)\}y_1^2-\theta(0)(y_1+y_2)^2\le-W_3(y).$$

7. When $\tilde{x}_1^N(t)>s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2\le s, \tilde{x}_1^N(t)+y_1\le s$, then $y_1<0, y_1+y_2<0$, and

$$\begin{aligned}
\tilde{g}_1^N(y) &= \mu(s-\tilde{x}_1^N(t)-y_1)+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)-s),\\
\tilde{g}_2^N(y) &= -\mu(\tilde{x}_2^N(t)+y_2)+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)\tilde{x}_2^N(t).
\end{aligned}$$

Then,

$$\begin{aligned}
y_1\tilde{g}_1^N(y) &\le(\theta(0)-\mu)(\tilde{x}_1^N(t)-s)y_1-\mu y_1^2\le-\min\{\mu,\theta(0)\}y_1^2.\\
(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) &= -\mu(y_1+y_2)^2+(y_1+y_2)\left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)-\mu\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)\\
&\le-\min\{\mu,\theta(0)\}(y_1+y_2)^2.
\end{aligned}$$

Thus,

$$\dot{V}(y)\le-\min\{\mu,\theta(0)\}y_1^2-\min\{\mu,\theta(0)\}(y_1+y_2)^2=-W_3(y).$$

8. When $\tilde{x}_1^N(t)>s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2>s, \tilde{x}_1^N(t)+y_1\le s$, then $y_1<0$, and

$$\begin{aligned}
\tilde{g}_1^N(y) &= \mu(s-\tilde{x}_1^N(t)-y_1)+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)-s),\\
\tilde{g}_2^N(y) &= -\mu(s-\tilde{x}_1^N(t)-y_1)+\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)\tilde{x}_2^N(t)\\
&\quad-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s).
\end{aligned}$$

Then,

$$y_1\tilde{g}_1^N(y)\le(\theta(0)-\mu)(\tilde{x}_1^N(t)-s)y_1-\mu y_1^2\le-\min\{\mu,\theta(0)\}y_1^2.$$

$$(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) = (y_1+y_2)\Bigg(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)$$
$$-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s)\Bigg)$$
$$\leq -\theta(0)(y_1+y_2)^2.$$

Thus,

$$\dot{V}(y) \leq -\min\{\mu,\theta(0)\}y_1^2 - \theta(0)(y_1+y_2)^2 \leq -W_3(y).$$

9. When $\tilde{x}_1^N(t) > s, \tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2 > s, \tilde{x}_1^N(t)+y_1 > s$,

$$\tilde{g}_1^N(y) = \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)-s)$$
$$-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+y_1-s),$$
$$\tilde{g}_2^N(y) = \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)\tilde{x}_2^N(t) - \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_2^N(t)+y_2).$$

(a) If $y_1(y_1+y_2) \geq 0$,

$$y_1\tilde{g}_1^N(y) \leq \left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right) - \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\right)(\tilde{x}_1^N(t)-s)y_1 - \theta(0)y_1^2$$
$$\leq -\theta(0)y_1^2.$$

$$(y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y)) = (y_1+y_2)\Bigg(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)$$
$$-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s)\Bigg)$$
$$\leq -\theta(0)(y_1+y_2)^2.$$

Thus,

$$\dot{V}(y) \leq -\theta(0)y_1^2 - \theta(0)(y_1+y_2)^2 \leq -W_3(y).$$

(b) If $y_1(y_1+y_2) < 0$, then

$$\dot{V}(y) = y_1\tilde{g}_1^N(y) + (y_1+y_2)(\tilde{g}_1^N(y)+\tilde{g}_2^N(y))$$
$$= \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(y_1(\tilde{x}_1^N(t)-s) + (y_1+y_2)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s))$$
$$-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(y_1(\tilde{x}_1^N(t)+y_1-s) + (y_1+y_2)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s))$$
$$= \left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right) - \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\right)$$
$$\times (y_1(\tilde{x}_1^N(t)-s) + (y_1+y_2)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s))$$
$$-\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(y_1^2+(y_1+y_2)^2).$$

i. If $2y_1+y_2 \leq 0$, then

$$\dot{V}(y) \leq \left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right) - \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\right)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)(2y_1+y_2)$$
$$-\theta(0)(y_1^2+(y_1+y_2)^2)$$
$$\leq \theta(0)(y_1^2+(y_1+y_2)^2) \leq -W_3(y).$$

ii. If $2y_1 + y_2 > 0$, then

$$\begin{aligned}
\dot{V}(y) \le & \left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right) - \theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\right) \\
& \times (y_1(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s)+(y_1+y_2)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)) \\
& -\left(\theta\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)-\theta(0)\right)(y_1^2+(y_1+y_2)^2)-\theta(0)(y_1^2+(y_1+y_2)^2) \\
& \le \theta'\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(-\beta\frac{y_1+y_2}{s}) \\
& \times (y_1(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s)+(y_1+y_2)(\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s)) \\
& -\theta'\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(y_1^2+(y_1+y_2)^2)-W_3(y) \\
& =\theta'\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)(-y_1(y_1+y_2)-y_1^2-(y_1+y_2)^2) \\
& -\theta'\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)-s}{s}\right)(y_1+y_2)^2-W_3(y) \\
& \le -\theta'\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\left(\beta\frac{\tilde{x}_1^N(t)+\tilde{x}_2^N(t)+y_1+y_2-s}{s}\right)\left(3y_1^2+3y_1y_2+y_2^2\right)-W_3(y) \\
& \le -W_3(y).
\end{aligned}$$

Note that we used above the fact that:

A. When $y_1(y_1+y_2) < 0$, we have $\tilde{x}_2^N(t)+y_1+y_2 \ge 0$. In particular, if $y_1 > 0, y_1+y_2 < 0$, then $\tilde{x}_2^N(t)+y_1+y_2 \ge \tilde{x}_2^N(t)+y_2 \ge 0$; if $y_1 < 0, y_1+y_2 > 0$, then $\tilde{x}_2^N(t)+y_1+y_2 \ge \tilde{x}_2^N(t) \ge 0$.

B. For concave and and differentiable function θ and $a, b \in \mathbb{R}$, we have $\theta(b) \le \theta(a) + \theta'(a)(b-a)$.

C. $3y_1^2 + 3y_1y_2 + y_2^2 = 3(y_1 + \frac{1}{2}y_2)^2 + \frac{1}{4}y_2^2 \ge 0$.

□

Appendix D: Supplementary Results and Proofs for Section 6

In this appendix, we provide the supplementary results and proofs for Section 6. In particular, in Appendices D.1 and D.2, we present concepts and theorems that facilitate the proofs in this section; in Appendices D.3–D.5, we provide the supplementary results and proofs for Sections 6.3–6.5, respectively.

D.1. Comparison Theorems of Ordinary Differential Equations

The following lemmas are useful for the comparison of the trajectories of the fluid model $x^I(t)$ under different information levels.

Lemma 1. ***Proposition 6.4 in Bagagiolo (2012)*** *Let $f, g : A \to \mathbb{R}, A \subseteq \mathbb{R}^2$ open, be Lipschitz continuous in x, such that $f(t,x) \le g(t,x), \forall (t,x) \in A$. Then, if $y, z : I \to \mathbb{R}$ are, respectively, the solutions of the following two initial value problems:*

$$y'(t) = f(t,y), y(t_0) = y_0,$$

$$z'(t) = g(t,z), z(t_0) = z_0,$$

where $y_0 \le z_0$ and I is the common interval of existence, then $y(t) \le z(t)$ for $t \in I, t \ge t_0$.

Lemma 2. ***Strong Comparison Theorem in McNabb (1986)*** *Suppose $y(t)$, $z(t)$ are continuous on $[a,b]$ and differentiable on $(a,b]$, f is a continuous mapping from $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ that satisfies a Lipschitz condition, and*

$$y(a)>z(a), \frac{dy}{dt}-f(t,y)\geq\frac{dz}{dt}-f(t,z) \text{ on } (a,b].$$

Then $y>z$ on $[a,b]$.

In the proofs we will rely on the following application of Lemma 2: Letting $f(t,y)=dy/dt, g(t,z)=dz/dt$, $a=t_0$, and $b=t_0+d$, then Lemma 2 implies that:

$$y(t_0)>z(t_0) \text{ and } f(t,z)\geq g(t,z)\Rightarrow y(t)>z(t) \quad \forall t\in[t_0,t_0+d]. \tag{34}$$

D.2. Continuity of $x^I(t)$ on parameters

Recall that $f_k^I(t,\tilde{x}^I(t))$ is the net flow rate of class k customers under information level I at time t in equilibrium, which depends on the arrival rates $\lambda_k(t)$. We obtain the continuity of $\tilde{x}^I(t)$ with respect to $\lambda_k(t)$ using the following theorem, and the continuity result is given in Corollary 1.

Before we proceed, we introduce the class of p**-norms** $||\cdot||_p$ for a vector $x\in\mathbb{R}$ in below:

$$||x||_p:=(|x_1|^p+\ldots+|x_n|^p)^{1/p}, \quad 1\leq p<\infty,$$

and

$$||x||_\infty:=\max_i|x_i|.$$

Theorem 7. *(Theorem 3.4 in Khalil (2002)) Let $f(t,x)$ be piecewise continuous in t and Lipschitz in x on $[t_0,t_1]\times W$ with a Lipschitz constant L, where $W\subset R^n$ is an open connected set. Let $y(t)$ and $z(t)$ be solutions of*

$$\dot{y}=f(t,y), \quad y(t_0)=y_0$$

and

$$\dot{z}=f(t,z)+g(t,z), \quad z(t_0)=z_0$$

such that $y(t),z(t)\in W$ for all $t\in[t_0,t_1]$. Suppose that

$$||g(t,x)||_p\leq\sigma, \quad \forall(t,x)\in[t_0,t_1]\times W$$

for some $\sigma>0$ and any p-norm. Then,

$$||y(t)-z(t)||_p\leq||y_0-z_0||_pe^{L(t-t_0)}+\frac{\sigma}{L}\left(e^{L(t-t_0)}-1\right), \quad \forall t\in[t_0,t_1].$$

Corollary 1. *For $t\geq 0$, $\tilde{x}^I(t)$ is continuous in $(\lambda_1(t),\lambda_2(t))$, for $I\in\{F,N\}$.*

Proof of Corollary 1. For simplicity, let $\lambda(t):=(\lambda_1(t),\lambda_2(t))$. Consider two systems under information level I starting from the same initial $\tilde{x}_0^I$ at $t=0$ with arrival rates $\lambda(t)$ and $\lambda'(t)$. Let $\tilde{x}^I(t)$ and $\tilde{x}'^I(t)$ be the corresponding number-in-system trajectories. We rewrite the net flow rate f_k^I as a function of $t,\tilde{x}^I(t)$, and $\lambda(t)$, for $k=1,2$, and denote $f^I:=(f_1^I,f_2^I)$. Then $\tilde{x}^I(t)$ and $\tilde{x}'^I(t)$ are the solutions of the following two initial value problems, respectively:

$$\dot{y}=f^I(t,y,\lambda(t)), \quad y(0)=\tilde{x}_0^I.$$

$$\dot{z} = f^I(t, z, \lambda'(t)), \quad z(0) = \tilde{x}_0^I.$$

In particular, $f^I(t,x,\lambda'(t)) - f^I(t,x,\lambda(t)) = \lambda'(t) - \lambda(t)$, for $\forall x \geq \mathbf{0}$. Recall that, f^I is Lipschitz in $\tilde{x}^I(t)$ with a Lipschitz constant $L \leq 2\sqrt{2}M$. Applying Theorem 7 with $f(t,y) = f^I(t,y,\lambda(t)), g(t,z) = \lambda'(t) - \lambda(t)$, we have:

$$\text{If } ||\lambda'(t) - \lambda(t)||_p \leq \sigma, \text{ then } ||\tilde{x}'^I(t) - \tilde{x}^I(t)||_p \leq \frac{\sigma}{L}\left(e^{Lt} - 1\right).$$

Thus, for any $\epsilon > 0$ and fixed t, L, there exists $\sigma := \frac{\epsilon L}{e^{Lt}-1}$ such that, $||\tilde{x}'^I(t) - \tilde{x}^I(t)||_p \leq \epsilon$ when $||\lambda'(t) - \lambda(t)||_p \leq \sigma$. Since L is finite, and we have shown the existence and stability of periodic equilibrium for $\tilde{x}^I(t)$ and $\tilde{x}'^I(t)$, the result follows. □

Lastly, we establish the continuity of $\tilde{x}^N(t,\boldsymbol{\beta})$ in $\boldsymbol{\beta}$ by rewriting Theorem 3.5 in Khalil (2002) as follows.

Proposition D.1. If $f(t,x,\boldsymbol{\beta})$ is continuous in $(t,x,\boldsymbol{\beta})$ and locally Lipschitz in x on $[t_0,t_1] \times R^n \times \{||\boldsymbol{\beta} - \boldsymbol{\beta}_0||_p \leq c\}$. Let $y(t,\boldsymbol{\beta}_0)$ be a solution of $\dot{x} = f(t,x,\boldsymbol{\beta}_0)$ with $y(t_0,\boldsymbol{\beta}_0) = y_0$. Then, given $\epsilon > 0$, there is $\delta > 0$ such that if

$$||z_0 - y_0||_p < \delta \text{ and } ||\boldsymbol{\beta} - \boldsymbol{\beta}_0||_p < \delta,$$

then there is a unique solution $z(t,\boldsymbol{\beta})$ of $\dot{x} = f(t,x,\boldsymbol{\beta})$ with $z(t_0,\boldsymbol{\beta}) = z_0$, and $z(t,\boldsymbol{\beta})$ satisfies

$$||z(t,\boldsymbol{\beta}) - y(t,\boldsymbol{\beta}_0)||_p < \epsilon, \quad \forall t \in [t_0,t_1].$$

D.3. Proofs for Section 6.3

For single-class non-stationary systems, (13), (15), and (16) reduce to:

$$\bar{A}^I = \frac{1}{d}\int_0^d \lambda(t)dt - \frac{\mu}{d}\int_0^d (\tilde{x}^I(t) \wedge s)dt, \text{ for } I \in \{F,N\}, \tag{35}$$

$$\bar{A}^N(\beta) = \frac{1}{d}\int_0^d \theta\left(\beta\frac{(\tilde{x}^N(t) - s)^+}{s}\right)(\tilde{x}^N(t) - s)^+ dt, \tag{36}$$

$$\bar{A}^F = \frac{1}{d}\int_0^d \int_0^{(\tilde{x}^F(t)-s)^+} \theta\left(u/s\right) du dt. \tag{37}$$

To facilitate the proof, we introduce the following useful lemma without a proof.

Lemma 3. *Let a_1, a_2, b_1, b_2 be any real numbers, then*

$$\min\{a_1,a_2\} - \min\{b_1,b_2\} \leq \max\{a_1 - b_1, a_2 - b_2\}.$$

Then, by (35) we have

$$\bar{A}^N(\beta) - \bar{A}^F \leq \frac{1}{d}\int_0^d \left(\tilde{x}^F(t) - \tilde{x}^N(t,\beta)\right)^+ dt, \tag{38}$$

$$\bar{A}^F - \bar{A}^N(\beta) \leq \frac{1}{d}\int_0^d \left(\tilde{x}^N(t,\beta) - \tilde{x}^F(t)\right)^+ dt. \tag{39}$$

We prove Propositions 4 (average number-in-system) and 5 (average system abandonment rates) by applying Lemma 3 and the following Lemma 4. Lemma 4 compares the equilibrium number-in-system processes.

Lemma 4. *For single-class systems with non-stationary periodic arrivals, the equilibrium number-in-system processes under no (N) and full (F) information compare as follows for all $t \geq 0$:*

1. *If $\bar{\rho} \leq 1$, then $\tilde{x}^N(t,\beta) = \tilde{x}^F(t) \leq s$.*

2. *If* $\underline{\rho} \geq 1$, *then* $s \leq \tilde{x}^N(t,\beta), s \leq \tilde{x}^F(t)$.

3. *If* $1 < \bar{\rho}$, *then* $\tilde{x}^N(t,\beta)$ *is continuously non-increasing in* β, *with strict decreasing if* $\max_{t\geq 0} \tilde{x}^N(t,0) > s$ *and:*

   (a) *If* $\beta \geq 0.5$, *then* $\tilde{x}^N(t,\beta) \leq \tilde{x}^F(t)$ $\forall t$;

   i. *If* $\max_{t\geq 0} \tilde{x}^N(t,\beta) > s$, *then the inequality is strict* $\forall t$,

   ii. *If* $\max_{t\geq 0} \tilde{x}^F(t) > s$, *then* $\max_{t\geq 0} \tilde{x}^N(t,\beta) > s$.

   (b) *If* $\beta = 0$, *then* $\tilde{x}^N(t,0) \geq \tilde{x}^F(t)$ $\forall t$;

   i. *If* $\max_{t\geq 0} \tilde{x}^F(t) > s$, *then the inequality is strict* $\forall t$,

   ii. *If* $\max_{t\geq 0} \tilde{x}^N(t,0) > s$, *then* $\max_{t\geq 0} \tilde{x}^F(t) > s$.

Proof of Lemma 4. To facilitate the proof, we first recall the net flow rate $f^I(t,x)$ for $I = N,F$ as follows:

$$f^N(t,x,\beta) = \lambda(t) - \mu(x(t)\wedge s) - \theta\left(\frac{\beta(x(t)-s)^+}{s}\right)(x(t)-s)^+,$$

$$f^F(t,x) = \lambda(t) - \mu(x(t)\wedge s) - \int_0^{(x(t)-s)^+} \theta(u/s)\,du.$$

Note that for $t,x \geq 0$, when $\beta \geq 0.5$, $f^N(t,x,\beta) \leq f^F(t,x)$; when $\beta = 0$, $f^N(t,x,\beta) \geq f^F(t,x)$. Also, by the balance equations (35)-(37) we have

$$0 = \int_0^d f^N(t,\tilde{x}^N,\beta)dt = \int_0^d f^F(t,\tilde{x}^F)dt. \tag{40}$$

1. If $\bar{\rho} \leq 1$, since the arrival rate $\lambda(t)$ is non-stationary, then $\lambda(t) \leq s\mu$, for $t \geq 0$, and $\lambda(t) < s\mu$ for some time interval(s) within each period. Therefore, we have $\frac{1}{d}\int_0^d \lambda(t)dt < s\mu$.

First, we show that $\tilde{x}^N(t,\beta) \leq s$. Note that, there must exist $t_1 \in [0,d)$ such that $\tilde{x}^N(t_1,\beta) < s$. If otherwise, i.e., $\tilde{x}^N(t,\beta) \geq s$ for $t \in [0,d)$, then (36) implies that

$$\bar{A}^N(\beta) = \frac{1}{d}\int_0^d \theta\left(\frac{\beta(\tilde{x}^N(t,\beta)-s)^+}{s}\right)(\tilde{x}^N(t,\beta)-s)^+dt \geq 0.$$

However, by (35) we have

$$\bar{A}^N(\beta) = \frac{1}{d}\int_0^d \lambda(t)dt - s\mu < 0,$$

which yields a contradiction. Therefore, we can always find such t_1 with $\tilde{x}^N(t_1,\beta) < s$. Moreover, if $\tilde{x}^N(t,\beta) < s$ and $\tilde{x}^N(t,\beta) \to s$, by (7) we have

$$\dot{\tilde{x}}(t) = \lambda(t) - \mu\tilde{x}(t) \to \lambda(t) - \mu s \leq 0.$$

That is, starting from t_1, $\tilde{x}^N(t,\beta)$ would never exceed s for $t \geq t_1$. Since $\tilde{x}^N(t,\beta)$ is a periodic equilibrium, we must have $\tilde{x}^N(t,\beta) \leq s$ for $t \geq 0$.

By a similar analysis, we can show that $\tilde{x}^F(t) \leq s$.

Once we obtain $\tilde{x}^N(t,\beta) \leq s, \tilde{x}^F(t) \leq s$, then by (36), we have $\bar{A}^I = 0$, for $I \in \{F,N\}$. By (7), $\tilde{x}^N(t,\beta)$ and $\tilde{x}^F(t)$ both uniquely solves the ODE

$$\dot{\tilde{x}}(t) = \lambda(t) - \mu\tilde{x}(t).$$

Hence, $\tilde{x}^N(t,\beta) = \tilde{x}^F(t) \leq s$.

2. If $\underline{\rho} \geq 1$, then $\lambda(t) \geq s\mu$, for $t \geq 0$, and $\lambda(t) > s\mu$ for some time interval(s) within each period. Therefore, we have $\frac{1}{d}\int_0^d \lambda(t)dt > s\mu$. By a similar analysis as we show $\tilde{x}^N(t,\beta) \leq s$ for $\bar{\rho} \leq 1$, we can obtain that $\tilde{x}^N(t,\beta) \geq s$, $\tilde{x}^F(t) \geq s$, and the inequality is strict for some time interval(s) within each period (i.e., $\max \tilde{x}^I(t) > s$), using (35)-(37).

3. If $1 < \bar{\rho}$, first we show that $\tilde{x}^N(t,\beta)$ is continuously non-increasing in β. The continuity part is given by Proposition D.1. We prove the monotonicity part by Lemma 1. Consider two systems under no information model, one with β^h and the other with β^l, where $\beta^h > \beta^l$. Then, $f^N(t,x,\beta^h) \leq f^N(t,x,\beta^l)$, for $t, x \geq 0$. Note that there must exist $t_2 \in [0,d)$ such that $\tilde{x}^N(t_2,\beta^h) \leq \tilde{x}^N(t_2,\beta^l)$. If otherwise, i.e., $\tilde{x}^N(t,\beta^h) > \tilde{x}^N(t,\beta^l)$ for $t \in [0,d)$, then $f^N(t,\tilde{x}^N(t,\beta^h),\beta^h) < f^N(t,\tilde{x}^N(t,\beta^l),\beta^l)$, for $t \in [0,d)$, which contradicts to (40). Let $y(t) = \tilde{x}^N(t,\beta^h)$, $z(t) = \tilde{x}^N(t,\beta^l)$, $f(t,y) = f^N(t,y,\beta^h)$, and $g(t,z) = f^N(t,z,\beta^l)$, then we have $\tilde{x}^N(t,\beta^h) \leq \tilde{x}^N(t,\beta^l)$ for $t \in [0,d)$ by Lemma 1. Due to the periodicity of $\tilde{x}^N$, we have $\tilde{x}^N(t,\beta^h) \leq \tilde{x}^N(t,\beta^l)$ for $t \geq 0$, i.e., $\tilde{x}^N(t,\beta)$ is non-increasing in β.

We show the strict monotonicity by Lemmas 3 and 2. To do so, we first show that $\tilde{x}^N(t,\beta^h) < \tilde{x}^N(t,\beta^l)$ for $t \geq 0$ when $\max_{t\geq 0} \tilde{x}^N(t,\beta^l) > s$. We need to show the existence of $t_2' \in [0,d)$ such that $\tilde{x}^N(t_2',\beta^h) < \tilde{x}^N(t_2',\beta^l)$. If otherwise, i.e., $\tilde{x}^N(t,\beta^h) \geq \tilde{x}^N(t,\beta^l)$ for $t \in [0,d)$, then since $\max_{t\geq 0} \tilde{x}^N(t,\beta^l) > s$, there must exists an interval $\eta_2 \subset [0,d)$ such that $\tilde{x}^N(t,\beta^h) \geq \tilde{x}^N(t,\beta^l) > s$ for $t \in \eta_2$. Then, $f^N(t,\tilde{x}^N(t,\beta^h),\beta^h) \leq f^N(t,\tilde{x}^N(t,\beta^l),\beta^l)$, for $t \in [0,d)$, and the inequality is strict for $t \in \eta_2$. This contradicts to $0 = \int_0^d f^N(t,\tilde{x}^N(t,\beta^h),\beta^h)dt = \int_0^d f^N(t,\tilde{x}^N(t,\beta^l),\beta^l)dt$ from (40). Thus, such t_2' exists and $\tilde{x}^N(t,\beta^h) < \tilde{x}^N(t,\beta^l)$ for all t when $\max_{t\geq 0} \tilde{x}^N(t,\beta^l) > s$ by Lemma 2.

Now we show that $\max_{t\geq 0} \tilde{x}^N(t,0) > s$ implies $\max_{t\geq 0} \tilde{x}^N(t,\beta^l) > s$, for any $\beta^l > 0$. Indeed, if $\max_{t\geq 0} \tilde{x}^N(t,0) > s$ and $\max_{t\geq 0} \tilde{x}^N(t,\beta^l) \leq s$, then by (36) we have $\bar{A}^N(0) > 0$ and $\bar{A}^N(\beta^l) = 0$. However, by Lemma 3 and (35) we have $\bar{A}^N(0) - \bar{A}^N(\beta^l) \leq \frac{1}{d}\int_0^d \left(\tilde{x}^N(t,\beta^l) - \tilde{x}^N(t,0)\right)^+ dt \leq 0$, which leads to a contradiction.

(a) If $\beta \geq 0.5$, recall that $f^N(t,x,\beta) \leq f^F(t,x)$, we show $\tilde{x}^N(t,\beta) \leq \tilde{x}^F(t)$ using Lemma 1. Let $y(t) = \tilde{x}^N(t,\beta)$, $z(t) = \tilde{x}^F(t)$, $f(t,y) = f^N(t,y,\beta)$, and $g(t,z) = f^F(t,z)$. Recall that $f^N(t,x,\beta) \leq f^F(t,x)$, to apply Lemma 1, we just need to identify a starting point $t_3 \in [0,d)$ such that $\tilde{x}^N(t_3,\beta) \leq \tilde{x}^F(t_3)$. If on the contrary, $\tilde{x}^N(t,\beta) > \tilde{x}^F(t)$ for $t \in [0,d)$, then

$$\begin{aligned} f^N(t,\tilde{x}^N,\beta) - f^F(t,\tilde{x}^F) = {} & \mu(\tilde{x}^F(t) \wedge s) + \int_0^{(\tilde{x}^F(t)-s)^+} \theta(t/s)\,dt \\ & - \mu(\tilde{x}^N(t,\beta) \wedge s) - \theta\left(\frac{(\tilde{x}^N(t,\beta)-s)^+}{s}\right)(\tilde{x}^N(t,\beta)-s)^+ < 0. \end{aligned} \tag{41}$$

However, the above inequality contradicts to the fact that $0 = \int_0^d f^N(t,\tilde{x}^N,\beta)dt = \int_0^d f^F(t,\tilde{x}^F)dt$. Therefore, there must exist such t_3 with $\tilde{x}^N(t_3,\beta) \leq \tilde{x}^F(t_3)$. Apply Lemma 1 we have $\tilde{x}^N(t,\beta) \leq \tilde{x}^F(t)$ for $t \geq t_2$. Since both $\tilde{x}^N(t,\beta)$ and $\tilde{x}^F(t)$ are periodic with period d, we have $\tilde{x}^N(t,\beta) \leq \tilde{x}^F(t)$ for $t \geq 0$.

i. Note that, (41) is strict when $\tilde{x}^N(t,\beta) \geq \tilde{x}^F(t)$ if and only if $\tilde{x}^N(t,\beta) > s$. Therefore, we can find the t_3 such that $\tilde{x}^N(t_3,\beta) < \tilde{x}^F(t_3)$ by contradiction if and only if $\max_{t\geq 0} \tilde{x}^N(t,\beta) > s$, and obtain $\tilde{x}^N(t,\beta) < \tilde{x}^F(t)$ for $t \geq 0$ by applying Lemma 2 with $y(t) = \tilde{x}^F(t)$, $z(t) = \tilde{x}^N(t,\beta)$, $f(t,z) = f^F(t,\tilde{x}^F)$, and $g(t,z) = f^N(t,\tilde{x}^N,\beta)$.

ii. Now we show that $\max_{t\geq 0}\tilde{x}^F(t)>s$ implies $\max_{t\geq 0}\tilde{x}^N(t,\beta)>s$. If not, i.e., if $\max_{t\geq 0}\tilde{x}^F(t)>s$ and $\max_{t\geq 0}\tilde{x}^N(t,\beta)\leq s$, then by (36) and (37) we have $\bar{A}^N(\beta)=0$ and $\bar{A}^F>0$. However, by (39), we have $\bar{A}^F-\bar{A}^N(\beta)\leq \frac{1}{d}\int_0^d(\tilde{x}^N(t,\beta)-\tilde{x}^F(t))^+\,dt\leq 0$, which yields a contradiction.

(b) If $\beta=0$, $f^N(t,x,\beta)\geq f^F(t,x)$, for $t,x\geq 0$. By a similar analysis as we show part (a), we can obtain the desired results. $\square$

D.3.1. Proof of Proposition 4. Case 1 directly follows from Lemma 4.1. For Case 2, $\max_{t\geq 0}\tilde{x}^N(t,0)>s$ implies that $\max_{t\geq 0}\tilde{x}^N(t,\beta)>s$ by Lemma 4.3. Moreover,

1. For $\beta=0$ we have $\bar{x}^N(0)>\bar{x}^F$: Since $\max_{t\geq 0}\tilde{x}^N(t,0)>s$, this follows from Lemma 4.3.(b).
2. For $\beta\geq 0.5$, since $\max_{t\geq 0}\tilde{x}^N(t,\beta)>s$, Lemma 4.3.(a) implies that $\bar{x}^N(\beta)<\bar{x}^F$.
3. Lemma 4.3 implies that $\bar{x}^N(\beta)$ continuously and strictly decreases in β as long as $\max_{t\geq 0}\tilde{x}^N(t,0)>s$.

Combine the three points above, we obtain the desired results. $\square$

D.3.2. Proof of Proposition 5. Case 1 follows from Lemma 4.1, 4.2, and (36)–(37). For Case 2, since $\max\limits_{t\geq 0}\tilde{x}^N(t,0)>s>\min_{t\geq 0}\tilde{x}^N(t,1)$, by Lemma 4.3 we have $\max_{t\geq 0}\tilde{x}^N(t,\beta)>s>\min_{t\geq 0}\tilde{x}^N(t,\beta)$, for any $\beta\in(0,1]$.

1. For $\beta=0$, $\bar{A}^N(0)<\bar{A}^F$: Since $\max_{t\geq 0}\tilde{x}^N(t,0)>s$, by Lemma 4.3.(b) we have $\tilde{x}^N(t,0)>\tilde{x}^F(t)$ $\forall t$ and $\max_{t\geq 0}\tilde{x}^F(t)>s$. Then, by (38) we have $\bar{A}^N(0)\leq\bar{A}^F$. Since $\min_{t\geq 0}\tilde{x}^N(t,0)<s$ implies $\min_{t\geq 0}\tilde{x}^F(t)<s$, there exists an interval $\eta_3\subseteq[0,d)$ such that for $t\in\eta_3$, $\tilde{x}^F(t)<s$ and $\tilde{x}^N(t,0)>\tilde{x}^F(t)$. Therefore, by (35), $\bar{A}^N(0)-\bar{A}^F=\frac{\mu}{d}\int_0^d(\tilde{x}^F(t)-(\tilde{x}^N(t,0)\wedge s))\,dt<0$.
2. For $\beta\geq 0.5$, $\bar{A}^N(\beta)>\bar{A}^F$: Since $\max_{t\geq 0}\tilde{x}^N(t,\beta)>s$, by Lemma 4.3.(a) we have $\tilde{x}^N(t,\beta)<\tilde{x}^F(t)$ $\forall t$. Then, by (39) we have $\bar{A}^F\leq\bar{A}^N(\beta)$. Since $\min_{t\geq 0}\tilde{x}^N(t,\beta)<s$, there exists an interval $\eta_3'\subseteq[0,d)$ such that for $t\in\eta_3'$, $\tilde{x}^N(t,\beta)<s$ and $\tilde{x}^N(t,\beta)<\tilde{x}^F$. Therefore, by (35), $\bar{A}^N(\beta)-\bar{A}^F=\frac{\mu}{d}\int_0^d((\tilde{x}^F(t)\wedge s)-\tilde{x}^N(t,\beta))\,dt>0$.
3. $\bar{A}^N(\beta)$ is increasing in β: Lemma 4.3 implies that $\tilde{x}^N(t,\beta)$ strictly decreases in β. For $\beta^h>\beta^l$, we have $\tilde{x}^N(t,\beta^h)<\tilde{x}^N(t,\beta^l)$. Since $\min_{t\geq 0}\tilde{x}^N(t,\beta)<s$, there exists an interval $\eta_3''\subseteq[0,d)$ such that for $t\in\eta_3''$, $\tilde{x}^N(t,\beta^h)<s$ and $\tilde{x}^N(t,\beta^h)<\tilde{x}^N(t,\beta^l)$. Therefore, by (35), $\bar{A}^N(\beta^h)-\bar{A}^N(\beta^l)=\frac{\mu}{d}\int_0^d((\tilde{x}^N(t,\beta^l)\wedge s)-\tilde{x}^N(t,\beta^h))\,dt>0$.

Combining the results above, we obtain the desired results. $\square$

D.4. Proof for Section 6.4: Proposition 6.

In this case, the flow balance equations (13) and (14) specialize to

$$\bar{A}_1^I=\lambda_1-\mu_1(\bar{x}_1^I\wedge s),\tag{42}$$

$$\bar{A}_2^I=\lambda_2-\mu_2\left((s-\bar{x}_1^I)^+\wedge\bar{x}_2^I\right),\tag{43}$$

for $I\in\{F,N\}$, and (15)-(18) specialize to

$$\bar{A}_1^N:=\theta\left(\beta_1\frac{(\bar{x}_1^N+\bar{x}_2^N-s)^+}{s}\right)(\bar{x}_1^N-s)^+,\tag{44}$$

$$\bar{A}_1^F:=\int_0^{(\bar{x}_1^F-s)^+}\theta\left(u/s\right)du,\tag{45}$$

$$\bar{A}_2^N:=\theta\left(\beta_2\frac{(\bar{x}_1^N+\bar{x}_2^N-s)^+}{s}\right)(\bar{x}_2^N-(s-\bar{x}_1^N)^+)^+,\tag{46}$$

$$\bar{A}_2^F:=\int_{(\bar{x}_1^F-s)^+}^{(\bar{x}_1^F+\bar{x}_2^F-s)^+}\theta\left(u/s\right)du.\tag{47}$$

To compare the equilibrium numbers-in-system and system abandonment rates under different information levels, we first characterize these two metrics under each of the information levels in Lemma 5. The following lemma follows from (42)–(47).

Lemma 5. *For two-priority systems with stationary arrivals, the fluid approximation (7)–(8) has the following equilibrium point:*

1. *If $\rho_1 \le 1$, then $\bar{x}_1^N(\boldsymbol{\beta}) = \bar{x}_1^F = s\rho_1$ and $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = 0$.*

   (i) If $\rho_1 + \rho_2 \le 1$, then $\bar{x}_2^N(\boldsymbol{\beta}) = \bar{x}_2^F = s\rho_2$ and $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = 0$.

   (ii) If $\rho_1 + \rho_2 > 1$, then $\bar{x}_1^N(\boldsymbol{\beta}) + \bar{x}_2^N(\boldsymbol{\beta}) > s$, $\bar{x}_1^F + \bar{x}_2^F > s$, $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = s\mu_2(\rho_1 + \rho_2 - 1)$, $\bar{x}_2^N(\boldsymbol{\beta})$ is the unique solution of

$$\mu_2(\rho_1 + \rho_2 - 1) = \theta\left(\beta_2(\frac{x_2}{s} + \rho_1 - 1)\right)\left(\frac{x_2}{s} + \rho_1 - 1\right), \tag{48}$$

and $\bar{x}_2^F$ is the unique solution of

$$\mu_2(\rho_1 + \rho_2 - 1) = \int_0^{\frac{x_2}{s} + \rho_1 - 1} \theta(u)du. \tag{49}$$

2. *If $\rho_1 > 1$, then $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = \lambda_1 - s\mu_1$ and $\bar{A}_2^N(\boldsymbol{\beta}) = \bar{A}_2^F = \lambda_2$. Further, $\bar{x}_1^N(\boldsymbol{\beta}) > s$, $\bar{x}_1^F > s$, $(\bar{x}_1^N(\boldsymbol{\beta}), \bar{x}_2^N(\boldsymbol{\beta}))$ is the unique solution of the following system of equations,*

$$\mu_1(\rho_1 - 1) = \theta\left(\beta_1 \frac{x_1 + x_2 - s}{s}\right)\left(\frac{x_1}{s} - 1\right), \tag{50}$$

$$\mu_2\rho_2 = \theta\left(\beta_2 \frac{x_1 + x_2 - s}{s}\right)\frac{x_2}{s}, \tag{51}$$

and $(\bar{x}_1^F, \bar{x}_2^F)$ is the unique solution of the following system of equations

$$\mu_1(\rho_1 - 1) = \int_0^{\frac{x_1}{s} - 1} \theta(u)du, \tag{52}$$

$$\mu_2\rho_2 = \int_{\frac{x_1}{s} - 1}^{\frac{x_1}{s} + \frac{x_2}{s} - 1} \theta(u)du. \tag{53}$$

Proof of Lemma 5. Divide both sides of (42) and (43) by $s\mu_1$ and $s\mu_2$, respectively, we can obtain that

$$\rho_1 = (\frac{\bar{x}_1^I}{s} \wedge 1) + \frac{\bar{A}_1^I}{s\mu_1}, \tag{54}$$

$$\rho_2 = ((1 - \frac{\bar{x}_1^I}{s})^+ \wedge \frac{\bar{x}_2^I}{s}) + \frac{\bar{A}_2^I}{s\mu_2}. \tag{55}$$

We provide the proof for $I = F$, and a similar analysis yields the proof for $I = N$.

1. When $\rho_1 \le 1$, we have $\bar{x}_1^F \le s$. If on the contrary, $\bar{x}_1^F > s$, then

$$\rho_1 = (\frac{\bar{x}_1^F}{s} \wedge 1) + \frac{\bar{A}_1^F}{s\mu_1} \ge \frac{\bar{x}_1^F}{s} > 1,$$

which contradicts to $\rho_1 \le 1$. By (45), when $\bar{x}_1^F \le s$, $\bar{A}_1^F = 0$, thus $\bar{x}_1^F = s\rho_1$ by (54).

   (a) When $\rho_1 + \rho_2 \le 1$, we show $\bar{x}_1^F + \bar{x}_2^F \le s$ by contradiction. If otherwise, $\bar{x}_1^F + \bar{x}_2^F > s$, then $\bar{x}_2^F > 0$, $\bar{A}_2^F > 0$, and by (55) we have

$$\rho_2 = (1 - \rho_1) + \frac{\bar{A}_2^F}{s\mu_2} > 1 - \rho_1,$$

which contradicts to $\rho_1 + \rho_2 \le 1$. Therefore, $\bar{x}_1^F + \bar{x}_2^F \le s$, which implies that $\bar{A}_2^F = 0$, and $\bar{x}_2^F = s\mu_2$.

(b) When $\rho_1+\rho_2>1$, we show $\bar{x}_1^F+\bar{x}_2^F>s$ by contradiction. If otherwise, $\bar{x}_1^F+\bar{x}_2^F\leq s$, then $\bar{x}_2^F\leq s(1-\rho_1)<s\rho_2$ and $\bar{A}_2^F=0$. However, (55) implies that $\rho_2=\frac{\bar{x}_2^F}{s}$, which leads to a contradiction. Therefore, $\bar{x}_1^F+\bar{x}_2^F>s$, and $\bar{x}_2^F$ is the solution of

$$\rho_2=1-\rho_1+\frac{\bar{A}_2^F}{s\mu_2}.$$

Plug in (46)–(47) into the above equation, we can obtain the desired results.

2. When $\rho_1>1$, similar to the previous analysis, we can show that $\bar{x}_1^F>s$ by contradiction. Then, (54) and (55) can be simplified as

$$\rho_1=1+\frac{\bar{A}_1^F}{s\mu_1},\quad \rho_2=\frac{\bar{A}_2^F}{s\mu_2}.$$

Plug in (44)–(47) into the above equations, we can obtain the desired results. □

The average abandonment rate ranking immediately follows from Lemma 5. As for the number-in-system rankings, part of the results in Case 1 of Proposition 6 is proved by Lemma 5. We complete the proof for Case 1 of Proposition 6 by the continuity of $\bar{x}^N(\boldsymbol{\beta})$ in $\boldsymbol{\beta}$ (i.e., Proposition D.1) and the following lemma.

LEMMA 6. *If $\rho_1\leq 1$, $\rho_1+\rho_2>1$, then $\bar{x}_2^N(\boldsymbol{\beta})$ is non-increasing in β_2, $\bar{x}_2^N(\beta_1,1)<\bar{x}_2^F$, and $\bar{x}_2^N(\beta_1,0)>\bar{x}_2^F$.*

Proof of Lemma 6. When $\rho_1\leq 1,\rho_1+\rho_2>1$, by Lemma 5, $\bar{x}_2^N(\boldsymbol{\beta})$ is the unique solution of equation (48). Differentiating both sides of (48) with respect to β_2 yields:

$$\frac{d\bar{x}_2^N(\boldsymbol{\beta})}{d\beta_2}=-\frac{s\theta'(\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}+\rho_1-1)(\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}+\rho_1-1)^2}{\theta(\beta_2(\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}+\rho_1-1))+\theta'(\beta_2(\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}+\rho_1-1))\beta_2(\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}+\rho_1-1)}<0.$$

Thus, $\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_2.

Next, we check the boundary case when $\beta_2=1$. Plugging $\beta_2=1$ in (48) we can obtain that:

$$\mu_2(\rho_1+\rho_2-1)=\theta(\frac{\bar{x}_2^N(1)}{s}+\rho_1-1)(\frac{\bar{x}_2^N(1)}{s}+\rho_1-1).$$

By Case 1.(ii) of Lemma 5, we have $\bar{x}_1^N(\beta_1,1)+\bar{x}_2^N(\beta_1,1)>s$. Then, $\bar{x}_2^N(\beta_1,1)<\bar{x}_2^F$ since if on the contrary $\bar{x}_2^N(\beta_1,1)\geq\bar{x}_2^F$, by (49) we have

$$\begin{aligned}\mu_2(\rho_1+\rho_2-1)&=\int_0^{\frac{\bar{x}_2^F}{s}+\rho_1-1}\theta(u)du\leq\int_0^{\frac{\bar{x}_2^N(\beta_1,1)}{s}+\rho_1-1}\theta(u)du\\&<\theta\left(\frac{\bar{x}_2^N(\beta_1,1)}{s}+\rho_1-1\right)\left(\frac{\bar{x}_2^N(\beta_1,1)}{s}+\rho_1-1\right)=\mu_2(\rho_1+\rho_2-1).\end{aligned}$$

Thus, by contradiction, we have $\bar{x}_2^N(\beta_1,1)<\bar{x}_2^F$.

Lastly, we check the other boundary case when $\beta_2=0$. Plugging $\beta_2=0$ in (48) we obtain that

$$\mu_2(\rho_1+\rho_2-1)=\theta(0)(\frac{\bar{x}_2^N(\beta_1,0)}{s}+\rho_1-1).$$

Then, $\bar{x}_2^N(\beta_1,0)>\bar{x}_2^F$ since otherwise, by (49),

$$\mu_2(\rho_1+\rho_2-1)=\int_0^{\frac{\bar{x}_2^F}{s}+\rho_1-1}\theta(u)du>\theta(0)(\frac{\bar{x}_2^F}{s}+\rho_1-1)\geq\theta(0)(\frac{\bar{x}_2^N(\beta_1,0)}{s}+\rho_1-1)=\mu_2(\rho_1+\rho_2-1)$$

yields a contradiction. □

Next, we consider the case when $\rho_1>1$. Case 2 of Proposition 6 is deduced by Lemma 5 and the following lemma:

LEMMA 7. *For two-priority systems with stationary arrivals, if $\rho_1 > 1$,*

1. *For HP customers: for any β_2, ρ_2, there exists $\beta_1^*(\beta_2) \in (0, 0.5)$ such that $\beta_1^*(\beta_2)$ is increasing in β_2, decreasing in ρ_2, and*
   (i) If $\beta_1 \in [\beta_1^*(\beta_2), 1]$, then $\bar{x}_1^N(\boldsymbol{\beta}) \le \bar{x}_1^F$.
   (ii) If $\beta_1 \in [0, \beta_1^*(\beta_2))$, then $\bar{x}_1^F < \bar{x}_1^N(\boldsymbol{\beta})$.
2. *For LP customers: for any β_1, ρ_2, there exists $\beta_2^*(\beta_1) \in (0, 1]$ such that $\beta_2^*(\beta_1)$ is increasing in β_1, decreasing in ρ_2, and*
   (i) If $\beta_2 \in [\beta_2^*(\beta_1), 1]$, then $\bar{x}_2^N(\boldsymbol{\beta}) \le \bar{x}_2^F$.
   (ii) If $\beta_2 \in [0, \beta_2^*(\beta_1))$, then $\bar{x}_2^F < \bar{x}_2^N(\boldsymbol{\beta})$.

Proof of Lemma 7. 1. We prove Part 1 of this lemma in two steps. First, we demonstrate the existence of such $\beta_1^*(\beta_2)$. Next, we prove its monotonicity with respect to the parameters.

Step 1: We prove the first step by showing that, for any fixed β_2, (a) $\bar{x}_1^N(\boldsymbol{\beta})$ is continuous and decreasing in β_1, (b) $\bar{x}_1^F < \bar{x}_1^N(0, \beta_2)$, and (c) $\bar{x}_1^F > \bar{x}_1^N(0.5, \beta_2)$.

(a) The continuity is given by Proposition D.1. To show that $\bar{x}_1^N(\boldsymbol{\beta})$ is decreasing in β_1, differentiating both sides of equations (50) and (51) with respect to β_1 yields:

$$\begin{aligned}
&\beta_1\theta'\left(\beta_1\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\left(\frac{\bar{x}_1^N(\boldsymbol{\beta})}{s}-1\right)\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}+\theta\left(\beta_1\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_1}\\
&=-s\theta'\left(\beta_1\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\left(\frac{\bar{x}_1^N(\boldsymbol{\beta})}{s}-1\right)\left(\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})}{s}-1\right)<0. \qquad (56)
\end{aligned}$$

$$\beta_2\theta'\left(\beta_2\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}+\theta\left(\beta_2\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}=0. \qquad (57)$$

By (57), there are three possible cases: (1) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}=\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}=0$; (2) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}>0$, $\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}<0$, and thus $\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_1}>0$; or (3) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}<0$, $\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_1}>0$, and thus $\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_1}<0$. Both cases (1) and (2) are impossible as they contradict (56). Therefore, case (3) must be true. That is, $\bar{x}_1^N(\boldsymbol{\beta})$ is decreasing in β_1,$\bar{x}_2^N(\boldsymbol{\beta})$ is increasing in β_1, and $\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_1.

(b) Since $\bar{x}_1^N(\boldsymbol{\beta}) > s$ and $\bar{x}_1^F > s$, plugging $\beta_1 = 0$ in (50) and together with (52) we have:

$$\mu_1(\rho_1-1)=\theta(0)(\frac{\bar{x}_1^N(0,\beta_2)}{s}-1)=\int_0^{\frac{\bar{x}_1^F}{s}-1}\theta(u)du>\theta(0)(\frac{\bar{x}_1^F}{s}-1)\Rightarrow\bar{x}_1^N(0,\beta_2)>\bar{x}_1^F.$$

(c) Plugging $\beta_1 = 0.5$ in (50) we have:

$$\theta\left(\frac{\bar{x}_1^N(0.5,\beta_2)-s}{2s}\right)(\frac{\bar{x}_1^N(0.5,\beta_2)}{s}-1)<\theta\left(\frac{\bar{x}_1^N(0.5,\beta_2)+\bar{x}_2^N(0.5,\beta_2)-s}{2s}\right)(\frac{\bar{x}_1^N(0.5,\beta_2)}{s}-1)=\mu_1(\rho_1-1).$$

Since θ is concave, then by (50) and (52) we have:

$$\theta(\frac{\bar{x}_1^F-s}{2s})(\frac{\bar{x}_1^F}{s}-1)>\int_0^{\frac{\bar{x}_1^F}{s}-1}\theta(u)du=\mu_1(\rho_1-1)>\theta\left(\frac{\bar{x}_1^N(0.5,\beta_2)-s}{2s}\right)(\frac{\bar{x}_1^N(0.5,\beta_2)}{s}-1).$$

Therefore, $\bar{x}_1^F > \bar{x}_1^N(0.5, \beta_2)$.

Combining (a)–(c), there exists a $\beta_1^*(\beta_2) \in (0, 0.5)$ such that $\bar{x}_1^N(\boldsymbol{\beta}) = \bar{x}_1^F$ when $\beta_1 = \beta_1^*(\beta_2)$, $\bar{x}_1^N(\boldsymbol{\beta}) < \bar{x}_1^F$ when $1 \ge \beta_1 > \beta_1^*(\beta_2)$ and $\bar{x}_1^F < \bar{x}_1^N(\boldsymbol{\beta})$ when $0 < \beta_1 < \beta_1^*(\beta_2)$.

Step 2: $\beta_1^*(\beta_2)$ is increasing in β_2 and decreasing in ρ_2. To facilitate the proof, we first show that $\bar{x}_1^N(\boldsymbol{\beta})$ is increasing in β_2, $\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_2, and $\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_2. Differentiating both sides of equations (50) and (51) with respect to β_2 yields:

$$\beta_1\theta'\left(\beta_1\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{\bar{x}_1^N(\boldsymbol{\beta})-s}{s}\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}+\theta\left(\beta_1\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_2}=0. \tag{58}$$

$$\begin{aligned}&\beta_2\theta'\left(\beta_2\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}+\theta\left(\beta_2\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}\\&=-s\theta'\left(\beta_2\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s}\right)\frac{\bar{x}_2^N(\boldsymbol{\beta})}{s}(\frac{\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta})-s}{s})<0.\end{aligned} \tag{59}$$

By (58), there are three possible cases: (1) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}=\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_2}=0$; (2) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}>0$, $\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_2}<0$, and thus $\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}>0$; or (3) $\frac{d(\bar{x}_1^N(\boldsymbol{\beta})+\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}<0$, $\frac{d(\bar{x}_1^N(\boldsymbol{\beta}))}{d\beta_2}>0$, and thus $\frac{d(\bar{x}_2^N(\boldsymbol{\beta}))}{d\beta_2}<0$. Both cases (1) and (2) are impossible as they contradict (59). Therefore, case (3) must be true.

By (50) and (52), when $\beta_1=\beta_1^*(\beta_2)$, then $\bar{x}_1^N(\beta_1^*(\beta_2),\beta_2)=\bar{x}_1^F=:x_1^*$, and

$$\mu_1(\rho_1-1)=\theta\left(\beta_1^*(\beta_2)\frac{x_1^*+\bar{x}_2^N(\beta_1^*(\beta_2),\beta_2)-s}{s}\right)\left(\frac{x_1^*}{s}-1\right)=\int_0^{\frac{x_1^*}{s}-1}\theta(u)du.$$

Since $\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_2, an increase in β_2 leads to a decrease in $\bar{x}_2^N(\beta_1^*(\beta_2),\beta_2)$. As $x_1^*=\bar{x}_1^F$ is independent of $\boldsymbol{\beta}$ and remains unchanged, $\beta_1^*(\beta_2)$ must increase to maintain the above equation.

Similarly, when ρ_2 increases, $\bar{x}_1^N(\beta_1^*(\beta_2),\beta_2)=x_1^*$ remains unchanged, $\bar{x}_2^N(\beta_1^*(\beta_2),\beta_2)$ must increase to maintain (51), and thus $\beta_1^*(\beta_2)$ must decrease to maintain (50).

2. Similar to Part 1, we prove Part 2 in two steps.

Step 1: We show the existence of $\beta_2^*(\beta_1)$ by showing the following: for any fixed β_1, (a) $\bar{x}_2^N(\boldsymbol{\beta})$ is decreasing in β_2; (b) when $\beta_2=0$, $\bar{x}_2^F<\bar{x}_2^N(\beta_1,0)$; and (c) when $\beta_2=1$, (i) if $\beta_1\le\beta_1^*(\beta_2)$, then $\bar{x}_2^F>\bar{x}_2^N(\beta_1,1)$; (ii) if $\beta_1>\beta_1^*(\beta_2)$, then there exist a threshold ρ_2^* such that $\bar{x}_2^F<\bar{x}_2^N(\boldsymbol{\beta})$ if $\rho_2<\rho_2^*$ and $\bar{x}_2^F\ge\bar{x}_2^N(\boldsymbol{\beta})$ if $\rho_2\ge\rho_2^*$. Then, by the continuity of $\bar{x}_2^N(\boldsymbol{\beta})$ on β_2, such $\beta_2^*(\beta_1)$ exists and when $\beta_1>\beta_1^*(\beta_2)$ and $\rho_2<\rho_2^*$, $\beta_2^*(\beta_1)=1$; otherwise, $0<\beta_2^*(\beta_1)<1$.

Since part (a) is already proven in step 2 of part 1, our focus shifts to parts (b) and (c).

(b) When $\beta_2=0$, by (51) and (53) we have

$$\mu_2\rho_2=\theta(0)\frac{\bar{x}_2^N(\beta_1,0)}{s}=\int_{\frac{\bar{x}_1^F}{s}-1}^{\frac{\bar{x}_1^F}{s}+\frac{\bar{x}_2^F}{s}-1}\theta(u)du>\int_{\frac{\bar{x}_1^F}{s}-1}^{\frac{\bar{x}_1^F}{s}+\frac{\bar{x}_2^F}{s}-1}\theta(0)du=\theta(0)\frac{\bar{x}_2^F}{s}.$$

Thus, $\bar{x}_2^N(\beta_1,0)>\bar{x}_2^F$.

(c.i) When $\beta_2=1$, if $\beta_1\le\beta_1^*(\beta_2)$, then by part 1 we have $\bar{x}_1^N(\boldsymbol{\beta})\ge\bar{x}_1^F$. By by (51) and (53) we have

$$\begin{aligned}\mu_2\rho_2&=\theta\left(\frac{\bar{x}_1^N(\beta_1,1)+\bar{x}_2^N(\beta_1,1)-s}{s}\right)\frac{\bar{x}_2^N(\beta_1,1)}{s}=\int_{\frac{\bar{x}_1^F}{s}-1}^{\frac{\bar{x}_1^F}{s}+\frac{\bar{x}_2^F}{s}-1}\theta(u)du\\&<\theta\left(\frac{\bar{x}_1^F+\bar{x}_2^F-s}{s}\right)\frac{\bar{x}_2^F}{s}\le\theta\left(\frac{\bar{x}_1^N(\beta_1,1)+\bar{x}_2^F-s}{s}\right)\frac{\bar{x}_2^F}{s}.\end{aligned}$$

Therefore, $\bar{x}_2^N(\beta_1,1)<\bar{x}_2^F$ since θ is increasing.

(c.ii) When $\beta_2=1$, if $\beta_1>\beta_1^*(\beta_2)$, then by part 1 we have $\bar{x}_1^N(\beta_1,1)<\bar{x}_1^F$. Let $\Delta=\bar{x}_1^F-\bar{x}_1^N(\beta_1,1)>0$. We will show that there exist a threshold ρ_2^* such that $\bar{x}_2^F<\bar{x}_2^N(\boldsymbol{\beta})$ if $\rho_2<\rho_2^*$ and $\bar{x}_2^F\geq\bar{x}_2^N(\boldsymbol{\beta})$ if $\rho_2\geq\rho_2^*$. Let

$$g(x)=\theta(\frac{\bar{x}_1^N(x)+x-s}{s})x-\int_{\bar{x}_1^F-s}^{\bar{x}_1^F+x-s}\theta(\frac{u}{s})du.$$

We first show that $g(x)$ has a unique positive root in two steps. In step A, we define two alternative systems of equations with solutions $(\underline{x}_{1,m}^N,\underline{x}_{2,m}^N)$, $(\bar{x}_{1,m}^N,\bar{x}_{2,m}^N)$ such that for any m, $\underline{x}_{1,m}^N$ is independent of $\underline{x}_{2,m}^N$, $\bar{x}_{1,m}^N$ is independent of $\bar{x}_{2,m}^N$, and $\underline{x}_{1,m}^N\uparrow\bar{x}_1^N(\beta_1,1),\bar{x}_{1,m}^N\downarrow\bar{x}_1^N(\beta_1,1)$ as $m\to\infty$. In step B, we identify functions $\bar{g}_m$ and $\underline{g}_m$ such that $\bar{g}_m$ and $\underline{g}_m$ have unique positive roots for $\forall m>0$, and $\underline{g}_m(x)\uparrow g(x),\bar{g}_m(x)\downarrow g(x)$ as $m\to\infty$.

Step A: Define alternative systems of equations, indexed by m where $m\to\infty$, as follows. Let $\bar{x}_{1,m}^N$ and $\bar{x}_{2,m}^N$ denote the solutions to the following system of equations:

$$\mu_1(\rho_1-1)=\theta(\beta_1\frac{x_{1,m}^N+\zeta_m-s}{s})(\frac{x_{1,m}^N}{s}-1), \tag{60}$$

$$\mu_2\rho_2=\theta(\frac{x_{1,m}^N+x_{2,m}^N-s}{s})\frac{x_{2,m}^N}{s}, \tag{61}$$

where we let $\{\zeta_m,m\geq 0\}$ be a sequence converging from below to $\bar{x}_2^N(\beta_1,1)$ i.e., $\zeta_m\leq\bar{x}_2^N(\beta_1,1)$ and $\zeta_m\uparrow\bar{x}_2^N(\beta_1,1)$. Similarly, let $\underline{x}_{1,m}^N$ and $\underline{x}_{2,m}^N$ denote the solutions to the following system of equations:

$$\mu_1(\rho_1-1)=\theta(\beta_1\frac{x_{1,m}^N+\xi_m-s}{s})(\frac{x_{1,m}^N}{s}-1), \tag{62}$$

$$\mu_2\rho_2=\theta(\frac{x_{1,m}^N+x_{2,m}^N-s}{s})\frac{x_{2,m}^N}{s}. \tag{63}$$

where we let $\{\xi_m,m\geq 0\}$ be a sequence converging from above to $\bar{x}_2^N(\beta_1,1)$ i.e., $\xi_m\geq\bar{x}_2^N(\beta_1,1)$ and $\xi_m\downarrow\bar{x}_2^N(\beta_1,1)$. Then, by comparing equations (50), (60), and (62), and the monotonicity of θ, we must have $\underline{x}_{1,m}^N\leq\bar{x}_1^N(\beta_1,1)\leq\bar{x}_{1,m}^N$ for every $m\geq 0$. Moreover, by the continuity and monotonicity of θ we can obtain that $\underline{x}_{1,m}^N\to\bar{x}_1^N(\beta_1,1)$ and $\bar{x}_{1,m}^N\to\bar{x}_1^N(\beta_1,1)$ as $m\to\infty$. To see this, suppose on the contrary that there exists $\epsilon>0$ such that $\bar{x}_1^N(\beta_1,1)-\underline{x}_{1,m}^N>\epsilon$ for $\forall m>0$. Then, since $\xi_m\downarrow\bar{x}_2^N(\beta_1,1)$, we have $\xi_m-\bar{x}_2^N(\beta_1,1)<\frac{\epsilon}{2}$ for m sufficiently large. Thus, $\underline{x}_{1,m}^N+\xi_m-\bar{x}_1^N(\beta_1,1)-\bar{x}_2^N(\beta_1,1)<-\frac{\epsilon}{2}<0$ for large enough m, and

$$\mu_1(\rho_1-1)=\theta(\beta_1\frac{\underline{x}_{1,m}^N+\xi_m-s}{s})(\frac{\underline{x}_{1,m}^N}{s}-1)<\theta(\beta_1\frac{\bar{x}_1^N(\beta_1,1)+\bar{x}_2^N(\beta_1,1)-s}{s})(\frac{\bar{x}_1^N(\beta_1,1)}{s}-1)=\mu_1(\rho_1-1),$$

which yields contradiction. Therefore, $\underline{x}_{1,m}^N\to\bar{x}_1^N(\beta_1,1)$ as $m\to\infty$. Similarly, we can obtain that $\bar{x}_{1,m}^N\to\bar{x}_1^N(\beta_1,1)$ as $m\to\infty$.

Step B: Define the functions $\bar{g}_m$ and $\underline{g}_m$ as follows:

$$\bar{g}_m(x)=\theta(\frac{\bar{x}_{1,m}^N+x-s}{s})x-\int_{\bar{x}_1^F-s}^{\bar{x}_1^F+x-s}\theta(\frac{u}{s})du,$$

and

$$\underline{g}_m(x)=\theta(\frac{\underline{x}_{1,m}^N+x-s}{s})x-\int_{\bar{x}_1^F-s}^{\bar{x}_1^F+x-s}\theta(\frac{u}{s})du.$$

It is clear that $\underline{g}_m(x)\leq g(x)\leq\bar{g}_m(x)$ for $\forall m>0,x\geq 0$. Furthermore, $\underline{g}_m(x)\to g(x)$ and $\bar{g}_m(x)\to g(x)$ as $m\to\infty$.

Now we show that $\underline{g}_m(x)$ has a unique positive root, denoted as $\underline{x}^2_{2,m}$, and $\underline{g}_m(x) < 0$ for $0 < x < \underline{x}^2_{2,m}$, $\underline{g}_m(x) > 0$ for $x > \underline{x}^2_{2,m}$.

Recall that $\Delta = \bar{x}^F_1 - \bar{x}^N_1(\beta_1, 1) > 0$, then $\Delta_1 := \bar{x}^F_1 - \underline{x}^N_{1,m} > \Delta > 0$. Note that, $\underline{g}_m(0) = 0$. When $0 < x \le \Delta_1$,

$$\theta(\frac{\underline{x}^N_{1,m} + x - s}{s})x \le \theta(\frac{\bar{x}^F_1 - s}{s})x < \int_{\bar{x}^F_1 - s}^{\bar{x}^F_1 + x - s} \theta(\frac{u}{s})du \Rightarrow \underline{g}_m(x) < 0.$$

When $x > \Delta_1$, since $\theta(\cdot)$ is increasing and concave, we have

$$\begin{aligned}\underline{g}'_m(x) &= \frac{x}{s}\theta'(\frac{\underline{x}^N_{1,m} + x - s}{s}) + \theta(\frac{\underline{x}^N_{1,m} + x - s}{s}) - \theta(\frac{\bar{x}^F_1 + x - s}{s}) \\ &> \frac{x}{s}\theta'(\frac{\underline{x}^N_{1,m} + x - s}{s}) - \frac{\Delta_1}{s}\theta'(\frac{\underline{x}^N_{1,m} + x - s}{s}) \ge 0.\end{aligned}$$

Thus $\underline{g}_m(x)$ is increasing when $x > \Delta_1$. Moreover,

$$\underline{g}_m(2\Delta_1) = \theta(\frac{\bar{x}^F_1 + \Delta_1 - s}{s})2\Delta_1 - \int_{\bar{x}^F_1 - s}^{\bar{x}^F_1 + 2\Delta_1 - s} \theta(\frac{u}{s})du > 0.$$

Therefore, there exists a unique $\underline{x}^2_{2,m} > 0$ such that $\underline{g}_m(\underline{x}^2_{2,m}) = 0$, and when $0 < x < \underline{x}^2_{2,m}$, $\underline{g}_m(x) < 0$; when $x > \underline{x}^2_{2,m}$, $\underline{g}_m(x) > 0$.

By a similar analysis, we can show that $\bar{g}_m(x)$ also has a unique positive root (denoted as $\bar{x}^2_{2,m}$), and $\bar{g}_m(x) < 0$ for $0 < x < \bar{x}^2_{2,m}$, $\bar{g}_m(x) > 0$ for $x > \bar{x}^2_{2,m}$. (Note that since $\bar{x}^N_{1,m} \to \bar{x}^N_1(\beta_1, 1)$ as $m \to \infty$, there exist $\bar{m}$ such that when $m > \bar{m}$, $\bar{x}^N_{1,m} - \bar{x}^N_1(\beta_1, 1) < \frac{1}{2}\Delta$. Then, $\Delta_2 := \bar{x}^F_1 - \bar{x}^N_{1,m} > \frac{1}{2}\Delta > 0$ when $m > \bar{m}$, and then the result follows from a similar analysis.)

Both $\underline{x}^2_{2,m}$ and $\bar{x}^2_{2,m}$ converge to some point $x^2_2 > 0$ by continuity of θ. Moreover, x^2_2 is the unique positive root of $g(x)$. First, we must have $g(x^2_2) = 0$. Otherwise, if $g(x^2_2) = \delta > 0$, then since $\bar{x}^2_{2,m} \to x^2_2$, we have $\bar{g}_m(x^2_2) - \bar{g}_m(\bar{x}^2_{2,m}) = \bar{g}_m(x^2_2) < \delta$ for m large enough, which contradicts to the fact that $g(x^2_2) \le \bar{g}_m(x^2_2)$. Also, x^2_2 is the unique positive root. Otherwise, if there exists another positive root of $g(x)$, denoted as $\hat{x}^2_2 \ne x^2_2$. Then, we must have either $\bar{g}_m(\hat{x}^2_2) > 0, \underline{g}_m(\hat{x}^2_2) > 0$ or $\bar{g}_m(\hat{x}^2_2) < 0, \underline{g}_m(\hat{x}^2_2) < 0$ for m sufficiently large, and thus contradicts with $\underline{g}_m(x) \le g(x) \le \bar{g}_m(x)$. In conclusion, x^2_2 is the unique positive root of $g(x)$. Also, we can show that $g(x) < 0$ for $0 < x < x^2_2$, and $g(x) > 0$ for $x > x^2_2$. Let

$$\rho^*_2 := \frac{1}{\mu_2}\int_{\frac{\bar{x}^F_1}{s} - 1}^{\frac{\bar{x}^F_1}{s} + \frac{x^2_2}{s} - 1} \theta(u)du,$$

then when $\rho_2 < \rho^*_2$, we have $\bar{x}^F_2 < x^2_2$, $g(\bar{x}^F_2) < 0$; when $\rho_2 \ge \rho^*_2$, we have $\bar{x}^F_2 \ge x^2_2$, $g(\bar{x}^F_2) \ge 0$. Then when $\rho_2 < \rho^*_2$, we must have $\bar{x}^F_2 < \bar{x}^N_2(\beta_1, 1)$ since otherwise, if $\bar{x}^F_2 \ge \bar{x}^N_2(\beta_1, 1)$,

$$\mu_2\rho_2 = \theta\left(\frac{\bar{x}^N_1(\beta_1, 1) + \bar{x}^N_2(\beta_1, 1) - s}{s}\right)\frac{\bar{x}^N_2(\beta_1, 1)}{s} \le \theta\left(\frac{\bar{x}^N_1(\beta_1, 1) + \bar{x}^F_2 - s}{s}\right)\frac{\bar{x}^F_2}{s} < \int_{\frac{\bar{x}^F_1}{s} - 1}^{\frac{\bar{x}^F_1}{s} + \frac{\bar{x}^F_2}{s} - 1} \theta(u)du = \mu_2\rho_2$$

yields a contradiction. Thus, when $\rho_2 < \rho^*_2$, $\bar{x}^F_2 < \bar{x}^N_2(\beta_1, 1)$. Similarly, when $\rho_2 \ge \rho^*_2$, $\bar{x}^F_2 \ge \bar{x}^N_2(\beta_1, 1)$.

Step 2: the proof is similar to to the method used in Step 2 for Case 1; therefore, we omit it here for brevity.

□

D.5. Supplementary Results and Proofs for Section 6.5

In this section, we provide supplementary results and proofs for the comparison results for systems under two-class and non-stationary periodic arrivals in Section 6.5.

For the ease of the analysis for the proofs in this section, we write down the net flow rate functions as follows:

$$\begin{aligned}
f_1^N(t,x,\boldsymbol{\beta}) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \theta\left(\beta_1 \frac{(x_1(t)+x_2(t)-s)^+}{s}\right)(x_1(t)-s)^+,\\
f_1^F(t,x) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \int_0^{(x_1(t)-s)^+} \theta\left(u/s\right)du,\\
f_2^N(t,x,\boldsymbol{\beta}) &= \lambda_2(t) - \mu_2\left((s-x_1(t))^+ \wedge x_2(t)\right) - \theta\left(\beta_2 \frac{(x_1(t)+x_2(t)-s)^+}{s}\right)(x_2(t)-(s-x_1(t))^+)^+,\\
f_2^F(t,x) &= \lambda_2(t) - \mu_2\left((s-x_1(t))^+ \wedge x_2(t)\right) - \int_{(x_1(t)-s)^+}^{(x_1(t)+x_2(t)-s)^+} \theta\left(u/s\right)du.
\end{aligned}$$

Also, by the balance equations (13)–(18) we have

$$0 = \int_0^d f_k^N(t,\tilde{x}^N,\boldsymbol{\beta})dt = \int_0^d f_k^F(t,\tilde{x}^F)dt,\ \text{for}\ k=1,2. \tag{64}$$

In Section D.5.1, we present proofs of Proposition 7. Proposition 8 follows from Proposition 7, and its proof is similar to the proof of Proposition 5, so we omit it here. Proposition 2 follows from Propositions 7 and 8, so we also omit its proof. In Section D.5.2, we provide supplementary results and proofs for two-priority systems with non-stationary periodic arrivals and uniformly underloaded HP class (i.e., $\bar{\rho}_1 \le 1$).

D.5.1. Proof of Proposition 7 Denote $\varphi_2(l(t))$ as the unique solution of (7) under N with $x_2(t) = l(t)$ as given, for any non-negative periodic function $l(t)$. That is, $\varphi_2(l)$ solves the following ODE:

$$\dot{x}_1(t) = \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \theta\left(\beta_1 \frac{(x_1(t)+l(t)-s)^+}{s}\right)(x_1(t)-s)^+.$$

Then, to facilitate the proof of Proposition 7, we introduce the following lemma. The proof of this lemma is similar to the proof of Lemma 7, thus we omit it here.

LEMMA 8. 1. *$\varphi_2(l)$ is non-increasing in $l(t)$, with strict decreasing if $\max_{t\ge 0}\varphi_2(l) > s$.*
2. *If $\max_{t\ge 0}\varphi_2(l_0) > s$, for some $l_0(t)$, then $\max_{t\ge 0}\varphi_2(l) > s$, for any $l(t)$.*
3. *When $\beta_1 \ge 0.5$, $\varphi_2(t,\boldsymbol{\beta},0) := \varphi_2(0) \le \tilde{x}_1^F(t)$, $\forall t$; if $\max_{t\ge 0}\tilde{x}_1^F(t) > s$, then the inequality is strict for all t.*

Proof of Proposition 7 When $\bar{\rho}_1 > 1$,

1. When $\beta_1 = 0$, for HP customers, $\tilde{x}_1^N(t,\boldsymbol{\beta})$ and $f_1^N(t,x,\boldsymbol{\beta})$ are independent of $\tilde{x}_2^N(t,\boldsymbol{\beta})$ and

$$\begin{aligned}
f_1^N(t,x_1,(0,\beta_2)) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \theta(0)(x_1(t)-s)^+,\\
f_1^F(t,x_1) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \int_0^{(x_1(t)-s)^+} \theta\left(u/s\right)du.
\end{aligned}$$

It is clear that $f_1^N(t,x_1,(0,\beta_2)) \ge f_1^F(t,x_1)$ for any $t \ge 0$. By a similar argument as we prove case 3 of Lemma 4 by applying Lemmas 1 and 2, we can obtain that $\tilde{x}_1^N(t,\boldsymbol{\beta}) \ge \tilde{x}_1^F(t)\ \forall t$, with strict inequality for all t if $\max_{t\ge 0}\tilde{x}_1^N(t,\boldsymbol{\beta}) > s$.

For LP customers: since $\tilde{x}_1^N(t,(0,\beta_2))$ and $\tilde{x}_1^F(t)$ are independent of $\tilde{x}_2^N(t,(0,\beta_2))$ and $\tilde{x}_2^F(t)$, we can rewrite $f_2^N(t,x,(0,\beta_2))$ and $f_2^F(t,x)$ as a function of $x_2(t)$ as follows:

$$\begin{aligned}
f_2^N(t,x_2,(0,\beta_2)) =& \lambda_2(t) - \mu_2\left((s-\tilde{x}_1^N(t,(0,\beta_2)))^+ \wedge x_2(t)\right) \\
&- \theta\left(\beta_2 \frac{(\tilde{x}_1^N(t,(0,\beta_2)) + x_2(t) - s)^+}{s}\right)(x_2(t) - (s - \tilde{x}_1^N(t,(0,\beta_2)))^+)^+, \\
f_2^F(t,x_2) =& \lambda_2(t) - \mu_2\left((s-\tilde{x}_1^F(t))^+ \wedge x_2(t)\right) - \int_{(\tilde{x}_1^F(t)-s)^+}^{(\tilde{x}_1^F(t)+x_2(t)-s)^+} \theta\left(u/s\right) du.
\end{aligned}$$

Define the difference between these two net flow rate function as follows:

$$\begin{aligned}
&\Delta f_2(t,x_2,(0,\beta_2)) := f_2^N(t,x_2,(0,\beta_2)) - f_2^F(t,x_2) \\
&= \mu_2\left((s-\tilde{x}_1^F(t))^+ \wedge x_2(t)\right) - \mu_2\left((s-\tilde{x}_1^N(t,(0,\beta_2)))^+ \wedge x_2(t)\right) \\
&+ \int_{(\tilde{x}_1^F(t)-s)^+}^{(\tilde{x}_1^F(t)+x_2(t)-s)^+} \theta\left(u/s\right) du - \theta\left(\beta_2 \frac{(\tilde{x}_1^N(t,(0,\beta_2)) + x_2(t) - s)^+}{s}\right)(x_2(t) - (s - \tilde{x}_1^N(t,(0,\beta_2)))^+)^+.
\end{aligned}$$

(a) When $\beta_2 = 0$, we show that if $\min_{t\geq 0} \tilde{x}_1^F(t) \geq s$, then $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,(0,0)), \forall t$. Note that $\min_{t\geq 0} \tilde{x}_1^F(t) \geq s$ and the HP ranking result imply that $\tilde{x}_1^N(t,(0,0)) > \tilde{x}_1^F(t) \geq s$, for all t. Therefore, for any $x_2(t) > 0$, $\Delta f_2(t,x_2,(0,\beta_2))$ can be simplified as follows:

$$\Delta f_2(t,x_2,(0,0)) = \int_{\tilde{x}_1^F(t)-s}^{\tilde{x}_1^F(t)+x_2(t)-s} \theta\left(u/s\right) du - \theta(0)x_2(t) > 0.$$

Using a similar argument as we prove Lemma 4, we can find $t_4 \in [0,d)$ such that $\tilde{x}_2^F(t_4) < \tilde{x}_2^N(t_4,(0,0))$ by contradiction via (64) with $k=2$. Apply Lemma 2 with $y(t) = \tilde{x}_2^N(t,(0,0)), z(t) = \tilde{x}_2^F(t)$, $f(t,y) = f_2^N(t,x_2,(0,0))$, and $g(t,z) = f_2^F(t,x_2)$, we obtain the desired result.

(b) When $\beta_2 = 1$, we show that if $\max_{t\geq 0} \tilde{x}_1^F(t) > s$ and $\mu_2 \leq \theta(0)$, then $\tilde{x}_2^F(t) > \tilde{x}_2^N(t,(0,1)), \forall t$. We first show that $\Delta f_2(t,x_2,(0,1)) \leq 0$, for all t. When $\max_{t\geq 0} \tilde{x}_1^F(t) > s$, we have $\tilde{x}_1^F(t) < \tilde{x}_1^N(t,(0,1))$ for all t. For $x_2(t) > 0$,

i. If $\tilde{x}_1^F(t) < \tilde{x}_1^N(t,(0,1)) \leq s$,

A. If $x_2(t) \leq s - \tilde{x}_1^N(t,(0,1))$, then the case is trivial with $\Delta f_2(t,x_2,(0,1)) = 0$.

B. If $s - \tilde{x}_1^N(t,(0,1)) < x_2(t) \leq s - \tilde{x}_1^F(t)$, then when $\mu_2 \leq \theta(0)$,

$$\begin{aligned}
\Delta f_2(t,x_2,(0,1)) &= \mu_2(\tilde{x}_1^N(t,(0,1)) + x_2(t) - s) - \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right)(\tilde{x}_1^N(t,(0,1)) + x_2(t) - s) \\
&< \mu_2(\tilde{x}_1^N(t,(0,1)) + x_2(t) - s) - \theta(0)(\tilde{x}_1^N(t,(0,1)) + x_2(t) - s) \leq 0.
\end{aligned}$$

C. If $x_2(t) > s - \tilde{x}_1^F(t)$, then when $\mu_2 \leq \theta(0)$,

$$\begin{aligned}
\Delta f_2(t,x_2,(0,1)) =& \mu_2(\tilde{x}_1^N(t,(0,1)) - \tilde{x}_1^F(t)) + \int_0^{\tilde{x}_1^F(t)+x_2(t)-s} \theta\left(u/s\right) du \\
&- \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right)(\tilde{x}_1^N(t,(0,1)) + x_2(t) - s) \\
\leq& \mu_2(\tilde{x}_1^N(t,(0,1)) - \tilde{x}_1^F(t)) + \int_0^{\tilde{x}_1^F(t)+x_2(t)-s} \theta\left(u/s\right) du - \int_0^{\tilde{x}_1^N(t,(0,1))+x_2(t)-s} \theta\left(u/s\right) du \\
=& \mu_2(\tilde{x}_1^N(t,(0,1)) - \tilde{x}_1^F(t)) - \int_{\tilde{x}_1^F(t)+x_2(t)-s}^{\tilde{x}_1^N(t,(0,1))+x_2(t)-s} \theta\left(u/s\right) du \\
<& \mu_2(\tilde{x}_1^N(t,(0,1)) - \tilde{x}_1^F(t)) - \int_{\tilde{x}_1^F(t)+x_2(t)-s}^{\tilde{x}_1^N(t,(0,1))+x_2(t)-s} \theta\left(0\right) du = (\mu_2 - \theta(0))(\tilde{x}_1^N(t,(0,1)) - \tilde{x}_1^F(t)) \leq 0.
\end{aligned}$$

ii. When $\tilde{x}_1^F(t) \le s < \tilde{x}_1^N(t,(0,1))$,

A. If $x_2(t) \le s - \tilde{x}_1^F(t)$, then when $\mu_2 \le \theta(0)$,

$$\Delta f_2(t,x_2,(0,1)) = \mu_2 x_2(t) - \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right) x_2(t) < (\mu_2 - \theta(0))x_2(t) \le 0.$$

B. If $x_2(t) > s - \tilde{x}_1^F(t)$, then when $\mu_2 \le \theta(0)$,

$$\begin{aligned}\Delta f_2(t,x_2,(0,1)) =& \mu_2(s - \tilde{x}_1^F(t)) + \int_0^{\tilde{x}_1^F(t)+x_2(t)-s} \theta(u/s)\,du - \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right) x_2(t)\\ <& \mu_2(s - \tilde{x}_1^F(t)) + \theta\left(\tilde{x}_1^F(t) + x_2(t) - s\right)\left(\tilde{x}_1^F(t) + x_2(t) - s\right) - \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right) x_2(t)\\ <& \left(\mu_2 - \theta\left(\tilde{x}_1^F(t) + x_2(t) - s\right)\right)(s - \tilde{x}_1^F(t)) \le 0.\end{aligned}$$

iii. When $\tilde{x}_1^F(t) > s$, then

$$\begin{aligned}\Delta f_2(t,x_2,(0,1)) =& \int_{\tilde{x}_1^F(t)-s}^{\tilde{x}_1^F(t)+x_2(t)-s} \theta(u/s)\,du - \theta\left(\frac{\tilde{x}_1^N(t,(0,1)) + x_2(t) - s}{s}\right) x_2(t)\\ <& \int_{\tilde{x}_1^F(t)-s}^{\tilde{x}_1^F(t)+x_2(t)-s} \theta(u/s)\,du - \theta\left(\frac{\tilde{x}_1^F(t) + x_2(t) - s}{s}\right) x_2(t) < 0.\end{aligned}$$

Thus, $\Delta f_2(t,x_2,(0,1)) \le 0$, i.e., $f_2^N(t,x_2,(0,1)) \le f_2^F(t,x_2)$, for all $t, x_2(t)$. Using a similar argument as we prove Lemma 4, we can find $t_5 \in [0,d)$ such that $\tilde{x}_2^N(t_5,(0,1)) < \tilde{x}_2^F(t_5)$ by contradiction via (64) with $k=2$. Apply Lemma 2 with $y(t) = \tilde{x}_2^F(t), z(t) = \tilde{x}_2^N(t,(0,1))$, $f(t,y) = f_2^F(t,x_2)$, and $g(t,z) = f_2^N(t,x_2,(0,1))$, we obtain the desired result.

2. When $\beta_1 \ge 0.5$, for HP customers, we prove the results by Lemmas 1 and 2. To apply these Lemmas, we need to rewrite the HP net flow rate functions f_1^I as a function of $x_1(t)$ (instead of the two-dimensional variable $x(t)$). By (7) and (9) we know that, $\tilde{x}_1^F(t)$ (so as $f_1^F(t,x)$) is independent of $\tilde{x}_2^F(t)$, while $\tilde{x}_1^N(t,\boldsymbol{\beta})$ (so as $f_1^N(t,x,\boldsymbol{\beta})$) depends on $\tilde{x}_2^N(t,\boldsymbol{\beta})$. Therefore, we can directly rewrite f_1^F as a function of $x_1(t)$. For $I=N$, denote $\varphi_1(l,\boldsymbol{\beta})$ as the unique solution of (8) under N with $x_1(t) = l(t)$ as given for any non-negative periodic function $l(t)$ with period d. Then, we can rewrite f_1^N as a function of $x_1(t)$ with $x_2(t,\boldsymbol{\beta}) = \varphi_1(x_1,\boldsymbol{\beta}) \ge 0$. Specifically, f_1^I can be rewritten as follows:

$$\begin{aligned}f_1^N(t,x_1,\boldsymbol{\beta}) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \theta\left(\beta_1\frac{(x_1(t) + \varphi_1(x_1(t),\boldsymbol{\beta}) - s)^+}{s}\right)(x_1(t) - s)^+,\\ f_1^F(t,x_1) &= \lambda_1(t) - \mu_1(x_1(t)\wedge s) - \int_0^{(x_1(t)-s)^+} \theta(u/s)\,du.\end{aligned}$$

Since θ is increasing and concave, $\varphi_1(x_1,\boldsymbol{\beta}) \ge 0$, it is obvious that $f_1^N(t,\boldsymbol{\beta}) \le f_1^F(t,x_1)$ for any $t \ge 0$ when $\beta_1 \ge 0.5$. Using a similar argument as we prove case 3 of Lemma 4 by applying Lemmas 1 and 2, we can obtain that $\tilde{x}_1^N(t,\boldsymbol{\beta}) \le \tilde{x}_1^F(t)\ \forall t$, and $\bar{x}_1^N(\boldsymbol{\beta}) < \bar{x}_1^F$ if $\max_{t\ge 0}\tilde{x}_1^F(t) > s$.

For LP customers: similar to how we handle the HP class, we can rewrite $f_2^N(t,x,\boldsymbol{\beta})$ as a function of $x_2(t)$ with $x_1(t) = \varphi_2(x_2(t))$, where $\varphi_2(l(t))$ denote the unique solution of (7) under N with $x_2(t) = l(t)$ as given, for any non-negative periodic function $l(t)$. Specifically,

$$f_2^N(t,x_2,\boldsymbol{\beta}) = \lambda_2(t) - \mu_2\left((s - \varphi_2(x_2(t)))^+ \wedge x_2(t)\right) - \theta\left(\beta_2\frac{(\varphi_2(x_2(t)) + x_2(t) - s)^+}{s}\right)(x_2(t) - (s - \varphi_2(x_2(t)))^+)^+.$$

We can also rewrite $f_2^F(t,x)$ as a function of $x_2(t)$ with $x_1(t) = \tilde{x}_1^F(t)$ and define $\Delta f_2(t,x_2,\boldsymbol{\beta}) := f_2^N(t,x_2,\boldsymbol{\beta}) - f_2^F(t,x_2)$ like we did in part 1 of this proof.

(a) When $\beta_2 = 0$, if $\max_{t\geq 0} \tilde{x}_1^F(t) > s$ and $\mu_2 \leq \theta(0)$, then $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,\boldsymbol{\beta})$ $\forall t$. Note that, Lemma 8 implies that $\varphi_2(x_2(t)) \leq \varphi_2(0) < \tilde{x}_1^F(t), \forall t,\ x_2(t) > 0$. Using this set of inequalities and a similar analysis as Case 1.(b), we can obtain the desired results by Lemma 2 with $y(t) = \tilde{x}_2^N(t,(\beta_1,0)), z(t) = \tilde{x}_2^F(t),\ f(t,y) = f_2^N(t,x_2,(\beta_1,0))$, and $g(t,z) = f_2^F(t,x_2)$.

(b) When $\beta_2 = 1$,

i. If $\min\limits_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$, we show that there exists a threshold $\tilde{\rho}_2^1$ such that if $\bar{\rho}_2 < \tilde{\rho}_2^1$, then $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,\boldsymbol{\beta}), \forall t$. Let $\delta_1(t) := \tilde{x}_1^F(t) - \varphi_2(t,\boldsymbol{\beta},0)$ and $\underline{\delta}_1 := \min_t \delta_1(t)$. Since $\varphi_2(x_2(t)) < \varphi_2(0) < \tilde{x}_1^F(t), \forall t$, we have $\tilde{x}_1^F(t) > \varphi_2(x_2(t)) + \underline{\delta}_1, \forall t$. We prove this result by three steps: (1) When $\tilde{x}_2^F(t) < \underline{\delta}_1$, for $t \geq 0$, we show that $\tilde{x}_2^F(t) < \tilde{x}_2^N(t,(\beta_1,1))$ by Lemma 2. (2) For the special case when the LP arrival rate is stationary with $\rho_2(t) = \bar{\rho}_2$, for $t \geq 0$, denote the LP periodic equilibrium of the corresponding system as $\hat{x}_2^F(t)$. Then, we show that there exists a threshold $\tilde{\rho}_2^1$ for the LP load such that $\hat{x}_2^F(t) < \underline{\delta}_1$ if $\bar{\rho}_2 < \tilde{\rho}_2^1$. (3) When the LP arrival rate is non-stationary with maximum LP load $\bar{\rho}_2$, we show that $\tilde{x}_2^F(t) \leq \hat{x}_2^F(t)$ by Lemma 1. Therefore, if $\bar{\rho}_2 < \tilde{\rho}_2^1$, then $\tilde{x}_2^F(t) < \underline{\delta}_1$, which further implies the desired result.

Step 1: Let $y(t) = \tilde{x}_2^N(t,(\beta_1,1))$, $z(t) = \tilde{x}_2^F(t)$, $f(t,z) = f_2^N(t,\tilde{x}_2^F(t),(\beta_1,1))$, and $g(t,z) = f_2^F(t,\tilde{x}_2^F(t))$. To apply Lemma 2, we need to show as follows that $f_2^F(t,\tilde{x}_2^F(t)) \leq f_2^N(t,\tilde{x}_2^F(t),(\beta_1,1))$, for $t \geq 0$, and there exists t_6 such that $\tilde{x}_2^F(t_6) < \tilde{x}_2^N(t_6,(\beta_1,1))$.

When $\min\limits_{t\geq 0} \tilde{x}_1^F(t) > \min\limits_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$, $\beta_2 = 1$, and $\tilde{x}_2^F(t) < \underline{\delta}_1 < \tilde{x}_1^F(t) - \varphi_2(\tilde{x}_2^F(t))$, we have

$$\begin{aligned}\Delta f_2(t,\tilde{x}_2^F(t),(\beta_1,1)) =& \int_{\tilde{x}_1^F(t)-s}^{\tilde{x}_1^F(t)+\tilde{x}_2^F(t)-s} \theta\left(u/s\right)du - \theta\left(\frac{\varphi_2(\tilde{x}_2^F(t)) + \tilde{x}_2^F(t) - s}{s}\right)\tilde{x}_2^F(t)(t)\\ >&\theta\left(\frac{\tilde{x}_1^F(t)-s}{s}\right)\tilde{x}_2^F(t) - \theta\left(\frac{\tilde{x}_1^F(t)-s}{s}\right)\tilde{x}_2^F(t) = 0.\end{aligned}$$

Next, we show by contradiction the existence of t_6. If on the contrary, $\tilde{x}_2^F(t) \geq \tilde{x}_2^N(t,(\beta_1,1))$, for all t, then $\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) + \tilde{x}_2^N(t,(\beta_1,1)) < \varphi_2(t,(\beta_1,1),0) + \tilde{x}_2^F(t) \leq \varphi_2(t,(\beta_1,1),0) + \underline{\delta}_1 \leq \tilde{x}_1^F(t)$, and

$$\begin{aligned}f_2^F(t,\tilde{x}_2^F) - f_2^N(t,\tilde{x}_2^N,(\beta_1,1)) =& \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) + \tilde{x}_2^N(t,(\beta_1,1)) - s)}{s}\right)\tilde{x}_2^N(t,(\beta_1,1)) - \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^F(t)-s)} \theta\left(u/s\right)du\\ \leq& \theta\left(\frac{(\tilde{x}_1^F(t)-s)}{s}\right)\tilde{x}_2^F(t) - \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^F(t)-s)} \theta\left(u/s\right)du < 0,\end{aligned}$$

which contradicts to (64).

Step 2: Since $\tilde{x}_1^F(t) > s$, when the LP arrival rate is stationary, by (8), $\hat{x}_2^F(t)$ is the unique periodic solution of

$$\dot{x}_2^F = \bar{\rho}_2\mu_2 s - \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+x_2^F(t)-s)} \theta\left(u/s\right)du := \hat{f}_2^F(t,x_2^F(t)).$$

By the nature of periodic equilibrium we have

$$\bar{\rho}_2\mu_2 s = \frac{1}{d}\int_0^d \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\hat{x}_2^F(t)-s)} \theta\left(u/s\right)dudt. \tag{65}$$

Thus, when $\bar{\rho}_2 \to 0$, we must have $\hat{x}_2^F(t) \to 0$ by continuity. Therefore, for any $\underline{\delta}_1 > 0$, there exists $\tilde{\rho}_2^1$ such that if $\bar{\rho}_2 < \tilde{\rho}_2^1$, we have $\hat{x}_2^F(t) < \underline{\delta}_1$ for $t \geq 0$.

Step 3: When the LP arrival rate is non-stationary with maximum LP load $\bar{\rho}_2$, we prove $\tilde{x}_2^F(t) \leq \hat{x}_2^F(t)$ by Lemma 1 with $y(t) = \tilde{x}_2^F(t)$, $z(t) = \hat{x}_2^F(t)$, $f(t,y) = f_2^F(t,y)$, and $g(t,z) = \hat{f}_2^F(t,z)$. It is clear that $f_2^F(t,y) \leq \hat{f}_2^F(t,z)$ since $f_2^F(t,x_2) - \hat{f}_2^F(t,x_2) = \lambda_2(t) - \bar{\rho}_2\mu_2 s \leq 0$. And we can show the existence of $t_{11} \in [0,d)$ such that $\tilde{x}_2^F(t_{11}) \leq \hat{x}_2^F(t_{11})$ by contradiction using (65) and $\int_0^d f_2^F(t, \tilde{x}_2^F(t)) = 0$ as follows:

$$\frac{1}{d}\int_0^d \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^F(t)-s)} \theta\left(u/s\right) du dt = \frac{1}{d}\int_0^d \lambda_2(t) dt < \bar{\rho}_2\mu_2 s = \frac{1}{d}\int_0^d \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\hat{x}_2^F(t)-s)} \theta\left(u/s\right) du dt.$$

ii. If $\min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$, or $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$ and $\mu_2 \geq \theta(\infty)$, there exists a threshold $\tilde{\rho}_2^2$ such that if $\underline{\rho}_2 > \tilde{\rho}_2^2$, then $\tilde{x}_2^F(t) > \tilde{x}_2^N(t,\boldsymbol{\beta}), \forall t$. By Lemma 8, when $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$, $\varphi_2(\infty) < \varphi_2(x_2(t)) < \tilde{x}_1^F(t)$, for $x_2(t) > 0$. Let $\delta_2(t) := \tilde{x}_1^F(t) - \varphi_2(\infty) > 0$ and $\bar{\delta}_2 := \max_t \delta_2(t)$. We prove this result by three steps: (1) When $\tilde{x}_2^N(t,(\beta_1,1)) > 2\bar{\delta}_2$, we show that $\tilde{x}_2^F(t) > \tilde{x}_2^N(t,(\beta_1,1))$ using Lemma 2. (2) For the special case when the LP arrival rate is stationary with $\rho_2(t) = \underline{\rho}_2$, for $t \geq 0$, denote the LP periodic equilibrium of the corresponding system as $\breve{x}_2^F(t)$. Then, we show that there exists a threshold $\tilde{\rho}_2^2$ for the LP load such that $\breve{x}_2^N(t,(\beta_1,1)) > 2\bar{\delta}_2$ if $\underline{\rho}_2 > \tilde{\rho}_2^2$. (3) When the LP arrival rate is non-stationary with minimum LP load $\underline{\rho}_2$, we show that $\tilde{x}_2^N(t,(\beta_1,1)) \geq \breve{x}_2^N(t,(\beta_1,1))$. Therefore, if $\underline{\rho}_2 > \tilde{\rho}_2^2$, then $\tilde{x}_2^N(t,(\beta_1,1)) > 2\bar{\delta}_2$, which further implies the desired result. The proofs of Steps 2 and 3 are similar to our proofs of Steps 2 and 3 for Case 2.(b).i, thus we omit it here for brevity.

Step 1: Let $y(t) = \tilde{x}_2^F(t)$, $z(t) = \tilde{x}_2^N(t,(\beta_1,1))$, $f(t,z) = f_2^F(t,\tilde{x}_2^N(t,(\beta_1,1)))$, and $g(t,z) = f_2^N(t,\tilde{x}_2^N(t,(\beta_1,1)))$. To apply Lemma 2, we need to show that when $\tilde{x}_2^N(t,(\beta_1,1)) \geq 2\bar{\delta}_2$, $\min_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \geq s$, or $\max_{t\geq 0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$ and $\mu_2 \geq \theta(\infty)$, the following two conditions are satisfied: $f_2^F(t,\tilde{x}_2^N) \geq f_2^N(t,\tilde{x}_2^N,(\beta_1,1))$ and there exists $t_7 \in [0,d)$ such that $\tilde{x}_2^F(t_7) > \tilde{x}_2^N(t_7,(\beta_1,1))$. For $\tilde{x}_2^N(t,(\beta_1,1)) > 0$,

- When $\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) < \tilde{x}_1^F(t) \leq s$,

  —If $\tilde{x}_2^N(t,(\beta_1,1)) \leq s - \tilde{x}_1^F(t)$, then $\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) = 0$ is trivial.

  —If $s - \tilde{x}_1^F(t) < \tilde{x}_2^N(t,(\beta_1,1)) \leq s - \varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))$, then when $\mu_2 \geq \theta(\infty)$,

$$\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) = \int_0^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta\left(u/s\right) du - \mu_2(\tilde{x}_1^F(t) + \tilde{x}_2^N(t,(\beta_1,1)) - s) < 0.$$

  —If $\tilde{x}_2^N(t,(\beta_1,1)) > s - \varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))$, then when $\mu_2 \geq \theta(\infty)$,

$$\begin{aligned}
\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) =& \mu_2(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) - \tilde{x}_1^F(t)) + \int_0^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta\left(u/s\right) du \\
& - \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) + \tilde{x}_2^N(t,(\beta_1,1)) - s)}{s}\right)(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) + \tilde{x}_2^N(t,(\beta_1,1)) - s) \\
<& \mu_2(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) - \tilde{x}_1^F(t)) \\
& + \int_0^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta\left(u/s\right) du - \int_0^{\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s} \theta\left(u/s\right) du \\
=& -\mu_2(\tilde{x}_1^F(t) - \varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))) + \int_{\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s}^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta\left(u/s\right) du \\
\leq& (\theta(\infty) - \mu_2)(\tilde{x}_1^F(t) - \varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))) \leq 0.
\end{aligned}$$

- When $\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) \leq s < \tilde{x}_1^F(t)$,

  —If $\tilde{x}_2^N(t,(\beta_1,1)) \leq s - \varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))$, then when $\mu_2 \geq \theta(\infty)$,

$$\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) = \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta\left(u/s\right) du - \mu_2\tilde{x}_2^N(t,(\beta_1,1)) < (\theta(\infty) - \mu_2)\tilde{x}_2^N(t,(\beta_1,1)) \leq 0.$$

—If $\tilde{x}_2^N(t,(\beta_1,1)) > s-\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))$, then when $\mu_2 \ge \theta(\infty)$ and $\tilde{x}_2^N(t,(\beta_1,1)) \ge 2\bar{\delta}_2$,

$$\begin{aligned}\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) =& -\mu_2(s-\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))) + \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta(u/s)\,du \\ &- \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right)(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s) \\ <& (\theta(\infty)-\mu_2)(s-\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))) \\ &+ \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta(u/s)\,du - \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right)\tilde{x}_2^N(t,(\beta_1,1)) \\ <& \left(\theta\left(\frac{(\tilde{x}_1^F(t)+\frac{1}{2}\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right) - \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right)\right)\tilde{x}_2^N(t,(\beta_1,1)) \le 0.\end{aligned}$$

Note that $\tilde{x}_2^N(t,(\beta_1,1)) \ge 2\bar{\delta}_2 > 2(\tilde{x}_1^F(t)-\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))))$ implies the last inequality.

- When $\varphi_2(\tilde{x}_2^N(t,(\beta_1,1))) > s$, since $\tilde{x}_2^N(t,(\beta_1,1)) \ge 2\bar{\delta}_2$,

$$\begin{aligned}\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) =& \int_{(\tilde{x}_1^F(t)-s)}^{(\tilde{x}_1^F(t)+\tilde{x}_2^N(t,(\beta_1,1))-s)} \theta(u/s)\,du - \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right)\tilde{x}_2^N(t,(\beta_1,1)) \\ <& \left(\theta\left(\frac{(\tilde{x}_1^F(t)+\frac{1}{2}\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right) - \theta\left(\frac{(\varphi_2(\tilde{x}_2^N(t,(\beta_1,1)))+\tilde{x}_2^N(t,(\beta_1,1))-s)}{s}\right)\right)\tilde{x}_2^N(t,(\beta_1,1)) \le 0.\end{aligned}$$

To summarize, we show that $\Delta f_2(t,\tilde{x}_2^N,(\beta_1,1)) < 0$ when $\tilde{x}_2^N(t,(\beta_1,1)) \ge 2\bar{\delta}_2$, $\min_{t\ge0} \tilde{x}_1^N(t,\boldsymbol{\beta}) \ge s$, or $\max_{t\ge0} \tilde{x}_1^N(t,\boldsymbol{\beta}) > s$ and $\mu_2 \ge \theta(\infty)$.

The existence of t_7 can be shown in a similar analysis as we prove Step 1 of Case 2.(b).i, therefore we omit it here for brevity. □

D.5.2. Supplementary Results When the HP Class Uniformly Underloaded. Proposition 9 summarizes how information affects the equilibrium average numbers-in-system and average abandonment rates.

PROPOSITION 9. *For two-priority systems with non-stationary periodic arrivals and uniformly underloaded HP class (i.e., $\bar{\rho}_1 \le 1$), the equilibrium average numbers-in-system and abandonment rates under no (N) and full (F) information compare as follows:*

1. *Numbers-in-system: $\bar{x}_1^N(\boldsymbol{\beta}) = \bar{x}_1^F \le s$. For LP, if $\max_{t\ge0}(\tilde{x}_1^N(t,(\beta_1,0))+\tilde{x}_2^N(t,(\beta_1,0))) > s$, there is a threshold $\beta_q^* \in (0,0.5)$ such that $\bar{x}_2^N(\beta_1,\beta_q^*) = \bar{x}_2^F$ and:*
   - (a) *If $\beta_2 \in [0,\beta_q^*)$, then $\bar{x}_2^N(\boldsymbol{\beta}) > \bar{x}_2^F$.*
   - (b) *If $\beta_2 \in (\beta_q^*,1]$, then $\bar{x}_2^N(\boldsymbol{\beta}) < \bar{x}_2^F$.*
2. *Abandonment: $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = 0$. For LP, if $\max_{t\ge0}(\tilde{x}_1^N(t,(\beta_1,0))+\tilde{x}_2^N(t,(\beta_1,0))) > s > \min_{t\ge0}(\tilde{x}_1^N(t,(\beta_1,1))+\tilde{x}_2^N(t,(\beta_1,1)))$, there is a threshold $\beta_a^* \in (0,0.5)$ such that $\bar{A}_2^N(\beta_1,\beta_a^*) = \bar{A}_2^F$ and:*
   - (a) *If $\beta_2 \in [0,\beta_a^*)$, then $\bar{A}_2^N(\boldsymbol{\beta}) < \bar{A}_2^F$.*
   - (b) *If $\beta_2 \in (\beta_a^*,1]$, then $\bar{A}_2^N(\boldsymbol{\beta}) > \bar{A}_2^F$.*

We note that the values of the thresholds β_q^* and β_a^* are close but need not coincide.

Proof of Proposition 9. The following lemma 9 and the equations (13) and (14) imply the rankings of the equilibrium average number-in-system and average abandonment rates shown in Proposition 9.

Lemma 9. *For two-priority systems with non-stationary periodic arrivals and uniformly underloaded HP class (i.e., $\bar{\rho}_1 \leq 1$), information has the following effects on the numbers-in-system:*

1. *For HP customers:* $\tilde{x}_1^N(t,\boldsymbol{\beta}) = \tilde{x}_1^F(t) \leq s$ *for all* $t \geq 0$.
2. *For LP customers:* $\tilde{x}_2^N(t,\boldsymbol{\beta})$ *is continuously non-increasing in* β_2*, with strict decreasing if* $\max_{t\geq 0}(\tilde{x}_1^N(t,(\beta_1,0)) + \tilde{x}_2^N(t,(\beta_1,0))) > s$ *and:*
   (i) If $\beta_2 \geq 0.5$*, then* $\tilde{x}_2^N(t,\boldsymbol{\beta}) \leq \tilde{x}_2^F(t)$ *for all* $t \geq 0$*;*
      i. *If* $\max_{t\geq 0}(\tilde{x}_1^N(t,\boldsymbol{\beta}) + \tilde{x}_2^N(t,\boldsymbol{\beta})) > s$*, then the inequality is strict for all* t*,*
      ii. *If* $\max_{t\geq 0}(\tilde{x}_1^F(t) + \tilde{x}_2^F(t,)) > s$*, then* $\max_{t\geq 0}(\tilde{x}_1^N(t,\boldsymbol{\beta}) + \tilde{x}_2^N(t,\boldsymbol{\beta})) > s$*.*
   (ii) If $\beta_2 = 0$*, then* $\tilde{x}_2^N(t,\boldsymbol{\beta}) \geq \tilde{x}_2^F(t)$ *for all* $t \geq 0$*;*
      i. *If* $\max_{t\geq 0}(\tilde{x}_1^F(t) + \tilde{x}_2^F(t)) > s$*, then the inequality is strict for all* t*,*
      ii. *If* $\max_{t\geq 0}(\tilde{x}_1^N(t,(\beta_1,0)) + \tilde{x}_2^N(t,(\beta_1,0))) > s$*, then* $\max_{t\geq 0}(\tilde{x}_1^F(t) + \tilde{x}_2^F(t)) > s$*.*

Proof of Lemma 9. When $\bar{\rho}_1 \leq 1$,

1. For HP customers, the proof of $\tilde{x}_1^N(t,\boldsymbol{\beta}) = \tilde{x}_1^F(t) \leq s$ is similar to that of case 1 in Lemma 4, so we omit it here. Moreover, combining with (13) and (15)–(16), we have $\bar{A}_1^N(\boldsymbol{\beta}) = \bar{A}_1^F = 0$ and $\int_0^d \lambda_1(t)dt = \mu_1 \int_0^d \tilde{x}_1^N(t,\boldsymbol{\beta})dt = \mu_1 \int_0^d \tilde{x}_1^F(t)dt$.

When $\tilde{x}_1^N(t,\boldsymbol{\beta}) = \tilde{x}_1^F(t) \leq s$, then $f_1^N(t,\tilde{x}^N,\boldsymbol{\beta}) = f_1^F(t,\tilde{x}^F) = \lambda_1(t) - \mu_1(\tilde{x}_1^F(t) \wedge s)$. That is, the solutions of $\tilde{x}_1^N(t,\boldsymbol{\beta})$, $\tilde{x}_1^F(t)$ are given by solving (7) and independent of $\boldsymbol{\beta}$, $\tilde{x}_2^N(t,\boldsymbol{\beta})$, and $\tilde{x}_2^F(t)$.

2. For LP customers, we compare the two metrics between N and F by Lemmas 1 and 2. Note that, Lemmas 1 and 2 only apply to one dimensional initial value problems. Since when $\bar{\rho}_1 \leq 1$, $\tilde{x}_1^N(t,\boldsymbol{\beta})$ is independent of $\tilde{x}_2^N(t,\boldsymbol{\beta})$ and $\tilde{x}_1^N(t,\boldsymbol{\beta}) = \tilde{x}_1^F(t)$. We can convert the two dimensional initial value problems into one dimensional ones for LP customers by rewriting $f_2^N(t,x,\boldsymbol{\beta})$ and $f_2^F(t,x)$ as a function of $x_2(t)$ with $x_1(t) = \tilde{x}_1^F(t)$ fixed and given by (7). Let $y(t) = \tilde{x}_2^N(t,\boldsymbol{\beta})$, $z(t) = \tilde{x}_2^F(t)$, $f(t,y) = f_2^N(t,(\tilde{x}_1^N,x_2),\boldsymbol{\beta}) =: f_2^N(t,x_2,\boldsymbol{\beta})$, and $g(t,z) = f_2^F(t,(\tilde{x}_1^F,x_2)) := f_2^F(t,x_2)$. The proof of $\tilde{x}_2^N(t,,\boldsymbol{\beta})$ being continuously decreasing in β_2 is similar to that of case 3 in Lemma 4, so we omit it here. Note that,

$$f_2^N(t,x_2,\boldsymbol{\beta}) - f_2^F(t,x_2) = \int_0^{(\tilde{x}_1^F(t)+x_2(t)-s)^+} \theta(u/s)\,du - \theta\left(\beta_2 \frac{(\tilde{x}_1^F(t)+x_2(t)-s)^+}{s}\right)(\tilde{x}_1^F(t)+x_2(t)-s)^+.$$

Thus, $f_2^N(t,x_2,\boldsymbol{\beta}) - f_2^F(t,x_2) \leq 0$ when $\beta_2 \geq 0.5$ and $f_2^N(t,x_2,\boldsymbol{\beta}) - f_2^F(t,x_2) \geq 0$ when $\beta_2 = 0$. Using a similar argument as in the proof of case 3 of Lemma 4, we can obtain the desired results. □

In systems with uniform HP underload, information clearly has no effect on HP customers because they never queue. Therefore, for LP customers the system is equivalent to a single-class non-stationary system with time-varying capacity. Proposition 9 shows that our results on the performance effects of information generalize naturally from the single-class non-stationary case with constant capacity (Propositions 4.2 and 5.2) to the LP class in two-class systems with uniform HP underload. Specifically, for sufficiently small β_2, no (N) information yields a longer LP queue length (at all times) and lower average LP abandonment rate, compared to full (F) information. Conversely, for sufficiently large β_2, no (N) information yields a shorter LP queue length (at all times) and higher average LP abandonment rate, compared to full (F) information (unless LP customers never queue even under no information, i.e., $\max_{t\geq 0}(\tilde{x}_1^N(t,\boldsymbol{\beta}) + \tilde{x}_2^N(t,\boldsymbol{\beta})) \leq s$). These results on the trade-off between less queueing and more abandonment under no vs. full information are consistent with those for single-class systems with non-stationary arrivals with constant capacity (Propositions 4.2 and 5.2) and also extend these results to systems with time-varying capacity.

Next, for the ranking of the total cost metric, we have the following proposition (from Proposition 9).

Proposition 10. *For two-priority systems with non-stationary periodic arrivals and uniformly under-loaded HP class (i.e., $\bar{\rho}_1 \leq 1$), the equilibrium average total costs under no (N) and full (F) information compare as follows: For HP, $TC_1^N(\boldsymbol{\beta}) = TC_1^F$. For LP, if $\max_{t\geq 0}(\tilde{x}_1^N(t,(\beta_1,0)) + \tilde{x}_2^N(t,(\beta_1,0))) > s$,*

1. *When $\beta_2 < \min\{\beta_q^*, \beta_a^*\}$, there is a threshold $\frac{V_k}{c_k}^l$ such that $TC_2^N(\boldsymbol{\beta}) = TC_2^F$ at $\frac{V_k}{c_k} = \frac{V_k}{c_k}^l(\boldsymbol{\beta})$ and*
   (a) *If $\frac{V_k}{c_k} < \frac{V_k}{c_k}^l(\boldsymbol{\beta})$, then $TC_2^N(\boldsymbol{\beta}) > TC_2^F$.*
   (b) *If $\frac{V_k}{c_k} > \frac{V_k}{c_k}^l(\boldsymbol{\beta})$, then $TC_2^N(\boldsymbol{\beta}) < TC_2^F$.*
2. *When $\beta_2 > \max\{\beta_q^*, \beta_a^*\}$, there is a threshold $\frac{V_k}{c_k}^h(\boldsymbol{\beta})$ such that $TC_2^N(\boldsymbol{\beta}) = TC_2^F$ at $\frac{V_k}{c_k} = \frac{V_k}{c_k}^h(\boldsymbol{\beta})$ and*
   (a) *If $\frac{V_k}{c_k} < \frac{V_k}{c_k}^h(\boldsymbol{\beta})$, then $TC_2^N(\boldsymbol{\beta}) < TC_2^F$.*
   (b) *If $\frac{V_k}{c_k} > \frac{V_k}{c_k}^h(\boldsymbol{\beta})$, then $TC_2^N(\boldsymbol{\beta}) > TC_2^F$.*

Recall that, the ratio $\frac{V_k}{c_k}$ represents the relative weight on the abandonments (i.e., throughput loss) over the waiting. Proposition 10 indicates that, when the system is uniformly underloaded with HP arrivals, information does not impact the HP total cost, and the LP total cost ranking depends on the performance measures rankings and the ratio $\frac{V_2}{c_2}$. When $\beta_2 < \min\{\beta_q^*, \beta_a^*\}$, by Proposition 9 we have that no information yields larger number-in-system but fewer abandonments; and there exists a threshold for $\frac{V_2}{c_2}$ such that when this relative weight is below the threshold, the LP total cost ranking aligns with the number-in-system ranking and when the relative weight is above the threshold, the LP total cost ranking align with the abandonment ranking. A similar analysis applies to the case when $\beta_2 > \max\{\beta_q^*, \beta_a^*\}$, except that this time the rankings of number-in-system and abandonments are reversed compared to the previous case. Intuitively, this is because when there is a trade-off between the two individual performance measures, if the impact of throughput loss outweighs the impact of waiting ($\frac{V_2}{c_2}$ is large), the abandonment ranking dominates the number-in-system ranking in its overall impact on the total cost ranking; and vice versa if the impact of waiting outweighs the throughput loss.

Appendix E: Robustness Checks

In this section, we establish the robustness of our key results if one relaxes important assumptions. In Section E.1, we show that our comparison results based on fluid approximations are also valid for small and moderately-sized systems. In Section E.2 we explain how our performance comparison results apply to the waiting time metric. Finally, in Section E.3 we discuss how our performance comparison results apply to settings with time-varying capacity.

E.1. Small Stochastic Systems

In Theorem 1, we establish the convergence of fluid-scaled stochastic processes to corresponding fluid limits, as the system size increases without bound. In this section, we use simulation experiments to illustrate the accuracy of the fluid approximations as well as our comparison results for systems with a moderate number of servers (e.g., less than 50).

(a) HP customers (b) LP customers

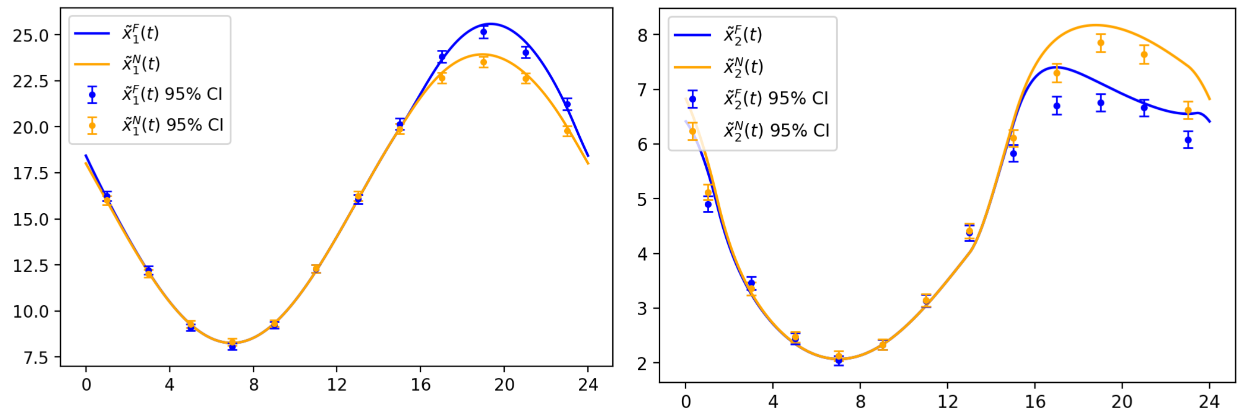


Figure 11 Comparisons of the number-in-system trajectories under different information levels for the simulated stochastic systems and the fluid models ($\mu_1 = \mu_2 = 1,$ $s = 20, \theta(x) = 1.5 - e^{-x}, \boldsymbol{\beta} = (1, 0.43), \rho_1 = 0.8, \rho_2 = 0.2, \lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$**).**

We consider two-class systems with time-varying arrival rates, under no and full information models. Specifically, we assume that $\mu_1 = \mu_2 = 1$, $s = 20$, $\theta(x) = 1.5 - e^{-x}$, and $\lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$. We consider two sets of system load when the system switches between over- and under-loaded: relatively low load with $\rho_1 = 0.8, \rho_2 = 0.2$, and high load with $\rho_1 = \rho_2 = 1$. For no information, we consider two $\boldsymbol{\beta}$ vectors, where $\boldsymbol{\beta} = (1, 0.43)$ is the best-fitting $\boldsymbol{\beta}^*$ for the class-independent $\beta_k(x) = e^{-x}$ at the low load $\rho_1 = 0.8, \rho_2 = 0.2$; and $\boldsymbol{\beta} = (0.43, 0.42)$ is the best-fitting $\boldsymbol{\beta}^*$ for the class-dependent $\beta_1(x) = 0.6 * e^{-2x}, \beta_2(x) = e^{-x}$ at the high load $\rho_1 = \rho_2 = 1$. For each information level, we estimate the expected average number of customers in periodic steady-state using simulation. In Figures 11 and 12, we plot 95% confidence intervals for the expected number-in-system process, under each information level, along with corresponding time-dependent fluid limits, at equilibrium, over one period, for $\rho_1 = 0.8, \rho_2 = 0.2$ and $\rho_1 = \rho_2 = 1$, respectively. Based on Figures 11 and 12, we make two observations:

1. For each information level, the periodic fluid equilibrium curve is very close to the corresponding simulation-based estimate of the expected number-in-system, which implies that our fluid approximations are fairly accurate even with a small (i.e., $s = 20$) number of servers.

2. The simulation-based ranking is consistent with the fluid-based ranking at each point in time. In particular, the number-in-system rankings in our numerical examples lie in Cases 2 of Proposition 7 and align with Figure 7(a).

For further robustness checks, we consider systems with alternative sizes, i.e., $s = 10$ or 50.

In Figures 13 and 14, we plot 95% confidence intervals for the expected number-in-system process, under each information level, along with their corresponding time-dependent fluid limits, at equilibrium, over one period, for small ($s = 10$) and large ($s = 50$) server sizes. Other parameters are consistent with our numerical examples in Figures 11 and 12.

As can be seen in Figures 13 and 14, the fluid approximations are effective to describe performance in a stochastic system even with small-sized system ($s = 10$), and is pretty precise with large-sized system

(a) HP customers (b) LP customers

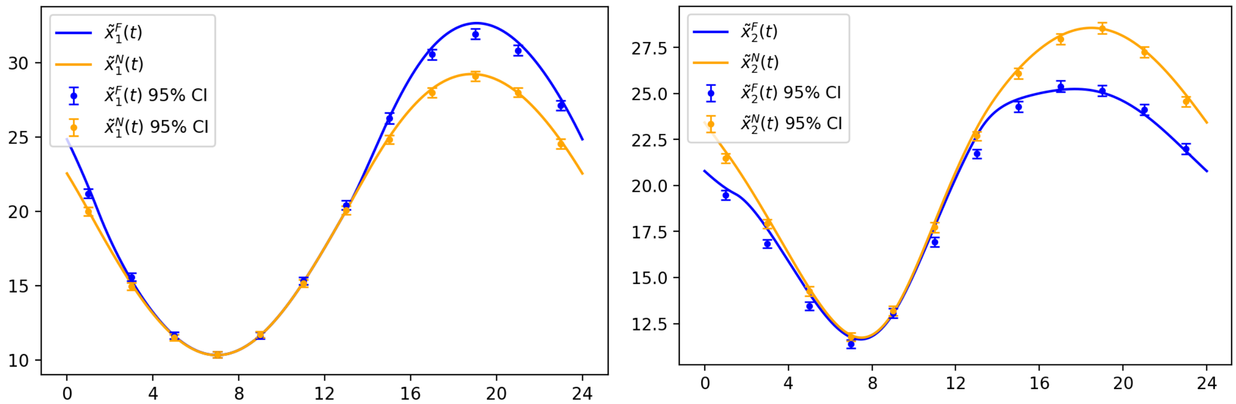
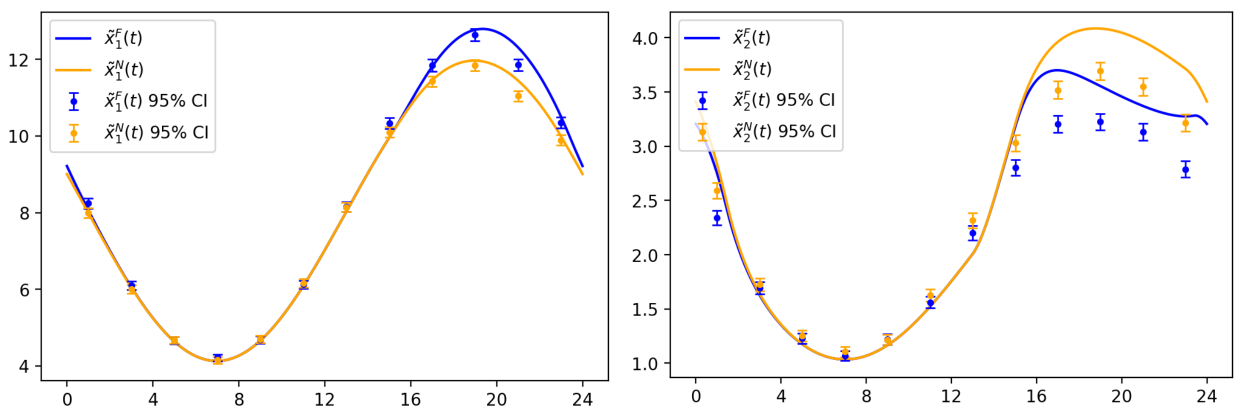


Figure 12 Comparisons of the number-in-system trajectories under different information levels for the simulated stochastic systems and the fluid models ($\mu_1 = \mu_2 = 1$, $s = 20, \theta(x) = 1.5 - e^{-x}, \boldsymbol{\beta} = (0.43, 0.42), \rho_1 = \rho_2 = 1, \lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$**).**

(a) HP customers, $\rho_1 = 0.8, \rho_2 = 0.2$ (b) LP customers, $\rho_1 = 0.8, \rho_2 = 0.2$



(c) HP customers, $\rho_1 = \rho_2 = 1$ (d) LP customers, $\rho_1 = \rho_2 = 1$

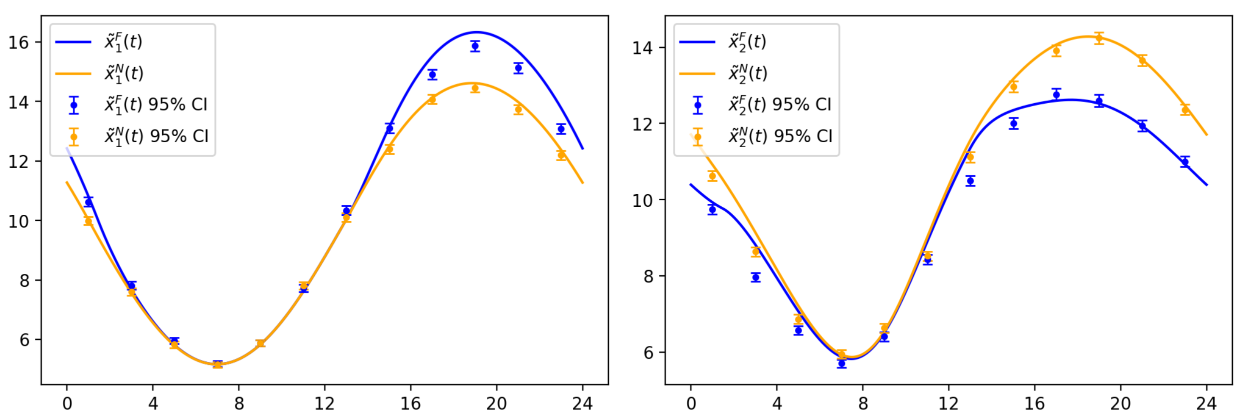


Figure 13 Comparisons of the number-in-system trajectories under different information levels for small size simulated stochastic systems and the fluid models ($s = 10$, $\mu_1 = \mu_2 = 1$, $\theta(x) = 1.5 - e^{-x}, \lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$**).**

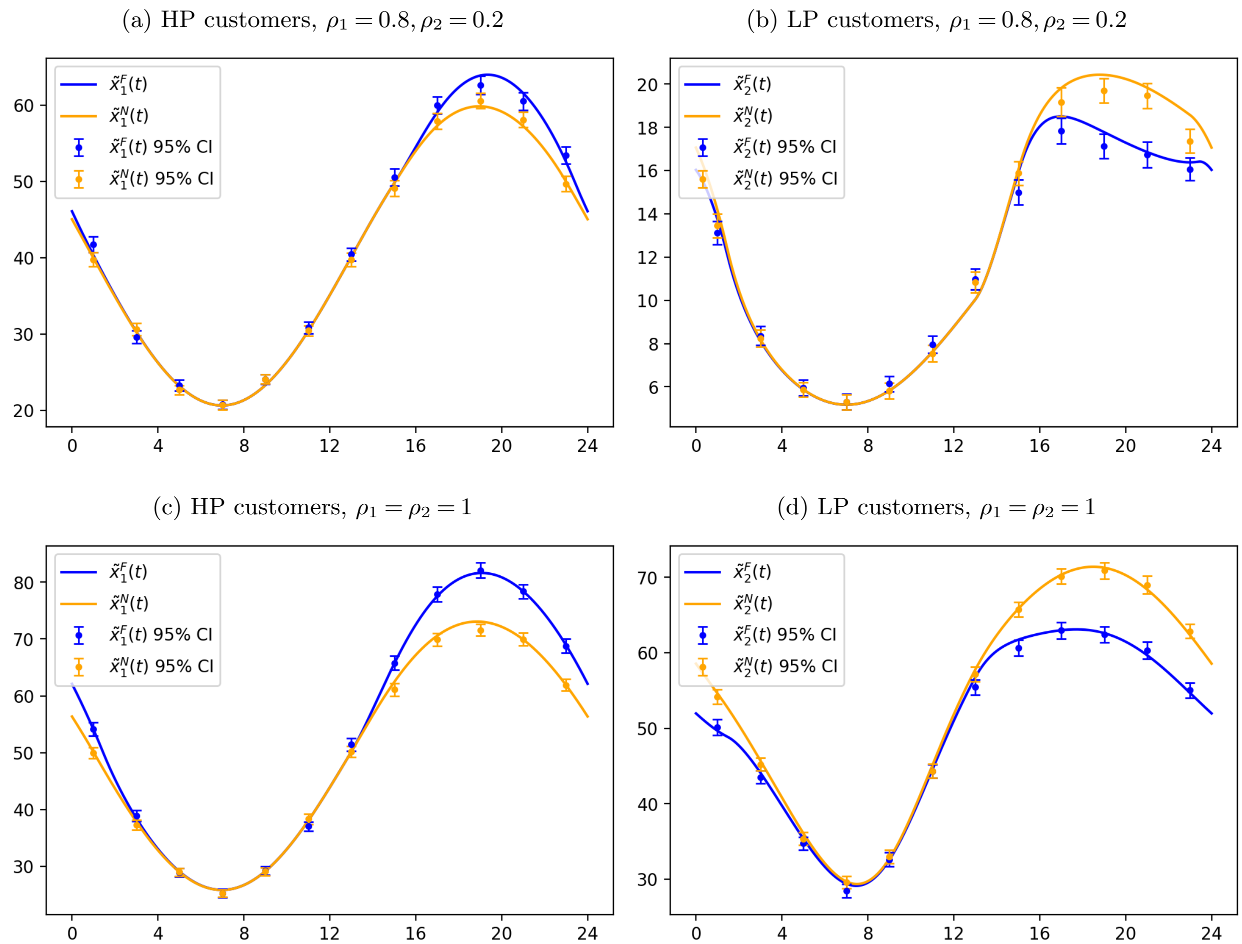


Figure 14 Comparisons of the number-in-system trajectories under different information designs for large size simulated stochastic systems and the fluid models ($s = 50$, $\mu_1 = \mu_2 = 1$, $\theta(x) = 1.5 - e^{-x}, \lambda_k(t) = \rho_k \mu_k s(1 - 0.5 \sin(\pi t/12))$).

($s = 50$). And the rankings of the simulated average number-in-system stochastic processes are consistent with the fluid-based ranking, at each point in time, for both system sizes.

As for the robustness of the fluid-based average system abandonment rate rankings for the stochastic systems, we estimate the expected average system abandonment rates of the stochastic systems for each information level and obtain their rankings (statistically significant at 95% confidence level) using simulation. The results are shown in Table 4.

Table 4 shows that full information yields lower HP average system abandonment and higher LP average system abandonment than no information in both simulation- and fluid-based systems, across all three cases where LP load ranges from low to high. In conclusion, Table 4 indicates that our fluid-based average system abandonment rankings are valid for small-sized stochastic systems.

Next, we examine the accuracy of the fluid-based average system abandonment rankings for the stochastic systems. In particular, for various sets of parameters, we estimate the expected average system abandonment rates of the stochastic systems under no and full information and obtain their rankings (statistically

Table 4 Comparison of the simulation- and fluid-based average system abandonment rankings ($\mu_1 = \mu_2 = 1$, $s = 20$, $\theta(x) = 1.5 - e^{-x}$, and $\lambda_k(t) = \rho_k \mu_k s(1 - 0.5\sin(\pi t/12))$).

(ρ_1, ρ_2)	$\boldsymbol{\beta}$	Class (k)	simulation-based ranking		fluid-based ranking	
			$\bar{\mathcal{A}}_k^N$	$\bar{\mathcal{A}}_k^F$	$\bar{A}_k^N$	$\bar{A}_k^F$
(0.8,0.2)	(1,0.43)	1	1.2536	1.2071	0.7939	0.7553
(1,1)	(0.43,0.42)	1	3.4263	3.3717	3.0750	3.0171
(0.8,0.2)	(1,0.43)	2	2.0048	2.1402	2.2226	2.2618
(1,1)	(0.43,0.42)	2	16.7253	16.8296	16.9240	16.9819

significant at 95% confidence level) using simulation. We observe that, when $s = 20$, the simulation-based average system abandonment rate rankings are consistent with the fluid-based rankings.

Overall, the results indicate that our fluid-based average number-in-system and abandonment rankings are also valid for small and moderately-sized stochastic systems.

E.2. Comparing information levels with respect to waiting time

Our results focus on the two key performance metrics of queue length and abandonment rate. Another important performance metric is the average waiting time. This metric is connected to the queue length through Little's Law. As such, the comparative results in Section 6 concerning long-run average performance also apply to the waiting time metric. Specifically, by Theorem LL.2 in John (2011) we have that, under information design I, $L_k^I = \lambda_k W_k^I$, where L_k^I is the equilibrium class k average queue length, λ_k is the class k average arrival rate, and W_k^I is the equilibrium class k average waiting time. Note that, by the remarks in Theorem LL.1 in John (2011), Little's law holds irrespective of the queue discipline and under non-stationary arrivals, as long as we are concerned with long-run average performance; see also Theorem 2.1. of Whitt (2015). Therefore, the waiting time rankings are consistent with the average queue length rankings presented in Section 6.

E.3. Non-stationary number of servers

Our paper focuses on systems with a static number of servers, s. In practice, however, the number of servers may be non-stationary and vary over the course of a week or a day to accommodate service providers' preferences or constraints. For cases where the number of servers is non-stationary and periodic, denoted as $s(t)$, our main results generalize as illustrated below.

For the single-class case, analyzing a system with a non-stationary number of servers $s(t)$ is equivalent to focusing on LP customers in a system with two classes, non-stationary arrivals, and a uniformly underloaded HP class, where the comparison results are provided in Proposition 9. This implies that our comparison results for single-class system with a stationary number of servers generalize to the system with a non-stationary number of servers.

For the two-class case, denote $\bar{s} := \max_{t \geq 0} s(t)$ and $\Delta s(t) = \bar{s} - s(t)$, then a two-class system with non-stationary number of servers $s(t)$ and HP arrival rate $\lambda_1(t)$ is equivalent to a system with a stationary number of servers $\bar{s}$ and HP arrival rate $\lambda_1(t) + \mu_1 \Delta s(t)$. As such, our results are also relevant for systems with non-stationary number of servers.