# Asymptotically optimal index policies for an abandonment queue with convex holding cost.

Maialen Larrañaga, Urtzi Ayesta, Ina Maria Maaike Verloop

# Asymptotically optimal index policies for an abandonment queue with convex holding cost*

M. Larrañaga[1,2,5], U. Ayesta[2,3,4,5], I.M. Verloop[1,5]
[1]CNRS, IRIT, 2 rue C. Carmichel, F-31071 Toulouse, France.
[2]CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France
[3]IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain
[4]UPV/EHU, University of the Basque Country, 20018 Donostia, Spain
[5]Univ. de Toulouse, INP, LAAS, F-31400 Toulouse, France

**Abstract**

We investigate a resource allocation problem in a multi-class server with convex holding costs and user impatience under the average cost criterion. In general, the optimal policy has a complex dependency on all the input parameters and state information. Our main contribution is to derive index policies that can serve as heuristics and are shown to give good performance. Our index policy attributes to each class an index, which depends on the number of customers currently present in that class. The index values are obtained by solving a relaxed version of the optimal stochastic control problem and combining results from restless multi-armed bandits and queueing theory. They can be expressed as a function of the steady-state distribution probabilities of a one-dimensional birth-and-death process. For linear holding cost, the index can be calculated in closed-form and turns out to be independent of the arrival rates and the number of customers present. In the case of no abandonments and linear holding cost, our index coincides with the $c\mu$-rule, which is known to be optimal in this simple setting. For general convex holding cost we derive properties of the index value in limiting regimes: we consider the behavior of the index (i) as the number of customers in a class grows large, which allows us to derive the asymptotic structure of the index policies, (ii) as the abandonment rate vanishes, which allows us to retrieve an index policy proposed for the multi-class M/M/1 queue with convex holding cost and no abandonments, and (iii) as the arrival rate goes to either 0 or $\infty$, representing light-traffic and heavy-traffic regimes, respectively. We show that Whittle's index policy is asymptotically optimal in both light-traffic and heavy-traffic regimes. To obtain further insights into the index policy, we consider the fluid version of the relaxed problem and derive a closed-form expression for the fluid index. The latter is shown to coincide with the index values for the stochastic model in asymptotic regimes. For arbitrary convex holding cost the fluid index can be seen as the $Gc\mu/\theta$-rule, that is, including abandonments into the generalized $c\mu$-rule ($Gc\mu$-rule). Numerical experiments for a wide range of parameters have shown that the Whittle index policy and the fluid index policy perform very well for a broad range of parameters.

## 1 Introduction

In this paper our objective is to develop a unifying framework to obtain well performing control policies in a multi-class single-server queue with convex holding costs and impatient customers.

*The PhD fellowship of Maialen Larrañaga is funded by a research grant of the Foundation Airbus Group (http://fondation.airbus-group.com/). A shorter version of this paper was published in the Proceedings of ACM Sigmetrics 2014 [30].

The single-server queue is the canonical model to study resource allocation problems and it can be considered as one of the most classical decision problems. It has been widely studied due to its applicability to any situation where a single-resource is shared by multiple concurrent customers. Abandonment or reneging takes place when customers, unsatisfied of their long waiting time, decide to voluntarily leave the system. It has a huge impact in various real life applications such as the Internet or call centers, where customers may abandon while waiting in the queue, or even while being served. In the presence of abandonments and/or convex holding cost, a characterization of the optimal control is out of reach, due to the curse of dimensionality.

When the holding costs are linear and customers are not impatient, a classical result shows that the celebrated $c\mu$-rule rule is optimal, that is, to serve the classes in decreasing order of priority according to the product $c_k\mu_k$, where $c_k$ is the holding cost per class-$k$ customer, and $\mu_k^{-1}$ is the mean service requirement of class-$k$ customers, [16, 22]. The $c\mu$-rule is a so-called *index policy*, that is, the solution to the stochastic control problem is characterized by an *index*, $c_k\mu_k$, which determines which customer is optimal to serve. This simple structure of the optimal policy vanishes however in the presence of convex costs and/or impatient customers. The optimal policy will in general be a complex function of all the input parameters function and the number of customers present in all the classes.

Optimality of index policies has enjoyed a great popularity. The solution to a complex control problem that, a priori, might depend on the entire state space, turns out to have a strikingly simple structure. For instance, in the case of the $c\mu$-rule, the solution does not depend on the number of customers in the various classes. Another classical result that can be seen as an index policy is the optimality of Shortest-Remaining-Processing-Time (SRPT), where the index of each customer is given by its remaining service time. Both examples fit the general context of Multi-Armed Bandit Problems (MABP). A MABP is a particular case of a Markov Decision Process: at every decision epoch the scheduler needs to select one *bandit*, and an associated reward is accrued. The state of this selected bandit evolves stochastically, while the state of all other bandits remain *frozen*. The scheduler knows the state of all bandits, the rewards in every state, and the transition probabilities, and aims at maximizing the total average reward. In a ground-breaking result Gittins showed that the optimal policy that solves a MABP is an index-rule, nowadays commonly referred to as Gittins' index [23]. Thus, for each bandit, one calculates an index that depends only on its own current state and stochastic evolution. The optimal policy activates in each decision epoch the bandit with highest current index.

Despite its generality, in multiple cases of practical interest the problem cannot be cast as a MABP. In a seminal work [40], Whittle introduced the so-called Restless Multi-Armed Bandit Problems (RMABP), a generalization of the standard MABP. In a RMABP all bandits in the system incur a cost. The scheduler selects a number of bandits to be made active. However, all bandits might evolve over time according to a stochastic kernel that depends on whether the bandit is selected for service or not. The objective is to determine a control policy that optimizes the average performance criterion. RMABP provides a more general modeling framework, but its solution has in general a complex structure that might depend on the entire state-space description. Whittle considered a relaxed version of the problem (where the restriction on the number of *active* bandits needs to be respected on average only, and not in every decision epoch), and showed that the solution to the relaxed problem is of index type, referred to as *Whittle's index.* Whittle then defined a heuristic for the original problem where in every decision epoch the bandit with highest Whittle index is selected. It has been shown that the Whittle index policy performs strikingly well, see [33] for a discussion, and can be shown to be asymptotically optimal, see [38, 36]. The latter explains the importance given in the literature to calculate Whittle's index. In order to calculate Whittle's index there are two main difficulties, first one needs to establish a technical property known as *indexability*, and second the calculation of the index might be involved or even infeasible.

In one of the main contributions of the paper, we verify indexability and calculate Whittle's index for the average cost criterion of the multi-class queue with abandonments and convex cost. In fact, our model can be written as a RMABP where each class is represented by a bandit and the state of a bandit describes the number of customers in that class. The evolution of the number of customers being birth-and-death, the bandit is of birth-and-death type. An important observation we make is that the Whittle index we obtain, which is expressed as a function of the steady-state probabilities, is in fact applicable to any birth-and-death bandit. This is a simple observation that has far reaching consequences since it allows to derive Whittle's index for a general class of control problems, as will be explained in the paper. Note that indexability would be needed to be established on a case-by-case basis. For the abandonment model with convex holding cost, we prove indexability by showing that threshold policies are optimal for the relaxed optimization problem and using properties of the steady-state distributions.

Having characterized Whittle's index in terms of steady-state distributions, we then apply it to various cases. In the case of linear holding cost, we show that the Whittle's index is a constant that is independent of the number of customers in the system and of the arrival rate. In fact, this index policy (with linear holding cost) coincides with the index policies as proposed in [9] and [7], for specific model assumptions, and is asymptotically optimal for a multi-server environment. For general convex holding cost we derive properties of the index value in limiting regimes: we consider the behavior of the index (i) as the number of customers in a class grows large, which allows us to derive the asymptotic structure of the index policies, ii) as the abandonment rate vanishes, which allows us to retrieve an index policy proposed for the multi-class M/M/1 queue with convex holding cost and no abandonments, and (iii) as the arrival rate goes to either 0 or $\infty$, representing light- and heavy-traffic regimes, respectively.

In another main result we show asymptotic optimality of Whittle's index policy in both light traffic and heavy traffic. We do so by establishing that for these two limiting regimes, the solution to the relaxed version of the optimization problem is a feasible policy for the original optimization problem.

Our index is expressed as a function of the steady-state probabilities and it can thus efficiently be calculated, but it does not always allow to obtain qualitative insights. We therefore formulate a fluid version of the relaxed optimization problem, where the objective is *bias optimality*, i.e., to determine the policy that minimizes the cost of bringing the fluid to its equilibrium. We show how to derive an index for the fluid model, and we compare it with Whittle's index as obtained for the stochastic model. The advantage of the fluid approach lies in its relatively simple expressions compared to the stochastic one. It shows equivalence with the $Gc\mu/\theta$-rule, that is, including abandonments into the generalized $c\mu$-rule ($Gc\mu$-rule) and provides useful insights on the dependence on the parameters. For linear holding cost the Whittle index and the fluid index are identical. In asymptotic regimes such as light-traffic, heavy-traffic, and as the value of the state grows large the Whittle index and the fluid index are equivalent.

Numerical experiments show that our index policies, in addition to being optimal in light traffic and heavy traffic, perform very well across all traffic loads.

In summary the main contributions of this paper are:

- Obtain Whittle's index for a multi-class queue with convex holding costs and abandonments under average cost criterion.
- Establish optimality of threshold policies and indexability for the relaxed optimization problem.
- For linear holding costs Whittle's index is independent on the arrival rate and number of

customers present in a class.

- Establish asymptotic optimality of Whittle's index policy in a light-traffic and heavy-traffic regime.
- Development of a fluid-based approach to derive a closed-form index policy for general holding cost.
- Establish equivalence of the fluid index and Whittle's index in the light-traffic regime and as the state of the system grows large.

The paper is organized as follows. In Section 2 we give an overview of related work and in Section 3 we describe the model. In Section 4 we present the relaxation of the original problem and show that threshold policies are optimal. We establish indexability and calculate Whittle's index under the average cost criterion. In Section 5 we explain a heuristic index policy, based on Whittle's index, for the original optimization problem. In Section 6 we calculate Whittle's index for linear holding cost and derive properties for general convex holding costs in several limiting regimes. In Section 7 we calculate the index for an $M/M/1$ queue without abandonments. Section 8 describes our asymptotic optimality results. In Section 9 we present the fluid model and derive the fluid index. Finally, in Section 10 we numerically evaluate the performance of Whittle's index policy and the fluid index policy. Most of the proofs are presented in the appendix.

## 2 Related Work

There are four main literature bodies that are relevant to our work: literature on (i) index policies for resource allocation problems, (ii) scheduling with convex costs, (iii) scheduling in the presence of impatient customers, and (iv) fluid-based scheduling. We provide below a brief summary of some of the main contributions in each of the domains.

(i) The seminal work on the optimality of index policies for MABP is in the book by Gittins et. al. [23]. The optimality of the $c\mu$-rule in a multi-class single server queue, i.e., strict priority is given according to the indices $c\mu$, is shown in [16, 22] in the preemptive and non-preemptive cases. Index policies for RMABP were introduced in the seminal paper [40]. In [33] the author develops an algorithm that allows to establish whether a problem is indexable, and if yes, to numerically calculate, in an efficient way, Whittle's index. Under the assumption that an ODE has an equilibrium point and that all bandits are symmetric, in [38] it is shown that Whittle's index policy is asymptotically optimal as the number of bandits and the number of bandits that can be made active grow to infinity, while their ratio is kept constant. This result is generalized in [36] to the case in which there are various classes of bandits, and new bandits can arrive over time. In addition to resource allocation problems, Whittle's index has been applied in a wide variety of cases, including opportunistic spectrum access, website morphing, pharmaceutical trials and many others, see for example [23, Chapter 6]. The recent survey paper [25] is a good up-to-date reference on the application of index policies in scheduling.

(ii) A seminal paper on scheduling in the presence of convex costs is [35], where the author introduced the Generalized-$c\mu$-rule ($Gc\mu$) and showed its optimality in heavy-traffic for convex delay cost. The $Gc\mu$-rule associates to each class-$i$ customer with experienced delay $d_i$ the index $C_i'(d_i)\mu_i$, where $C_i(\cdot)$ denotes the class-$i$ delay cost. The optimality of the $Gc\mu$-rule in a heavy-traffic setting with multiple servers was established in [32]. In [3] the authors calculate Whittle's index policy for a multi-class queue with general holding cost functions. In [15], convex holding costs are considered as well and, taking a stochastic approach, the author obtains an index rule that consists on first-order differences of the cost function, rather than on its derivatives.

(iii) The impact of abandonments has attracted considerable interest from the research community, with a surge in recent years. To illustrate the latter, we can mention the recent Special Issue in Queueing Systems on queueing systems with abandonments [27] and the survey paper [19] on abandonments in a many-server setting. Related literature that is more close to our present work consists of papers that deal with optimal scheduling or control aspects of multi-class queueing systems in the presence of abandonments, see for instance [24, 4, 6, 20, 7, 9, 5, 28, 29, 13]. Note that, with the exception of [5], these papers consider linear holding cost. In the case of one server, the authors of [20, 13] show that (for exponential distributed service requirements and impatience times) under an additional condition on the ordering of the abandonment rates, an index policy is optimal for linear holding cost. In the case of no arrivals and non-preemptive service, the authors of [4] provide partial characterizations of the optimal policy and show that an optimal policy is typically state dependent. As far as the authors are aware, the above two settings are the only ones for which structural optimality results have been obtained. State-dependent heuristics for the multi-class queue are proposed in [4] for two classes and no arrivals and in [24] for an arbitrary number of classes including new arrivals. In [9] the authors obtain Whittle's index for a multi-class abandonment queue without arrivals, that is, each customer is a bandit and the state of a bandit is either *present* or *departed*. In an overload setting the abandonment queue has been studied under a fluid scaling in [6, 7], where the authors scale the number of servers and the arrival rate and show that an index rule is asymptotically fluid optimal. In our analysis we will show how the indices of [9] and [6, 7] coincide with the Whittle's index rule in the case of linear holding costs and in the presence of arrivals. In [29] the optimal policy is obtained for two classes of customers for a fluid approximation of the stochastic model, which allows to propose a heuristic for the stochastic model for an arbitrary number of classes. We finally mention [5, 28] where the authors derive index policies by studying the Brownian control problem arising in heavy traffic. In [5] general delay costs are considered while in [28] the impatience of customers has a general distribution with increasing failure rate.

(iv) The approach of using the fluid control model to find an approximation for the stochastic optimization problem finds its roots in the pioneering works by Avram et al. [8] and Weiss [39]. It is remarkable that in some cases the optimal control for the fluid model coincides with the optimal solution for the stochastic problem. See for example [8] where this is shown for the $c\mu$-rule in a multi-class single-server queue and [11] where this is shown for Klimov's rule in a multi-class queue with feedback. For other cases, researchers have aimed at establishing that the fluid control is asymptotically optimal, that is, the fluid-based control is optimal for the stochastic optimization problem after a suitable scaling, see for example [10, 21, 37]. We conclude by mentioning that the fluid approach owes its popularity to the groundbreaking result stating that if the fluid model drains in finite time, the stochastic process is stable, see [18].

# 3 Model Description

We consider a multi-class single-server queue with $K$ classes of customers. Class-$k$ customers arrive according to a Poisson process with rate $\lambda_k$ and have an exponentially distributed service requirement with mean $1/\mu_k$, $k = 1, \ldots, K$. We denote by $\rho_k := \lambda_k/\mu_k$ the traffic load of class $k$, and by $\rho := \sum_{k=1}^{K} \rho_k$ the total load to the system. We model abandonments of customers in the following way:

- Any class-$k$ customer not served abandons after an exponentially distributed amount of time with mean $1/\theta_k$, $k = 1, \ldots, K$, with $\theta_k > 0$.

- A class-$k$ customer that is being served abandons after an exponentially distributed amount of time with mean $1/\theta'_k$, $k = 1, \ldots, K$, with $\theta'_k \geq 0$.

The server has capacity 1 and can serve at most one customer at a time, where the service can be preemptive. We make the following natural assumption:

$$\mu_k + \theta'_k \geq \theta_k, \text{ for all } k.$$

That is, the departure rate of a class-$k$ customer is higher when being served than when not being served.

At each moment in time, a policy $\varphi$ decides which class is served. Because of the Markov property, we can focus on policies that only base their decisions on the current number of customers present in the various classes. For a given policy $\varphi$, $N_k^\varphi(t)$ denotes the number of class-$k$ customers in the system at time $t$, (hence, including the one in service), and $\vec{N}^\varphi(t) = (N_1^\varphi(t), \ldots, N_K^\varphi(t))$. Let $S_k^\varphi(\vec{n}) \in \{0, 1\}$ represent the service capacity devoted to class-$k$ customers at time $t$ under policy $\varphi$ in state $\vec{N}(t) = \vec{n}$. The constraint on the service amount devoted to each class is $S_k^\varphi(\vec{n}) = 0$ if $n_k = 0$ and

$$\sum_{k=1}^{K} S_k^\varphi(\vec{n}) \leq 1, \tag{1}$$

and we denote by $\mathcal{U}$ the set of admissible control strategies that satisfy this constraint.

The above describes a birth-and-death process that makes a transitions

$$\vec{n} \to \vec{n} + \vec{e}_k \text{ with rate } \lambda_k, \text{ and,}$$

$$\vec{n} \to \vec{n} - \vec{e}_k \text{ with rate } \mu_k S_k^\varphi(\vec{n}) + \theta_k(n_k - S_k^\varphi(\vec{n})) + \theta'_k S_k^\varphi(\vec{n}),$$

for $n_k > 0$, with $\vec{e}_k$ a $K$-dimensional vector with all zeros except for the $k$-th component which is equal to 1.

Let $C_k(n, a)$ denote the cost per unit of time when there are $n$ class-$k$ customers in the system and when either class $k$ is not served (if $a = 0$), or when class $k$ is served (if $a = 1$). We assume $C_k(\cdot, 0)$ and $C_k(\cdot, 1)$ are convex and non-decreasing functions and satisfy

$$C_k(n, 0) - C_k((n-1)^+, 0) \leq C_k(n+1, 1) - C_k(n, 1) \leq C_k(n+1, 0) - C_k(n, 0), \tag{2}$$

for all $n \geq 0$. Observe that if $C_k(0, 0) \geq C_k(0, 1)$, then (2) implies that, for all $n$, $C_k(n, 0) \geq C_k(n, 1)$. We also note that (2) is always satisfied when (i) $C_k(n, a) = C_k(n)$, or when (ii) $C_k(n, a) = C_k((n-a)^+)$. Case (i) represents holding costs for customers in the *system*, while (ii) represents holding costs for customers in the *queue*.

We further introduce a cost $d_k$ for every class-$k$ customer that abandons the system when not being served and a cost $d'_k$ for a class-$k$ customer that abandons the system while being served.

The objective of the optimization is to find the optimal scheduling policy, denoted by $OPT$, under the average-cost criteria, that is, find the policy $\varphi$ that minimizes

$$\mathcal{C}^\varphi := \limsup_{T \to \infty} \sum_{k=1}^{K} \frac{1}{T} \mathbb{E}\bigg( \int_0^T C_k(N_k^\varphi(t), S_k^\varphi(\vec{N}^\varphi(t))) \, \mathrm{d}t + d_k R_k^\varphi(T) + d'_k R_k'^\varphi(T) \bigg), \tag{3}$$

where $R_k^\varphi(T)$ and $R_k'^\varphi(T)$ denote the number of class-$k$ customers that abandoned the queue while waiting and while being served, respectively, in the interval $[0, T]$ under policy $\varphi$. We denote by $\mathcal{C}^{OPT} = \inf_{\varphi \in \mathcal{U}} \mathcal{C}^\varphi$ the average cost under the optimal policy.

We have

$$\mathbb{E}(R_k^{\varphi}(T)) = \theta_k \mathbb{E}\left(\int_0^T (N_k^{\varphi}(t) - S_k^{\varphi}(\vec{N}^{\varphi}(t)))\mathrm{d}t\right)$$

and

$$\mathbb{E}(R_k'^{\varphi}(T)) = \theta_k' \mathbb{E}\left(\int_0^T S_k^{\varphi}(\vec{N}^{\varphi}(t))\mathrm{d}t\right),$$

by Dynkin's formula [2, Chapter 6.5]. We introduce the following notation:

$$\tilde{C}_k(n_k, a) := C_k(n_k, a) + d_k\theta_k(n_k - a)^+ + d_k'\theta_k' \min(a, n_k), a \in \{0, 1\} \tag{4}$$

so that the objective (3) can be equivalently written as

$$\limsup_{T\to\infty} \sum_{k=1}^{K} \frac{1}{T}\mathbb{E}\left(\int_0^T \tilde{C}_k(N_k^{\varphi}(t), S_k^{\varphi}(\vec{N}^{\varphi}(t)))\ \mathrm{d}t\right). \tag{5}$$

The above described stochastic control problems have proved to be very difficult to solve. Already for the special case of linear holding cost, deriving structural properties of optimal policies is extremely challenging. For example, in [20] optimal dynamic scheduling is studied for two classes of customers (K = 2), with $d_k = d_k'$, $\theta_k = \theta_k'$, $\mu_1 = \mu_2 = 1$, and linear holding cost, $C_k(n, a) = c_k n$. Define $\tilde{c}_k := c_k + d_k\mu_k$. For the special case where $\tilde{c}_1 \geq \tilde{c}_2$ and $\theta_1 \leq \theta_2$, the authors show that it is optimal to give strict priority to class 1, see [20, Theorem 3.5]. It is intuitively clear that giving priority to class 1 is the optimal thing to do, since serving class 1 myopically minimizes the (holding and abandonment) cost and in addition it is advantageous to keep the maximum number of class-2 customers in the system (without idling), since they have the highest abandonment rate. In [13] optimal dynamic scheduling is studied for $C_k(n, a) = c_k n$, $d_k = d_k'$, and either $\theta_k = \theta_k'$ or $\theta_k' = 0$. For the special case where the classes can be ordered such that $\tilde{c}_1 \geq \cdots \geq \tilde{c}_K$, $\tilde{c}_1(\mu_1 + \theta_1' - \theta_1) \geq \cdots \geq \tilde{c}_K(\mu_K + \theta_K - \theta_K')$, and $\tilde{c}_1(\mu_1 + \theta_1' - \theta_1)/\theta_1 \geq \ldots \geq \tilde{c}_K(\mu_K + \theta_K' - \theta_K)/\theta_K$, the authors show that it is optimal to give strict priority according to the ordering $1 > 2 > \ldots > k$.

Outside these special parameter settings, or for convex holding cost, an optimal policy is expected to be state dependent, and as far as the authors are aware, no (structural) results exist for this stochastic optimal control problem.

In order to obtain insights into optimal control for *convex* holding cost, in this paper we will solve a relaxed version of the optimization problem. The latter allows us to propose a heuristic for the original model, which we will prove to be optimal in light and heavy traffic. The details of the relaxation technique are described in the next section.

# 4 Relaxation and Indexability

The solution to (5) under constraint (1) cannot be solved in general. Following Whittle [40], we study the relaxed problem in which the constraint on the service devoted to each class must be satisfied on *average*, and not in every decision epoch. The control policy must thus satisfy

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^{K} S_k^{\varphi}(\vec{N}^{\varphi}(t))\mathrm{d}t\right) \leq 1, \tag{6}$$

or equivalently $\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^{K}(1 - S_k^{\varphi}(\vec{N}^{\varphi}(t)))\mathrm{d}t\right) \geq K - 1$. We denote by $\mathcal{U}^{REL}$ the set of policies that satisfy (6), and we note that $\mathcal{U} \subseteq \mathcal{U}^{REL}$.

The objective of the relaxed problem is hence to determine the policy that solves (5) under constraint (6). A standard Lagrangian argument shows that this problem can be solved by considering the following unconstrained control problem: find a policy $\varphi$ that minimizes

$$\mathcal{C}^{\varphi}(W) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \left(\sum_{k=1}^{K} \tilde{C}_k(N_k^{\varphi}(t), S_k^{\varphi}(\vec{N}^{\varphi}(t))) - W(1-K+\sum_{k=1}^{K}(1-S_k^{\varphi}(\vec{N}^{\varphi}(t))))\right) \mathrm{d}t\right), \quad (7)$$

where $W$ is the Lagrange multiplier. For a given $W$, let $REL(W)$ denote a policy that minimizes (7), and let $\mathcal{C}^{REL(W)}(W) := \min_{\varphi\in\mathcal{U}^{REL}} \mathcal{C}^{\varphi}(W)$ denote the optimal performance of the relaxed problem. For any value of the multiplier $W \geq 0$, it holds that $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$. To see this, note that for a given $W \geq 0$ and $\varphi \in \mathcal{U}$ it holds that

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{\varphi}(W) \leq \mathcal{C}^{\varphi}.$$

The first inequality follows by definition of $REL(W)$, and the second inequality follows from the fact that $1-K+\sum_{k=1}^{K}(1-S_k^{\varphi}(\vec{N}^{\varphi}(t))) \geq 0$ for a policy $\varphi \in \mathcal{U}$.

The key observation made by Whittle is that problem (7) can be decomposed into $K$ subproblems, each corresponding to a different class (or bandit when using terminology from the RMABP literature). Thus, the solution to (7) is obtained by combining the solution to $K$ separate optimization problems. For the remainder of this section we focus on the optimization problem of one class and drop the dependency on the class from the notation. For a given $W$ we hence consider the individual optimization problem for a given class, that is, minimize

$$g^{\varphi}(W) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \left(\tilde{C}(N^{\varphi}(t), S^{\varphi}(N^{\varphi}(t))) - W(1-S^{\varphi}(N^{\varphi}(t)))\right) \mathrm{d}t\right), \quad (8)$$

where now $N^{\varphi}(t)$ is the state of a given class at time $t$. Under a stationarity assumption, we can invoke ergodicity to show that (8) is equivalent to minimizing

$$g^{\varphi}(W) = \mathbb{E}(\tilde{C}(N^{\varphi}, S^{\varphi}(N^{\varphi})) - W\mathbb{E}(\mathbb{1}_{S^{\varphi}(N^{\varphi})=0}), \quad (9)$$

where $N^{\varphi}$ denotes the steady-state number of customers in a class under policy $\varphi$. We observe that the multiplier $W$ can be interpreted as a *subsidy for passivity.*

In summary, the relaxed optimization problem can be written as $K$ independent one-dimensional Markov Decision Problems (8). In Section 4.1 we will determine the structure of the optimal control of the relaxed problem (8). In Section 4.2 and Section 4.3 we derive Whittle's index and describe the optimal solution of the relaxed problem.

## 4.1 Threshold policies

In the following proposition we show that an optimal solution of the relaxed problem (8) is of threshold type, i.e., when the number of customers is above a certain threshold $n$, the class is served, and not served otherwise. We denote by $\varphi = n$, $n = -1, 0, 1, \ldots$, the threshold policy with threshold $n$, that is, $S^n(m) = 1$ if $m > n$, and $S^n(m) = 0$ otherwise.

**Proposition 1** *There is an $n = -1, 0, 1, \ldots,$ such that the policy $\varphi = n$ is an optimal solution of the relaxed problem (8).*

**Proof.** The value function $V(n)$ satisfies the Bellman optimality equation for average cost models [34], that is,

$$(\mu+\theta'+m\theta+\lambda)V(m)+g=\lambda V(m+1)+\theta(m-1)V((m-1)^+)$$
$$+\min\{\tilde{C}(m,0)-W+(\mu+\theta')V(m)+\theta V((m-1)^+),\tilde{C}(m,1)+(\mu+\theta')V((m-1)^+)+\theta V(m)\},\tag{10}$$

where $g$ is the average cost incurred under an optimal policy. Proving optimality of a threshold policy is hence equivalent to showing that if it is optimal in (10) for state $m+1$, $m\geq 0$ to be passive, then it is also optimal in (10) for state $m$ to be passive, i.e., $\tilde{C}(m+1,0)-W+(\mu+\theta'-\theta)V(m+1)\leq \tilde{C}(m+1,1)+(\mu+\theta'-\theta)V(m)$, implies $\tilde{C}(m,0)-W+(\mu+\theta'-\theta)V(m)\leq \tilde{C}(m,1)+(\mu+\theta'-\theta)V((m-1)^+)$. A sufficient condition for the above to be true is (2) together with the inequality $V(m+1)+V((m-1)^+)\geq 2V(m)$, for $m\geq 0$. The latter condition, convexity of the value function, will be proved below, which concludes the proof.

In case of bounded transition rates, one can uniformize the system and use value iteration in order to prove convexity. However, our transition rates are unbounded. We therefore consider the truncated space, truncated by $L>1$, and smooth the arrival transition rates from $m$ to $m+1$ as follows:

$$q^{\varphi,L}(m,m+1):=\lambda\left(1-\frac{m}{L}\right)^+=\lambda\max\left(0,1-\frac{m}{L}\right),$$

$m=0,\ldots,L$. Denote by $V^L(m)$ the value function of the $L$-truncated system. After verifying two conditions, (as done in Appendix A.1), we have by [14, Theorem 3.1] that $V^L(m)\to V(m)$ as $L\to\infty$. Hence, convexity of the function $V$ is implied by convexity of $V^L$ for all $L$, and we are left with proving the latter. The latter is uniformizable, hence we can use the value iteration technique in order to prove convexity of $V^L$. This proof is available in Appendix A.2. ■

Below we write the steady-state distribution of threshold policy $\varphi=n$. We denote the steady-state probability of being in state $i$ under policy $\varphi=n$ by $\pi^n(i)$. We have

$$\pi^n(i)=\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\pi^n(0),\quad i=1,2,\ldots,\tag{11}$$

where $\pi^n(0)=\left(1+\sum_{i=1}^{\infty}\prod_{m=1}^{i}\dfrac{q^n(m-1,m)}{q^n(m,m-1)}\right)^{-1}$ and

$$q^n(m,m-1):=\begin{cases}\theta m & \text{for all } m\leq n,\\ \mu+\theta'+\theta(m-1) & \text{for all } m>n,\end{cases}$$
$$q^n(m,m+1):=\lambda,\qquad \text{for all } m.\tag{12}$$

**Remark 1** *In Proposition 1 we established optimality of threshold policies for problem (8) in the case $\mu+\theta'\geq\theta$ and when (2) is satisfied.If instead $\mu+\theta'<\theta$, and in addition $\tilde{C}(m,1)>\tilde{C}(m,0)$ for all $m$ (but without requiring (2) to hold), then (for $W\geq 0$) the optimal policy is to be passive in all states $m$. This can be easily seen from Equation* (10)*, since being always passive is optimal if for all $m$*

$$\tilde{C}(m,0)-W+(\mu+\theta'-\theta)V(m)\leq\tilde{C}(m,1)+(\mu+\theta'-\theta)V((m-1)^+).$$

*The latter follows from the above assumptions and the fact that the value function $V$ is non-decreasing. The proof of $V$ being a non-decreasing function follows as in Appendix A.2.*

*In other cases, we have numerically observed that threshold policies are optimal, but we have not established this formally.*

### 4.2 Indexability and Whittle's index

Indexability is the property that allows to develop a heuristic for the original problem. This property requires to establish that as the subsidy for passivity, $W$, increases, the collection of states in which the optimal action is *passive* increases. It was first introduced by Whittle [40] and we formalize it in the following definition.

**Definition 1** *A class is indexable if the set of states in which passive is an optimal action (denoted by $D(w)$) increases in $W$, that is, $W' < W \Rightarrow D(W') \subseteq D(W)$.*

Note that an optimal solution of problem (8) is a threshold policy, or more specifically, if it is optimal to be passive in state $m$, $m \geq 1$, then it is also optimal to be passive in state $m-1$, see the proof of Proposition 1. We can therefore equivalently write the following definition for indexability.

**Definition 2** *Let $n(W)$ denote the largest value of $n$ such that the threshold policy $n$ minimizes (8). A class is indexable if $n(W)$ is non-decreasing in $W$, that is, $W' < W \Rightarrow n(W') \leq n(W)$.*

Provided we can establish indexability, the Whittle index in a state $m$ is defined as the smallest value for the subsidy such that it is optimal to be passive in state $m$. Formally:

**Definition 3** *When a class is indexable, the Whittle index in state $m$ is defined by $W(m) := \inf\{W : m \leq n(W)\}$.*

The solution to the relaxed control problem (7), *i.e.*, $REL(W)$, will then be to activate all classes $k$ that are in a state $n_k$ such that their Whittle's index exceeds the subsidy for passivity, i.e., $W_k(n_k) > W$. A standard Lagrangian argument shows that there exists a value of $W$ (possibly negative) for which the constraint (6) is binding, i.e., the optimal policy $\varphi$ that solves Problem (7) will on average activate one class.

Obviously, the solution to the relaxed optimization problem is not feasible for the original problem. Following Whittle, we use Whittle's index to construct the following heuristic for the original problem (5) under the constraint (1): select in every decision epoch the class with *largest* Whittle index. We will formally describe this in Section 5.

To conclude this subsection we show that for the model under consideration, the classes are indexable.

**Proposition 2** *All classes are indexable.*

**Proof.** Since an optimal policy for (8) is of threshold type, for a given subsidy $W$ the optimal average cost is given by $g(W) := \min_n\{g^{(n)}(W)\}$, where

$$g^{(n)}(W) := \sum_{m=0}^{\infty} \tilde{C}(m, S^n(m))\pi^n(m) - W \sum_{m=0}^{n} \pi^n(m), \tag{13}$$

is the average cost under threshold policy $n$. The function $g(W)$ is a lower envelope of affine non-increasing functions in $W$ (see Figure 1, where we depict the lower-envelope for the case of quadratic cost). It thus follows that $g(W)$ is a concave non-increasing function.

It follows directly that the right-derivative of $g(W)$ in $W$ is given by $-\sum_{m=0}^{n(W)} \pi^{n(W)}(m)$. Moreover, we will prove below that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in $n$. Since $g(W)$ is concave in $W$, its first derivative $-\sum_{m=0}^{n(W)} \pi^{n(W)}(m)$ is non-increasing in $W$. It hence follows that $n(W)$ is non-decreasing in $W$, that is, this class is indexable (see Definition 2).

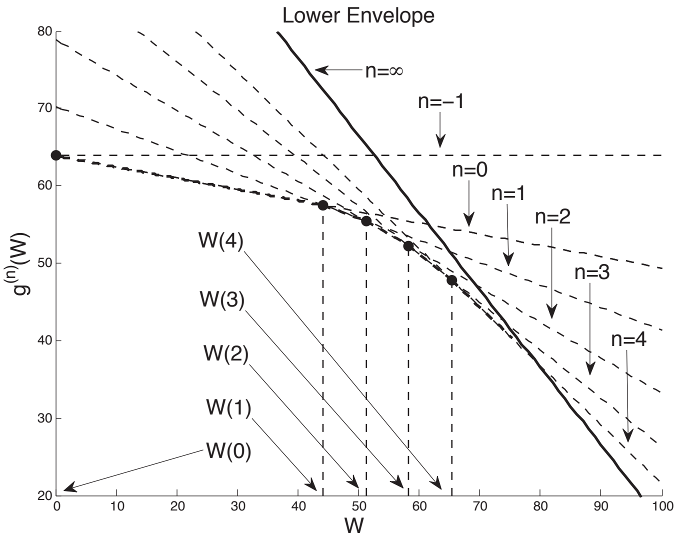


Figure 1: Lower envelop $g = \min_n\{g^{(n)}\}$ when $\tilde{C}(n,a) = (1+2\theta)n + 3n^2$, for $a = 0, 1$, and $\theta = 6, \lambda = 23, \mu = 10$.

We now prove that $\sum_{i=0}^{n} \pi^n(i)$ is strictly increasing in $n$, or equivalently, that $1 - \sum_{i=n+1}^{\infty} \pi^n(i)$ is strictly decreasing in $n$. Using (11), the latter is equivalent to verifying that

$$\frac{\sum_{m=n+1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{\sum_{m=n}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}} < \frac{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}}, \tag{14}$$

holds for all $n$, where $q^n(\cdot,\cdot)$ are defined in (12). Note that $q^n(m-1,m) = q^{n-1}(m-1,m)$ for all $m$ and $q^n(m,m-1) = q^{n-1}(m,m-1)$ for all $m \neq n$. From the assumption $\mu + \theta' \geq \theta$ we have $q^n(n,n-1) \leq q^{n-1}(n,n-1)$. Hence, the left-hand-side of (14) is strictly less than 1, while the right-hand-side is larger than or equal to 1. This proves (14). ■

### 4.3 Derivation of Whittle's index

We are now in position of deriving Whittle's index. An optimal policy is fully characterized by a threshold $n$ such that the passive action is prescribed for states $m \leq n$, and the active action for states $m > n$. Our key observation to derive Whittle's index is that it is not necessary to solve the optimality equation (10), but that it suffices to determine the average cost for threshold policies. In turn, the average reward $g$ can be expressed as a function of the steady-state probabilities, which in the case of birth-and-death processes has a well-known solution.

We can now state one of the main results of the paper, which describes the steps to obtain Whittle's index. The proof of Theorem 1 can be found in Appendix B.

**Theorem 1** *Whittle's index values are computed by the following steps:*

- **Step 0** *Compute*

$$W_0 = \inf_{n \in \mathbb{N}\cup\{0\}} \frac{\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - \mathbb{E}(\tilde{C}(N^{-1}, S^{-1}(N^{-1})))}{\sum_{m=0}^{n} \pi^n(m)},$$

*and name by $n_0$ the largest minimizer. Then, define $W(n) := W_0$ for all $n \leq n_0$. If $n_0 = \infty$ define $W(n) := W_0$ for all $n > n_0$, otherwise go to Step 1.*

- **Step $j$** *Compute*

$$W_j = \inf_{n\in\mathbb{N}\setminus\{0,\dots,n_{j-1}\}} \frac{\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - \mathbb{E}(\tilde{C}(N^{n_{j-1}}, S^{n_{j-1}}(N^{n_{j-1}})))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n_{j-1}} \pi^{n_{j-1}}(m)}, j \geq 1,$$

*and name by $n_j$ the largest minimizer. Then, define $W(n) := W_j$ for all $n_{j-1} < n \leq n_j$. If $n_j = \infty$ then $W(n) = W_j$ for all $n > n_j$, otherwise go to step $j+1$.*

In the next corollary we characterize the Whittle index in the particular case in which $n_i = i$ for all $i \in \mathbb{N} \cup \{0\}$, with $n_i$ as defined in Theorem 1.

**Corollary 1** *If*

$$\frac{\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - \mathbb{E}(\tilde{C}(N^{n-1}, S^{n-1}(N^{n-1})))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)}, \tag{15}$$

*is non-decreasing in $n$, then Whittle's index $W(n)$ is given by (15). In particular, $W(0) = \tilde{C}(0,0) - \tilde{C}(0,1)$.*

**Proof.** Let $\tilde{W}(n)$ be the value for the subsidy such that the average cost under threshold policy $n$ is equal to that under policy $n-1$. Hence, using (9), we have that for all $n \geq 1$, $\mathbb{E}(\tilde{C}(N^n, S^n(N^n)) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^n(N^n)=0})$ is equal to $\mathbb{E}(\tilde{C}(N^{n-1}, S^{n-1}(N^{n-1})) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^{n-1}(N^{n-1})=0})$. For threshold policy $n$ we have $\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}) = \sum_{m=0}^{n} \pi^n(m)$, hence $\tilde{W}(n)$ is given by (15).

A direct consequence of Theorem 1 is that $\tilde{W}(n)$ being non-decreasing, implies that $g(\tilde{W}(n)) = g^{(n)}(\tilde{W}(n)) = g^{(n-1)}(\tilde{W}(n))$. We therefore have $g(W) = g^{(n-1)}(W)$ for $\tilde{W}(n-1) \leq W \leq \tilde{W}(n)$. This implies that Whittle's index is given by $W(n) = \tilde{W}(n)$.

To show that $W(0) = \tilde{C}(0,0) - \tilde{C}(0,1)$, observe that $\pi^0(m) = \pi^{-1}(m)$ for all $m$. Hence,

$$\begin{aligned} W(0) &= \frac{\mathbb{E}(\tilde{C}(N^0, S^0(N^0))) - \mathbb{E}(\tilde{C}(N^{-1}, S^{-1}(N^{-1})))}{\pi^0(m)} = \frac{\tilde{C}(0,0)\pi^0(0) - \tilde{C}(0,1)\pi^{-1}(0)}{\pi^0(0)} \\ &= \tilde{C}(0,0) - \tilde{C}(0,1), \end{aligned}$$

where the first equality holds due to $W(n)$ being non-decreasing. ■

Whittle's index as defined in Theorem 1 and Equation (15) can be numerically computed, since the cost function and the steady-state probabilities are known. In Section 6 closed-form expressions and limiting properties for Whittle's index will be derived for special cases.

We could not prove that Whittle's index $W(n)$ as given in (15) is non-decreasing in $n$. However, in many particular cases this property can be established. For instance,

- In the case $\mu + \theta' = \theta$, we have $\pi^n(m) = \pi^{n-1}(m)$ for all $m$. Hence (15) can be written as

$$\frac{\tilde{C}(n,0)\pi^n(n) - \tilde{C}(n,1)\pi^{n-1}(n)}{\pi^n(n)} = \tilde{C}(n,0) - \tilde{C}(n,1). \tag{16}$$

  By condition (2) we have that $\tilde{C}(n,0) - \tilde{C}(n,1) \leq \tilde{C}(n+1,0) - \tilde{C}(n+1,1)$, hence (15) is non-decreasing in $n$. This implies that Whittle's index is given by (16).

- In Proposition 3 it will be proved that when $C(n,a)$ is linear in $n$, (15) is a constant and therefore non-decreasing in $n$. Hence, Whittle's index is given by (15).

A few comments are in order. The first concerns the form of (15). The numerator in (15) can be interpreted as the increase in cost by deciding to become passive in state $n$ and keeping all other actions unchanged, and similarly, the denominator can be understood as the corresponding increase of passivity rate for the process, measured by the additional probability in which a subsidy is received. Thus, $W(n)$ can be interpreted as a measure of increased cost per unit of increased passivity, a term coined as Marginal Productivity Index by Niño-Mora [33].

The second comment regards the applicability of Whittle's index (15) in other contexts. Indeed we can outline a general recipe to develop Whittle's indices for bandits whose evolution can be described by general birth-and-death processes:

(i) Establish optimality of monotone policies (as in Proposition 1).
(ii) Establish indexability (as in Proposition 2).
(iii) If (i) and (ii) can be established, then Whittle's index is given by Proposition 1, where the steady-state probabilities are as in (11).

Steps (i) and (ii) are model dependent. Step (iii) is immediate and the index will always be given by Proposition 1.

To the best of our knowledge, it has not been reported previously that for bandits whose evolution can be described by a birth-and-death process, one can get an explicit closed-form expression for Whittle's index. Perhaps a reason for this lies in the difficulty to solve the optimality equation (10), which has two unknowns $g$ and $V(m)$, This has led researchers to circumvent this difficulty by considering the discounted cost first, equating the total discounted costs as done in Theorem 1 for average cost and then taking the limit in order to retrieve an index for the average cost case. This is for instance the approach taken in [3] to derive an index for convex costs without abandonments or in [23, Section 6.5] for bi-directional bandits in which the *active* and *passive* actions push the process in opposite directions. In [26] the authors develop an algorithm to calculate an index in a multi-class queue with admission control. All these models have in common that after the relaxation, the bandits are birth-and-death, and the obtained Whittle's index is thus equal to (15). We will explain in Section 7 how to derive the index of [3] using the approach as taken in our paper. Regarding the bi-directional bandit it can be directly checked that index (15) is equivalent to the index of [23, Theorem 6.4]. Finally, we note that by adapting the cost structure we obtain that index (15) is equivalent to that of [26, Theorem 2].

Having made this remark on the applicability of (15) in a wider context, in the remainder of the paper we will discuss the properties of Whittle's index (15) in the context of a queue with convex costs and abandonments.

## 5 Whittle's index policy

In this section we describe how the solution to the relaxed optimization problem is used to obtain a heuristic for the original stochastic model. The optimal control for the relaxed problem is not feasible for the original stochastic model, since in the latter at most one class can be served at a time. Whittle [40] therefore proposed the following heuristic, which is nowadays known as Whittle's index policy:

**Definition 4 (Whittle's index policy)** *Assume at time $t$ we are in state $\vec{N}(t) = \vec{n}$. Whittle's index policy prescribes to serve the class $k$ having currently the highest non-negative Whittle's index $W_k(n_k)$, as defined in Proposition 1 (after adding subscript $k$). We refer to this policy as $WI$.*

Note that in case all classes have a negative index, we define that Whittle's index policy will keep the server idle (until there is a class having a positive value for its index). This follows, since, when the Whittle's index is negative, in the relaxed problem you will keep the class passive even though a negative subsidy is given. A formal explanation is given in [36] by the introduction of dummy bandits.

When $\tilde{C}_k(m_k, 0) \geq \tilde{C}_k(m_k, 1)$ for all $m_k$, the Whittle index $W_k(n_k)$ will always be positive. This can be seen as follows. Recall that $W_k(n_k)$ refers to the value of $W$ such that a threshold policy $n_k$ is an optimal solution of the relaxed problem. Hence, for all $m_k \leq n_k$, it is optimal to keep the class passive, that is, $\tilde{C}_k(m_k, 0) - W_k(n_k) + (\mu_k + \theta'_k - \theta_k)V(m_k) \leq \tilde{C}_k(m_k, 1) + (\mu_k + \theta'_k - \theta_k)V(m_k - 1)$, as we saw in the proof of Proposition 1. Since $\tilde{C}_k(m_k, 0) \geq \tilde{C}_k(m_k, 1)$, $\mu_k + \theta'_k \geq \theta_k$, and $V(\cdot)$ is non-decreasing (see proof of Proposition 1), it follows that $W_k(n_k) \geq 0$.

Instead, when $\tilde{C}(m_k, 0) < \tilde{C}(m_k, 1)$ for an $m_k$, $W_k(n_k)$ can be negative for certain states $n_k$. For example, when $\theta'_k = \theta_k$ and $d'_k \gg d_k$. Then, even though the total departure rate of class-$k$ customers is highest when serving class $k$ ($\mu_k + \theta'_k \geq \theta_k$), for certain states $n_k$ it might be better not to serve class $k$. The latter follows since having a class-$k$ customer abandon while being served, will incur a much higher cost than when it abandons while waiting. Hence, a negative subsidy, that is, a cost, is needed in order for it to be optimal to serve class $k$.

From the practical point of view, the interest of Whittle's index $W_k(n_k)$ as defined in Theorem 1 (after adding subscript $k$) lies in the fact that the index of class $k$ does not depend on the number of customers present in the other classes $j$, $j \neq k$. Hence, it provides a systematic way to derive implementable policies which we will show perform very well, see Section 10, and are asymptotically optimal in certain settings, see Section 8.

# 6 Case studies

In this section we further investigate properties of the obtained Whittle's index in Theorem 1. In Section 6.1 we obtain that the index is state-independent for linear holding cost. In Section 6.2 we derive asymptotic properties of the index for general convex holding cost functions.

## 6.1 Linear holding cost

We consider here linear holding cost, that is, $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. Hence, under this function, any class-$k$ customer in the queue contributes with $c_k$ to the cost, and a class-$k$ customer in service contributes with $c'_k$ to the cost. In particular, if $c'_k = c_k$, then $C_k$ represents the linear holding cost of customers in the *system* and if $c'_k = 0$ then $C_k$ represents the linear holding cost of customers in the *queue*. These two holding cost functions have been considered in the literature in the context of abandonments, for example [9] considers the former, while [6] takes the latter. From our formula (15) we will be able to obtain a full characterization of Whittle's index. Interestingly, we show that the Whittle's index becomes state-independent and does not depend on the arrival rate $\lambda_k$.

It will be convenient to define $\tilde{c}_k := c_k + d_k\theta_k$, $k = 1, \ldots, K$, which can be interpreted as the total cost per unit of time incurred by a class-$k$ customer in the queue. Similarly, $\tilde{c}'_k := \tilde{c}'_k + d'_k\theta'_k$ denotes the total cost per unit of time incurred by a class-$k$ customer in service.

We now present the Whittle index for linear holding cost. The proof can be found in Appendix C.

**Proposition 3** *Assume linear holding cost $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. Then, the Whittle index for class $k$ is*

$$W_k(n_k) = \frac{\tilde{c}_k(\mu_k + \theta'_k)}{\theta_k} - \tilde{c}'_k, \text{ for all } n_k. \tag{17}$$

| $W_k(n_k)$ | $\theta'_k = \theta_k, d'_k = d_k$ | $\theta'_k = 0$ |
|---|---|---|
| $c'_k = c_k$ | $\frac{\tilde{c}_k \mu_k}{\theta_k}$ | $\frac{\tilde{c}_k \mu_k}{\theta_k} - c_k$ |
| $c'_k = 0$ | $\frac{\tilde{c}_k \mu_k}{\theta_k} + c_k$ | $\frac{\tilde{c}_k \mu_k}{\theta_k}$ |

Table 1: $W_k(n_k)$ for linear holding cost as in Proposition 3

An interesting feature of (17) is that it is independent of the arrival rate $\lambda_k$ and of the number of class-$k$ customers present, $n_k$. In Section 6.2 we will show that this observation only holds for linear holding costs.

The index (17) allows for the following interpretation. Consider there is only one class-$k$ customer in the system and no future arrivals, we then have $\tilde{C}_k(1,1) = \tilde{c}'_k$, $\tilde{C}(1,0) = \tilde{c}_k$, $q^1_k(1,0) = \theta_k$, $q^0_k(1,0) = \mu_k + \theta'_k$. Index (17) can equivalently be written as $(\mu_k + \theta'_k)\left(\frac{\tilde{c}_k}{\theta_k} - \frac{\tilde{c}'_k}{\mu_k + \theta'_k}\right)$, which is equal to $q^0_k(1,0)\left(\frac{\tilde{C}(1,0)}{q^1_k(1,0)} - \frac{\tilde{C}_k(1,1)}{q^0_k(1,0)}\right)$. Hence, the index can be interpreted as the reduction in cost when making a class-$k$ bandit active instead of keeping him passive (the term within the brackets) during a time lag equal to the departure time in the active phase.

We now consider some particular cases that have been studied in the literature, see also Table 1. For example, let us consider first the case in which all customers can abandon the system, i.e., $\theta'_k = \theta_k$, for $k = 1, \ldots, K$, and that the cost for abandonment is the same for both *active* and *passive*, so $d_k = d'_k$. We first assume that all customers in the system incur a holding cost. This implies that $c_k = c'_k$, and thus $\tilde{c}_k = \tilde{c}'_k$. Substituting into (17) we get $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k}$. In the case where only customers in the queue incur a holding cost, *i.e.* $c'_k = 0$, we have $\tilde{c}_k - \tilde{c}'_k = c_k$, and upon substitution in (17) we get the index $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k} + c_k$.

We now assume that only customers in the queue can abandon, that is, the customer in service will not abandon, hence $\theta'_k = 0$, for $k = 1, \ldots, K$. This is the model assumption of [9] and [6]. We first assume that all customers in the system incur a holding cost, that is, $c_k = c'_k$, and we thus get $\tilde{c}'_k = c_k$. From (17) we get $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k} - c_k$. We can similarly calculate the index in the case in which only customers in the queue incur a holding cost, i.e., $c'_k = 0$, to obtain the index $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k}$. These two last indices have been derived in [9] and [6], respectively. More specifically, [9] derives the index $\frac{\tilde{c}_k \mu_k}{\theta_k} - c_k$ when studying one customer and no future arrivals. Interestingly, we observe that the index remains the same in the presence of random arrivals as considered in this paper. When the customer in service does not contribute to the holding cost, our model coincides with that analyzed in [6], where it is shown that the index rule $\frac{\tilde{c}_k \mu_k}{\theta_k}$ is asymptotically fluid optimal in a multi-server queue in overload ($\rho > 1$). We therefore conclude that the Whittle's index, we have derived, retrieves index policies that had been proposed in the literature when studying the system in special parameter regimes.

To finish this subsection we now provide an intuition to understand the result of Proposition 3 in the case $\theta'_k = \theta_k$ and $c_k = c'_k$. In this setting, at any moment in time, all customers in the system incur a holding cost $c_k$ and can abandon at rate $\theta_k$. Substituting $\mathbb{E}(\tilde{C}_k(N^{n_k}_k, S^{n_k}_k(N^{n_k}_k))) = \tilde{c}_k \mathbb{E}(N^{n_k}_k)$ and $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k}$ in (15), we get the relation

$$\theta_k(\mathbb{E}(N^{n_k-1}_k) - \mathbb{E}(N^{n_k}_k)) = \mu_k \left( \sum_{m=n_k}^{\infty} \pi^{n_k-1}_k(m) - \sum_{m=n_k+1}^{\infty} \pi^{n_k}_k(m) \right),$$

which can be seen as a rate conservation. Indeed, the term on the left-hand-side represents the difference in the average number of customers that abandons the system per time unit when comparing both policies $n_k$ and $n_k - 1$. The right-hand-side represents the difference in the average number of customers that is served per time unit when comparing both policies $n_k$ and $n_k - 1$. The left-hand-side being equal to the right-hand-side is exactly the rate conservation.

## 6.2 Convex holding cost

In this section we characterize Whittle's index, assuming that $W_k(n)$ is given by Equation (15), for general convex non-decreasing holding cost functions. We note that the cost associated to abandonments of customers are linear functions. We can thus use the result of Proposition 3 to rewrite Whittle's index as

$$W_k(n_k) = d_k(\mu_k + \theta'_k) - d'_k\theta'_k + W^c_k(n_k), \tag{18}$$

where

$$W^c_k(n_k) := \frac{\mathbb{E}(C_k(N_k^{n_k}, S^{n_k}(N_k^{n_k}))) - \mathbb{E}(C_k(N_k^{n_k-1}, S^{n_k-1}(N_k^{n_k-1})))}{\sum_{m=0}^{n_k} \pi_k^{n_k}(m) - \sum_{m=0}^{n_k-1} \pi_k^{n_k-1}(m)}$$

is the term corresponding to the holding cost. In the remainder of this section, we will focus on $W^c_k(n_k)$.

In Section 6.2.1 we characterize Whittle's index for large state values. In Section 6.2.2 and Section 6.2.3 we obtain Whittle's index as $\lambda_k \downarrow 0$ and $\lambda_k \uparrow \infty$, representing a light-traffic and heavy-traffic regime, respectively. For all cases, we will observe that for non-linear holding cost Whittle's index is dependent on $n_k$, that is, is state-dependent.

### 6.2.1 Whittle's index for large states

In this section we assume that the holding costs $C_k(n_k, 1)$ and $C_k(n_k, 0)$ are upper bounded by polynomials of finite degrees $P_k < \infty$ and $Q_k < \infty$, respectively. Hence, we can write $C_k(n_k, a) = E_k(n_k, a) + o(1)$, for large values of $n_k$, where $E_k(n_k, 1) = \sum_{i=0}^{P_k} C_k^{(P_k,i)} n_k^i$, with

$$C_k^{(P_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 1) - \sum_{j=i+1}^{P_k} C_k^{(P_k,j)} n_k^j}{n_k^i},$$

and $E_k(n_k, 0) = \sum_{i=0}^{Q_k} E_k^{(Q_k,i)} n_k^i$, with

$$E_k^{(Q_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 0) - \sum_{j=i+1}^{Q_k} E_k^{(Q_k,j)} n_k^j}{n_k^i}.$$

We assume w.l.o.g. that $P_k$ is such that $C_k^{(P_k,P_k)} > 0$ and $Q_k$ is such that $E_k^{(Q_k,Q_k)} > 0$.

In the following proposition we give the expression for Whittle's index for large states. The proof can be found in Appendix D.

**Proposition 4** *Assume Whittle's index is given as in* (15)*. Let $C_k(n_k, 1)$ and $C_k(n_k, 0)$ be upper bounded by a polynomial of degree $P_k$ and $Q_k$ respectively. Then, we have $W_k(n_k) = W_k^\infty(n_k) + o(1)$, as $n_k \to \infty$, where $W_k^\infty(n_k) := d_k(\mu_k + \theta'_k) - d'_k\theta'_k + \tilde{W}^c_k(n_k)$ and*

$$\begin{aligned}\tilde{W}^c_k(n_k) :=& (E_k(n_k, 0) - E_k(n_k, 1)) + (\mu_k + \theta'_k - \theta_k)/\theta_k \\ &\cdot \left( \sum_{i=1}^{Q_k} E_k^{(Q_k,i)} n_k^{i-1} + \sum_{i=2}^{P_k} C_k^{(P_k,i)} \sum_{j=0}^{i-2} n_k^{i-2-j} \left(\frac{\lambda_k}{\theta_k}\right)^{j+1} \right).\end{aligned} \tag{19}$$

*The index $W_k^\infty(n_k)$ is a non-decreasing function.*

Assume $C_k(n_k,a) = C_k(n_k)$ or $C_k(n_k,a) = C_k((n_k-a)^+)$ with $P_k \geq 2$. In that case, $P_k = Q_k$ and $C_k^{(P_k,P_k)} = E_k^{(Q_k,Q_k)}$. For states that are large enough, the value of $W_k^\infty(n_k)$ is determined by the highest polynomial, which is given by

$$\left(E_k^{(P_k,P_k-1)} - C_k^{(P_k,P_k-1)} + \frac{\mu_k + \theta'_k - \theta_k}{\theta_k} E_k^{(P_k,P_k)}\right) n_k^{P_k-1}. \tag{20}$$

The latter is independent of the arrival rate $\lambda_k$, and hence, so is $W_k^\infty$ for large enough states. This robust index (20) can serve as an approximation for Whittle's index policy when there are a large number of customers in the system. In Section 10 we numerically assess the performance under this index policy $W^\infty(\cdot)$.

#### 6.2.2 Light-traffic indices

We present in the following proposition the expression for Whittle's index as $\lambda_k \downarrow 0$, also referred to as the light-traffic regime. The proof can be found in Appendix E. Under the light-traffic assumption, the index can be given in closed form. In Section 8 we will use this expression to show that Whittle's index is asymptotically optimal in light traffic.

**Proposition 5** *Assume Whittle's index $W_k(n_k)$ is as given in* (15)*. Then, $W_k(n_k) = d_k(\mu_k + \theta'_k) - d'_k\theta'_k + W_k^c(n_k)$, where*

$$\lim_{\lambda_k \downarrow 0} W_k^c(n_k) = C_k(n_k,0) - C_k(n_k,1) + (C_k(n_k,0) - C_k(0,0))\frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k n_k}.$$

Assuming $C_k(0,0) = 0$, the above index can be rewritten as follows:

$$\lim_{\lambda_k \downarrow 0} W_k^c(n_k) = (\mu_k + \theta'_k + \theta_k(n_k-1))\left(\frac{C_k(n_k,0)}{\theta_k n_k} - \frac{C_k(n_k,1)}{\mu_k + \theta'_k + \theta_k(n_k-1)}\right).$$

This allows us for the following interpretation in light traffic. Given that there are $n_k$ class-$k$ customers, and there are no future arrivals, the index measures the reduction in cost when making a class-$k$ bandit active instead of keeping him passive (the term within the brackets) during a time lag equal to the departure time in the active phase.

#### 6.2.3 Heavy-traffic indices

We present in the following proposition the expression for Whittle's index as $\lambda_k \uparrow \infty$, also referred to as the heavy-traffic regime. The proof can be found in Appendix F. Under the heavy-traffic assumption, the index can be given in closed form.

**Proposition 6** *Assume Whittle's index $W_k(n_k)$ is as given in* (15)*. Define*

$$W_k^{HT}(n) := C_k(n,0) - C_k(n,1) + \frac{\mu_k + \theta'_k - \theta_k}{\theta_k}\frac{\mathbb{E}(C_k(N_k^{n-1},1))}{\lambda_k/\theta_k},$$

*where $N_k^{n-1}$ denotes the steady-state number of class-$k$ customers under threshold policy $n-1$, and is defined by the transition rates given in* (12)*. If there exists $z \geq 1$ such that $\frac{\mathbb{E}(C_k(N_k^{n-1},1))}{\lambda_k^z} \to 0$, as $\lambda_k \to \infty$, then $W_k(n) = d_k(\mu_k + \theta'_k) - d'_k\theta'_k + W_k^{HT}(n) + o(1)$ as $\lambda_k \to \infty$.*

# 7 M/M/1 multi-class queue

The multi-class M/M/1 queue without abandonments has received lot of attention from the research community. In the case of linear holding cost, the $c\mu$-index rule has been proved to be optimal in two main settings: (i) with exponential distributed service times and preemptive scheduling [16], and (ii) general service time distributions and non-preemptive scheduling [22]. A brief explanation of the optimality of an index rule is that having a linear holding cost $c_k$ for a class-$k$ customer per unit of time is equivalent to a problem where a reward $c_k$ is received upon service completion (and no holding cost) [23, Section 4.9]. The latter can be seen as a MABP, for which an index rule (in this case $c\mu$) is optimal[1]. However, this equivalence holds only for linear holding costs, which explains why for general cost functions the structure of the optimal scheduling policy is no longer of index type. In that context, a fruitful approach has been to derive scheduling policies with near-optimal performance or asymptotically optimal performance in a limiting regime, see the references as stated in Section 2.

In this section, we derive an index policy for the multi-class M/M/1 system by considering the limit of our Whittle index as the abandonment rate tends to 0. Note that the Whittle's index $W_k(n_k)$ goes to $\infty$ as $\theta_k \to 0$, and it turns out that when scaling the index by $\theta_k$ we get a non-trivial limit. The proof of the next proposition may be found in Appendix G.

**Proposition 7** *Assume $C_k(n_k, a) = C_k(n_k)$, $a = 0, 1$, $\theta'_k = \theta_k$, and $d_k = d'_k = 0$. Then,*

$$\lim_{\theta_k \to 0} \theta_k W_k(n_k) = \frac{\mu_k(1-\rho_k)}{\rho_k} \cdot \left[\sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) C_k(n_k - 1 + m) - C_k(n_k - 1)\right]. \tag{21}$$

Observe that convexity of the function $C(\cdot)$ implies that (21) is a non-decreasing function.

A heuristic for the M/M/1 queue with as objective to minimize the holding cost can now be derived as follows. Set $\theta_k = \theta'_k$ for all $k$ and consider the index multiplied by $\theta_k$ as $\theta_k \to 0$. A heuristic is then to give priority according to the index as given in (21).

In case of linear holding costs $C_k(n_k) = c_k n_k$, the index (21) coincides with the $c_k\mu_k$-rule. For general holding cost the index in (21) was also obtained in Glazebrook *et al.* [3] (see also Section [23, Section 6.5]) by carrying out a model-dependent analysis, which consists in considering first the total discounted holding cost criterion, calculating the corresponding Whittle's index, and afterwards taking the limit in the discounting factor. In that case too, indexability needs to be established.

As pointed out in [23, Section 6.5] applying directly the average cost criteria to the M/M/1 queue without abandonments gives no meaningful index. Consider a single server queue with threshold policy $n$, where the taken action is *passive* for all states below and equal to $n$, and *active* for all states above $n$. This system is equivalent to the classical M/M/1 queue where state $m$ corresponds to $m - n$. A classical result shows that in the absence of abandonments the probability that the stationary process is in state 0 in an M/M/1 queue is $1 - \rho$, and therefore in a single server queue under policy $n$ the probability of being in state $n$ will be $1 - \rho$, i.e., independently of where the threshold is set. Hence, the subsidy obtained is $W(1 - \rho)$, which is independent of the policy $n$, and therefore, the subsidy does not allow us to "calibrate" the states. In our approach this is circumvented by obtaining an index for the, well-defined, case with abandonments and then letting $\theta_k \to 0$, while in [23, Section 6.5] this is circumvented by looking at the discounted problem and scaling the immediate cost.

[1] This is known as the *tax* formulation of a MABP, see [23, Section 4.9].

| $\rho$ | 0.11 | 0.21 | 0.31 | 0.41 |
|---|---|---|---|---|
| (21) | 4.25e-06 | 1.51e-05 | 6.07e-06 | 5.02e-07 |
| $C'(n)\mu$ | 0.0072 | 0.0636 | 0.1002 | 0.1320 |
| $\rho$ | 0.51 | 0.61 | 0.71 | 0.81 |
| (21) | 0.008 | 0.0291 | 0.0919 | 1.7129 |
| $C'(n)\mu$ | 0.1689 | 0.3616 | 1.8280 | 4.9539 |

Table 2: Suboptimality gap

For large values of $n_k$, the index (21) is approximately equal to $C'_k(n_k)\mu_k$, which we refer to as the $Gc\mu$-rule. This rule was introduced in [35] for convex delay cost. The equivalence with the $Gc\mu$ rule can be seen as follows. We have for $n_k$ large,

$$\begin{aligned}
&\sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) C_k(n_k - 1 + m) - C_k(n_k - 1) \sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) \\
&= (1-\rho_k) \sum_{m=0}^{\infty} \rho_k^m (C(n_k - 1 + m) - C(n_k - 1)) \\
&\approx (1-\rho_k) \sum_{m=0}^{\infty} m \rho_k^m C'(n_k - 1) = C'(n_k) \frac{\rho_k}{(1-\rho_k)},
\end{aligned}$$

where we used that for $n_k$ large with respect to $m$, we have $\frac{C(n_k-1+m)-C(n_k-1)}{m} \approx C'(n_k)$ and that large values of $m$ have a negligible weight on the summation. Hence, it follows from (21) that $\lim_{\theta_k \to 0} \theta_k W_k(n_k) \approx C'_k(n_k)\mu_k$.

**Numerical example.** In Table 2 we compare the suboptimality of the $C'(n)\mu$-rule and index-rule (21) in an M/M/1 queue without abandonments. Note that when $\theta_k = 0$, for all $k$, we need to assume $\sum_{k=1}^{K} \rho_k < 1$ in order to assure stability of the system. Consider 4 classes of customers with the following parameters: $\mu_1 = 16, \mu_2 = 27, \mu_3 = 12$ and $\mu_4 = 21$, $\rho_1 = 3\rho/9, \rho_2 = \rho/9$, $\rho_3 = 5\rho/9$ and $\rho_4 = \rho/9$. The holding cost of each class are cubic, $C_k(n_k) := \alpha_k + \beta_k n_k + \gamma_k n_k^2 + \delta_k n_k^3$, for which (21) simplifies to: $\beta_k\mu_k + \gamma_k\mu_k\left(\frac{3\rho_k - 1}{1-\rho_k} + 2n_k\right) + \delta_k\mu_k\left(3n_k^2 + 3\left(\frac{2\rho_k - 1}{1-\rho_k}\right)n_k + \frac{4\rho_k^2 + \rho_k + 1}{(1-\rho_k)^2}\right)$. We take the particular example: $C_1(n_1) = 6n_1 + 2n_1^2 + 2n_1^3$, $C_2(n_2) = 2n_2 + 2n_2^2 + 2n_2^3$, $C_3(n_3) = n_3 + n_3^2 + 3n_3^3$ and $C_4(n_4) = 8n_4 + 2n_4^3$. We observe that for this example the $C'(n)\mu$-rule is outperformed by the index-rule (21), but both policies give nearly optimal performance.

# 8 Asymptotic optimality

In this section we will discuss various notions of asymptotic optimality of Whittle's index policy. Section 8.1 deals with the optimality of Whittle's index policy in a multi-server setting, and Section 8.2 proves Whittle's index policy to be optimal in light-traffic and heavy-traffic regimes.

## 8.1 Multi-server setting

For linear holding cost, asymptotic optimality in a multi-server setting can be directly derived from [36]. Assume there are $M$ servers and the arrival rate of class-$k$ customers is $M\lambda_k$. Let $W_k$ be the state-independent index as given in (17). In [36, Proposition 6.2] it is shown that the

Whittle index policy ($WI$), where at each moment in time a server serves a customers having highest non-negative index $W_k$, is asymptotically optimal in the following sense: for any policy $\varphi$,

$$\lim_{M\to\infty} \mathcal{C}^{WI}(M) \leq \liminf_{M\to\infty} \mathcal{C}^{\varphi}(M),$$

where $\mathcal{C}^{WI}(M)$ denotes the average cost incurred by Whittle's index, and $\mathcal{C}^{\varphi}(M)$ denotes the average cost incurred by policy $\varphi$ when there are $M$ servers in the system.

For general holding cost, we can not derive asymptotic optimality. We do expect however that under certain conditions one would have the following. Assume there are $M$ servers and $x_k M$ queues where class-$k$ customers arrive with rate $\lambda_k$, $k = 1, \ldots, K$[2]. A queue can be served by at most one server. In bandit terminology this represents having $x_k M$ class-$k$ bandits whose state (that is, the number of customers in the queue) has values in $\mathcal{S} := \{0, 1, \ldots\}$, and the scheduler needs to decide which $M$ bandits to make active (so which $M$ queues to serve). In case the state space $\mathcal{S}$ would have been finite, the result in [38, 36] implies (under certain conditions) asymptotic optimality of Whittle's index policy as $M \to \infty$. However, for infinite state space, as is the case for our model, no result is known so far.

## 8.2 Light-traffic and heavy-traffic regimes

Light traffic and heavy traffic refer to the situations in which the total arrival rate goes to 0 and $\infty$, respectively. Note that due to abandonments, our model is *stable* for any value of the arrival rate. In this section we will show that Whittle index policy is optimal in these two limiting regimes. In order to take the limits we will modify the total arrival rate while keeping constant the proportion of traffic of each class. To do so, we assume that $\lambda_k = \gamma_k \lambda$, where $\lambda$ denotes the total arrival rate, and $\sum_{k=1}^{K} \gamma_k = 1$.

We recall that $\mathcal{U}$ and $\mathcal{U}^{REL}$ refer to the set of admissible policies in the original and relaxed problem, respectively, and that $\mathcal{U} \subseteq \mathcal{U}^{REL}$. As we argued in Section 4, for any value of the multiplier $W \geq 0$, $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$, where $\mathcal{C}^{REL(W)}(W)$ and $\mathcal{C}^{OPT}$ are the minimum cost in the relaxed and original problems, respectively. We also recall that $\mathcal{C}^{REL(W)}(W)$ is achieved by a policy that serves all the classes with current Whittle's index larger than $W$. We denote by $\mathcal{C}^{WI}$ the performance in the original problem under *the admissible* Whittle index policy and we set $\mathcal{C}^* = \sup_W \mathcal{C}^{REL(W)}(W)$. It then trivially holds that

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^* \leq \mathcal{C}^{OPT} \leq \mathcal{C}^{WI}. \tag{22}$$

We now argue that if either

*(i)* $REL(0) \in \mathcal{U}$, or,

*(ii)* $REL(W) \in \mathcal{U}$ and the constraint (6) is satisfied with equality,

then it holds that, for that choice of $W$, $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$, and hence in those cases Whittle's index policy is optimal for the original policy. This can be seen as follows. First we observe that if $REL(W) \in \mathcal{U}$, then $REL(W)$ coincides with Whittle's index policy. Hence, for $W = 0$ we have $\mathcal{C}^{REL(0)}(0) = \mathcal{C}^{REL(0)} = \mathcal{C}^{WI}$, where the first equality holds by definition since $W = 0$. Now assume $W > 0$, then since (6) holds with equality, we have again $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^{REL(W)} = \mathcal{C}^{WI}$. In both cases, we use (22) to conclude that $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$. We note that the same approach is described in [23, Chapter 6] and [26, Section 5].

[2]This can represent for example a setting where there are $x_k M$ class-$k$ flows having newly arriving packets (represented by customers).

We can now use the above in order to show Whittle's index to be asymptotically optimal in both light traffic and heavy traffic. In the light-traffic regime, we will consider the case *(i)*, and in the heavy-traffic regime we will consider case *(ii)*. In light traffic, most of the time the system is empty or at most there is one customer in the system. This implies that as $\lambda \to 0$, $REL(0)$ becomes admissible for the original problem, that is, $REL(0) \in \mathcal{U}$. Hence, we are in case (i), which will allow us to conclude for asymptotic optimality of Whittle's index policy. In heavy traffic, we will prove that for the correct choice of $W$, under the Whittle index policy, constraint (6) is satisfied with equality, and $REL(W) \in \mathcal{U}$. Hence, we are in case (ii) and we deduce that asymptotic optimality holds in heavy-traffic regime.

We present the asymptotic optimality result for the light-traffic regime in Theorem 2 and in Theorem 3 for the heavy-traffic regime. The proofs can be found in Appendix H and Appendix I.

**Theorem 2** *Assume Whittle's index $W_k(n)$ is as given in* (15). *Assume $C_k(0,0) \geq C_k(0,1)$, $\forall k$. The Whittle index policy (WI) is asymptotically optimal in light traffic, that is,*

$$\lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0,$$

*with $\lambda_k = \lambda \gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.*

**Theorem 3** *Assume Whittle's index $W_k(n)$ is as given in* (15). *Assume there exists a $\bar{k} \in \{1, \ldots, K\}$ such that*

$$\lim_{\lambda \uparrow \infty} \frac{W_{\bar{k}}(\lambda \gamma_{\bar{k}} / \theta_{\bar{k}})}{W_k(\lambda \gamma_k / \theta_k)} > 1,$$

*for all $k \neq \bar{k}$. Then, the Whittle index policy (WI) is asymptotically optimal in heavy traffic, that is,*

$$\lim_{\lambda \to \infty} (\mathcal{C}^{WI} - \mathcal{C}^{OPT}) = 0,$$

*with $\lambda_k = \lambda \gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.*

Whittle's index policy gives strict priority to class $\bar{k}$. In fact, we can see from the proof of Theorem 3 that any policy that gives strict priority to class $\bar{k}$ will be optimal as $\lambda \uparrow \infty$.

# 9 Fluid index

In Section 4 we derived the optimal policy of the relaxed optimization problem (8), which was described by the index value as given in Proposition 1 and Corollary 1. Unfortunately, for non-linear holding cost the index could not be written in closed-form. In this section we will therefore solve the fluid version of the relaxed optimization problem (8), that is, we only take into account the average behavior of the system. This will allow us to obtain a closed-form expression for the fluid index. In Section 9.1 we describe the fluid control problem we need to solve and in Section 9.2 we obtain the solution and the fluid index. In addition, in Section 9.3 we compare the fluid index with the index for the stochastic model.

## 9.1 Fluid model description

We approximate the stochastic model as presented in Section 3 by a deterministic fluid model, where only the mean dynamics are taken into account. Let $m_k(t) \geq 0$ be the amount of class-$k$ fluid and let $s_k(t) \in \{0,1\}$ be the control parameter. Let $u$ denote a fluid control that determines $s_k^u(t)$. The fluid dynamics under control $u$ is given by $\frac{\mathrm{d}m_k^u(t)}{\mathrm{d}t} = \lambda_k - \theta_k m_k^u(t)$ if the chosen action is passive, that is, $s^u(t) = 0$, and is given by $\frac{\mathrm{d}m_k^u(t)}{\mathrm{d}t} = \lambda_k - \mu_k - \theta_k' - \theta_k(m_k^u(t) - 1)$, if the chosen action is active, that is, $s^u(t) = 1$. Hence, the dynamics can be written as

$$\begin{aligned}\frac{\mathrm{d}m_k^u(t)}{\mathrm{d}t} =& \lambda_k - s_k^u(t)(\mu_k + \theta_k' + \theta_k(m_k^u(t) - 1)) - (1 - s_k^u(t))\theta_k m_k^u(t)\\ =& \lambda_k - (\mu_k + \theta_k' - \theta_k)s_k^u(t) - \theta_k m_k^u(t),\end{aligned}$$

where the control $u$ is such that $m_k^u(t) \geq 0$ for all t.

At time $t$, the cost for the fluid model under the relaxed problem is written as

$$(1 - s_k(t))\tilde{C}_k(m_k(t), 0) + s_k(t)\tilde{C}_k(m_k(t), 1) - W(1 - s_k(t)).$$

The cost functions $C_k(m,0)$ and $C_k(m,1)$ are assumed to be continuous in $m$. Note that we have used the same notation as in the stochastic model where the cost functions were discrete in $m$ (slight abuse of notation). Assume

$$\frac{\mathrm{d}C_k(m,1)}{\mathrm{d}m} \leq \frac{\mathrm{d}C_k(m,0)}{\mathrm{d}m}, \tag{23}$$

which is the continuous equivalence of the RHS of (2).

An equilibrium point $(\bar{m}_k, \bar{s}_k)$ of $m_k(t)$ is such that $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = 0$, that is,

$$0 = \lambda_k - (\mu_k + \theta_k' - \theta_k)\bar{s}_k - \theta_k \bar{m}_k,$$

with $\bar{s}_k \in [0, \min\{1, \lambda_k/(\mu_k + \theta_k' - \theta_k)\}]$ and $\bar{m}_k \in [\max(0, (\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k), \lambda_k/\theta_k]$.

In the stochastic model the aim is to minimize (8), that is, to minimize the time-average cost minus the subsidy obtained. In equilibrium, $\bar{s}_k$ is the average amount of time the system is active, hence, the fluid version of (8) will be to find the equilibrium point that minimizes the cost in equilibrium, that is, to minimize

$$EC(\bar{s}_k, W) := (1 - \bar{s}_k)\tilde{C}_k(\bar{m}_k, 0) + \bar{s}_k\tilde{C}_k(\bar{m}_k, 1) - W(1 - \bar{s}_k).$$

We denote by $(m_k^*, s_k^*)$ an optimal equilibrium point and define the optimal equilibrium cost under subsidy $W$ by

$$\begin{aligned}EC_k^*(W) \quad &:= \quad \min_{\bar{s}_k \in [0, \min\{1, \lambda_k/(\mu_k + \theta_k' - \theta_k)\}]} EC_k(\bar{s}_k, W) &(24)\\ &= \quad (1 - s_k^*)\tilde{C}_k(m_k^*, 0) + s_k^*\tilde{C}_k(m_k^*, 1) - W(1 - s_k^*). &(25)\end{aligned}$$

Since the time-average criteria might be attained by several controls, in the next section we will study controls that are *bias-optimal*. That is, among all controls that reach the optimal equilibrium point, a bias-optimal control is the one that minimizes the cost to get to this equilibrium point.

## 9.2 Fluid index for bias optimality

Having characterized the optimal equilibrium point in the previous section, the question is which control minimizes the cost to get to this equilibrium, referred to as *bias-optimality.* Hence, our aim is to find the control $u$ that minimizes

$$\int_0^\infty \left(\tilde{C}_k(m_k^u(t), s_k^u(t)) - W(1 - s_k^u(t)) - EC_k^*(W)\right)\mathrm{d}t. \tag{26}$$

That is, minimize the total cost over time minus the optimal cost in equilibrium.

The optimal solution to the fluid bias optimal problem is stated below.

**Theorem 4** *An optimal control for the relaxed fluid problem* (26) *is* $s_k^*(t) = 1$ *if* $w_k(m_k(t)) > W$ *and* $s_k^*(t) = 0$ *otherwise, with*

$$\begin{aligned} w_k(m_k) \;&:=\; C_k(m_k, 0) - C_k(m_k, 1) + d_k(\mu_k + \theta_k') - d_k'\theta_k' \\ &\quad + \begin{cases} w_k^{(1)}(m_k) & \text{if } 0 \le m_k < \max\left(0, \frac{\lambda_k - (\mu_k + \theta_k' - \theta_k)}{\theta_k}\right), \\ w_k^{(2)}(m_k) & \text{if } \max\left(0, \frac{\lambda_k - (\mu_k + \theta_k' - \theta_k)}{\theta_k}\right) \le m_k \le \frac{\lambda_k}{\theta_k}, \\ w_k^{(3)}(m_k) & \text{if } m_k > \frac{\lambda_k}{\theta_k}, \end{cases} \end{aligned}$$

*where*

$$\begin{aligned} w_k^{(1)}(m_k) =& \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{\left(C\left(\frac{\lambda_k - (\mu_k + \theta_k' - \theta_k)}{\theta_k}, 1\right) - C(m_k, 1)\right)}{(\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k - m_k}, \\ w_k^{(2)}(m_k) =& \frac{(\lambda_k - \theta_k m_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k, 1) + (\theta_k m_k + \mu_k + \theta_k' - \theta_k - \lambda_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k, 0)}{\theta_k}, \\ w_k^{(3)}(m_k) =& \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{\left(C_k(m_k, 0) - C_k\left(\frac{\lambda_k}{\theta_k}, 0\right)\right)}{m_k - \lambda_k/\theta_k}. \end{aligned}$$

*The fluid index $w_k(m_k)$ is non-decreasing and continuous.*

The proof of Theorem 4 can be found in Appendix K.

Having solved the fluid version of the relaxed problem, we propose the following heuristic for the stochastic model.

**Definition 5 (Fluid index policy)** *Assume at time $t$ we are in state $\vec{N}(t) = \vec{n}$. The fluid index policy prescribes to serve the class $k$ having currently the highest non-negative fluid index $w_k(n_k)$.*

We directly observe that for linear holding cost, the fluid index is state-independent and coincides with that of the stochastic model as stated in Proposition 3. Now assume $C_k(m_k, a_k) = C_k(m_k)$, that is, holding cost for customers in the system. In that case, the fluid index simplifies as follows:

$$w_k^{(2)}(m_k) = \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{\mathrm{d}}{\mathrm{d}m_k} C_k(m_k),$$

which corresponds to the $C'(m)\mu/\theta$-rule when $\theta_k' = \theta_k$. We refer to this rule as the Generalized $c\mu/\theta$-rule ($Gc\mu/\theta$). The terms $w_k^{(1)}(m_k)$ and $w_k^{(3)}(m_k)$ reduce to

$$\begin{aligned} w_k^{(1)}(m_k) &= \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{(C_k((\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k) - C_k(m_k))}{(\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k - m_k}, \\ w_k^{(3)}(m_k) &= \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{(C_k(m_k) - C_k(\lambda_k/\theta_k))}{m_k - \lambda_k/\theta_k}. \end{aligned}$$

We refer to [15] where index policies based on first-order difference have also been proposed and are shown to empty the system with the lowest cost possible in a single server multi-class queue without abandonments and no future arrivals.

### 9.3 Asymptotic equivalence of stochastic index and fluid index

In this section we discuss the relation between the Whittle index as obtained for the original stochastic problem and the fluid index. As mentioned in the previous section, for linear holding cost both indices coincide. Here we consider general holding cost and we study the equivalence of both indices in asymptotic regimes.

We first consider the light-traffic scenario, that is, $\lambda_k \downarrow 0$.

**Proposition 8** *Let $W_k(\cdot)$ be given as in* (15). *Then as $\lambda_k \downarrow 0$,*

$$W_k(n_k) = w_k(n_k) + o(1).$$

**Proof.** The fluid index as $\lambda_k \downarrow 0$ reduces to

$$\lim_{\lambda_k \downarrow 0} w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + d_k(\mu_k + \theta_k') - d_k'\theta_k' + \frac{\mu_k + \theta_k' - \theta_k}{\theta_k} \frac{(C_k(m_k, 0) - C_k(0, 0))}{m_k}.$$

The latter coincides with the Whittle index as given in Proposition 5. ■

We now focus on the indices for large values of the state. In the next proposition we show that the fluid index $w_k(n_k)$ coincides with Whittle's index as given in (15), when the cost functions are upper bounded by polynomial functions.

**Proposition 9** *Assume that $C_k(n_k, 1)$ and $C_k(n_k, 0)$ are upper bounded by a polynomial of degree $P_k$ and $Q_k$, respectively, with $Q_k > P_k$. Then,*

$$\lim_{n_k \to \infty} \frac{W_k(n_k)}{w_k(n_k)} = 1. \tag{27}$$

*If we further assume $P_k = Q_k$ and $C_k^{(P_k,i)} = E_k^{(P_k,i)}$ for all $i \in \{2, \dots, P_k\}$, then as $n_k \to \infty$,*

$$W_k(n_k) = w_k(n_k) + o(1). \tag{28}$$

As an example we consider $C_k(n_k, a) = C_k(n_k)$ or $C_k(n_k, a) = C_k((n_k - a)^+)$. Then $Q_k = P_k$, and hence (27) holds. In case, $C_k(n_k, a) = C_k(n_k)$, then in addition (28) holds.

## 10 Numerical Results

The objective of the present section is to show in which regimes the Whittle index policy $W(n)$ (Equation (15)) performs well. We will focus on holding cost functions of the shape $C_k(n_k, a) = C_k(n_k)$ or $C_k(n_k, a) = C_k((n_k - a)^+)$, that is, the holding cost is a function of the number of class-$k$ customers in the *system* or *queue* respectively. Hence, $\tilde{C}_k(n_k, a)$ reduces to $C_k(n_k) + d_k\theta_k n_k$ or $C_k((n_k - a)^+) + d_k\theta_k(n_k - a)^+ + d_k'\theta_k' \min(a, n_k)$, respectively.

In Section 10.1 we compare the structure of Whittle's index policy with the structure of the optimal policy, numerically. In Section 10.2 we then numerically compare the performance of the index policies with that of the optimal policy.

## 10.1 Structure of different policies

We compare the structure of the different index policies and the optimal policy for linear and convex holding cost.

### 10.1.1 Linear holding cost

By value iteration [34] we observed that for a wide range of parameters the optimal policy, under linear holding cost, is of the following structure: when $(N_1, \ldots, N_K)$ is close enough to the origin (and $N_i$ denotes the number of class-$i$ customers in the system), it is optimal to prioritize classes according to the $\tilde{c}\mu$-rule, otherwise to prioritize classes according to the $\tilde{c}\mu/\theta$-rule, where $\tilde{c}_k := c_k + d_k\theta_k$, see Figure 2 (left) with $\epsilon = 0$ as described in the next section. Hence, the Whittle's index (which corresponds to the $\tilde{c}\mu/\theta$-rule in the linear case) captures the optimal action for states that are not too close to the origin.

### 10.1.2 General holding cost

To discuss the structure of index policies for general holding cost, we focus on two classes of customers ($K = 2$). In a state $(N_1, N_2)$, the action taken by Whittle's index rule is to serve the class having highest value $W_k(N_k)$. Since $W_k(N_k)$ is an non-decreasing function, this implies that there is an increasing switching curve (SC) such that when $(N_1, N_2)$ is below the SC, Whittle's index policy serves class 1 and for any state $(N_1, N_2)$ above the curve the policy serves class 2. Note that for linear holding cost this switching curve collapses either to the vertical or horizontal axis.

By value iteration we observed that an optimal policy is as well of switching curve type. For example, in Figure 2 (left) we plot the switching curve of the optimal policy with the following holding cost: $C_1(n) = n + \epsilon n^2$ and $C_2(n) = n$ (parameters $\theta = \theta'$ and $\lambda = [9, 10], \mu = [14, 16], \theta = [2, 0.05], d = [4, 0.3]$). When $\epsilon = 0$, we obtain a decreasing switching curve, which describes the behavior of the optimal policy for linear cost as explained in Section 10.1.1. As $\epsilon$ becomes positive, the switching curve becomes increasing. In addition, $\epsilon$ becomes larger, and hence the quadratic cost of class 1 increases, and therefore, class 1 gets priority in a larger region.

We now compare the actions taken under Whittle's index policy and the optimal policy. We consider an example with quadratic costs $C_1(n) = (c_{11} + d_1\theta_1)n + c_{21}n^2$ and $C_2(n) = (c_{12} + d_2\theta_2)n + c_{22}n^2$, and set the following parameters $\theta = \theta'$ and $\mu = [15, 18]; \theta = [4, 7]; c_1 = [1, 4]; c_2 = [2, 1]; d = [8, 6.5]$. In Figures 2 (middle and right) we plot the optimal actions (obtained by value iteration) for load $\rho = 0.8$ and $\rho = 2.5$, respectively, and compare it to the actions taken under Whittle's index policy. We observe that the optimal policy can be described by a switching curve. In addition the optimal policy coincides with that of Whittle's index $W(n)$ in almost all the state space as the workload increases. We also plot the switching curve corresponding to the fluid index $w(n)$ and observe a very good fit.

## 10.2 Performance evaluation

In this section we evaluate numerically the performance of the index policies. This is carried out by computing the relative sub optimality gap between the average cost of the optimal solution and an index policy. In order to compute this we use the Value Iteration algorithm [34].

We saw in Section 7 that the index policy with index (21) performs very well in an M/M/1 multi-class systems (when there are no abandonments). We considered cubic costs and 4 classes

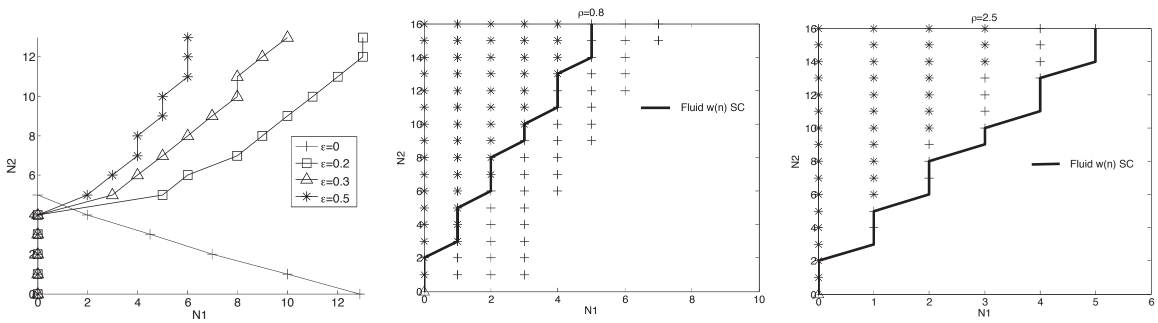


Figure 2: (Left:) Switching curves of the optimal policy for varying holding cost (from linear to quadratic). (Middle and right:) Actions under the optimal policy, the index policy $W(n)$, and the fluid index policy for quadratic holding cost. Area with "+": $W(n)$ serves class 1 while it is optimal to serve class 2, Area with "*": $W(n)$ serves class 2, which is also optimal, and in the white area $W(n)$ serve class 1, which is also optimal.

of customers and compared the Generalized index rule ($Gc\mu$) and the index-rule of (21) and we observed there that the latter performs slightly better than the $Gc\mu$-rule.

In this section we will consider scenarios allowing abandonments. We will evaluate the following indices: (i) the Whittle index $W(n)$ (Equation (15)), (ii) the Whittle index for large states $W^\infty(n)$ and (iii) the fluid index $w(n)$. We compare these to the two index policies proposed for a multi-class queue without abandonments: the $Gc\mu$-rule, and the index-rule corresponding to (21) which is an approximation of $W(n)$ for $\theta$ close to zero. We will analyze two different scenarios: (1) varying the workload $\rho$, and (2) varying the abandonment rates $\theta_k$.

### 10.2.1 Varying Workload

In this section we aim at observing the behavior of index policies for varying workload.

**Example with linear holding cost** ($\theta = \theta'$)**:** We set $C_k(n,a) = c_k n$, $\mu = [15, 25]$, $\theta' = \theta = [4, 2]$, $c = [1, 1]$, $d = [5, 3.2]$, and let $\rho = \sum_{k=1}^{2} \lambda_k/\mu_k$ vary in the interval $[0, 2.6]$, with $\lambda_1/\mu_1 = \lambda_2/\mu_2$. For linear holding costs, the indices $W(n), W^\infty(n)$ and $w(n)$ reduce to the $\tilde{c}\mu/\theta$-rule and the indices $Gc\mu$ and (21) reduce to the $\tilde{c}\mu$-rule, with $\tilde{c}_k = c_k + d_k\theta_k$.

**Example with linear holding cost** ($\theta \neq \theta'$)**:** We set $C_k(n,a) = c_k(n-a)^+$, $\mu = [15, 25]$, $\theta = [4, 2]$, $\theta' = [3, 2]$, $c' = c = [1, 1]$, $d = [5, 3.2]$, $d' = [2, 1]$, and let $\rho = \sum_{k=1}^{2} \lambda_k/\mu_k$ vary in the interval $[0, 2.6]$, with $2\lambda_1/\mu_1 = \lambda_2/\mu_2$. For linear holding costs and $\theta \neq \theta'$, the indices $W(n), W^\infty(n)$ and $w(n)$ reduce to the $\tilde{c}(\mu+\theta')/\theta - \tilde{c}'$-rule and the indices $Gc\mu$ and (21) reduce to the $\tilde{c}\mu$-rule, with $\tilde{c}_k = c_k + d_k\theta_k$.

In Figure 3 we observe for both cases that the $\tilde{c}\mu$-rule is optimal in underload, while the performance of the index $W(n)$ is nearly optimal in overload, as expected from Theorem 3. As discussed in Section 10.1.1, in a state far from the origin, the optimal action is to serve according to $\tilde{c}\mu/\theta$, which is the region in which the process will live in overload, explaining why the $\tilde{c}\mu/\theta$-rule and the $\tilde{c}(\mu+\theta')/\theta - \tilde{c}'$-rule perform well in this case. In underload, the effect of abandonments is not that important and the $\tilde{c}\mu$-rule performs very well.

**Example with quadratic holding cost** ($\theta = \theta'$)**:** Consider the following parameters: $\mu = [15, 18]$, $\theta' = \theta = [4, 7]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $d = [8, 6.5]$, and we let $\lambda$ vary, but keeping $\lambda_1/\mu_1 = \lambda_2/\mu_2$. We assume quadratic costs $C_1(n) = (c_{11}+d_1\theta_1)n + c_{21}n^2$ and $C_2(n) = (c_{12}+d_2\theta_2)n + c_{22}n^2$. See Figure 4 (left) for the sub-optimality gap and Table 3 for the absolute errors.

We observe that for low load the $Gc\mu$-rule and the index-rule (21) behave very well. However,

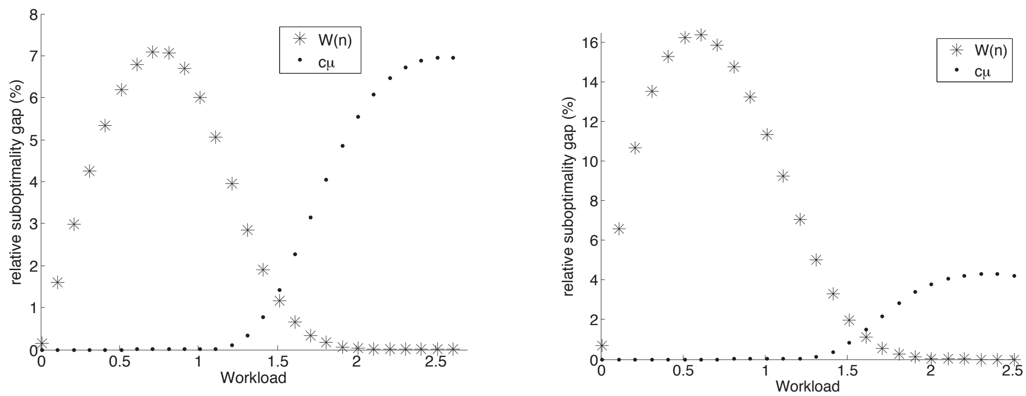


Figure 3: Left: sub optimality for linear holding cost, as $\rho$ increases when $\theta = \theta'$. Right: sub optimality for linear holding cost, as $\rho$ increases when $\theta \neq \theta'$.

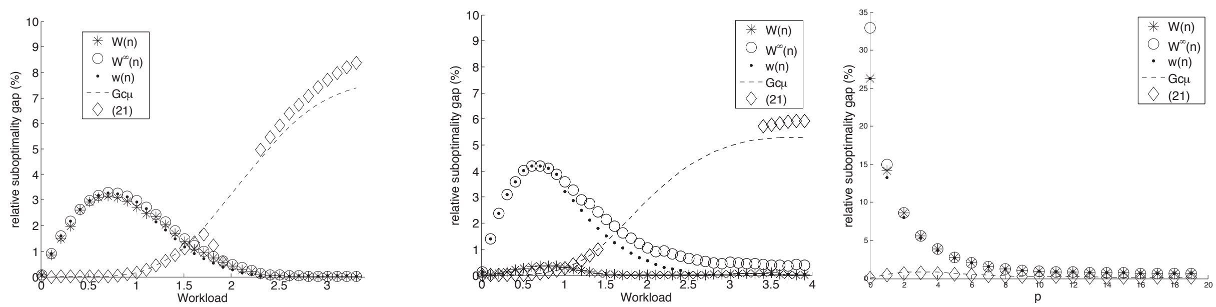


Figure 4: Left: sub-optimality for linear holding cost, as $\rho$ increases. Middle: sub optimality for quadratic holding cost as $\rho$ increases. Right: sub optimality for quadratic holding cost as $p$ ($\theta_i = p\epsilon_i, i \in \{1,2\}$) varies.

as the load grows larger, the sub-optimality gap of these $\theta$-independent policies grows large, while our Whittle index policy $W(n)$, the Whittle index policy for large states $W^\infty(n)$ and the fluid index policy $w(n)$ become near optimal. In Table 3 we observe that the convergence towards optimality is reached very fast as the absolute error ($\mathcal{C}^{WI} - \mathcal{C}^{OPT}$) of the $W(n), W^\infty(n)$ and $w(n)$ indices is of order $10^{-4}$ when $\rho = 5.25$. On the other hand, both (21) and the $Gc\mu$-rule perform very bad in overload. Hence, our index policies are very suitable for the overload setting, which are from a practical point of view of main importance.

Note that the jump around $\rho = 2$ for the index-rule (21) is a result of undefined values around $\lambda_k = \mu_k$.

| Workload | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 5.25 |
|---|---|---|---|---|---|---|---|
| $W(n)$ | 1.3089 | 1.4608 | 0.8055 | 0.1094 | 0.0185 | 0.0065 | 0.00017 |
| $W^\infty(n)$ | 1.4028 | 1.5596 | 0.8902 | 0.1732 | 0.0614 | 0.0329 | 0.0007 |
| $w(n)$ | 1.3823 | 1.2885 | 0.5534 | 0.0026 | 0.0771 | 0.0904 | 0.0004 |
| (21) | 0.0409 | 0.7327 | 0.8010 | 11.2134 | 20.5851 | 28.3926 | 50.0996 |
| $Gc\mu$ | 0.0409 | 0.7483 | 3.9951 | 10.4111 | 18.7237 | 25.0454 | 42.5645 |

Table 3: Absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ that corresponds to the example in Figure 4 (left).

| Workload | 1 | 1.5 | 2.5 | 3 | 3.5 | 5.25 | 7.25 | 10 | 16 |
|---|---|---|---|---|---|---|---|---|---|
| $W(n)$ | 0.1332 | 0.0664 | 0.0098 | 0.1260 | 0.2874 | 0.2448 | 0.1404 | 0.0486 | 0.0061 |
| $W^{\infty}(n)$ | 1.4817 | 1.9167 | 1.4429 | 1.1485 | 1.4243 | 1.7296 | 1.4784 | 0.7977 | 0.1012 |
| $w(n)$ | 1.4817 | 1.4157 | 0.3397 | 0.0382 | 0.1288 | 0.5125 | 04383 | 0.1542 | 0.0093 |
| (21) | 0.0720 | - | - | - | 19.3226 | 35.5180 | 48.5766 | 66.1024 | 91.4859 |
| $Gc\mu$ | 0.0720 | 0.7896 | 7.7697 | 12.8528 | 17.6942 | 31.1417 | 43.3748 | 59.7161 | 99.4344 |

Table 4: Absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ that corresponds to the example in Figure 4 (middle).

**Example with quadratic holding cost** ($\theta \neq \theta'$)**:** Consider the following parameters: $\mu = [15, 18]$, $\theta = [4, 7]$, $\theta' = [3, 4]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $d = [8, 6.5]$, $d' = [7, 7]$ and we let $\lambda$ vary, but keeping $2\lambda_1/\mu_1 = \lambda_2/\mu_2$. We assume quadratic costs $\widetilde{C}_1(n,a) = c_{11}(n-a)^+ + c_{21}((n-a)^+)^2 + d_1\theta_1(n-a)^+ + d_1'\theta_1'a$ and $\widetilde{C}_2(n,a) = c_{12}(n-a)^+ + c_{22}((n-a)^+)^2 + d_2\theta_2(n-a)^+ + d_2'\theta_2 a$. See Figure 3 for the sub-optimality gap and Table 4 for the absolute errors.

We observe that for low loads the $Gc\mu$-rule and the index-rule (21) behave very well. In this example, also the Whittle index policy performs close to optimal for low loads, while $W^{\infty}(n)$ and $w(n)$ do not. As the load grows larger, Whittle's index policy $W(n)$, and the fluid index policy $w(n)$ become near optimal. However, in this example the convergence towards optimality in absolute terms is much slower than for the previous example. The absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ is of order $10^{-3}$ for the indices $W(n)$ and $w(n)$ and of order $10^{-1}$ for $W^{\infty}(n)$ when $\rho = 16$. This phenomena is explained by the fact that the process lives around an area where the optimal policy prescribes to serve class-2 customers and the index policies prescribe to serve class-1 customers. As the workload increases this phenomena disappears.

The jump around the interval $\rho = [1.5, 3]$ for the index-rule (21) is a result of undefined values around $\lambda_k = \mu_k$.

### 10.2.2 Varying abandonment rates

In this section we evaluate the performance of the index policies for varying abandonment rates.

**Linear holding cost:** In this case, the five index policies mentioned above reduce to the $\tilde{c}\mu/\theta$-rule and the $\tilde{c}\mu$-rule, as explained in Section 10.2.1. As $\theta_k \to 0$, we observed in the numerical experiments that the $\tilde{c}\mu$-rule performs optimal, while the $\tilde{c}\mu/\theta$-rule might perform very bad when the abandonment rates are negligibly small. It is known that for the non-reneging case, the $\tilde{c}\mu$-rule is optimal in underload (the celebrated $c\mu$-rule for a multi-class M/M/1 queue). The $\tilde{c}\mu/\theta = (c + d\theta)\mu/\theta$ index might however give an opposite priority rule when $\theta$'s are very small, which explains the non-optimality of the $\tilde{c}\mu/\theta$-rule when $\theta_k$'s are very small.

**Quadratic holding cost:** Consider a system with two classes of customers. We assume quadratic holding costs $C_1(n) = \tilde{c}_{11}n + c_{21}n^2$ where, $\tilde{c}_{11} = (c_{11} + d_1\theta_1)$, and $C_2(n) = \tilde{c}_{21}n + c_{22}n^2$, where $\tilde{c}_{21} = (c_{21} + d_2\theta_2)$ and fix the following parameters: $\lambda = [4, 5]$, $\mu = [15, 17]$, $c_1 = [1, 4]$, $c_2 = [5, 1]$, $d = [2, 3]$, $\theta_1 = \epsilon_1 p$ and $\theta_2 = \epsilon_2 p$, where $\epsilon_1 = 0.05$ and $\epsilon_2 = 0.01$, and let $p$ vary. Hence, $\rho = \sum_k \rho_k < 1$ so that the stability of the system is assured as $\theta_k \to 0$.

In Figures 4 (right) we plot the sub-optimality gap as $p$ varies from 0 to 200, hence $\theta_1$ and $\theta_2$ range from $[0, 10]$ and $[0, 2]$, respectively. We observe for the $\theta$-dependent indices a sub-optimality gap of 25% around $p = 0$. As $\theta$ grows large, this gap disappears however very fast. Note that the $\theta$-independent indices work well, as we are in an underload scenario.

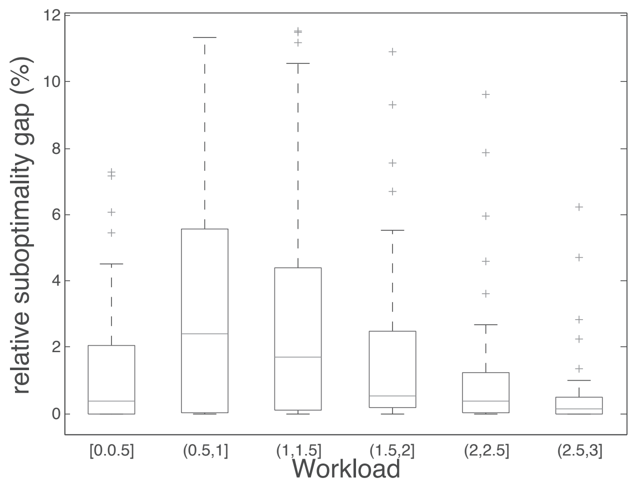


Figure 5: Sub-optimality gap of Whittles index policy, for randomly generated parameters. The edges of the box represent the 25th and 75th percentile, the line inside the box the mean value corresponding to all values in that box and the "+"s are the outliers.

#### 10.2.3 Example with random samples

In this section we assume two classes of customers and quadratic holding costs of type $\tilde{C}_1(n,a) = c_{11}(n-a)^+ + c_{21}((n-a)^+)^2 + d_1\theta_1(n-a)^+ + d_1'\theta_1'a$ and $\tilde{C}_2(n,a) = c_{12}(n-a)^+ + c_{22}((n-a)^+)^2 + d_2\theta_2(n-a)^+ + d_2'\theta_2 a$. We consider 360 samples with randomly generated values (in the interval $[0,1]$) for $\lambda_k, \mu_k, \theta_k', \theta_k$ and $c_k = [c_{k1}, c_{k2}]$ for $k = 1,2$. We compute the relative sub-optimality gap of Whittle's index policy, see Figure 5. We group the results in workload intervals of length 0.5, where for each interval we computed the sub optimality gap of 60 samples. In Figure 5 we plot for each interval the 25th and 75th percentiles, the average value with an horizontal line, and the outliers with "+". We observe that the average performance of Whittle's index policy is nearly optimal for high workloads, whereas the sub-optimality gap is largest for values of the workload in the interval $(0.5, 1]$.

## 11 Conclusions

In one of the main contributions of the paper we have derived a closed-form expression for Whittle's index for a multi-class queue with abandonments and convex holding cost. We have observed that in particular instances we can obtain simple expressions that enable to understand how the Whittle index policy depends on the input parameters. This was the case for linear holding cost, for convex holding cost as $\theta \to 0$ and also for convex holding cost for large values of the state. In the second main contribution we have established that in light-traffic and heavy-traffic regimes the Whittle index policy is asymptotically optimal. Finally, we have developed a fluid-based index policy, which is easy to implement and is equivalent to the Whittle index in limiting regimes. Numerical experiments for a wide range of parameters have shown that the Whittle index policy and the fluid index policy perform very well for a broad range of parameters.

This study opens several interesting research directions. The model considered in this paper could be generalized by considering a multi-server setting. All the results up to Section 5 can easily be adapted to the case with $M > 1$ servers. As explained in Section 8.1, in the linear holding cost case, the Whittle index policy, *i.e.*, the policy that serves the $M$ users with highest Whittle's index, is asymptotically optimal. For the general holding cost case, one would need to define how to use the state-dependent Whittle indices in a multi-server setting, and study the asymptotic behavior accordingly. Another interesting problem would be to develop the fluid index approach in a general setting. Preliminary results for birth-and-death processes can be found in [31]. This can be a fruitful approach to derive well performing index policies.

## 12 Aknowledgments

The authors would like to thank O.J. Boxma and A.J.E.M. Janssen for the proof of Lemma 1. The authors are grateful to the two anonymous referees for their valuable comments which helped improve the readability and focus of the paper.

## References

[1] M. Abramowitz and A.I. Stegun. *Handbook of Mathematical Functions, with Formulas, Graphs and Mathematical Tables.* Dover Publications, Inc., New York, 1965.

[2] D.F. Anderson. *Introduction to Stochastic Processes with Applications in the Biosciences.* University of Wisconsin at Madison, Available at www.math.wisc.edu/∼anderson/605F13/Notes/StochBio.pdf, 2013.

[3] P. S. Ansell, K. D. Glazebrook, J. Niño-Mora, and M. O'Keeffe. Whittle's index policy for a multi-class queueing system with convex holding costs. *Mathematical Methods of Operations Research*, 57(1):21–39, 2003.

[4] N.T. Argon, S. Ziya, and R. Righter. Scheduling impatient jobs in a clearing system with insights on patient triage in mass-casualty incidents. *Prob. Eng. Inf. Sci.*, 22(3):301–332, 2010.

[5] B. Ata and M.H. Tongarlak. On scheduling a multiclass queue with abandonments under general delay costs. *Queueing Systems*, 74:65–104, 2013.

[6] R. Atar, C. Giat, and N. Shimkin. The $c\mu/\theta$ rule for many-server queues with abandonment. *Operation Research*, 58(5):1427–1439, 2010.

[7] R. Atar, C. Giat, and N. Shimkin. On the asymptotic optimality of the $c\mu/\theta$ rule under ergodic cost. *Queueing Systems*, 67:127–144, 2011.

[8] F. Avram, D. Bertsimas, and M. Richard. Optimization of multiclass queuing networks: a linear control approach. *Stochastic Networks, eds. F.P. Kelly and R.J. Williams*, pages 199–234, 1995.

[9] U. Ayesta, P. Jacko, and V. Novak. A nearly-optimal index rule for scheduling of users with abandonment. In *IEEE Infocom*, 2011.

[10] N. Bäuerle. Asymptotic optimality of tracking policies in stochastic networks. *Ann. Appl. Probab*, 10:1065–1083, 2000.

[11] N. Bäuerle and U. Rieder. Optimal control of single-server fluid networks. *Queueing Syst.*, 35:185–200, 2000.

[12] D.P. Bertsekas. *Dynamic programming and optimal control.* Athena Scientific, 2005.

[13] S. Bhulai, H. Blok, and F.M. Spieksma. $k$ computing queues with customer abandonment: optimality of a generalized $c\mu$-rule by the smoothed rate truncation method. *Work in progress*, 2014.

[14] S. Bhulai, A.C. Brooms, and F.M. Spieksma. On structural properties of the value function for an unbounded jump markov process with an application to a processor sharing retrial queue. *Queueing Systems*, 76:425–446, 2014.

[15] C.F. Bispo. The single-server scheduling problem with convex costs. *Queueing Systems*, 73:261–294, 2013.

[16] C. Buyukkoc, P. Varaya, and J. Walrand. The $c\mu$ rule revisited. *Adv. Appl. Prob.*, 17:237–238, 1985.

[17] M. Chaichian and M. Demichev. *Path Integrals in Physics Volume 1: Stochastic Process and Quantum Mechanics*. Taylor & Francis, 2001.

[18] J.G. Dai. On positive harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, 5:49–77, 1995.

[19] J.G. Dai and S. He. Many-server queues with customer abandonment: A survey of diffusion and fluid approximations. *J. Syst. Sci. Syst. Eng.*, 21:1–36, 2012.

[20] D.G. Down, G. Koole, and M.E. Lewis. Dynamic control of a single server system with abandonments. *Queueing Systems*, 67:63–90, 2011.

[21] A. Gajrat and A. Hordijk. Fluid approximation of a controlled multiclass tandem network. *Queueing Systems*, 35:349–380, 2000.

[22] E. Gelenbe and I. Mitrani. *Analysis and Synthesis of Computer Systems*. London: Academic Press, 1980.

[23] J.C. Gittins, K. Glazebrook, and R. Weber. *Multi-armed Bandit Allocation Indices*. Wiley, 2011.

[24] K.D. Glazebrook, P.S. Ansell, R.T. Dunn, and R.R. Lumley. On the optimal allocation of service to impatient tasks. *J. Appl. Prob.*, 41:51–72, 2004.

[25] K.D. Glazebrook, D.J. Hodge, C. Kirkbride, and R.J. Minty. Stochastic scheduling: A short history of index policies and new approaches to index generation for dynamic resource allocation. *J. of Sched.*, 2013.

[26] K.D. Glazebrook, C. Kirkbride, and J. Ouenniche. Index policies for the admission control and routing of impatient customers to heterogeneous service stations. *Operations Research*, 57:975–989, 2009.

[27] J. Hasenbein and D. Perry (Eds). Special issue on queueing systems with abandonments. *Queueng Systems*, 75(2–4):111–384, 2013.

[28] J. Kim and A. Ward. Dynamic scheduling of a GI/GI/1+GI queue with multiple customer classes. *Queueing Systems*, 75(2-4):339–384, 2013.

[29] M. Larrañaga, U. Ayesta, and I.M. Verloop. Dynamic fluid-based scheduling in a multi-class abandonment queue. *Performance Evaluation*, 70:841–858, 2013.

[30] M. Larrañaga, U. Ayesta, and I.M. Verloop. Index policies for a multi-class queue with convex holding cost and abandonment. In *Proceedings of ACM SIGMETRICS*, 2014.

[31] M. Larrañaga, U. Ayesta, and I.M. Verloop. Stochastic and fluid index policies for resource allocation problems. In *Proceedings of IEEE Infocom*, 2015.

[32] A. Mandelbaum and S. Stolyar. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized $c\mu$-rule. *Operations Research*, 52(6):836–855, 2004.

[33] J. Niño-Mora. Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15(2):161–198, 2007.

[34] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[35] J. A. van Mieghem. Dynamic scheduling with convex delay costs: The generalized $c\mu$ rule. *The Annals of Applied Probability*, 5(3):808–833, 1995.

[36] I.M. Verloop. Asymptotically optimal priority policies for indexable and non-indexable restless bandits. CNRS Technical Report, hal-00743781, September 2015.

[37] I.M. Verloop and R. Núñez-Queija. Asymptotically optimal parallel resource assignment with interference. *Queueing Systems*, 65(1):43–92, 2010.

[38] R.R. Weber and G. Weiss. On an index policy for restless bandits. *J. Appl. Prob.*, 27:637–648, 1990.

[39] G. Weiss. On optimal draining of reentrant fluid lines. *Stochastic Networks, eds. F.P. Kelly and R.J. Williams*, pages 91–103, 1995.

[40] P. Whittle. Restless bandits: Activity allocation in a changing world. *J. Appl. Prob.*, 25:287–298, 1988.

# Appendix

## A Proof of Proposition 1

In Proposition 1 we aim at proving that threshold policy $\varphi = n$ is an optimal solution of the relaxed problem (8). In order to do so, we are left to prove the convexity of the value function $V$. We will therefore prove that the value function that corresponds to the truncated system $V^L(m)$ (truncated by $L > 1$) is convex. Having done this, due to the result in [14, Th. 3.1] we have that $V^L(m) \to V(m)$ as $L \to \infty$ and hence, the convexity of $V^L$ for all $L$ will imply convexity of the function $V$. In order to apply [14, Th. 3.1] we need to make sure that the conditions required are satisfied. We therefore check the conditions required by [14, Th. 3.1] in A.1, and prove the convexity of $V^L$ in A.2.

### A.1 Conditions to be checked for [14, Th. 3.1]

Let us first present the following definition:

**Definition 6** *A function $f : E \longrightarrow \mathbb{R}_+$ is a moment function if there exists an increasing sequence of finite sets $E_r \uparrow E$, $r \to \infty$, such that* $\inf\{f(m) : m \notin E_r\} \to \infty$ *as $r \to \infty$. (Where $E$ is the state space).*

Let us define $q^{\varphi,L}(m, m-1) = \mu S^\varphi(m) + \theta' S^\varphi(m) - \theta(m - S^\varphi(m))$, and recall that $q^{\varphi,L}(m, m+1) = \lambda\left(1 - \frac{m}{L}\right)$. The conditions to be checked in [14, Th. 3.1] are the following:

1. There exists a moment function $f : \mathbb{N} \cup \{0\} \longrightarrow \mathbb{R}_+$, constants $\alpha, \beta > 0$ and $M > 0$ such that

$$\sum_{\tilde{m}=0}^{\infty} q^{\varphi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m) + \beta \mathbb{1}_{\{m<M\}}(m), \text{ for all } \varphi, L,$$

   where $\varphi$ defines the policy followed, $L$ is the truncating parameter and $q^{\varphi,L}(m, \tilde{m})$ the transition rate from $m$ to $\tilde{m}$ under $\varphi$ and $L$.

2. $(S^\varphi(m), L) \mapsto q^{\varphi,L}(m, \tilde{m})$ and $(S^\varphi(m), L) \mapsto \sum_{\tilde{m}} q^{\varphi,L}(m, \tilde{m}) f(\tilde{m})$ are continuous functions in $S^\varphi(m)$ and $L$ for all $m$ and $\tilde{m}$.

We define $f(m) := \mathrm{e}^{\epsilon m}$, where $\epsilon > 0$. We can construct $E_r = \{0, \ldots, r\}$ such that $E_r$ is finite, $E_r \uparrow \mathbb{N} \cup \{0\}$ as $r \to \infty$ and $\inf\{f(m) : m \notin E_r\} \to \infty$. The objective is then to see, that there exists $\epsilon > 0$, an $M > 0$ and a constant $\alpha > 0$, such that

$$\sum_{\tilde{m}=0}^{\infty} q^{\varphi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m), \text{ for all } m \geq M,$$

that is,

$$\begin{aligned}
&\lambda\left(1 - \frac{m}{L}\right) \mathrm{e}^{\epsilon(m+1)} + \left((\mu + \theta')S^\varphi(m) + \theta(m - S^\varphi(m))\right) \mathrm{e}^{\epsilon(m-1)} \\
&- \left(\left(1 - \frac{m}{L}\right) + (\mu + \theta')S^\varphi(m) + \theta(m - S^\varphi(m))\right) \mathrm{e}^{\epsilon m} \leq -\alpha \mathrm{e}^{\epsilon m}, \text{ for all } m \geq M.
\end{aligned}$$

After some algebra we get

$$\lambda\left(1-\frac{m}{L}\right)(\mathrm{e}^{\epsilon}-1)+\big((\mu+\theta'-\theta)S^{\varphi}(m)+\theta m\big)(\mathrm{e}^{-\epsilon}-1)\le-\alpha,\ \text{for all } m\ge M.$$

Note that $\lambda(1-m/L)(\mathrm{e}^{-\epsilon}-1)$ can be upper bounded by a constant, $\kappa_1$, and $(\mu+\theta'-\theta)S^{\varphi}(m)(\mathrm{e}^{\epsilon}-1)$ can be upper bounded by $\kappa_2$. Besides, $\theta m(\mathrm{e}^{-\epsilon}-1)<0$. Hence, we can find $M$ large enough so that $-\theta m(\mathrm{e}^{-\epsilon}-1)\ge\kappa_1+\kappa_2$ for all $m\ge M$. This proves that condition (1) is satisfied.

Condition (2), i.e., the continuity of the functions $(S^{\varphi}(m),L)\mapsto q^{\varphi,L}(m,\tilde m)$ and $(S^{\varphi}(m),L)\mapsto\sum_{\tilde m}q^{\varphi,L}(m,\tilde m)f(\tilde m)$ in $L$ and $S^{\varphi}(m)$ is satisfied by definition of transition rates.

## A.2 Convexity of $V^L$

For the ease of clarity we define $\omega:=\mu+\theta'-\theta$ throughout this proof. W.l.o.g. assume $\lambda+\mu+\theta'+\theta L=1$. For $n\in\{0,1,\dots,L\}$ we define $V_t^L(n)$ by $V_0^L(n)=0$ and

$$\begin{aligned}V_{t+1}^L(n)=&\lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\})\\&+\min\{-W+\tilde C(n,0)+\omega V_t^L(n),\tilde C(n,1)+\omega V_t^L((n-1)^+)\}\\&+\theta nV_t^L((n-1)^+)+\lambda\frac{n}{L}V_t^L(n)+(L-n+1)\theta V_t^L(n).\end{aligned}$$

We will prove that $V_t^L$ is a convex function for $n\le L-1$, that is,

$$2V_t^L(n)\le V_t^L((n-1)^+)+V_t^L(n+1),\ \text{for } n\le L-1. \tag{29}$$

The function $V_t^L$ being convex, for any $t$, implies convexity of $V^L$ and concludes the proof.

In order to prove convexity of $V_t^L$ we first prove that $V_t^L(\cdot)$ is a non-decreasing function. The proof follows by induction: $V_0^L(n)=0$ is non-decreasing for $t=0$, then we assume $V_t^L(n)$ is non-decreasing and we prove that

$$V_{t+1}^L(n+1)-V_{t+1}^L(n)\ge 0 \text{ for all } n\le L-1. \tag{30}$$

Let us first consider the terms multiplied by $\lambda$ in $V_{t+1}^L(n+1)-V_{t+1}^L(n)$, that is,

$$\begin{aligned}&\lambda\left(1-\frac{n+1}{L}\right)V_t^L(\min\{n+2,L\})+\lambda\frac{n+1}{L}V_t^L(\min\{n+1,L\})\\&-\lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\})-\lambda\frac{n}{L}V_t^L(n)\\&\ge\lambda\left(1-\frac{n+1}{L}\right)(V_t^L(\min\{n+2,L\})-V_t^L(\min\{n+1,L\}))\\&\quad+\lambda\frac{n}{L}(V_t^L(\min\{n+1,L\}-V_t^L(n)))\ge 0,\end{aligned}$$

where the last inequality holds due to the non-decreasingness of $V_t^L(n)$. Let us now consider the terms multiplied by $\theta$ in $V_{t+1}^L(n+1)-V_{t+1}^L(n)$, namely,

$$\begin{aligned}&\theta(n+1)V_t^L(n)+(L-n-1)\theta V_t^L(\min\{n+1,L\})-\theta nV_t^L((n-1)^+)-(L-n)\theta V_t^L(n)\\&\ge\ \theta n(V_t^L(n)-V_t^L((n-1)^+))+(L-n-1)(V_t^L(\min\{n+1,L\})-V_t^L(n))\ge 0,\end{aligned}$$

where, again, the last inequality holds due to $V_t^L(n)$ being non-decreasing for all $n \leq L-1$. Finally, let us consider the min-terms in $V_{t+1}^L(n+1) - V_{t+1}^L(n)$. It is straightforward that

$$\begin{aligned}
&\min\{-W + \tilde{C}(\min\{n+1, L\}, 0) + (\mu+\theta')V_t^L(\min\{n+1, L\}), \\
&\qquad \tilde{C}(\min\{n+1, L\}, 1) + (\mu+\theta')V_t^L(n)\} \\
&- \min\{-W + \tilde{C}(n, 0) + (\mu+\theta')V_t^L(n), \\
&\qquad \tilde{C}(n, 1) + (\mu+\theta')V_t^L((n-1)^+)\} \geq 0,
\end{aligned}$$

due to $\tilde{C}$ and $V_t^L$ being non-decreasing. This proves (30) and hence we showed that $V_t^L(n)$ is non-decreasing.

Equation (29) for $n = 0$ follows directly from $V_t^L(\cdot)$ being non-decreasing. In the remainder of the proof we therefore prove Equation (29) for $n \geq 1$.

We will prove convexity (29) by induction on $t$. Since $V_0^L(n) = 0$, it holds for $t = 0$. Now assume $V_t^L$ is convex. For $1 \leq n \leq L-1$ we have

$$\begin{aligned}
2V_{t+1}^L(n) = &2\lambda\left(1 - \frac{n}{L}\right)V_t^L(n+1) + 2\lambda\frac{n}{L}V_t^L(n) + 2\theta n V_t^L(n-1) + 2(L-n+1)\theta V_t^L(n) \\
&+ 2\min\{-W + \tilde{C}(n, 0) + \omega V_t^L(n), \tilde{C}(n, 1) + \omega V_t^L(n-1)\}. \qquad (31)
\end{aligned}$$

We need to show that this is less than or equal to $V_{t+1}^L(n-1) + V_{t+1}^L(n+1)$, which is given by

$$\begin{aligned}
&\lambda\left(1 - \frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1 - \frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1) \\
&+ \theta(n-1)V_t^L((n-2)^+) + \theta(n+1)V_t^L(n) + (L-n+2)\theta V_t^L(n-1) + (L-n)\theta V_t^L(n+1) \\
&+ \min\{-W + \tilde{C}(n-1, 0) + \omega V_t^L(n-1), \tilde{C}(n-1, 1) + \omega V_t^L((n-2)^+)\} \\
&+ \min\{-W + \tilde{C}(n+1, 0) + \omega V_t^L(n+1), \tilde{C}(n+1, 1) + \omega V_t^L(n)\}. \qquad (32)
\end{aligned}$$

We first consider the two terms multiplied by $\lambda$ in (31) and show that they are smaller than or equal to

$$\lambda\left(1 - \frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1 - \frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1). \quad (33)$$

When $1 \leq n < L-1$, then for the terms multiplied by $\lambda$ in (31) we can write

$$\begin{aligned}
&2\left(1 - \frac{n}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) = 2\left(1 - \frac{n+1}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1) \\
&\leq \left(1 - \frac{n-1}{L}\right)V_t^L(n) - \frac{2}{L}V_t^L(n) + \left(1 - \frac{n+1}{L}\right)V_t^L(n+2) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1), \quad (34)
\end{aligned}$$

by convexity of $V_t^L$. Since by convexity $2\frac{n-1}{L}V_t^L(n) \leq \frac{n-1}{L}(V_t^L(n-1) + V_t^L(n+1))$, we obtain that (34) is smaller than or equal to (33). When $n = L-1$, it reduces to verifying $2(1-2/L)V_t^L(L-1) \leq (1-2/L)(V_t^L(L-2) + V_t^L(L))$, which follows from convexity of $V_t^L$.

For the terms multiplied by $\theta$, we need to show for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2nV_t^L(n-1) + 2V_t^L(n) + 2(L-n)V_t^L(n) \\
&\leq (n-1)V_t^L((n-2)^+) + (n+1)V_t^L(n) + 2V_t^L(n-1) + (L-n)(V_t^L(n-1) + V_t^L(n+1)).
\end{aligned}$$

We apply the inequality $2V_t^L(n-1) \leq V_t^L((n-2)^+) + V_t^L(n)$ on the right hand side and the whole initial inequality reduces to

$$2nV_t^L(n-1) + 2(L-n)V_t^L(n) \leq n(V_t^L((n-2)^+) + V_t^L(n)) + (L-n)(V_t^L(n-1) + V_t^L(n+1)),$$

which holds by convexity of $V_t^L$.

We now consider the min-terms. We will condition on the possible optimal actions in states $n-1$ and $n+1$. Since at time $t$ we have that $V_t^L$ is convex, the optimal actions satisfy the monotonicity property. Denote by $a_n^* \in \{0,1\}$ the optimal action in state $n$, with action 0 (1) being passive (active). Then, by monotonicity there are the following three possibilities: $(a_{n-1}^*, a_{n+1}^*)$ equals $(0,0), (0,1)$ or $(1,1)$. First assume $a^* = (0,1)$. Then, we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq -W + \tilde{C}(n,0) + \omega V_t^L(n) + \tilde{C}(n,1) + \omega V_t^L(n-1) \\
&\leq -W + \tilde{C}(n-1,0) + \omega V_t^L(n) + \tilde{C}(n+1,1) + \omega V_t^L(n-1) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\},
\end{aligned} \tag{35}$$

where in the second inequality we used that $C$ and hence $\tilde{C}$ satisfies (2). In the case $a^* = (1,1)$ we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq 2\tilde{C}(n,1) + 2\omega V_t^L(n-1) \\
&\leq \tilde{C}(n-1,1) + \tilde{C}(n+1,1) + \omega(V_t^L((n-2)^+) + V_t^L(n)) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\}.
\end{aligned} \tag{36}$$

In the second inequality we used the convexity of $C$ (and hence of $\tilde{C}$) and the convexity of $V_t^L$.

When $a^* = (0,0)$ we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq -2W + 2\tilde{C}(n,0) + 2\omega V_t^L(n) \\
&\leq -2W + \tilde{C}(n-1,0) + \tilde{C}(n+1,0) + \omega V_t^L(n-1) + \omega V_t^L(n+1) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\}.
\end{aligned} \tag{37}$$

In the second inequality we used the convexity of $C$ (and hence of $\tilde{C}$) and the convexity of $V_t^L$.

Hence, we have that (31) is less than or equal to $V_{t+1}^L(n-1) + V_{t+1}^L(n+1)$, hence $V_{t+1}^L$ is convex. This concludes the proof for convexity of $V_t^L(\cdot)$. Since $V_t^L \to V^L$ as $t \to \infty$ [34, Chap. 9.4], convexity of $V_t^L(\cdot)$ implies convexity of $V^L(\cdot)$.

# B Proof of Theorem 1

In this section we prove that the steps described in Theorem 1 indeed defines Whittle's index correctly. To do so let us assume that the steps stop at iteration $J \in \mathbb{N} \cup \{\infty\}$, and hence

$n_J = \infty$. We further set $W_i := W_J$ and $n_i = \infty$ for all $i \in \{J+1, \ldots\} \cup \{\infty\}$. We will prove that $W_0 < W_1 < W_2 < \ldots$, and note that by construction $n_i$ for $i \in \mathbb{N} \cup \{0, \infty\}$ is an increasing sequence. Let us prove $W_i < W_{i+1}$ for all $i \in \{0, 1, 2, \ldots\} \cup \{\infty\}$. We have from the characterization of $W_i$ that

$$\begin{aligned}
&\frac{\mathbb{E}(\tilde{C}(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(\tilde{C}(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} \\
&> \frac{\mathbb{E}(\tilde{C}(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(\tilde{C}(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} \\
&\implies \left(\mathbb{E}(\tilde{C}(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(\tilde{C}(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))\right) \left(\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)\right) \\
&> \left(\mathbb{E}(\tilde{C}(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(\tilde{C}(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))\right) \left(\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)\right),
\end{aligned}$$

and adding $\mathbb{E}(\tilde{C}(N^{n_i}, S^{n_i}(N^{n_i}))(\sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m))$ on both sides of the inequality, after some algebra we obtain

$$\begin{aligned}
W_{i+1} &= \frac{\mathbb{E}(\tilde{C}(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(\tilde{C}(N^{n_i}, S^{n_i}(N^{n_i})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m)} \\
&> \frac{\mathbb{E}(\tilde{C}(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(\tilde{C}(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} = W_i.
\end{aligned}$$

To prove that the steps given in Theorem 1 indeed define the Whittle index we have to show that,

1. the threshold policy $-1$ is optimal for problem (8) for all $W$ such that $W < W_0$.
2. The threshold policy $n_i < \infty$ is optimal for problem (8) for all $W$ such that $W_i < W < W_{i+1}$.
3. And finally that the policy $\infty$, is optimal for problem (8) for all $W$ such that $\infty > W > W_J$ and $J < \infty$.

To show 1., note that for all $W < W_0$

$$\begin{aligned}
&W \sum_{m=0}^{n} \pi^n(m) < \mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - \mathbb{E}(\tilde{C}(N^{-1}, S^{-1}(N^{-1}))), \\
&\implies \mathbb{E}(\tilde{C}(N^{-1}, S^{-1}(N^{-1}))) < \mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m), \forall n,
\end{aligned}$$

that is, $g^{(-1)}(W) < g^{(n)}(W)$ for all $n \in \mathbb{N}^0$, and hence $g(W) = g^{(-1)}(W)$. Policy $-1$ is therefore optimal for problem (8) for $W < W_0$.

We will prove 2. by induction, observe from the definition of $n_0$ that for all $n \geq 0$

$$\mathbb{E}(\tilde{C}(N^{n_0}, S^{n_0}(N^{n_0}))) - W_0 \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - W_0 \sum_{m=0}^{n} \pi^n(m),$$

that is, $g^{(n_0)}(W_0) \leq g^{(n)}(W_0)$, for all $n \geq 0$. Besides, we trivially have that $g^{(n_0)}(W_0) \leq g^{(-1)}(W_0)$. We have proven in the proof of Proposition 2 that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in $n$, and therefore for all $n \leq n_0$ and $W_0 < W$

$$\mathbb{E}(\tilde{C}(N^{n_0}, S^{n_0}(N^{n_0}))) - W \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(\tilde{C}(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m)$$
$$\implies g^{(n_0)}(W) \leq g^{(n)}(W).$$

In particular, $g^{(n_0)}(W) \leq g^{(n)}(W)$ is satisfied for all $W_0 < W < W_1$ and $n \leq n_0$. Similarly, from the definition of $W_1$ we have that $g^{(n_0)}(W_1) \leq g^{(n)}(W_1)$ for all $n \geq n_0 + 1$, and again using that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing we obtain $g^{(n_0)}(W) \leq g^{(n)}(W)$ for all $W_0 < W < W_1$ and $n \geq n_0 + 1$.

We have therefore proven that $g^{(n_0)}(W) \leq g^{(n)}(W)$ for all $n$ and $W_0 < W < W_1$, that is, policy $n_0$ is optimal for all $W$ such that $W_0 < W < W_1$. This establishes the first step of the induction $i = 0$. Let us now assume that it holds for step $i - 1 \geq 0$, that is, $n_i$ is an optimal policy for problem (8), given $W$ such that $W_{i-1} < W < W_i$. And let us assume $n_i < \infty$. The definition of $W_i$ together with the fact that $n_{i-1}$ is optimal for the choice of $W$ such that $W_{i-1} < W < W_i$, imply

$$g^{(n_{i-1})}(W_i) = g^{(n_i)}(W_i) \leq g^{(n)}(W_i), n \geq 0.$$

Recall that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in $n$ and therefore

$$g^{(n_i)}(W) \leq g^{(n)}(W), n \leq n_i, W_i < W < W_{i+1}.$$

Besides, from the definition of $W_{i+1}$ we have

$$g^{(n_i)}(W) \leq g^{(n)}(W), n \geq n_i + 1, W_i < W < W_{i+1}.$$

We therefore have obtained that threshold policy $n_i$ is optimal for problem (8) given $W$ such that $W_i < W < W_{i+1}$.

Finally, we prove 3. for $J < \infty$, note that from the induction followed in the previous point we have that

$$g^{(n_{J-1})}(W_J) = g^{(n_J)}(W_J) \leq g^{(n)}(W_J), n \geq 0,$$

and the fact that $\sum_{m=0}^{n} \pi^n(m)$ is increasing in $n$ gives that

$$g^{(n_J)}(W) < g^{(n)}(W), n \leq n_J = \infty, W_J < W.$$

Which concludes the proof of the theorem.

# C Proof of Proposition 3

For ease of notation we omit subscript $k$ from the notation in the proof. To calculate Whittle's index as in Theorem 1 we need to consider the monotone policies $n$ and $n-1$ in which the server is active in states $m \geq n+1$ and $m \geq n$, respectively.

Let us consider the policy $n$ first. Let $f^n(ab)$ and $f^n(ser)$ denote the fraction of customers that end up abandoning and being served, respectively. A rate conservation argument implies that all arriving users either abandon or are served, thus $\lambda = \lambda f^n(ab) + \lambda f^n(ser)$. Conditioning on the state, the rate of abandonment from the system can be written as $\sum_{m=0}^{\infty} \theta m \pi^n(m) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m)$, and equating the rates we get the relation

$$\theta \mathbb{E}(N^n) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m) = \lambda f^n(ab) = \lambda(1 - f^n(ser)). \tag{38}$$

Conditioning on the state, the rate of service is given by $\sum_{m=n+1}^{\infty} \mu\pi^n(m)$, and we get the relation

$$\lambda f(ser) = \mu \sum_{m=n+1}^{\infty} \pi^n(m),$$

and substituting in (38) we get

$$\mathbb{E}(N^n) = \frac{\lambda}{\theta} + \frac{\theta - \theta' - \mu}{\theta} \sum_{m=n+1}^{\infty} \pi^n(m),$$

where $N^n$ denotes the stationary number of class-$k$ customers in the system under the threshold policy $n$. We calculate now the average holding cost. Plugging the holding cost $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k a$ in the total cost relation (4) we get $\tilde{C}(n, a) = \tilde{c}n + a(\tilde{c}' - \tilde{c})$, where the constants $\tilde{c}$ and $\tilde{c}'$ are defined in the statement. The average cost then becomes

$$\begin{aligned}
&\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) = \tilde{c}\mathbb{E}(N^n) + (\tilde{c}' - \tilde{c}) \sum_{m=n+1}^{\infty} \pi^n(m) \\
&= \tilde{c}\frac{\lambda}{\theta} + \left(\frac{\tilde{c}(\theta - \theta' - \mu)}{\theta} + \tilde{c}' - \tilde{c}\right) \sum_{m=n+1}^{\infty} \pi^n(m) = \tilde{c}\frac{\lambda}{\theta} + \left(\tilde{c}' - \frac{\tilde{c}(\theta' + \mu)}{\theta}\right) \sum_{m=n+1}^{\infty} \pi^n(m).
\end{aligned}$$

We substitute now all the terms in (15) to get

$$W(n) = \frac{\tilde{c}(\mu + \theta')}{\theta} - \tilde{c}', \tag{39}$$

which concludes the proof.

# D Proof of Proposition 4

For ease of notation we drop the dependency on $k$ throughout the proof.

The index in the case $\mu + \theta' = \theta$ was obtained in (16), therefore we assume $\mu + \theta' > \theta$ throughout the proof. First of all recall that the steady-state probabilities $\pi^n(i)$ for policy $n$ and state $i$ are given by (11). To compute Whittle's index for large values of $n$, we need to compute $\pi^n(i) - \pi^{n-1}(i), \forall i \geq 0$. Let us start by $i = 0$, that is,

$$\begin{aligned}
\pi^n(0) - \pi^{n-1}(0) =& \frac{\left(\pi^{n-1}(0)\right)^{-1} - \left(\pi^n(0)\right)^{-1}}{\left(\pi^n(0)\pi^{n-1}(0)\right)^{-1}} \\
=& \left(\sum_{i=1}^{\infty} \prod_{m=1}^{i} \frac{q^{n-1}(m-1, m)}{q^{n-1}(m, m-1)} - \sum_{i=1}^{\infty} \prod_{m=1}^{i} \frac{q^n(m-1, m)}{q^n(m, m-1)}\right) \pi^n(0)\pi^{n-1}(0).
\end{aligned}$$

The following observations on the transition rates will be used throughout the proof:

$$q^n(m, m-1) = q^{n-1}(m, m-1), \;\; \forall\, m \neq n, m \geq 1, \tag{40}$$

$$q^n(m-1, m) = q^{n-1}(m-1, m), \;\; \forall\, m \geq 1. \tag{41}$$

Taking these relations into account together with the fact that $q^n(n,n-1)-q^{n-1}(n,n-1)=\theta-\mu-\theta'$, we get after some calculations

$$\pi^n(0)-\pi^{n-1}(0)=\pi^n(0)\pi^{n-1}(0)\sum_{i=n}^{\infty}\prod_{\substack{m=1\\m\neq n}}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\left(\frac{1}{q^{n-1}(n,n-1)}-\frac{1}{q^n(n,n-1)}\right)$$
$$=\pi^n(0)\pi^{n-1}(0)\frac{\theta-\mu-\theta'}{q^{n-1}(n,n-1)}\sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}.$$

Since $q^n(m-1,m)=\lambda$ for all $m\geq 1$, $q^n(m,m-1)=\theta m$ for all $1\leq m\leq n-1$ and $q^n(m,m-1)=\mu+\theta'+\theta(m-1)$ for all $m\geq n$, together with $\pi^n(0)$ given as in (11), we observe that

$$\frac{\pi^n(0)\pi^{n-1}(0)}{q^{n-1}(n,n-1)}\in\mathcal{O}\left(\frac{1}{n}\right)\qquad\text{and}\qquad\sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\in\mathcal{O}\left(\frac{1}{n!}\right).$$

We then get that

$$\pi^n(0)-\pi^{n-1}(0)\in\mathcal{O}\left(\frac{1}{nn!}\right).\tag{42}$$

We can now compute $\pi^n(i)-\pi^{n-1}(i)$, for all $0<i\leq n-1$. Using (41), we obtain for $i\leq n-1$,

$$\pi^n(i)-\pi^{n-1}(i)=\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}(\pi^n(0)-\pi^{n-1}(0)).$$

Due to (42) and since $q^n(m,m-1)=\theta m$, and $q^n(m-1,m)=\lambda$ for all $m\leq n-1$, we obtain for $i\leq n-1$

$$\pi^n(i)-\pi^{n-1}(i)=\frac{\lambda^i}{i!\theta^i}(\pi^n(0)-\pi^{n-1}(0))\in\mathcal{O}\left(\frac{1}{nn!}\right),\tag{43}$$

For states $i\geq n$, with $n$ sufficiently large, we have the following:

$$\pi^n(i)-\pi^{n-1}(i)=\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\pi^n(0)-\prod_{m=1}^{i}\frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)}\left(\pi^n(0)-\pi^n(0)+\pi^{n-1}(0)\right).$$

From observation (42), together with $\prod_{m=1}^{i}\dfrac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)}\in\mathcal{O}\left(\frac{1}{i!}\right)$, we obtain

$$\pi^n(i)-\pi^{n-1}(i)=\mathcal{O}\left(\frac{1}{i!n!n}\right)+\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\pi^n(0)-\prod_{m=1}^{i}\frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)}\pi^n(0).$$

After some calculations and by observations (40) and (41) we obtain

$$\pi^n(i)-\pi^{n-1}(i)=\left(\frac{1}{q^n(n,n-1)}-\frac{1}{q^{n-1}(n,n-1)}\right)\prod_{\substack{m=1\\m\neq n}}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}+\mathcal{O}\left(\frac{1}{i!n!n}\right)$$
$$=\frac{\mu+\theta'-\theta}{q^{n-1}(n,n-1)}\pi^n(i)+\mathcal{O}\left(\frac{1}{i!n!n}\right),\tag{44}$$

for $i \geq n$. Recall from (18) that Whittle's index can be written as $d(\mu+\theta') - d'\theta' + W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. $W^c(n)$ can be written as follows

$$W^c(n) = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\mathcal{O}(1/n!n)}, \tag{45}$$

with

$$\begin{aligned}
\xi_1(n) &:= \sum_{i=1}^{n-1} C(i,0)(\pi^n(i)-\pi^{n-1}(i)),\\
\xi_2(n) &:= C(n,0)\pi^n(n) - C(n,1)\pi^{n-1}(n),\\
\xi_3(n) &:= \sum_{i=n+1}^{\infty} C(i,1)(\pi^n(i)-\pi^{n-1}(i)).
\end{aligned} \tag{46}$$

Recall now the assumption that the holding costs $C(n,1)$ and $C(n,0)$ are upper bounded by polynomials of finite degrees $P<\infty$ and $Q<\infty$, respectively. Hence, we can write $C(n,a) = E(n,a) + o(1)$, for large values of $n$, where $E(n,1) = \sum_{i=0}^{P} C^{(P,i)} n^i$, with $C^{(P,i)} := \lim_{n\to\infty} \frac{C(n,1)-\sum_{j=i+1}^{P} C^{(P,j)} n^j}{n^i}$, and $E(n,0) = \sum_{i=0}^{Q} E^{(Q,i)} n^i$, with $E^{(Q,i)} := \lim_{n\to\infty} \frac{C(n,0)-\sum_{j=i+1}^{Q} E^{(Q,j)} n^j}{n^i}$. We assume w.l.o.g. that $P$ is such that $C^{(P,P)}>0$ and $Q$ such that $E^{(Q,Q)}>0$. We then have

$$\begin{aligned}
\xi_1(n) &= \sum_{i=1}^{n-1} E(i,0)(\pi^n(i)-\pi^{n-1}(i)) + o(1),\\
\xi_2(n) &= E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n) + o(1),\\
\xi_3(n) &= \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i)-\pi^{n-1}(i)) + o(1).
\end{aligned}$$

We now define $\hat{\xi_1} := \sum_{i=1}^{n-1} E(i,0)(\pi^n(i)-\pi^{n-1}(i))$, and with the result obtained in Equation (43) we have that for large values of $n$ $\hat{\xi_1}(n) \in \mathcal{O}\left(\frac{n^{Q-1}}{n!}\right) \subset o(1)$. Hence, for large values of $n$, $\xi_1(n) \in o(1)$. Let us now define $\hat{\xi_2}(n) := E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n)$. Using (40) and (41) we have after some calculations,

$$\hat{\xi_2}(n) = \frac{\prod_{m=1}^{n} q^n(m-1,m)}{\prod_{m=1}^{n-1} q^n(m,m-1)}\left(\frac{E(n,0)\pi^n(0)}{q^n(n,n-1)} - \frac{E(n,1)\pi^{n-1}(0)}{q^{n-1}(n,n-1)}\right).$$

We recall that $q^{n-1}(n,n-1) = \mu+\theta'+\theta(n-1)$ and $q^n(n,n-1) = \theta n$, which together with (42) give, after some calculations,

$$\begin{aligned}
\hat{\xi_2}(n) = \prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)} &\left(\left(E(n,0)-E(n,1)\right)\pi^n(0) + \mathcal{O}\left(\frac{n^{P-1}}{n!}\right)\right)\\
&+ \pi^n(n)(\mu+\theta'-\theta)\frac{E(n,0)}{q^{n-1}(n,n-1)}.
\end{aligned}$$

Since for large values of $n$

$$\prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)} \cdot \mathcal{O}\left(\frac{n^{P-1}}{n!}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{(n!)^2}\right) \subset o(1),$$

we conclude that

$$\xi_2(n) = \frac{\pi^n(n)}{q^{n-1}(n,n-1)} \Big( \theta n (E(n,0) - E(n,1)) + (\mu + \theta' - \theta) E(n,0) \Big) + o(1). \tag{47}$$

Finally, we compute $\hat{\xi}_3(n) := \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i) - \pi^{n-1}(i))$. From (44) we see that

$$\hat{\xi}_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + \sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right).$$

Since for large values of $n$ $\sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{i!n!}\right) \subset o(1)$, we obtain

$$\xi_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + o(1). \tag{48}$$

Now using $\xi_1 \in o(1)$, the expression of $\xi_2(n)$ in (47) and (48) and letting $n$ be large, we see that $\frac{\xi_1(n)}{\pi^n(n)} \in o(1)$, and,

$$\begin{aligned}
\frac{\xi_2(n)}{\pi^n(n)} &= \frac{\theta n (E(n,0) - E(n,1))}{\mu + \theta' + \theta(n-1)} + \frac{(\mu + \theta' - \theta)E(n,0)}{\mu + \theta' + \theta(n-1)} + o(1) \\
&= E(n,0) - E(n,1) + \frac{(\mu + \theta' - \theta)}{\theta n} E(n,0) + o(1) \\
&= E(n,0) - E(n,1) + \frac{(\mu + \theta' - \theta)}{\theta} \sum_{j=1}^{Q} E^{(P,j)} n^{j-1} + o(1),
\end{aligned}$$

and

$$\begin{aligned}
\frac{\xi_3(n)}{\pi^n(n)} &= \frac{\mu + \theta' - \theta}{\mu + \theta' + \theta(n-1)} \cdot \sum_{i=n+1}^{\infty} E(i,1) \prod_{m=n+1}^{i} \frac{\lambda}{\mu + \theta' + \theta(m-1)} + o(1) \\
&= \frac{\mu + \theta' - \theta}{\theta n} \sum_{i=n+1}^{\infty} \sum_{j=0}^{P} C^{(P,j)} i^j \left(\frac{\lambda}{\theta m}\right)^{i-n} + o(1).
\end{aligned}$$

Define $\tilde{W}^c(n)$ as $W^c(n)$ for large values of $n$. Substituting the expressions for $\xi_1(n)/\pi^n(n), \xi_2(n)/\pi^n(n)$ and $\xi_3(n)/\pi^n(n)$ in Equation (45), we obtain

$$\begin{aligned}
&\tilde{W}^c(n) = (E(n,0) - E(n,1)) + (\mu + \theta' - \theta)/\theta \\
&\times \left( \sum_{j=1}^{Q} E^{(Q,j)} n^{j-1} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + o(1),
\end{aligned}$$

as $n \to \infty$, that is, the expression in Equation (19). $E(n,a)$ being non-decreasing together with Condition 2 implies that $\tilde{W}^c$ is non-decreasing, and hence $W^\infty$ as well, which concludes the proof.

# E Proof of Propostion 5

For ease of notation we drop the dependency on $k$ throughout the proof.

The index in the case $\mu+\theta' = \theta$ was obtained in (16), therefore we assume $\mu+\theta' > \theta$ throughout the proof. Recall from (18) that Whittle's index can be written as $d(\mu+\theta') - d'\theta' + W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. Recall from (45) that $W^c(n)$ can be written as

$$W^c(n) = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} \tag{49}$$

with $\xi_i(n)$ for $i \in \{1,2,3\}$ as given in Equation (46).

Let us first compute $\lim_{\lambda\to 0}\pi^{n-1}(0)/\pi^n(0)$, since this result will be used later in the proof. Recall the expression of the steady-state probabilities as defined in (11). Using this together with (40) and (41) we obtain

$$\begin{aligned}\lim_{\lambda\to 0}\frac{\pi^{n-1}(0)}{\pi^n(0)} &= \lim_{\lambda\to 0}\frac{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = 1+\lim_{\lambda\to 0}\frac{\sum_{m=n}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}-\sum_{m=n}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}\\ &= 1+\lim_{\lambda\to 0}\frac{\sum_{m=n}^{\infty}\left(\frac{\lambda^m(\mu+\theta'+\theta(n-1))}{\theta n\prod_{i=1}^m q^{n-1}(i,i-1)}-\frac{\lambda^m\theta n}{\theta n\prod_{i=1}^m q^{n-1}(i,i-1)}\right)}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = 1+\frac{(\mu+\theta'-\theta)}{\theta n}\cdot\lim_{\lambda\to 0}\frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1.\end{aligned} \tag{50}$$

From this last result we observe the following

$$\begin{aligned}\lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)} &= \lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{-\frac{(\mu+\theta'-\theta)}{\theta n}\left(\frac{\frac{\lambda^n}{(\mu+\theta'+\theta(n-1))\theta^{n-1}(n-1)!}+\mathcal{O}(\lambda^{n+1})}{1+\mathcal{O}(\lambda)}\right)}\\ &= \lim_{\lambda\to 0}-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+o(1) = -\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}.\end{aligned} \tag{51}$$

Let us now consider the first term in (49), that is,

$$\begin{aligned}\frac{\sum_{m=0}^{n-1}C(m,0)(\pi^n(m)-\pi^{n-1}(m))}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} &= \frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}{\pi^n(n)+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}\\ &= \frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}{\frac{\pi^n(n)}{\pi^n(0)-\pi^{n-1}(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}} = \frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}.\end{aligned} \tag{52}$$

where in the first inequality we used the conditions (40) and (41). In order to obtain the limit of (52) as $\lambda\to 0$ we substitute the result obtained in (51), and we obtain the following

$$\begin{aligned}\lim_{\lambda\to 0}\frac{\xi_1(n)}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} &= \lim_{\lambda\to 0}\frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}\\ &= \lim_{\lambda\to 0}\frac{C(0,0)+\mathcal{O}(\lambda)}{-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+1+\mathcal{O}(\lambda)} = -C(0,0)\frac{(\mu+\theta'-\theta)}{\theta n}.\end{aligned} \tag{53}$$

Let us now consider the second term in (49), that is,

$$\begin{aligned}\frac{C(n,0)\pi^n(n)-C(n,1)\pi^{n-1}(n)}{\pi^n(n)+\sum_{m=0}^{n-1}\pi^n(n)-\sum_{m=0}^{n-1}\pi^n(n-1)} &= \frac{C(n,0)-C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1+\frac{1}{\pi^n(n)}(\pi^n(0)-\pi^{n-1}(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}\\ &= \frac{C(n,0)-C(n,1)\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}}{1+\frac{\theta^n n!}{\lambda^n}(1-\pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}.\end{aligned} \tag{54}$$

Substituting the results obtained in (50) and (51) in the expression of Equation (54) we obtain

$$\lim_{\lambda\to 0}\frac{\xi_2(n)}{\sum_{m=0}^{n}\pi^n(n)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}=\lim_{\lambda\to 0}\frac{C(n,0)-C(n,1)\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\right)(1+\mathcal{O}(\lambda^n))}{1-\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}(1+\mathcal{O}(\lambda))}$$
$$=\lim_{\lambda\to 0}\frac{C(n,0)(\mu+\theta'+\theta(n-1))-C(n,1)\theta n+\mathcal{O}(\lambda^n)}{\theta n(1+\mathcal{O}(\lambda))}$$
$$=C(n,0)-C(n,1)+C(n,0)\frac{(\mu+\theta'-\theta)}{\theta n}+\mathcal{O}(\lambda). \tag{55}$$

To conclude the proof we need to analyze the third term in (49), that is,

$$\frac{\sum_{m=n+1}^{\infty}C(m,1)\pi^n(m)-\sum_{m=n+1}^{\infty}C(m,1)\pi^{n-1}(m)}{\pi^n(n)+\sum_{m=0}^{n-1}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}$$
$$=\frac{\lambda^n\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=1}^{n-1}q^n(i,i-1)\prod_{i=n+1}^{m}q^n(i,i-1)}\left(\frac{\pi^n(0)}{q^n(n,n-1)}-\frac{\pi^{n-1}(0)}{q^{n-1}(n,n-1)}\right)}{\lambda^n\left(\frac{\pi^n(0)}{\theta^n n!}+\frac{1}{\lambda^n}\left(\pi^n(0)-\pi^{n-1}(0)\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}$$
$$=\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}\right)}{\left(1+\frac{\theta^n n!}{\lambda^n}\left(1-\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}.$$

In the last expression we substitute the results obtained in (50) and (51), and we show that

$$\lim_{\lambda\to 0}\frac{\xi_3(n)}{\sum_{m=0}^{n}\pi^n(n)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}=\lim_{\lambda\to 0}\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}\right)}{\left(1+\frac{\theta^n n!}{\lambda^n}\left(1-\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}$$
$$=\lim_{\lambda\to 0}\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n}{\mu+\theta'+\theta(n-1)}(1+\mathcal{O}(\lambda^n))\right)}{\left(1-\left(\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}+\mathcal{O}(\lambda)\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}=\lim_{\lambda\to 0}\frac{\mathcal{O}(\lambda)}{\frac{\theta n}{\mu+\theta'+\theta(n-1)}+\mathcal{O}(\lambda)}=0. \tag{56}$$

We now substitute the results obtained in Equations (53), (55) and (56) in $\lim_{\lambda\to 0}W^c(n)$, and we obtain

$$\lim_{\lambda\to 0}W^c(n)=C(n,0)-C(n,1)+\frac{(\mu+\theta'-\theta)}{\theta n}(C(n,0)-C(0,0)).$$

# F Proof of Proposition 6

For ease of notation we drop the dependency on $k$ throughout the proof.

The index in the case $\mu+\theta'=\theta$ was obtained in (16), therefore we assume $\mu+\theta'>\theta$ throughout the proof. Recall from (18) that Whittle's index can be written as $d(\mu+\theta')-d'\theta'+W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. Recall from (45) that $W^c(n)$ can be written as

$$W^c(n)=\frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} \tag{57}$$

with $\xi_i(n)$ for $i\in 1,2,3$ as given by Equation (46)

We first compute $\pi^{n-1}(0)/\pi^n(0)$, which will be used later in the proof;

$$\frac{\pi^{n-1}(0)}{\pi^n(0)} = \frac{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = \left(1 + \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)} - \sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}\right)$$

$$= 1 + \frac{(\mu + \theta' - \theta)}{\theta n} \cdot \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} \tag{58}$$

$$= 1 + \frac{\mu + \theta' - \theta}{\theta n}(1 + o(1)). \tag{59}$$

We now proceed to compute (57) as $\lambda \to \infty$. Let us begin by computing the term that corresponds to $\xi_1(n)$. We have after some algebra

$$\begin{aligned}\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} &= \frac{\sum_{m=0}^{n-1} C(m,0)(\pi^n(m) - \pi^{n-1}(m))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} \\ &= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m}{\theta^m m!}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)} + \sum_{m=0}^{n-1} \frac{\lambda^m}{\theta^m m!}},\end{aligned} \tag{60}$$

which after substitution of (59) reduces to

$$\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \mathcal{O}\left(\frac{1}{\lambda}\right), \tag{61}$$

as $\lambda \uparrow \infty$, for all $n$. We are now interested in computing the second term in (57) as $\lambda \to \infty$. Using (59) we obtain

$$\begin{aligned}\lim_{\lambda\to\infty} \frac{\xi_2(n)}{\sum_{m=0}^{n} \pi^n(n) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} &= \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1 + \frac{\pi^n(0)-\pi^{n-1}(0)}{\pi^n(n)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} \\ = \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\theta n}{\mu+\theta'+\theta(n-1)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}{1 + \frac{1-\pi^{n-1}(0)/\pi^n(0)}{\lambda^n/(\theta^n n!)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} &= C(n,0) - C(n,1),\end{aligned} \tag{62}$$

for all $n$. We are left with the third term in (57), that is,

$$\begin{aligned}&\frac{\xi_3(n)}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} \\ &= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1)}\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta'+\theta(n-1)}\right)}{\pi^n(n) + (\pi^n(0) - \pi^{n-1}(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} \\ &= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^n(i,i-1)}\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)}\right) + \frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}\right)}{\lambda^n/(\theta^n n!) + (1 - \pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}},\end{aligned} \tag{63}$$

where in the second step we used that $\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1) = \prod_{i=1}^{m} q^n(i,i-1)/\theta n$. After substituting (58) in the latter equation and some algebra, we obtain that (63) can be written as

$$\frac{\mu + \theta' - \theta}{\theta n} \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta n}\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}(1 + o(1))}.$$

Hence the third term as $\lambda \to \infty$ simplifies to

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta}\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}+o(1)=\frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}+o(1). \tag{64}$$

The latter equality follows due to $\pi^{n-1}(0) = (\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)})^{-1}$. We now write (64) as follows

$$\begin{aligned}
&\frac{\mu+\theta'-\theta}{\theta}\left(\frac{\sum_{m=0}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}-\frac{\sum_{m=0}^{n} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}\right)\\
&=\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}\left(1-\frac{\sum_{m=0}^{n} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\mathcal{O}(\lambda^n)+\sum_{m=n+1}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}\right),
\end{aligned} \tag{65}$$

where

$$\mathbb{E}(C(N^{n-1},1))=\frac{\sum_{m=0}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}.$$

We then have that if there exists $z \geq 1$ such that $\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda^z} \to 0$, as $\lambda \to \infty$, then (65) reduces to

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}+o(1),$$

Hence, together with Equations (57), (61) and (62) we obtain that

$$W^c(n)=C(n,0)-C(n,1)+\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}+o(1),$$

as $\lambda \to \infty$. This concludes the proof.

# G Proof of Proposition 7

For ease of notation, we omit the class index $k$ in the proof.

Since $\theta' = \theta$ we have $\mu + \theta' > \theta$. Since $d' = d = 0$, $\theta' = \theta$ and $C(n,a) = C(n)$, we can write $\tilde{C}(n,a) = C(n)$. Hence, we are interested in the following limit

$$\begin{aligned}
\lim_{\theta\to 0}\theta W(n) &= \lim_{\theta\to 0}\frac{\theta\sum_{m=0}^{\infty} C(m)\left(\pi^n(m)-\pi^{n-1}(m)\right)}{\sum_{m=1}^{n-1}\left(\pi^n(m)-\pi^{n-1}(m)\right)+\pi^n(n)}\\
&=\varepsilon_1(n)\varepsilon_2(n),
\end{aligned}$$

with

$$\varepsilon_1(n)=\lim_{\theta\to 0}\frac{\theta}{\sum_{m=1}^{n-1}\left(\pi^n(m)-\pi^{n-1}(m)\right)+\pi^n(n)},$$

and

$$\varepsilon_2(n)=\lim_{\theta\to 0}\sum_{m=0}^{\infty} C(m)\left(\pi^n(m)-\pi^{n-1}(m)\right).$$

Consider $\varepsilon_2(n)$. We have

$$\pi^n(0) - \pi^{n-1}(0) \xrightarrow{\theta \to 0} 0.$$

hence

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta \to 0} 0, \qquad \forall m < n-1,$$
$$\pi^n(n-1) - \pi^{n-1}(n-1) \xrightarrow{\theta \to 0} (\rho - 1),$$

and

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta \to 0} \rho^{m-n}(1-\rho)^2, \qquad \forall m \geq n.$$

This gives,

$$\begin{aligned}\varepsilon_2(n) &= -C(n-1)(1-\rho) + \frac{(1-\rho)}{\rho}\sum_{m=n}^{\infty} C(m)(1-\rho)\rho^{m-n+1} \\ &= \frac{(1-\rho)}{\rho}(-C(n-1) + \sum_{m=0}^{\infty} C(m+n-1)(1-\rho)\rho^m).\end{aligned}$$

After some algebra and using that $\pi^n(n) \xrightarrow{\theta \to 0} (1-\rho)^{-1}$ (as pointed out in Section 7), we obtain $\varepsilon_1(n) = 1/\mu$. This concludes the proof.

# H Proof of Theorem 2

We first assume there exists a $k$ such that $C_k(0,1) > 0$. Let us consider that $W = 0$, and from (22) we know that necessarily $\mathcal{C}^{REL}(0) \leq \mathcal{C}^{OPT}$. We also consider the policy $\bar{u} \in \mathcal{U}$ that takes active action when the total number of customers in the system is 0, and is passive otherwise. Note that policy $\bar{u}$ does not take any *scheduling* decision. Since $\mu_k + \theta'_k \geq \theta_k$, for all $k$, the queue length under policy $\bar{u}$ stochastically upper bounds any policy $u \in \mathcal{U}$. Note that under the assumption $C_k(0,0) \geq C_k(0,1)$, $\forall k$, it holds from (2) that, for all $n$, $C_k(n,0) \geq C_k(n,1)$, which implies that $W_k(n)$ is always positive, see Section 5. Hence, it follows $\mathcal{C}^{WI} \leq \mathcal{C}^{\bar{u}}$. We will now show that $\frac{\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL}(0)}{\mathcal{C}^{OPT}} \to 0$ as $\lambda \to 0$, which in view of (22) implies the optimality of Whittle's index policy.

We have $W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$, for all $k$. Setting $W = 0$, it follows that for every class $REL(0)$ is the threshold policy with threshold $-1$, that is, class-$k$ always activates for any state $n_k > -1$. Hence, under policy $REL(0)$ the steady-state probabilities for class-$k$ are given by (11) with threshold $n = -1$. It then follows that

$$\begin{aligned}\mathcal{C}^{REL(0)}(0) &= \sum_{k=1}^{K}\sum_{m=0}^{\infty} C_k(m,1)\pi_k^{-1}(m) \\ &= \sum_{k=1}^{K} C_k(0,1)\pi_k^{-1}(0) + \sum_{k=1}^{K} C_k(1,1)\frac{\lambda\gamma_k}{\mu_k + \theta'_k}\pi_k^{-1}(0) + \mathcal{O}(\lambda^2), \qquad (66)\end{aligned}$$

as $\lambda \downarrow 0$. We have $\pi_k^{-1}(0) = (1 + \mathcal{O}(\lambda))^{-1}$, hence $\mathcal{C}^{REL(0)}(0) = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$.

Under policy $\bar{u} \in \mathcal{U}$, every class $k$ behaves as an $M/M/\infty$ queue with arrival rate $\lambda\gamma_k$ and departure rate $\theta_k n_k$. We then have $\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(0,1)\mathrm{e}^{-\lambda\gamma_k/\theta_k} + \sum_{k=1}^{K}\sum_{m=1}^{\infty} C_k(m,0)\frac{(\lambda\gamma_k)^m}{\theta_k^m m!}\mathrm{e}^{-\lambda\gamma_k/\theta_k} = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$.

Hence,

$$\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) \;\; = \;\; \mathcal{O}(\lambda). \qquad (67)$$

We now note that in the limit $\lambda \to 0$, $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(1)$. Together with (22) and (67), we thus conclude that

$$\lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} \leq \lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0)}{\mathcal{C}^{OPT}} = 0.$$

In case $C_k(0,1) = 0$, then $C_k(0,0) \geq C_k(0,1)$ for all $k$ and $W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$, for all $k$. Setting $W = 0$, it follows that $REL(0)$ is the policy that activates class-$k$ for any $n_k \geq 0$. We consider $\bar{u}$ to be the policy that takes active action when the total number of customers in the system is 0 or 1, and is passive otherwise. Then

$$\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(1,1) \frac{\lambda \gamma_k}{\mu_k + \theta'_k} \pi_k^{\bar{u}}(0) + \sum_{k=1}^{K} \sum_{m=2}^{\infty} C_k(m,0) \frac{(\lambda \gamma_k)^m}{(\mu_k + \theta'_k)\theta^{m-1} m!} \pi_k^{\bar{u}}(0),$$

and $\pi_k^{\bar{u}}(0) = \left(1 + \frac{\lambda \gamma_k}{\mu_k + \theta'_k} + \frac{(\lambda \gamma_k)^2}{(\mu_k + \theta'_k) 2\theta_k} + \mathcal{O}(\lambda^3)\right)^{-1}$ as $\lambda \to 0$. We have that $\pi_k^{-1}(0) = \pi_k^{\bar{u}}(0) + \mathcal{O}(\lambda^2)$ as $\lambda \to 0$. Then the term that corresponds to $C(1,1)$ in $\mathcal{C}^{\bar{u}}$ and $\mathcal{C}^{REL(0)}(0)$ as given in (66) coincide up to a $\mathcal{O}(\lambda^2)$ term. Hence, $\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda^2)$ and $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda)$, which lead to the desired result $\lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0$.

# I Proof of Theorem 3

Let us assume for the sake of clarity that there are only 2 classes of customers. It extends trivially to the general case of $k$ classes. We further assume w.l.o.g. $\bar{k} = 2$, hence

$$\lim_{\lambda \to \infty} \frac{W_2(\lambda \gamma_2 / \theta_2)}{W_1(\lambda \gamma_1 / \theta_1)} > 1.$$

We prove Theorem 3 as follows:

- **Step 1:** We assume that Whittle's index is either constant (linear holding cost case), or strictly increasing, the general case follows similarly. We prove that there exists $\bar{W}(\lambda)$ such that $\lim_{\lambda \to \infty} \mathcal{C}^{REL(\bar{W}(\lambda))}(\bar{W}(\lambda)) - \mathcal{C}^{REL(\bar{W}(\lambda))} = 0$, *i.e.*, the optimal solution of the relaxed problem is feasible for the original problem.

- **Step 2:** From Step 1 we deduce that in the limit, $REL(\bar{W}(\lambda))$ will only serve class $\bar{k}$ with probability 1, it becomes feasible for the original problem, i.e., $REL(\bar{W}(\lambda)) \in \mathcal{U}$ as $\lambda \to \infty$, and hence equivalent to Whittle's index policy. Therefore $\lim_{\lambda \to \infty} \mathcal{C}^{REL(\bar{W}(\lambda))} - \mathcal{C}^{WI} = 0$.

- **Step 3:** Applying the result in Equation (22) we obtain $\lim_{\lambda \to \infty} \mathcal{C}^{WI} - \mathcal{C}^{OPT} = 0$.

Let us first assume that $W_k(n_k)$ is constant for $k = 1, 2$, that is the case for linear holding cost. Later on we solve the case in which $W_k(n_k)$ in strictly increasing, following the steps above. If $W_k(n_k)$ is constant for $k = 1, 2$, from the assumption in the statement, we can find $\bar{W}$ constant such that $W_1(n_1) \leq \bar{W} \leq W_2(n_2)$. It then follows trivially that the relaxed policy becomes feasible for the original problem taking $W = \bar{W}$ and equivalent to the Whittle index policy. In view of (22) this in particular implies $\lim_{\lambda \to \infty} \mathcal{C}^{WI} - \mathcal{C}^{OPT} = 0$.

We now assume $W_k(n_k)$ is strictly increasing for $k = 1, 2$. We will denote by $b_k = \gamma_k / \theta_k$ for $k = 1, 2$.

**Step 1.** From the assumption $\lim_{\lambda\to\infty} W_2(\lambda b_2)/W_1(\lambda b_1) > 1$, where $W_k(\lambda b_k)$ is a continuous non-decreasing function in $\lambda b_k$. It then follows that there exists $\bar{W}(\lambda)$ a continuous non-decreasing function in $\lambda$, such that

$$\lim_{\lambda\uparrow\infty} \frac{W_1(\lambda b_1)}{\bar{W}(\lambda)} < 1, \lim_{\lambda\uparrow\infty} \frac{W_2(\lambda b_2)}{\bar{W}(\lambda)} > 1.$$

We have assumed $W_k(\lambda b_k)$ to be increasing, and hence it is invertible. We then obtain

$$\lim_{\lambda\uparrow\infty} \frac{\lambda b_1}{(W_1)^{-1}(\bar{W}(\lambda))} < 1 \tag{68}$$

$$\lim_{\lambda\uparrow\infty} \frac{\lambda b_2}{(W_2)^{-1}(\bar{W}(\lambda))} < 1. \tag{69}$$

The optimal policy of the relaxed problem is to serve all customers whose index is greater than $\bar{W}(\lambda)$. Together with (68) and (69) we will now prove that the optimal policy for the relaxed problem becomes feasible for the original problem taking $W = \bar{W}(\lambda)$ as $\lambda \to \infty$. Hence,

$$\lim_{\lambda\uparrow\infty} \mathcal{C}^{REL(\bar{W}(\lambda))}(\bar{W}(\lambda)) - \mathcal{C}^{REL(\bar{W}(\lambda))} = 0.$$

From (7) we have

$$\begin{aligned}
&\mathcal{C}^{REL(\bar{W})}(\bar{W}(\lambda)) \\
&= \sum_{k=1}^{2} \mathbb{E}(\tilde{C}(N_k^{\bar{W}(\lambda)}, S^{\bar{W}(\lambda)})(N_k^{\bar{W}(\lambda)})) - \bar{W}(\lambda)\left(1 - 2 + \sum_{k=1}^{2}\left(1 - \mathbb{E}\left(\mathbb{1}_{S^{\bar{W}(\lambda)}(N_k^{\bar{W}(\lambda)})=1}\right)\right)\right).
\end{aligned} \tag{70}$$

Due to the independence of the classes of customers in the relaxed problem (note that in the relaxed problem serving one of the classes does not mean we can not serve the other) we can write

$$\begin{aligned}
&\lim_{\lambda\uparrow\infty} \sum_{k=1}^{2}\left(1 - \mathbb{E}\left(\mathbb{1}_{S^{\bar{W}(\lambda)}(N_k^{\bar{W}(\lambda)})=1}\right)\right) = 2 - \left(\lim_{\lambda\uparrow\infty} \mathbb{P}(W_1(N_1) > \bar{W}(\lambda)) + \mathbb{P}(W_2(N_2) > \bar{W}(\lambda))\right) \\
&= 2 - \left(\lim_{\lambda\uparrow\infty} \mathbb{P}(N_1 > (W_1)^{-1}(\bar{W}(\lambda))) + \mathbb{P}(N_2 > (W_2)^{-1}(\bar{W}(\lambda)))\right).
\end{aligned} \tag{71}$$

Let us then compute $\lim_{\lambda\uparrow\infty} \mathbb{P}(N_k > (W_k)^{-1}(\bar{W}(\lambda))) = \lim_{\lambda\uparrow\infty} \mathbb{P}(N_k > \lfloor (W_k)^{-1}(\bar{W}(\lambda))\rfloor)$. To do so we first note that for a given $f(\lambda)$

$$\begin{aligned}
\mathbb{P}(N_k > f(\lambda)) = &\sum_{m=f(\lambda)}^{\infty} \frac{(\lambda\gamma_k)^m}{\theta^{f(\lambda)} f(\lambda)! \prod_{j=f(\lambda)+1}^{m}(\mu_k + \theta_k' + \theta_k(j-1))} \\
&\cdot \frac{1}{\sum_{r=0}^{f(\lambda)} \frac{(\lambda\gamma_k)^r}{\theta_k^r r!} + \sum_{r=f(\lambda)+1}^{\infty} \frac{(\lambda\gamma_k)^r}{\theta_k^{f(\lambda)} f(\lambda)! \prod_{j=f(\lambda)+1}^{r}(\mu_k+\theta_k'+\theta_k(j-1))}}.
\end{aligned}$$

Assume $f(\lambda)$ is a positive non-decreasing function in $\lambda$. Then, in the limit as $\lambda \to \infty$ we have

$$\lim_{\lambda\uparrow\infty} \mathbb{P}(N_k > f(\lambda)) = \lim_{\lambda\uparrow\infty} \sum_{m=f(\lambda)}^{\infty} \frac{(\lambda\gamma_k)^m}{\theta_k^m m! + \mathcal{O}(\theta_k^{m-1}(m-1)!)} \frac{1}{\sum_{j=0}^{f(\lambda)} \frac{(\lambda\gamma_k)^j}{\theta_k^j j!} + \sum_{j=f(\lambda)+1}^{\infty} \frac{(\lambda\gamma_k)^j}{\theta_k^j j! + \mathcal{O}(\theta_k^{j-1}(j-1)!)}}$$

$$= \lim_{\lambda\uparrow\infty} \sum_{m=f(\lambda)}^{\infty} \frac{(\lambda b_k)^m}{m!} \frac{1}{\sum_{j=0}^{\infty} \frac{(\lambda b_k)^j}{j!}} = \lim_{\lambda\uparrow\infty} \frac{\sum_{m=0}^{\infty} \frac{(\lambda b_k)^m}{m!} - \sum_{m=0}^{f(\lambda)-1} \frac{(\lambda b_k)^m}{m!}}{\sum_{j=0}^{\infty} (\lambda b_k)^j \frac{1}{j!}}$$

$$= 1 - \lim_{\lambda\uparrow\infty} \mathrm{e}^{-\lambda b_k} \sum_{m=0}^{f(\lambda)-1} \frac{(\lambda b_k)^m}{m!} = 1 - \lim_{\lambda\uparrow\infty} \mathrm{e}^{-\lambda b_k} \sum_{m=0}^{f(\lambda)} \frac{(\lambda b_k)^m}{m!} = P(f(\lambda), \lambda b_k)$$

$$= \begin{cases} 0, & \text{if } \lim_{\lambda\uparrow\infty} \frac{\lambda b_k}{f(\lambda)} < 1, \\ 1, & \text{if} \lim_{\lambda\uparrow\infty} \frac{\lambda b_k}{f(\lambda)} > 1. \end{cases} \tag{72}$$

The last equality follows from the auxiliary Lemma 1 (see Appendix J). We now take $f^k(\lambda) = \lfloor (W_k)^{-1}(\bar{W}(\lambda)) \rfloor$, for $k = 1, 2$. Then from (72) together with (68) and (69) we obtain that $\mathbb{P}(N_1 > f^1(\lambda)) = 0$, and $\mathbb{P}(N_2 > f^2(\lambda)) = 1$. Hence, (71)=1, which implies

$$1 - 2 + \sum_{k=1}^{2} \left(1 - \mathbb{E}\left(\mathbb{1}_{S^{\bar{W}(\lambda)}(N_k^{\bar{W}(\lambda)})=1}\right)\right) = 0. \tag{73}$$

From (70) we then obtain

$$\lim_{\lambda\to\infty} \mathcal{C}^{REL(\bar{W}(\lambda))}(\bar{W}(\lambda)) = \lim_{\lambda\to\infty} \sum_{k=1}^{2} \mathbb{E}(\tilde{C}(N_k^{\bar{W}(\lambda)}, S^{\bar{W}(\lambda)}(N_k^{\bar{W}(\lambda)}))) = \lim_{\lambda\to\infty} \mathcal{C}^{REL(\bar{W}(\lambda))}.$$

**Step 2.** Since $REL(\bar{W}(\lambda))$ will only serve class 2 with probability 1, it becomes feasible for the original problem, i.e., $REL(\bar{W}(\lambda)) \in \mathcal{U}$ as $\lambda \to \infty$, and hence equivalent to Whittle's index policy, that is,

$$\lim_{\lambda\to\infty} \mathcal{C}^{REL(\bar{W}(\lambda))} = \lim_{\lambda\to\infty} \mathcal{C}^{WI}.$$

**Step 3.** In view of Equation (22) and the result in Step 2 we obtain

$$\lim_{\lambda\to\infty} \mathcal{C}^{WI} - \mathcal{C}^{OPT} = 0.$$

Which concludes the proof.

# J Auxiliary Lemma 1

**Lemma 1** *Let $f(\lambda)$ be a positive continuous non-decreasing function in $\lambda$, and let $b > 0$ be some constant. We further define $P(y, \tilde{\lambda}) := 1 - \mathrm{e}^{-\tilde{\lambda}} \sum_{m=0}^{y} \frac{\tilde{\lambda}^m}{m!}$. Then*

$$\lim_{\lambda\to\infty} P(f(\lambda), \lambda b) = \begin{cases} 0, & \text{if } \lim_{\lambda\to\infty} \frac{\lambda b}{f(\lambda)} < 1, \\ 1, & \text{if } \lim_{\lambda\to\infty} \frac{\lambda b}{f(\lambda)} > 1. \end{cases}$$

**Proof.** Let us first note that $P(f(\lambda),\lambda)=\frac{1}{f(\lambda)!}\int_0^{\lambda b}\mathrm{e}^{-t}t^{f(\lambda)}\mathrm{d}t$, see [1]. That is,

$$\begin{aligned}P(f(\lambda),\lambda)&=\frac{\mathrm{e}^{-f(\lambda)}}{f(\lambda)!}\int_0^{\lambda b}\mathrm{e}^{f(\lambda)-t}t^{f(\lambda)}\mathrm{d}t=\frac{f(\lambda)^{f(\lambda)+1}\mathrm{e}^{-f(\lambda)}}{f(\lambda)!}\int_0^{\lambda b}\left(\mathrm{e}^{1-\frac{t}{f(\lambda)}}\frac{t}{f(\lambda)}\right)^{f(\lambda)}\mathrm{d}\left(\frac{t}{f(\lambda)}\right)\\&=\frac{f(\lambda)^{f(\lambda)+1}\mathrm{e}^{-f(\lambda)}}{f(\lambda)!}\int_0^{\lambda b/f(\lambda)}(\mathrm{e}^{1-u}u)^{f(\lambda)}\mathrm{d}u. \end{aligned}\tag{74}$$

We recall Stirling's formula $f(\lambda)!=\sqrt{2\pi}f(\lambda)^{f(\lambda)+1/2}\mathrm{e}^{-f(\lambda)}(1+\mathcal{O}(\frac{1}{f(\lambda)}))$, from where we obtain

$$\frac{f(\lambda)^{f(\lambda)+1}\mathrm{e}^{-f(\lambda)}}{f(\lambda)!}=\frac{\sqrt{\frac{f(\lambda)}{2\pi}}}{1+\mathcal{O}(\frac{1}{f(\lambda)})}=\sqrt{\frac{f(\lambda)}{2\pi}}\left(1+\mathcal{O}\left(\frac{1}{f(\lambda)}\right)\right).\tag{75}$$

Let us first analyze the case $\lim_{\lambda\to\infty}(\lambda b)/f(\lambda)<1$. Then there exists $\epsilon>0$ such that $0\leq(\lambda b)/f(\lambda)\leq 1-\epsilon$ for large enough $\lambda$. Hence, $0\leq\mathrm{e}^{1-u}u\leq\mathrm{e}^{\epsilon}(1-\epsilon)<1$ for all $0\leq u\leq(\lambda b)/f(\lambda)$, and therefore from Equation (74) and (75) we obtain

$$P(f(\lambda),\lambda)\leq\sqrt{\frac{f(\lambda)}{2\pi}}\left(1+\mathcal{O}\left(\frac{1}{f(\lambda)}\right)\right)\cdot\frac{\lambda b}{f(\lambda)}\left(\mathrm{e}^{\epsilon}(1-\epsilon)\right)^{f(\lambda)}=\mathcal{O}\left(\frac{1}{\sqrt{f(\lambda)}}(\mathrm{e}^{\epsilon}(1-\epsilon))^{f(\lambda)}\right),\tag{76}$$

for $\lambda$ large enough. Since, $\mathrm{e}^{\epsilon}(1-\epsilon)<1$, we have $\lim_{\lambda\to\infty}\mathcal{O}\left(\frac{1}{\sqrt{f(\lambda)}}(\mathrm{e}^{\epsilon}(1-\epsilon))^{\kappa_\lambda}\right)=0$. Hence, from (76) we obtain $P(f(\lambda),\lambda)=0$.

We now analyze the case $\lim_{\lambda\to\infty}(\lambda b)/f(\lambda)>1$. Then there exists $\epsilon>0$ such that $0\leq(\lambda b)/f(\lambda)\geq 1+\epsilon$ for $\lambda$ large enough. The function $\mathrm{e}^{1-u}u$ can also be written as

$$\mathrm{e}^{1-u}u=\mathrm{e}^{1-u-\log(1-(1-u))}=\mathrm{e}^{-\sum_{i=2}^{\infty}\frac{1}{i}(1-u)^i}.$$

From the latter and the saddle point method [17, p. 174] we have that for $\lambda$ large enough

$$\int_0^{\lambda b/f(\lambda)}(\mathrm{e}^{1-u}u)^{f(\lambda)}\mathrm{d}u=\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{1}{2}f(\lambda)(1-u)^2}\mathrm{d}u+\mathcal{O}\left(\frac{1}{f(\lambda)}\right)=\sqrt{\frac{2\pi}{f(\lambda)}}+\mathcal{O}\left(\frac{1}{f(\lambda)}\right).$$

From Equation (74) together with Equation (75), we then obtain

$$P(f(\lambda),\lambda)=\sqrt{\frac{f(\lambda)}{2\pi}}\left(1+\mathcal{O}\left(\frac{1}{f(\lambda)}\right)\right)\left(\sqrt{\frac{2\pi}{f(\lambda)}}+\mathcal{O}\left(\frac{1}{f(\lambda)}\right)\right)=1+\mathcal{O}\left(\frac{1}{\sqrt{f(\lambda)}}\right),$$

for $\lambda$ large enough. From the latter we obtain, $\lim_{\lambda\to\infty}P(f(\lambda),\lambda)=1$, if $\lim_{\lambda\to\infty}\lambda b/f(\lambda)>1$. ■

# K Proof of Theorem 4

Throughout the proof we drop the dependency on $k$.

We first prove that $w^{(1)},w^{(2)}$ and $w^{(3)}$ are non-decreasing and continuous functions. For that recall that the function $C(m,a)$ is convex, which implies

$$\begin{aligned}&tC(m,a)+(1-t)C(m',a)\geq C(tm+(1-t)m'),\forall t\in[0,1],\\&\Longrightarrow C(m,a)-C(m',a)\geq\frac{C(tm+(1-t)m')-C(m',a)}{t}\\&\Longrightarrow C(m,a)-C(m',a)\geq\lim_{t\to 0}\frac{C(tm+(1-t)m')-C(m',a)}{t}=(m-m')\frac{\mathrm{d}C(m',a)}{\mathrm{d}m'}.\end{aligned}$$

From the latter we deduce

$$\frac{C(m,a)-C(m',a)}{m-m'} \geq \frac{\mathrm{d}C(m',a)}{\mathrm{d}m'} \geq \frac{C(m',a)-C(m'',a)}{m'-m''},$$

for all $m'' \leq m' \leq m$. Then $(C(m,a)-C(m,a))(m'-m'') \geq (C(m',a)-C(m',a))(m-m')$. Adding and subtracting $C(m'',a)(m'-m'')$ in the LHS of the inequality and after some algebra, we obtain

$$\frac{C(m,a)-C(m'',a)}{m-m''} \geq \frac{C(m',a)-C(m'',a)}{m'-m''}.$$

Hence, $\frac{C(m,a)-C(m'',a)}{m-m''}$ is non-decreasing in $m$. Similarly in $m''$. The latter directly implies that functions $w^{(1)}$ and $w^{(3)}$ under the assumption $\mu+\theta' \geq \theta$ are non-decreasing. To prove that $w^{(2)}$ is also non-decreasing, let us prove that $\mathrm{d}w^{(2)}(m)/\mathrm{d}m > 0$ for all $\max(0, \lambda/(\mu+\theta'-\theta)) \leq m \leq \lambda/\theta$. We write

$$\frac{\mathrm{d}w^{(2)}(m)}{\mathrm{d}m} = 2\left(\frac{\mathrm{d}C(m,0)}{\mathrm{d}m} - \frac{\mathrm{d}C(m,1)}{\mathrm{d}m}\right) + \frac{(\lambda-\theta m)}{\theta}\frac{\mathrm{d}^2C(m,1)}{\mathrm{d}m^2} + \frac{(\theta m+\mu+\theta'-\theta-\lambda)}{\theta}\frac{\mathrm{d}^2C(m,0)}{\mathrm{d}m^2}.$$

The first term is positive because of Equation (23). Convexity of $C(\cdot,\cdot)$ implies that the second and the third terms are positive in the interval $[\max(0, \lambda/(\mu+\theta'-\theta)), \lambda/\theta]$. This implies that the function $w^{(2)}$ is also non-decreasing in $m$. Continuity of $w^{(1)}, w^{(2)}$ and $w^{(3)}$ follows from the fact that

$$\lim_{m\uparrow(\lambda-(\mu+\theta'-\theta))/\theta} \frac{C\left(\frac{\lambda-(\mu+\theta'-\theta)}{\theta}, 1\right) - C(m,1)}{(\lambda-(\mu+\theta'-\theta))/\theta - m} = \frac{\mathrm{d}C(m,1)}{\mathrm{d}m},$$

hence $\lim_{m\uparrow(\lambda-(\mu+\theta'-\theta))/\theta} w^{(1)}(m) = w^{(2)}((\lambda-(\mu+\theta'-\theta))/\theta)$, and

$$\lim_{m\downarrow\lambda/\theta} \frac{C(m,0)-C(\lambda/\theta,0)}{m-\lambda/\theta} = \frac{\mathrm{d}C(m,0)}{\mathrm{d}m},$$

hence $\lim_{m\downarrow\lambda/\theta} w^{(3)}(m) = w^{(2)}(\lambda/\theta)$.

Having proved that $w(\cdot)$ is non-decreasing and continuous, we are left to prove that the optimal control for problem (26) is $s^*(t) = 1$ when $W < w(m(t))$ and $s^*(t) = 0$ when $W \geq w(m(t))$. In order to do so, we start by characterizing the optimal equilibrium point. Recall that an equilibrium point $(\bar{m}, \bar{s})$ of $\frac{\mathrm{d}m(t)}{\mathrm{d}t}$ is such that

$$0 = \lambda - (\mu+\theta'-\theta)\bar{s} - \theta\bar{m},$$

with $\bar{s} \in [0, \min\{1, \frac{\lambda}{\mu+\theta'-\theta}\}]$ and $\bar{m} = (\lambda - \bar{s}(\mu+\theta'-\theta))/\theta$, hence $\bar{m} \in [\max(0, (\lambda-(\mu+\theta'-\theta))/\theta), \lambda/\theta]$. The optimal equilibrium point $(m^*, s^*)$ minimizes $EC(\bar{s}, W)$. We first prove that $EC(\bar{s}, W)$ is a convex function in $\bar{s} \in [0, \min\{1, \frac{\lambda}{\mu+\theta'-\theta}\}]$, by checking that $\frac{\mathrm{d}}{\mathrm{d}\bar{s}}\left(\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}}\right) > 0$. After some algebra, we obtain that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\bar{s}}\left(\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}}\right) \\
&= \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{\mathrm{d}\tilde{C}(\bar{m},0)}{\mathrm{d}\bar{m}} - \frac{\mathrm{d}\tilde{C}(\bar{m},1)}{\mathrm{d}\bar{m}}\right) \\
&\quad + \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{\mathrm{d}^2\tilde{C}(\bar{m},0)}{\mathrm{d}m^2}\frac{(-\lambda+(\mu+\theta'-\theta)+\theta\bar{m})}{\theta} + \frac{\mathrm{d}^2\tilde{C}(\bar{m},1)}{\mathrm{d}m^2}\frac{(\lambda-\theta\bar{m})}{\theta}\right) > 0.
\end{aligned}$$

The inequality follows from $\bar{m} \in [(\lambda - (\mu + \theta' - \theta))/\theta, \lambda/\theta]$, $\mu + \theta' \geq \theta$ and convexity of $C(\cdot, \cdot)$.

Let us assume from now on that $\min\{1, \lambda/(\mu+\theta'-\theta)\} = 1$. The proof when $\min\{1, \lambda/(\mu+\theta'-\theta)\} = \lambda/(\mu+\theta'-\theta)$ follows similarly. Having proved convexity of $EC(\bar{s}, W)$, we can distinguish the following three cases:

(1) **Case 1:** $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \leq 0$ for all $\bar{s} \in [0,1]$, hence the optimal equilibrium point satisfies $s^* = 1$, $m^* = \lambda/(\mu + \theta' - \theta)$.

(2) **Case 2:** $\frac{\mathrm{d}EC(s^*,W)}{\mathrm{d}s^*} = 0$, hence the optimal equilibrium point satisfies $s^* \in [0,1]$, $m^* \in [\frac{\lambda}{\mu+\theta'-\theta}, \frac{\lambda}{\theta}]$.

(3) **Case 3:** $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0$ for all $\bar{s} \in [0,1]$, hence the optimal equilibrium point satisfies $s^* = 0$, $m^* = \lambda/\theta$.

In the case $\min\{1, \lambda/(\mu+\theta'-\theta)\} = \lambda/(\mu+\theta'-\theta)$, only Case 2 and 3 hold.

Now note that

$$\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0 \Leftrightarrow W \geq \tilde{C}(\bar{m},0) + \tilde{C}(\bar{m},1) + (1-\bar{s})\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}\tilde{C}(\bar{m},0)}{\mathrm{d}\bar{m}} + \bar{s}\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}\tilde{C}(\bar{m},1)}{\mathrm{d}\bar{m}}, \tag{77}$$

which after substitution of $\bar{s} = \frac{\lambda - \theta\bar{m}}{\mu+\theta'-\theta}$ and the expression for $\mathrm{d}\bar{m}/\mathrm{d}\bar{s} = -\frac{\mu+\theta'-\theta}{\theta}$, gives that (77) is equivalent to

$$W \geq \tilde{C}(\bar{m},0) - \tilde{C}(\bar{m},1) + \frac{(\lambda - \theta\bar{m})\frac{\mathrm{d}\tilde{C}(\bar{m},1)}{\mathrm{d}\bar{m}} - (\lambda - (\mu+\theta'-\theta) - \theta\bar{m})\frac{\mathrm{d}\tilde{C}(\bar{m},0)}{\mathrm{d}\bar{m}})}{\theta},$$

that is,

$$W \geq \tilde{C}(\bar{m},0) - \tilde{C}(\bar{m},1) + w^{(2)}(\bar{m}).$$

Hence, in Case 3 the $W$ is such that $W \geq \tilde{C}(\bar{m},0) - \tilde{C}(\bar{m},1) + w^{(2)}(\bar{m})$ for all $\bar{m} \in [\frac{\lambda}{\mu+\theta'-\theta}, \frac{\lambda}{\theta}]$, and in particular $W \geq w(\lambda/\theta)$.

Similarly, being in Case 1 implies $W \leq w(\lambda/(\mu+\theta'-\theta))$.

In Case 2, from $\mathrm{d}EC(s^*, W)/\mathrm{d}s^* = 0$ we obtain, $W = \tilde{C}(m^*,0) - \tilde{C}(m^*,1) + w^{(2)}(m^*) = w(m^*)$, for $m^* \in [\frac{\lambda}{\mu+\theta'-\theta}, \frac{\lambda}{\theta}]$, since $EC^*(W) = (1-s^*)(\tilde{C}(m^*,0) - W) + s^*\tilde{C}(m^*,1)$, $s^* = (\lambda - \theta m^*)/(\mu + \theta' - \theta)$ and $\mathrm{d}m^*/\mathrm{d}s^* = -(\mu+\theta'-\theta)/\theta$. The function $w(m)$ being non-decreasing in particular implies that in Case 2, $W$ is such that $w(\lambda/(\mu+\theta'-\theta)) \leq W \leq w(\lambda/\theta)$.

The objective is to find the control $u$ that minimizes the total bias cost, that is, the cost and subsidy obtained over time minus the optimal cost in equilibrium, denoted as

$$J^u(m(0), W) := \int_0^\infty \left(\tilde{C}(m(t), s^u(t)) - W(1 - s^u(t)) - EC^*(W)\right)\mathrm{d}t. \tag{78}$$

We define $J(m(0), W) = \min_u J^u(m(0), W)$. The theory of optimal control shows that a sufficient condition in order for a control to be bias optimal is that it solves the Hamilton-Jacobi-Bellman (HJB) equation, [34]:

$$0 = \min\{\mathcal{J}_0(m,W), \mathcal{J}_1(m,W)\}, \text{ for all } m, \tag{79}$$

where

$$\mathcal{J}_0(m,W) = \tilde{C}(m,0) - W - EC^*(W) + (\lambda - \theta m)\frac{\partial J(m,W)}{\partial m}, \tag{80}$$

$$\mathcal{J}_1(m,W) = \tilde{C}(m,1) - EC^*(W) + (\lambda - (\mu + \theta' - \theta) - \theta m)\frac{\partial J(m,W)}{\partial m}, \tag{81}$$

and the function $J(m,W)$ is continuous and differentiable. The reader is referred to [12] for a derivation of the HJB. For a given $W$, we consider the policy that prescribes to be passive, $s(t)=0$, in all states $m$ for which $W \geq w(m)$, and active, $s(t)=1$, in all states $m$ for which $W < w(m)$. Observe that due to $w(m)$ being non-decreasing, this will be a threshold policy. That is, there exists $n(W) \in \mathbb{Z}_+$ for which $W > w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$. We refer to this policy as threshold policy $n(W)$. We want to prove that the policy $n(W)$ satisfies the HJB. To do so let us define $J^{n(W)}(m,W)$ for a given $W$ as the cost under policy $n(W)$, starting at state $m$ and up to equilibrium, that is,

$$\begin{aligned}J^{n(W)}(m,W) = &\int_0^{t_0} \tilde{C}(m^{n(W)}(t), s_0) - W(1-s_0) - EC^*(W)\mathrm{d}t \\ &+ \int_{t_0}^{\infty} \tilde{C}(m^{n(W)}(t), s_1) - W(1-s_1) - EC^*(W)\mathrm{d}t,\end{aligned} \tag{82}$$

where $s_0 = s(0)$, $s_1 = 1 - s_0$, and $t_0 \geq 0$, the time at which threshold $n(W)$ is reached. Note that $s_0 = 0$ if $m(0) = m \leq n(W)$ and $s_0 = 1$ otherwise. The function $J^{n(W)}(m,W)$ can be written as the sum of two terms, the first term corresponding to the phase from the starting point $m$ up to the time the threshold is reached, $t_0$. In this phase the control equals $s_0$. Once the threshold is reached, a switch in the control happens and therefore the second term corresponds to the phase from the switch time $t_0$ until the equilibrium in reached. In this phase the control equals $s_1$. This is due to threshold policies having at most one switch in the control.

Let us assume $m(0) = m \leq n(W)$, which implies $s_0 = 0$ and $s_1 = 1$, then from (82)

$$\begin{aligned}\frac{\mathrm{d}J^{n(W)}(m,W)}{\mathrm{d}m} =&\frac{\mathrm{d}t_0}{\mathrm{d}m}\Big(\tilde{C}(n(W),0) - W - EC^*(W) - \tilde{C}(n(W),1) + EC^*(W)\Big) \\ &+ \int_0^{t_0} \frac{\mathrm{d}\tilde{C}(m^{n(W)}(t),0)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t + \int_{t_0}^{\infty} \frac{\mathrm{d}\tilde{C}(m^{n(W)}(t),1)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t.\end{aligned} \tag{83}$$

Policy $n(W)$ implies $\mathrm{d}m^{n(W)}(t)/\mathrm{d}t = \lambda - \theta m^{n(W)}(t)$ for all $t \in [0, t_0]$. Then,

$$\begin{aligned}&\frac{\mathrm{d}m^{n(W)}(t)}{\mathrm{d}t} = \lambda - \theta m^{n(W)}(t) \Rightarrow m^{n(W)}(t) = \left(m - \frac{\lambda}{\theta}\right)\mathrm{e}^{-\theta t} + \frac{\lambda}{\theta} \\ &\Rightarrow t_0 = -\frac{1}{\theta}\log\left(\frac{n(W) - \lambda/\theta}{m - \lambda/\theta}\right) \Rightarrow \frac{\mathrm{d}t_0}{\mathrm{d}m} = \frac{1}{\theta m - \lambda},\end{aligned}$$

and $\mathrm{d}t/\mathrm{d}m = 1/f^0(m)$. Substituting the latter in Equation (83), we obtain for all $m \leq n(W)$

$$\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{\tilde{C}(m,0) - W - EC^*(W)}{\theta m - \lambda},$$

and similarly for all $m > n(W)$

$$\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{\tilde{C}(m,1) - EC^*(W)}{\mu + \theta' - \theta + \theta m - \lambda}.$$

For all $m \leq n(W)$, the action under policy $n(W)$ is to keep the bandit passive. In addition, when substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ in (80), we obtain $\mathcal{J}_0(m,W)=0$. In order for the threshold policy $n(W)$ to satisfy the HJB in (79), we therefore need to prove that $\mathcal{J}_1(m,W) \geq 0$. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ in (81) we obtain that this is equivalent to

$$\begin{aligned}
&\mathcal{J}_1(m,W) \geq 0\\
&\Leftrightarrow W \geq \tilde{C}(m,0)-\tilde{C}(m,1)+\frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta-\theta m}(\tilde{C}(m,1)-EC^*(W)),
\end{aligned} \tag{84}$$

for all $m \notin [\frac{\lambda}{\mu+\theta'-\theta},\frac{\lambda}{\theta}]$ with $m \leq n(W)$, and

$$\begin{aligned}
&\mathcal{J}_1(m,W) \geq 0\\
&\Leftrightarrow W \leq \tilde{C}(m,0)-\tilde{C}(m,1)+\frac{-(\mu+\theta'-\theta))}{\lambda-\mu-\theta'+\theta-\theta m}(\tilde{C}(m,1)-EC^*(W)),
\end{aligned} \tag{85}$$

for all $m \in [\frac{\lambda}{\mu+\theta'-\theta},\frac{\lambda}{\theta}]$ with $m \leq n(W)$. If (84) is satisfied for $m \notin [\frac{\lambda}{\mu+\theta'-\theta},\frac{\lambda}{\theta}]$ and (85) for $m \in [\frac{\lambda}{\mu+\theta'-\theta},\frac{\lambda}{\theta}]$ then, the action under policy $n(W)$ is to keep the bandit passive.

Assume now $m > n(W)$. Hence, action under policy $n(W)$ is to keep the bandit active. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W)-\tilde{C}(m,1)}{\lambda-\mu-\theta'+\theta-\theta m}$ in (81) we then obtain $\mathcal{J}_1(m,W)=0$. In order for the threshold policy $n(W)$ to satisfy the HJB in (79), we need therefore to prove that $\mathcal{J}_0(m,W)\geq 0$. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W)-\tilde{C}(m,1)}{\lambda-\mu-\theta'+\theta-\theta m}$ in (80), this is equivalent to

$$\begin{aligned}
&\mathcal{J}_0(m,W) \geq 0\\
&\Leftrightarrow W \leq \tilde{C}(m,0)-\tilde{C}(m,1)+\frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta-\theta m}(\tilde{C}(m,1)-EC^*(W)),
\end{aligned} \tag{86}$$

for all $m > n(W)$. If (86) is satisfied in $m > n(W)$ then action under policy $n(W)$ is to keep the bandit active.

Hence, if conditions (84)$-$(86) are satisfied, then threshold policy $n(W)$ is optimal. It remains to be proved that conditions (84)$-$(86) are satisfied. This will be done in the remainder of the proof for the three different cases.

Let us first assume that $m^* = \lambda/(\mu+\theta'-\theta)$ and $W \leq w(\lambda/(\mu+\theta'-\theta))$, that is, Case 1, then $EC^*(W)=\tilde{C}(\frac{\lambda}{\mu+\theta'-\theta},1)$. Recall that threshold policy $n(W)$ implies that $W \geq w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$. Hence, $W \leq w(\lambda/(\mu+\theta'-\theta))$ and $w(m)$ being non-decreasing imply that $n(W) \leq \lambda/(\mu+\theta'-\theta)$. Conditions (84)$-$(86) reduce then to the following: the HJB is satisfied if and only if $W \geq (\leq)\tilde{C}(m,0)-\tilde{C}(m,1)+w^{(1)}(m)$ for all $m \leq (\geq) n(W)$. This is equivalent to $W \geq (\leq) w(m)$ for all $m \leq (\geq) n(W)$, since $w^{(1)}(m)$ is non-decreasing and $W \leq w(\lambda/(\mu+\theta'-\theta))$. Hence, in Case 1 the threshold policy $n(W)$ satisfies the HJB and is hence optimal.

Similarly, if $m^* = \lambda/\theta$ and $W \geq w(\lambda/\theta)$, that is, Case 3, then $EC^*(W) = \tilde{C}(\lambda/\theta,0)-W$. Since under threshold policy $n(W)$, $W \geq w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$, $w(m)$ being non-decreasing implies $n(W) \geq \lambda/\theta$. Using $EC^*(W)=\tilde{C}(\lambda/\theta,0)-W$, we obtain that conditions (84)$-$(86) simplify to $W \geq (\leq)\tilde{C}(m,0)-\tilde{C}(m,1)+w^{(3)}(m)$, for all $m \leq (\geq) n(W)$. This is equivalent to $W \geq (\leq) w(m)$ for all $m \leq (\geq) n(W)$, due to $w^{(3)}(m)$ being non-decreasing and $W \geq w(\lambda/\theta)$. Hence, in Case 3, threshold policy $n(W)$ satisfies the HJB and is hence optimal.

We are left with Case 2 in which $W$ is such that $\frac{\mathrm{d}E(s^*,W)}{\mathrm{d}s^*}=0$, and $s^* \in [0,1]$, that is, $w(\lambda/(\mu+\theta'-\theta)) \leq W \leq w(\lambda/\theta)$. In addition $W = w(m^*)$, hence $n(W)=m^*$, by definition of $n(W)$. In this setting we have that

$$EC^*(W) = (1-s^*)(\tilde{C}(m^*,0) - W) + s^*\tilde{C}(m^*,1).$$

Substituting the latter in Conditions (84) and (86) the conditions simplify to

$$\begin{aligned} W &\geq (\leq)\tilde{C}(m,0) - \tilde{C}(m,1) \\ &+ \frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta-\theta m}\Big(\tilde{C}(m,1) - (1-s^*)(\tilde{C}(m^*,0)-W) - s^*\tilde{C}(m^*,1)\Big), \end{aligned} \tag{87}$$

for all $m \leq \lambda/(\mu+\theta'-\theta)(\ m \geq \lambda/\theta)$.

Condition (85) and (86) reduce to

$$\begin{aligned} W &\leq \tilde{C}(m,0) - \tilde{C}(m,1) \\ &+ \frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta-\theta m}\Big(\tilde{C}(m,1) - (1-s^*)(\tilde{C}(m^*,0)-W) - s^*\tilde{C}(m^*,1)\Big), \end{aligned} \tag{88}$$

for all $m \in [\frac{\lambda}{\mu+\theta'-\theta}, m^*]$ and

$$\begin{aligned} W &\leq \tilde{C}(m,0) - \tilde{C}(m,1) \\ &+ \frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta-\theta m}\Big(\tilde{C}(m,1) - (1-s^*)(\tilde{C}(m^*,0)-W) - s^*\tilde{C}(m^*,1)\Big), \end{aligned} \tag{89}$$

for all $m \in [m^*, \lambda/\theta]$.

Taking into account that $\lambda-\mu-\theta'+\theta-\theta m \geq 0$, for all $m < \lambda/(\mu+\theta'-\theta)$, and $\lambda-\mu-\theta'+\theta-\theta m \leq 0$, otherwise, and that $\lambda - s^*(\mu+\theta'-\theta) - \theta m \geq 0$, for all $m \leq m^*$, and $\lambda - s^*(\mu+\theta'-\theta) - \theta m \leq 0$, otherwise, Conditions (87)–(89) reduce to the following:

$$\begin{aligned} W \geq (\leq)\Bigg(&\tilde{C}(m,0) - \tilde{C}(m,1) + \frac{-(\mu+\theta'-\theta)}{\lambda-\mu-\theta'+\theta+\theta m} \\ &\cdot \Big(\tilde{C}(m,1) - (1-s^*)\tilde{C}(m^*,0) - s^*\tilde{C}(m^*,1)\Big)\Bigg) \\ &\cdot \frac{\lambda-\mu-\theta'+\theta-\theta m}{\lambda - s^*(\mu+\theta'-\theta) - \theta m}, \end{aligned} \tag{90}$$

for all $m \leq m^*(m \geq m^*)$. After some algebra, the latter gives

$$\begin{aligned} W \geq (\leq)\tilde{C}(m,0) - \tilde{C}(m,1) &+ \frac{(\lambda-\theta m^*)}{\theta}\frac{(\tilde{C}(m,1) - \tilde{C}(m^*,1))}{m-m^*} \\ &+ \frac{(\mu+\theta'+\theta(m^*-1)-\lambda)}{\theta}\frac{(\tilde{C}(m^*,0) - \tilde{C}(m,0))}{m^*-m}, \end{aligned} \tag{91}$$

for all $m \leq (\geq)m^*$. Since $w^{(2)}(\cdot)$ and $w(\cdot)$ are non-decreasing, in order to prove (90) it therefore suffices to prove that the RHS in (91) is a non-decreasing function and that RHS in $(91) \xrightarrow{m\to m^*} C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$. Let us denote the RHS in (91) by $\tilde{W}(m)$. Convexity of $\tilde{C}(\cdot,\cdot)$ and $\tilde{C}(\cdot,\cdot)$ being non-decreasing imply $\tilde{W}(m)$ to be non-decreasing. Now note that

$$\begin{aligned} &\lim_{m\to m^*}\tilde{W}(m) \\ &= C(m^*,0) - C(m^*,1) + d(\mu+\theta') - d'\theta' + \frac{(\mu+\theta'-\theta(m^*-1)-\lambda)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*}}{\theta}\frac{(\lambda-\theta m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}}{\theta} \\ &= C(m^*,0) - C(m^*,1) + d(\mu+\theta') - d'\theta' + w^{(2)}(m^*) = W. \end{aligned} \tag{92}$$

Hence, for all $m \leq (\geq)m^*$, we have $\tilde{W}(m) \leq (\geq)C(m^*,0) - C(m^*,1) + d(\mu+\theta') - d'\theta' + w^{(2)}(m^*)$. In other words,threshold policy $n(W) = m^*$ satisfies the HJB and is hence optimal.

# L Proof of Proposition 9

We drop the dependency on $k$ throughout the proof.

As $n \to \infty$, then the fluid index is given by $w(n) = C(n,0) - C(n,1) + d(\mu+\theta') - d'\theta' + w^{(3)}(n)$. We have assumed that $C(n,a)$, $a = 0, 1$, are upper bounded by a polynomial of degree $P$. Therefore, we can write $C(n,a) = E(n,a) + o(1)$, for large values of $n$, where $E(n,1) = \sum_{i=0}^{P} C^{(P,i)} n^i$, with

$$C^{(P,i)} := \lim_{n\to\infty} \frac{C(n,1) - \sum_{j=i+1}^{P} C^{(P,j)} n^j}{n^i},$$

and $E(n,0) = \sum_{i=0}^{Q} E^{(Q,i)} n^i$, with

$$E^{(Q,i)} := \lim_{n\to\infty} \frac{C(n,0) - \sum_{j=i+1}^{Q} E^{(Q,j)} n^j}{n^i},$$

Then, as $n \to \infty$, $w(n) = w^{\infty}(n) + o(1)$, where $w^{\infty}(n) = d(\mu+\theta') - d'\theta' + w^c(n) + o(1)$, and

$$w^c(n) = E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \frac{(E(n,0) - E(\lambda/\theta,0))}{n - \lambda/\theta}.$$

Note that $(E(n,0) - E(\lambda/\theta,0))/(n-\lambda/\theta)$ for large values of $n$ can equivalently be written as

$$\begin{aligned}
&\frac{\sum_{i=0}^{Q} E^{(Q,i)} n^i - \sum_{i=0}^{Q} E^{(Q,i)} (\lambda/\theta)^i}{n - \lambda/\theta} = \sum_{i=0}^{Q} E^{(Q,i)} \frac{(n^i - (\lambda/\theta)^i)}{n-\lambda/\theta} = \sum_{i=2}^{Q} E^{(Q,i)} \left( \sum_{j=0}^{i} \left(\frac{\lambda}{\theta}\right)^j n^{i-1-j} \right) \\
&= \frac{E(n,0)}{n} + \frac{E^{(Q,1)}\left(\frac{\lambda}{\theta}\right) + E^{(Q,2)}\left(\frac{\lambda}{\theta}\right)^2 + \ldots + E^{(Q,Q)}\left(\frac{\lambda}{\theta}\right)^Q}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \\
&= \frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} + o(1). \qquad (93)
\end{aligned}$$

We then compute $\lim_{n\to\infty} W(n)/w(n)$, which by the result in (19) is equivalent to

$$\begin{aligned}
&\lim_{n\to\infty} \frac{W(n)}{w(n)} = \lim_{n\to\infty} \frac{W^{\infty}(n) + o(1)}{w^{\infty}(n) + o(1)} = \lim_{n\to\infty} \frac{d(\mu+\theta') - d'\theta' + W^c(n) + o(1)}{d(\mu+\theta') - d'\theta' + w^c(n) + o(1)} \\
&= \lim_{n\to\infty} \frac{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \left( \frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + \mathcal{O}(1)}{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \left( \frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + \mathcal{O}(1)} \\
&= 1 + o(1),
\end{aligned}$$

which follows from the fact that both in the denominator and numerator the highest term comes from $E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \frac{E(n,0)}{n}$. This concludes the proof for the expression in (27).

Let us now obtain the expression in (28) with the extra assumptions $P = Q$ and $C^{(P,i)} = E^{(P,i)}$ for all $i \in \{2, \ldots, P\}$. Observe that under this assumption we obtain from (93) that $(E(n,0) - E(\lambda/\theta,0))/(n-\lambda/\theta)$, for large values of $n$, can be written as

$$\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} + o(1),$$

and hence,

$$
\begin{aligned}
w^{\infty}(n) =& d_k(\mu+\theta') - d'\theta' + E(n,0) - E(n,1) \\
&+ \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1}\right) + o(1).
\end{aligned}
$$

Then, by the result in (19) we have $W^{\infty}(n) = w^{\infty}(n) + o(1)$, and hence $W(n) = w(n) + o(1)$ for large values of $n$ which concludes the proof for (28).