MATHEMATICS OF OPERATIONS RESEARCH
Vol. 31, No. 1, February 2006, pp. 50–84
ISSN 0364-765X | EISSN 1526-5471 | 06 | 3101 | 0050

DOI 10.1287/moor.1050.0165

Restless Bandit Marginal Productivity Indices, Diminishing Returns, and Optimal Control of Make-to-Order/Make-to-Stock $M/G/1$ Queues

José Niño-Mora
Department of Statistics, Universidad Carlos III de Madrid, C/Madrid 126, E-28903 Getafe (Madrid), Spain,
jnimora@alum.mit.edu, http://alum.mit.edu/www/jnimora

This paper presents a framework grounded on convex optimization and economics ideas to solve by index policies problems of optimal dynamic allocation of effort to a discrete-state (finite or countable) binary-action (work/rest) semi-Markov restless bandit project, elucidating issues raised by previous work. Its contributions include: (i) the concept of a restless bandit's marginal productivity index (MPI), characterizing optimal policies relative to general cost and work measures; (ii) the characterization of indexable restless bandits as those satisfying diminishing marginal returns to work, consistently with a nested family of threshold policies; (iii) sufficient indexability conditions via partial conservation laws (PCLs); (iv) the characterization of the MPI as an optimal marginal productivity rate relative to feasible active-state sets; (v) application to semi-Markov bandits under several criteria, including a new mixed average-bias criterion; and (vi) PCL-indexability analyses and MPIs for optimal service control of make-to-order/make-to-stock queues with convex holding costs, under discounted and average-bias criteria.

Key words: restless bandits; stochastic scheduling; index policies; indexability; control by price; semi-Markov decision processes; dynamic resource allocation; diminishing returns; marginal productivity; efficient frontier; convex optimization; bias; mixed criteria; make to order; make to stock; control of queues; production-inventory control; partial conservation laws; achievable performance
MSC2000 subject classification: Primary: 90B36, 90C40; secondary: 90B05, 90B22, 90B30, 90C25
OR/MS subject classification: Primary: dynamic programming/optimal control: semi Markov; secondary: queues: optimization; inventory/production: policies
History: Received February 16, 2004; revised October 27, 2004.

1. Introduction. This paper presents a unifying framework based on convex optimization and economics ideas to formulate and solve by index policies problems of optimal dynamic allocation of effort to a generic discrete-state (finite or countable) binary-action (work/rest) semi-Markov *restless bandit project*, elucidating in a broader setting a host of issues raised by Whittle's [47] index policy and later developments. The framework is deployed to solve by index policies problems of optimal service control of make-to-order (MTO) and make-to-stock (MTS) $M/G/1$ queues with convex backorder and stock holding cost rates, under discounted and mixed (average-bias) criteria. In Niño-Mora [35, 36] the results obtained herein are used to obtain new hedging point and index policies for scheduling a multiclass combined MTO-MTS $M/G/1$ queue.

The proposed framework draws on and combines ideas from relatively autonomous areas including: (i) convex optimization in mathematical programming; (ii) the theory of optimal resource allocation in economics; (iii) index policies for scheduling multiclass queues; (iv) index policies for bandit problems; and (v) conservation laws and polyhedral methods in stochastic scheduling. To put our contributions in context, we start by discussing the relevant background.

1.1. Solution approaches to resource allocation problems. The prevailing solution approaches in the domains of static/deterministic and of dynamic/stochastic resource allocation problems are remarkably distinct. In the former, the concern is to find a fixed allocation of resources optimizing a cost/reward objective. Formulation and solution methods are those of *mathematical programming*. The concepts of *convexity* and *duality* play central roles, both as analysis tools and as insightful bridges with economic interpretation. Convexity is the mathematical counterpart of the economics *law of diminishing marginal returns/productivity*, stating that, as usage of a resource increases, its *marginal productivity* diminishes. Duality relates to the *resource valuation/pricing problem*, which is to find a resource's *shadow price*. Two fundamental results holding under diminishing marginal returns are: (i) a resource's shadow price at a given usage level equals its marginal productivity; and (ii) to achieve an optimal allocation, usage of a resource must be increased as long as its price is lower than its marginal productivity, until both coincide. The classic texts of Kantorovich [21] and Koopmans [25] provide insightful accounts of such ideas.

In the latter domain, the concern is to design a *policy* for dynamic resource allocation to competing activities, in a *system* whose *state* evolves randomly over time. The objective is to optimize a measure of expected cost/reward performance. The main modeling paradigm is furnished by *Markov decision processes* (MDPs), especially in the *discrete-state and discrete-action* case, to which we restrict attention. See, e.g., Puterman [39]. The prevailing solution paradigm has been the *dynamic programming* technique, which draws on the *principle of optimality* to formulate a set of *optimality equations*. When solved, the latter yield an *optimal policy*. Major research efforts have been devoted to the equations' analytical solution in relatively simple models, typically by ad hoc approaches, and to their computational solution by general algorithms, such as *value/policy iteration*. A deep connection between the mathematical programming and the dynamic programming approaches was revealed by D'Epénoux [11] and Manne [27], who showed that the optimality equations for a finite MDP can be formulated as a *linear programming* (LP) problem. The current status of the field remains, however, unsatisfactory. Thus, no unifying analytical solution method has emerged. Also, application of general algorithms is hindered by their excessive computational demands (*curse of dimensionality*). Further, even when a solution is available, it is often not clear how to gain from it qualitative insights of the kind provided by convexity and duality.

1.2. Index policies and mathematical programming in dynamic resource allocation. Limited research efforts have explored use of mathematical programming methods in dynamic and stochastic resource allocation problems, focusing mostly in the area of *stochastic scheduling* (cf. Niño-Mora [33]). The latter is concerned with problems of optimal allocation of service to a collection of stochastic *projects*, which can represent a variety of entities, e.g., jobs or queues.

A notorious phenomenon across a wide range of such models is the optimality of *index policies*. In a typical result, to every project k is attached an *index* $\nu_k(i_k)$ which is a function of its state i_k, such that the policy assigning dynamically higher service priority to projects with larger index values is optimal.

While most of such results have been first derived through ad hoc arguments, some later proofs have used LP methods. These are based on formulating LP constraints on *performance measures* (e.g., mean delays). In tractable models, such constraints characterize the *achievable performance region* spanned by performance vectors under all *admissible policies*. This is a polytope, whose vertices are achieved by priority policies. The optimal vertex/policy is characterized by indices, which emerge from an optimal *dual* solution. In intractable models, available constraints give a tractable *relaxation*, whose dual solution may suggest a heuristic index policy. See, e.g., Dacre et al. [10].

Table 1 highlights selected results in such vein, pointing out the evolution of ideas used to obtain LP constraints, which we review next. The ground-breaking work is due to Klimov [24], who used *aggregate flow balance* to formulate as an LP the problem of scheduling a multiclass $M/G/1$ queue with feedback to minimize average linear holding costs. He further introduced an *adaptive-greedy algorithm* to construct an optimal dual solution, given in terms of a *static index* ν_k^{K} attached to classes k. Finally, he used strong duality to establish optimality of his *Klimov index* policy.

Coffman and Mitrani [8] introduced a different LP formulation of the no-feedback case of Klimov's model. Constraints are based on Kleinrock's [22] *work conservation laws*. They characterize the achievable performance region of mean delays as a *polymatroid*, a fundamental polytope in polyhedral combinatorics due to Edmonds [14]. Optimality of the *cμ-rule* (cf. Cox and Smith [9]) thus follows from that of the *greedy algorithm* for LP over polymatroids.

Table 1. LP formulations yielding index policies for stochastic scheduling problems.

LP constraints	Models and papers
Aggregate flow balance	Multiclass (MC) queues (feedback): Klimov [24]
Strong conservation laws Polymatroids	MC queues (no feedback): Coffman and Mitrani [8], Federgruen and Groenevelt [15], Shanthikumar and Yao [41]
Generalized conservation laws Extended polymatroids	Klimov's model: Tsoucas [42] Klimov's model & branching bandits: Bertsimas and Niño-Mora [4]
Approximate conservation laws Extended polymatroids	MC queues (feedback & parallel servers): Glazebrook and Niño-Mora [19]
Flow balance & average activity Lagrangian relaxation	Restless bandits (RBs): Whittle [47], Bertsimas and Niño-Mora [5]
Partial conservation laws (PCLs) $\mathscr{F}$-extended polymatroids	RBs & MC queues (convex costs, finite-state): Niño-Mora [32, 34]
Diminishing returns & PCLs Work-cost efficient frontier	RBs & MC queues (convex costs, countable-state): this paper

This line of research was continued by Federgruen and Groenevelt [15], and by Shanthikumar and Yao [41]. The latter paper unified previous results via the framework of *strong* (*work*) *conservation laws*.

Tsoucas [42] used work conservation laws to give a new LP formulation of Klimov's model, characterizing its achievable performance region as a new polytope (*extended polymatroid*). Bertsimas and Niño-Mora [4] extended such work by introducing and deploying the framework of *generalized conservation laws*, yielding a new LP proof of optimality of index policies for Weiss' [45] *branching bandit problem*. This encompasses Klimov's model and the classic multiarmed bandit problem.

The multiarmed bandit problem concerns the optimal dynamic allocation of effort to a collection of projects, modeled as discounted binary-action (active/passive) discrete-state and discrete-time MDPs which can only change state when active, and one of which must be engaged at each time. In a celebrated result, Gittins and Jones [18] and Gittins [17] introduced an index $\nu_k^{\mathrm{G}}(i_k)$ for each project k, depending on its state i_k, and established optimality of the resulting *Gittins index* policy.

The parallel-server version of Klimov's model was addressed in Glazebrook and Niño-Mora [19] via approximate conservation laws. The latter furnish a tractable LP relaxation yielding suboptimality bounds on Klimov's policy, which imply its asymptotic optimality in heavy traffic.

1.3. Restless bandits, indexability, and queueing control. The *restless bandit problem* is the extension of the classic multiarmed bandit problem where passive projects can change state. It provides a powerful modeling paradigm at the expense of tractability, being *P-space hard* (see Papadimitriou and Tsitsiklis [37]). Whittle [47] introduced an index $\nu_k^{\mathrm{W}}(i_k)$ attached to a *restless bandit project* k, and proposed as a heuristic the resulting index policy. The *Whittle index* emerges from the solution of a *relaxed problem*, as the state-dependent critical value of the Lagrange multiplier for an average-activity constraint. Whittle's policy recovers Gittins' in the classic (nonrestless) case.

Yet the Whittle index does not exist for all restless projects: only for so-called *indexable* projects. Whittle [47] stated that ". . . one would very much like to have simple sufficient conditions for indexability; at the moment, none are known." Such scope limitation prompted Bertsimas and Niño-Mora [5] to introduce a different LP-based index policy, applicable to arbitrary projects, along with a sequence of increasingly stronger LP relaxations.

The indexability issue was taken up in Niño-Mora [32], where we introduced the framework of *partial conservation laws* (PCLs). Satisfaction of PCLs by the performance measures of a stochastic scheduling problem ensures optimality of index policies *with a postulated structure*, under *admissible objectives*. Use of PCLs yielded tractable sufficient conditions for indexability (*PCL indexability*), and a variant of Klimov's algorithm for computing the Whittle index. The PCL framework's polyhedral foundation was laid out in Niño-Mora [34], yielding extensions of the Whittle index with a significantly expanded scope.

Yet the framework in Niño-Mora [32, 34], relying on polyhedral methods, applies only to finite-state projects. Also, it only gives *sufficient conditions* for indexability, instead of a complete characterization. The first limitation renders it inapplicable to restless bandit problems arising from countable-state queueing system models. Thus, e.g., the problem of scheduling a multiclass $M/G/1$ queue to minimize average holding costs is readily formulated as a restless bandit problem, where projects correspond to queues for each class. Yet the Whittle index does not exist in such setting, as noticed by Whittle [48, Chapter 14.7] himself, and by Veatch and Wein [43]. The latter authors state:

> "In contrast, the backorder problem is not indexable. $\nu(x)$ does not exist (i.e., equals $-\infty$) for all x. The difficulty is that ν is a Lagrange multiplier for the constraint on the time-average number of active arms. For the backorder problem, any stable policy must serve a time-average of ρ classes, so relaxing this constraint does not change the optimal value, and the Lagrange multiplier does not exist. In fact, no scheduling problem with a fixed utilization will be indexable."

We first proposed in Niño-Mora [30, 31] to overcome such difficulty by noting and exploiting the fact that, though the *average-criterion* Whittle index does not exist, its *discounted* counterpart does. Taking the limit of the latter scaled by the discount factor as this vanishes gives a convenient heuristic index. Such approach is deployed in Ansell et al. [1] in an MTO queue via an ad hoc dynamic programming analysis, under the assumption that holding cost rates are *convex increasing* in the queue length. The results in that paper are, however, hindered by the following limitations: (i) the convex increasing holding cost rate assumption is violated in an MTS queue with backorders, where holding cost rates are U-shaped in the natural state of *net backorder* levels; (ii) the approach relies on first establishing indexability under the discounted criterion, which can be significantly more demanding (as the results in this paper illustrate); and (iii) the limiting index does not emerge from an autonomous concept of indexability, as it is not shown that it characterizes optimal policies for an appropriate problem.

Dusonchet and Hongler [13] have calculated the discounted Whittle index, though without actually establishing indexability, in an MTS $M/M/1$ queue with *linear* backorder and stock holding cost rates.

1.4. Contributions. This paper is motivated by the overall goal of developing a unifying and applicable theory of restless bandit indices, which elucidates and resolves the issues raised by previous work, as discussed above. We next briefly summarize and discuss its main contributions towards such goal.

We address the issue of indexability in the setting of a general discrete-state continuous-time restless bandit project, so that the analysis can focus on fundamental principles while ignoring model-specific distracting details. This allows us to introduce the broader, unifying concept of a project's *marginal productivity index* (MPI), defined relative to general cost and work measures. The Gittins index, the Whittle index, and the new indices introduced in Niño-Mora [34] and herein are special cases of the MPI.

We elucidate the property of *indexability* (existence of the MPI) in qualitative terms. We present a *complete and intuitive characterization* of the class of indexable projects, as projects obeying the economics *law of diminishing marginal returns to work*, consistently with a nested family of *threshold policies*.

We give sufficient conditions for indexability based on satisfaction of PCLs, which apply to countable-state projects. We had previously given PCL-based sufficient indexability conditions for finite-state projects, grounded on LP. While our approach herein can be readily cast in terms of infinite-dimensional LP, we have chosen to avoid the latter's machinery by relying only on first-principles analyses.

We present a characterization of the MPI for PCL-indexable projects as an *optimal marginal productivity rate relative to feasible active-state sets*, under a condition (wedge-shaped marginal workloads) holding in the motivating model. Such characterization adapts to restless bandits Gittins' [17] characterization of his index for nonrestless bandits as an *optimal average productivity rate relative to active-state sets*.

We discuss applicability of the PCL framework to semi-Markov projects, under discounted and average criteria, and a new *mixed average-bias criterion*. The latter, which is motivated by nonexistence of the average-criterion Whittle index in a service-controlled $M/G/1$ queue, measures cost performance under the average criterion, while measuring work performance under Blackwell's [6] *bias* criterion.

Finally, we carry out a *PCL-indexability analysis* of a service-controlled MTO/MTS $M/G/1$ queue with convex backorder and stock holding cost rates, under discounted and mixed average-bias criteria. This yields new MPIs, given in terms of relatively simple and insightful expressions.

1.5. Structure of the paper. The rest of the paper is organized as follows. Section 2 introduces our motivating model, concerning the optimal control of service in an MTO/MTS $M/G/1$ queue, along with notation to be used throughout the paper. Section 3 addresses a generic discrete-state restless project control problem, introducing and characterizing the concept of MPI policy. Section 4 develops sufficient indexability conditions for discrete-state projects based on satisfaction of PCLs. Section 5 applies the PCL framework to semi-Markov projects under discounted, average, and average-bias criteria. Section 6 carries out a PCL-indexability analysis of a service-controlled pure MTO $M/G/1$ queue, while §7 analyzes the corresponding MTS model with backorders.

2. Motivating model. Consider the single-machine and single-product MTO/MTS production-inventory queueing control model portrayed in Figure 1, where unit-size customer orders arrive as a Poisson process with

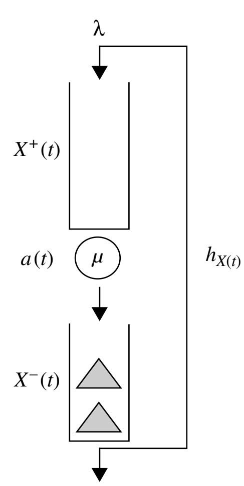

FIGURE 1. The MTO/MTS production-inventory queueing control model.

rate λ. To complete a finished product, the machine expends a *production time* distributed as a random variable with *Laplace-Stieltjes transform* (*LST*) $\psi(\cdot)$, having finite mean $1/\mu$ and variance σ^2. Arrivals and production times are mutually independent. Letting $\rho = \lambda/\mu$ be the *traffic intensity*, we assume the *stability condition* $\rho < 1$.

Production orders can be initiated in advance of demand (MTS mode), in which case completed products are stored in a *finished goods stock* of finite storage capacity $s \geq 0$. An arriving order is immediately filled from stock, if available; otherwise, it is placed in a *backorder queue* (MTO mode) of unlimited size. We denote by $X^-(t)$ (resp. $X^+(t)$) the size of the finished goods stock (resp. the backorder queue) at times $t \geq 0$, and take the system *state* to be the *net backorder queue's size* $X(t) = X^+(t) - X^-(t)$. The *state space* is thus $N = \{-s, \ldots, 0, 1, \ldots\}$. Setting $s = 0$ yields the *pure MTO case*.

A controller governs the system evolution through adoption of a *policy* π. This prescribes, based on past and present information, the *action* $a_n \in \{0, 1\}$ to be taken at each of a sequence of *decision epochs* t_n, with $0 = t_0 < t_1 < t_2 < \cdots$, consisting of all order arrival epochs to an *idle* system and all product completion epochs. Taking the *active action* $a_n = 1$ at epoch t_n corresponds to having the machine work uninterruptedly on a production order until its completion. Such action can only be taken at epochs when the stock is not full, i.e., when the system lies in a *controllable state* $X_n \triangleq X(t_n) \in N^{\{0,1\}} = \{-s+1, \ldots\}$. Each of the random number $O_n \geq 0$ of orders arriving *during* the ensuing *active period* $[t_n, t_{n+1})$ causes an upwards unit jump in the *natural process* $X(t)$, so that the *embedded process* X_n (observed at decision epochs) evolves by $X_{n+1} = X_n + O_n - 1$. Taking the *passive action* $a_n = 0$ at epoch t_n corresponds to letting the machine sit idle, i.e., rest, during the *passive period* $[t_n, t_{n+1})$ until the next order arrival, so that $X_{n+1} = X_n + 1$ in such case. The passive action can be chosen in any state, but it *must* be chosen when the stock is full, i.e., when the embedded process lies in the *uncontrollable state* $X_n = -s$. Further, the adopted policy π must be *stable*, i.e., its induced process $X(t)$ must be ergodic. We denote by Π the class of *admissible policies* satisfying all the required properties.

The system incurs backorder and/or stock holding costs, accruing continuously at rate h_i per unit time while natural process $X(t)$ occupies state i. We denote first- and second-order differences by $\Delta h_i \triangleq h_i - h_{i-1}$ and $\Delta^2 h_i \triangleq \Delta h_i - \Delta h_{i-1}$. In the analyses in §§6 and 7, we will refer to the *shifted cost rate sequences* $\mathbf{h}^l$, for $l \geq 0$, where $\mathbf{h} \triangleq (h_{-s}, h_{-s+1}, \ldots)$ is the original cost sequence and $\mathbf{h}^l \triangleq (h_{-s+l}, h_{-s+l+1}, \ldots)$, i.e., $h_j^l \triangleq h_{j+l}$. We will further refer to sequences of first- and second-order differences $\Delta\mathbf{h}^l \triangleq \mathbf{h}^l - \mathbf{h}^{l-1} = (\Delta h_{-s+l}, \ldots)$, for $l \geq 1$, and $\Delta^2\mathbf{h}^l \triangleq \Delta\mathbf{h}^l - \Delta\mathbf{h}^{l-1} = (\Delta^2 h_{-s+l}, \ldots)$, for $l \geq 2$. We impose the following requirements on cost rates.

Assumption 2.1. *The following conditions hold*:
(i) $\mathbf{h}$ *is bounded below*: $\inf\{h_i\colon i \in N\} > -\infty$.
(ii) $\mathbf{h}$ *is convex*: $\Delta^2 h_i \geq 0$, *for* $i \geq -s+2$.
(iii) *If service times have finite moments of order* $m+1$, *then* $h_i = O(i^m)$ *as* $i \to +\infty$.

To evaluate the value of costs incurred under a policy π, given the initial state $X(0)$'s distribution, we can use the (*total expected*) *discounted cost measure* for a given discount factor $\alpha > 0$,

$$f^{\pi,\alpha} \triangleq \mathbb{E}^{\pi}\left[\int_0^{\infty} h_{X(t)} e^{-\alpha t}\, dt\right], \tag{1}$$

or the (*long-run*) *average cost measure*

$$f^{\pi} \triangleq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{\pi}\left[\int_0^{T} h_{X(t)}\, dt\right]. \tag{2}$$

We will be investigate the *average cost problem*, which is to find a policy $\pi^* \in \Pi$ attaining the minimum average rate f^* of expected costs incurred per unit time:

$$\text{Find } \pi^* \in \Pi\colon\ f^{\pi^*} = f^* \triangleq \inf\{f^{\pi}\colon \pi \in \Pi\}. \tag{3}$$

We will further consider the corresponding *discounted cost problem*.

Intuition suggests that it might suffice to look for optimal policies to problem (3) and its discounted counterpart within the family of *threshold policies*. The latter make the server work when the state lies above a critical state, and let it rest otherwise.

Previous work on problem (3) has focused on the case of linear backorder and stock holding cost rates, where threshold policies are known to be optimal. See Buzacott and Shanthikumar [7, Chapter 4] and references therein. In the pure MTO $M/M/1$ discounted cost case, Bertsekas [3, Chapter 6.7] proves optimality of threshold policies under the stronger requirement of *convex nondecreasing* cost rates.

3. Optimal control of a generic restless project by MPI policies. In this section we generalize the motivating problem above into a generic restless project's control problem. This will allow us to focus on fundamental problem features without being distracted by ancillary model-specific details. Subsequent sections will gradually include more detail on the project's description, as required by the theory. We will thus develop a unifying solution approach based on the broad concept of a project's MPI.

3.1. Project control problem. Consider a generic *semi-Markov* restless project, such as that in §2, whose state $X(t)$ evolves over continuous time $t \geq 0$ across the *discrete* (finite or countable) *state space* $N \subseteq \mathbb{Z}$. Control is exercised by a *central planner*, who observes the state $X_n \triangleq X(t_n)$ at an embedded sequence of *decision epochs* $t_0 = 0 < t_1 < t_2 < \cdots$, and decides at each epoch whether a *single operator* is allocated to work on the project (*active* action: $a_n \triangleq a(t_n) = 1$), or is let to rest (*passive* action: $a_n = 0$), over the ensuing *period* ($a(t) = a_n$ for $t \in [t_n, t_{n+1})$). We partition the state space N into the *controllable state space* $N^{\{0,1\}}$, where both actions can be chosen, and the *uncontrollable state space* $N^{\{0\}}$, where only the passive action is available. We assume that both N and $N^{\{0,1\}}$ are consecutive-integer sets *bounded below*, with $N^{\{0\}}$ consisting of the smallest state, so that for given $-\infty < \ell^0 < \ell^1 \leq +\infty$,

$$N \triangleq \{j \in \mathbb{Z}: \ell^0 \leq j \leq \ell^1\}, \qquad N^{\{0,1\}} \triangleq \{j \in \mathbb{Z}: \ell^0 < j \leq \ell^1\}, \qquad N^{\{0\}} = \{\ell^0\}.$$

We will refer to $X(t)$ and $a(t)$ as the *natural state* and *action processes*, and to X_n and a_n as the corresponding *embedded processes*. Typically, the system state can change between decision epochs, though action choice is only allowed at decision epochs. We will further call a period $[t_n, t_{n+1})$ where $X_n = i$ and $a_n = a$ an (i, a)-*period*, or an i-*period*, as convenient. We defer a description of specific process dynamics to §5, as such details are not needed for this section's discussion.

Action choice results from the planner's adoption of a *policy* π, chosen from a given class Π of *admissible policies*. A *manager* is in charge of policy implementation. We assume Π to be *closed under randomization*. Namely, given policies $\pi, \pi' \in \Pi$, and $q \in [0, 1]$, the policy resulting from a random draw of π or π' with respective probabilities q and $1 - q$, denoted by $q\pi + (1-q)\pi'$, also belongs in Π. We will refer to the class $\Pi^{\mathrm{SD}} \subset \Pi$ (resp. $\Pi^{\mathrm{SR}} \subset \Pi$) of *admissible stationary deterministic* (resp. *randomized*) policies, where the chosen action at a decision epoch is a deterministic (resp. random) function of the state. A policy $\pi \in \Pi^{\mathrm{SD}}$ is thus represented by the *active-state set* $S \subseteq N^{\{0,1\}}$ where it prescribes to work on the project. We will then term it the S-*active policy* and denote it by S, writing, e.g., $S \in \Pi^{\mathrm{SD}}$.

The project accrues *holding costs* continuously over time at rate h_i^a while $X(t)$ occupies state i and action a prevails. We assume such rates to be bounded below:

$$\inf_{i,a} h_i^a > -\infty. \tag{4}$$

The value of costs accrued under a policy $\pi \in \Pi$ is evaluated by a finite *cost measure* f^π. See, e.g., (1) and (2) for two specific examples (where $h_i^a \equiv h_i$). We will address the *project's control problem*, which is to find an admissible dynamic resource allocation policy π^* minimizing the value of costs incurred:

$$\text{Find } \pi^* \in \Pi: \quad f^{\pi^*} = f^* \triangleq \inf\{f^\pi: \pi \in \Pi\}. \tag{5}$$

3.2. Solution by threshold policies. Motivated by applications such as that in §2, we aim to find an optimal policy to problem (5) within the class of *threshold policies*, which prescribe to take the active action at decision epochs where the state lies above a critical, *threshold state*. We will represent the threshold policy with threshold state i by its corresponding active-state set

$$S_i \triangleq \{j \in N^{\{0,1\}}: j > i\}, \quad i \in N.$$

This allows us to represent the family of threshold policies by the *active-state set family*

$$\mathscr{F} \triangleq \{S_i: i \in N\}.$$

We will henceforth refer to such threshold policies as $\mathscr{F}$-*policies*, writing, e.g., f^S for $S \in \mathscr{F}$.

In such setting, we will seek to (i) elucidate conditions for existence of an optimal $\mathscr{F}$-policy to problem (5), and (ii) find an optimal $\mathscr{F}$-policy when such conditions hold.

We next develop a solution approach hinging on existence of an appropriate *work measure* g^π. This is a measure of the time expended by the operator working on the project under policy π. We will consider, in §5, three specific work measures corresponding to discounted, average, and bias criteria.

We assume that $\mathscr{F}$-policies, work, and cost measures satisfy the following regularity conditions.

ASSUMPTION 3.1.
(i) *For* $S \in \mathscr{F}$, $j_1 \in S$, $j_2 \in S^c \triangleq N^{\{0,1\}} \setminus S$, *it holds that* S, $S\setminus\{j_1\}$, $S \cup \{j_2\} \in \Pi^{\mathrm{SD}}$.
(ii) *Work measure* g^π *is bounded above*: $\sup\{g^\pi: \pi \in \Pi\} < +\infty$.
(iii) *Cost measure* f^π *is bounded below*: $\inf\{f^\pi: \pi \in \Pi\} > -\infty$.
(iv) *Work measure* g^{S_i} *is strictly decreasing in threshold state* i:

$$\Delta g^{S_i} \triangleq g^{S_i} - g^{S_{i-1}} < 0, \quad i \in N^{\{0,1\}}. \tag{6}$$

(v) *The achievable work performance region is the interval spanned by threshold policies*:

$$\{g^\pi: \pi \in \Pi\} = \bigcup_{i \in N^{\{0,1\}}} [g^{S_i}, g^{S_{i-1}}].$$

Notice that Assumption 3.1(i) needs only be checked in the countable-state case. We will further refer to first-order differences of cost measure f^{S_i}:

$$\Delta f^{S_i} \triangleq f^{S_i} - f^{S_{i-1}}, \quad i \in N^{\{0,1\}}. \tag{7}$$

3.3. Reformulation as a convex resource allocation problem. The solution approach we deploy below is grounded on deterministic convex optimization and the economics theory of optimal resource allocation. Consider the *achievable work-cost performance region*

$$\mathbb{H} \triangleq \{(b, z) \in \mathbb{R}^2: (b, z) = (g^\pi, f^\pi) \text{ for some } \pi \in \Pi\}.$$

Its projections over the coordinate axes give the *achievable work performance region*

$$\mathbb{B} \triangleq \{b \in \mathbb{R}: b = g^\pi \text{ for some } \pi \in \Pi\},$$

and the *achievable cost performance region*

$$\mathbb{V} \triangleq \{z \in \mathbb{R}: z = f^\pi \text{ for some } \pi \in \Pi\}.$$

Convexity of such regions, and hence of their closures $\overline{\mathbb{H}}$, $\overline{\mathbb{B}}$, and $\overline{\mathbb{V}}$, follows from the requirement that Π be closed under randomization.

The *efficient work-cost performance frontier* is

$$\underline{\partial}\mathbb{H} \triangleq \{(b, z) \in \overline{\mathbb{H}}: b \in \mathbb{B} \text{ and } z \le f^\pi \text{ for any } \pi \in \Pi \text{ with } g^\pi = b\}.$$

We can readily characterize $\underline{\partial}\mathbb{H}$ as the graph of *optimal-cost function*

$$C(b) \triangleq \inf\{f^\pi: g^\pi = b,\ \pi \in \Pi\} = \inf\{z: (b, z) \in \mathbb{H}\}, \quad b \in \mathbb{B}, \tag{8}$$

whose convexity follows from that of region $\mathbb{H}$, so that

$$\underline{\partial}\mathbb{H} = \{(b, C(b)): b \in \mathbb{B}\}.$$

Notice that $C(b)$ is the optimal cost achievable by policies that expend a *supply* of b work units.

We can thus reformulate project control problem (5) as the *convex resource allocation problem*

$$\text{Find } b^* \in \mathbb{B}: \ C(b^*) = f^* \triangleq \inf\{C(b): b \in \mathbb{B}\}, \tag{9}$$

which is to find the optimal supply of work (as measured by g^π) to expend on the project.

To evaluate optimal-cost function $C(b)$ we will further address the following *b-work problem*:

$$\text{Find } \pi^* \in \Pi \text{ with } g^{\pi^*} = b: \ f^{\pi^*} = C(b) \triangleq \inf\{f^\pi: g^\pi = b, \pi \in \Pi\}. \tag{10}$$

Henceforth, we will say that a policy π is *b-work feasible* if it allocates a supply of b work units ($g^\pi = b$).

3.4. Lagrangian multiplier analysis and decentralization. To address b-work problem (10) we adapt the method of Lagrange multipliers. Dualizing *work supply constraint* $g^\pi = b$ by multiplier $\nu \in \mathbb{R}$ gives the *Lagrangian function*

$$\mathcal{L}_b^\pi(\nu) \triangleq f^\pi + \nu[g^\pi - b] = v^\pi(\nu) - \nu b, \tag{11}$$

defined for $\pi \in \Pi$ and $\nu \in \mathbb{R}$, where

$$v^\pi(\nu) \triangleq f^\pi + \nu g^\pi.$$

We interpret multiplier ν in economics terms as the *wage* earned by the operator per unit work performed. Thus, $v^\pi(\nu)$ is the value of holding and labor costs incurred, and $\mathcal{L}_b^\pi(\nu)$ is the adjusted cost value when work expended above (resp. below) the target b units supply is paid (resp. sold) at wage ν.

The (unconstrained) *Lagrangian problem* is:

$$\text{Find } \pi^* \in \Pi\text{:} \quad \mathcal{L}_b^{\pi^*}(\nu) = \mathcal{L}_b^*(\nu) \triangleq \inf\{\mathcal{L}_b^\pi(\nu)\colon \pi \in \Pi\}. \tag{12}$$

The latter is clearly equivalent to the following *ν-wage problem*, which is to find a policy minimizing the project's combined value of holding and working costs:

$$\text{Find } \pi^* \in \Pi\text{:} \quad v^{\pi^*}(\nu) = v^*(\nu) \triangleq \inf\{v^\pi(\nu)\colon \pi \in \Pi\}. \tag{13}$$

The optimal values of problems (12) and (13) are related by

$$\mathcal{L}_b^*(\nu) = v^*(\nu) - \nu b. \tag{14}$$

We will investigate the structure of solutions to ν-wage problem (13) as ν varies. Notice that the original control problem (13) is recovered as the case $\nu = 0$.

Drawing on economic interpretation, we view Lagrangian problem (12) as a *decentralized planning relaxation* of *centrally planned* b-work problem (10), where (i) the planner quotes the *prevailing wage* ν to the manager, and (ii) the latter is left to autonomously solve ν-wage problem (13). This suggests the possibility, which we discuss next, that the planner might decentralize the b-work problem's solution through an appropriate wage choice.

3.5. Duality-based optimality conditions and shadow wages. To price the value of work in b-work problem (10) we will use the *dual* (or *pricing*) *problem*, which is to find a wage that maximizes the (concave) objective $\mathcal{L}_b^*(\nu)$:

$$\text{Find } \nu^* \in \mathbb{R}\text{:} \quad \mathcal{L}_b^*(\nu^*) = Q(b) \triangleq \sup\{\mathcal{L}_b^*(\nu)\colon \nu \in \mathbb{R}\}. \tag{15}$$

We will further refer to the *duality gap* for a policy π and a wage ν, given by

$$\Delta_b^\pi(\nu) \triangleq f^\pi - \mathcal{L}_b^*(\nu) = v^\pi(\nu) - v^*(\nu). \tag{16}$$

The next result follows immediately.

Lemma 3.1 (Weak Duality).
(a) *Let policy $\pi \in \Pi$ be b-work feasible and let $\nu \in \mathbb{R}$. Then, $\mathcal{L}_b^*(\nu) \le f^\pi$.*
(b) $Q(b) \le C(b)$.

Lemma 3.1 readily yields the following result, which gives a sufficient optimality condition for the b-work problem and its dual, via *strong duality* ($Q(b) = C(b)$).

Lemma 3.2 (Sufficient Optimality Condition). *Let $\pi^* \in \Pi$ be a b-work feasible policy for which there is a wage $\nu^* \in \mathbb{R}$ with $\Delta_b^{\pi^*}(\nu^*) = 0$. Then*:
(a) *Policy π^* is optimal for b-work problem* (10).
(b) *Wage ν^* is optimal for its dual problem* (15).
(c) *Strong duality holds*: $Q(b) = C(b) = f^{\pi^*}$.

We will refer to a wage ν^* satisfying the sufficient optimality condition in Lemma 3.2 as a *shadow wage* for the b-work problem. In the present setting, its existence is also necessary for optimality.

Lemma 3.3 (Necessary Optimality Condition). *Let π^* be an optimal policy for the b-work problem. Then, there exists a corresponding shadow wage ν^*.*

PROOF. The assumed optimality of policy π^* implies that point $(g^{\pi^*}, f^{\pi^*}) = (b, f^{\pi^*})$ lies in the efficient frontier $\underline{\partial}\mathbb{H}$ of work-cost performance region $\mathbb{H}$. The latter's convexity allows us to invoke the *separating hyperplane theorem* (see, e.g., Weitzman [46]), to ensure existence of a shadow wage ν^* such that the line $\{(b', z') \in \mathbb{R}^2: z' + \nu^* b' = v^{\pi^*}(\nu^*)\}$ *supports* point (b, f^{π^*}) relative to $\mathbb{H}$. □

REMARK 3.1. If the sufficient optimality condition in Lemma 3.2 holds, then:

(i) The b-work problem's solution can be decentralized. The central planner needs only quote the shadow wage ν^* to the manager, and then let him solve autonomously the ν^*-wage problem.

(ii) If there is a unique shadow wage ν^* and optimal-cost function $C(\cdot)$ is differentiable at b, then

$$\nu^* = -\frac{dC}{db}(b). \tag{17}$$

Namely, shadow wage ν^* represents the marginal rate of decrease in cost incurred per unit increase in work expended, or *marginal productivity of work*, in the b-work problem.

3.6. Indexability and the MPI. To incorporate $\mathcal{F}$-policies into the preceding analysis, we next introduce a tractable project class, based on the structure of solutions to ν-wage problem (13) as ν varies.

DEFINITION 3.1 ($\mathcal{F}$-INDEXABILITY: MPI). We say that the project is *$\mathcal{F}$-indexable* (relative to f^π and g^π) if there exists a nondecreasing *index* $\nu_i^* \in \mathbb{R}$ for $i \in N^{\{0,1\}}$, which we term the project's MPI, such that, for each state $\ell^0 < i < \ell^1$, the S_i-active policy is optimal for the ν-wage problem iff $\nu \in [\nu_i^*, \nu_{i+1}^*]$.

Under $\mathcal{F}$-indexability, we see—recall that active-set S_i does not include state i—that it is optimal for the ν-wage problem to work on the project in state $i \in N^{\{0,1\}}$ iff wage ν does not exceed MPI ν_i^*. Definition 3.1 generalizes into a broader setting previous, particular definitions of indexability and restless bandit indices introduced by Whittle [47] and extended by Niño-Mora [34].

In economics terms, the indexability property characterizes the structure of the *demand curve for work*. Namely, for each wage (price of work) ν it gives a quantity (or a range thereof) of work units b, which it is optimal for the project manager to demand in order to solve his ν-wage problem. Thus, e.g., to any wage $\nu \in (\nu_i^*, \nu_{i+1}^*)$ there corresponds a demand of $b = g^{S_i}$ work units (see Figure 2).

The next result follows immediately.

LEMMA 3.4. *If the project is $\mathcal{F}$-indexable then*:

(a) *The ν-wage problem's optimal value can be represented as*

$$v^*(\nu) = \inf\{f^{S_i} + \nu g^{S_i}: i \in N\}, \quad \nu \in \mathbb{R}, \tag{18}$$

and is hence piecewise linear and concave in ν.

(b) *The project's MPI is given by*

$$\nu_i^* = -\frac{\Delta f^{S_i}}{\Delta g^{S_i}}, \quad i \in N^{\{0,1\}}. \tag{19}$$

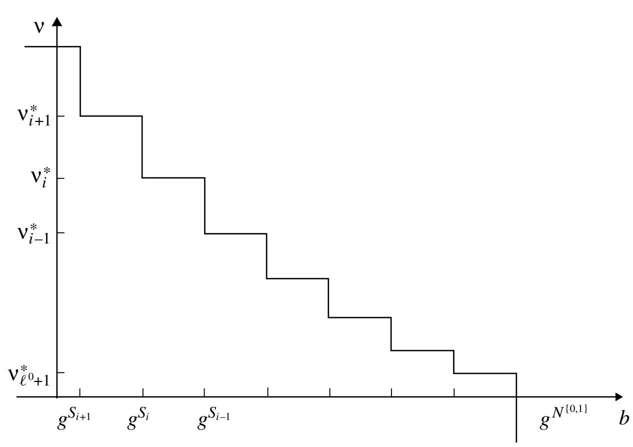

FIGURE 2. The demand curve for work of an $\mathcal{F}$-indexable project.

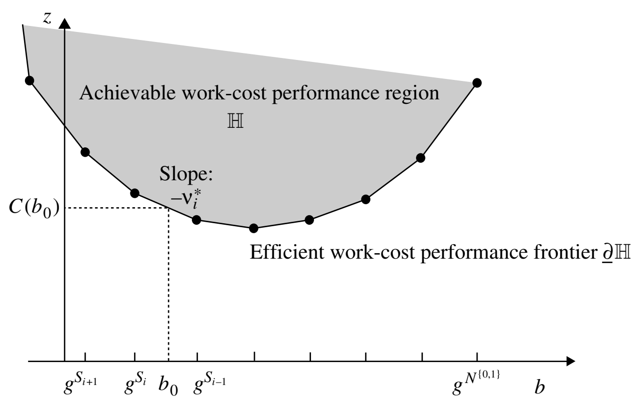

FIGURE 3. The efficient work-cost frontier of a project obeying $\mathcal{F}$-diminishing marginal returns to work.

3.7. $\mathcal{F}$-Diminishing marginal returns to work. To prepare the ground for the intuitive characterization of indexability to be given in the next section, we introduce below the class of projects obeying the economics *law of diminishing marginal returns to work*, in a form consistent with $\mathcal{F}$-policies.

Let us define index ν_i^* by (19), *without assuming $\mathcal{F}$-indexability*. Define now cost function $C^{\mathcal{F}}(\cdot)$: $\mathbb{B} \to \mathbb{R}$ by linear interpolation on work-cost pairs (g^S, f^S), for $S \in \mathcal{F}$. Namely, for $b \in [g^{S_i}, g^{S_{i-1}}]$, let

$$\begin{aligned} C^{\mathcal{F}}(b) &\triangleq f^{S_i} + \nu_i^*[g^{S_i} - b] = q f^{S_{i-1}} + (1-q) f^{S_i} \\ &= f^{qS_{i-1}+(1-q)S_i} = f^{S_i^q}, \end{aligned} \tag{20}$$

where

$$q = \frac{b - g^{S_i}}{g^{S_{i-1}} - g^{S_i}} \in [0, 1]; \tag{21}$$

and $S_i^q \in \Pi^{\mathrm{SR}}$ is the policy that, at state $j \in N^{\{0,1\}}$, prescribes to (i) engage the project if $j > i$, (ii) rest it if $j < i$, and (iii) engage it with probability (w.p.) q and rest it w.p. $1-q$ if $j = i$. Thus, for such b, $C^{\mathcal{F}}(b)$ is the cost performance achieved both by policy $qS_{i-1} + (1-q)S_i$ and by policy S_i^q.

DEFINITION 3.2 ($\mathcal{F}$-DIMINISHING MARGINAL RETURNS). We say that the project obeys *$\mathcal{F}$-diminishing marginal returns to work* if the b-work problem's optimal-cost function is $C(b) = C^{\mathcal{F}}(b)$.

Notice that the property of $\mathcal{F}$-diminishing marginal returns to work characterizes the structure of the project's efficient work-cost performance frontier (see Figure 3). The next result follows immediately.

LEMMA 3.5. *If the project obeys $\mathcal{F}$-diminishing marginal returns to work, then*:
(a) *Index ν_i^* is nondecreasing.*
(b) *Optimal-cost function $C(b)$ is the upper envelope of Lagrangians $\mathcal{L}_b^{S_i}(\nu_i^*)$ for $i \in N^{\{0,1\}}$, i.e.,*

$$C(b) = \max\{f^{S_i} + \nu_i^*[g^{S_i} - b]\colon i \in N^{\{0,1\}}\}, \quad b \in \mathbb{B}, \tag{22}$$

and is hence piecewise linear and convex in the work supply b.
(c) *The b-work problem, for $b \in [g^{S_i}, g^{S_{i-1}}]$, is solved by policy S_i^q, with q as given in* (21).

3.8. Intuitive characterization of indexability via diminishing marginal returns. Intuition suggests that the classes of $\mathcal{F}$-indexable projects and of projects obeying $\mathcal{F}$-diminishing marginal returns to work should be closely related. We establish next that they are, indeed, equivalent, thus providing an economically intuitive characterization of the indexability property.

THEOREM 3.1. *The project is $\mathcal{F}$-indexable iff it obeys $\mathcal{F}$-diminishing marginal returns to work.*

PROOF. Suppose the project obeys $\mathcal{F}$-diminishing marginal returns. Then, for $\ell^0 < i < \ell^1$:

$$\begin{aligned} \nu_i^* \le \nu \le \nu_{i+1}^* &\iff \frac{f^{S_i} - f^{S_{i-1}}}{g^{S_{i-1}} - g^{S_i}} \le \nu \le \frac{f^{S_{i+1}} - f^{S_i}}{g^{S_i} - g^{S_{i+1}}} \\ &\iff v^{S_i}(\nu) = \min\{v^{S_{i+k}}(\nu)\colon k = -1, 0, 1\} \\ &\iff C(g^{S_i}) + \nu g^{S_i} = \min\{C(g^{S_{i+k}}) + \nu g^{S_{i+k}}\colon k = -1, 0, 1\} \end{aligned}$$

$$\iff C(g^{S_i}) + \nu g^{S_i} = \min\{C(b) + \nu b\colon b \in [g^{S_{i+1}}, g^{S_{i-1}}]\}$$
$$\iff C(g^{S_i}) + \nu g^{S_i} = \min\{C(b) + \nu b\colon b \in \mathbb{B}\}$$
$$\iff v^{S_i}(\nu) = \min\{z + \nu b\colon (b, z) \in \mathbb{H}\}$$
$$\iff v^*(\nu) = v^{S_i}(\nu).$$

This establishes that the project is $\mathcal{F}$-indexable.

Suppose now that the project is $\mathcal{F}$-indexable. Let $b \in [g^{S_{i+1}}, g^{S_i}]$, with $\ell^0 < i < \ell^1$. Letting q be given by (21), we can write

$$\big(g^{S_i^q}, f^{S_i^q}\big) = \big(b, f^{S_i^q}\big) = (b, C^{\mathcal{F}}(b)). \tag{23}$$

Hence, S_i^q is a feasible policy for the b-work problem. Furthermore, the duality gap (cf. (16)) associated to policy S_i^q and wage rate ν_i^* is

$$\Delta_b^{S_i^q}(\nu_i^*) = v^{S_i^q}(\nu_i^*) - v^*(\nu_i^*) = 0,$$

where the last identity follows by $\mathcal{F}$-indexability. Hence, the sufficient optimality condition in Lemma 3.2 holds, which gives, using (23), that $C(b) = C^{\mathcal{F}}(b)$. This shows that the project obeys $\mathcal{F}$-diminishing marginal returns, and completes the proof. □

4. Sufficient indexability conditions via PCLs. Suppose we aim to establish $\mathcal{F}$-indexability of a restless bandit project model fitting in the above setting. To do so we must calculate index ν_i^* by (19) and show that it is nondecreasing. Yet this is only a *necessary*, but *not sufficient*, condition for $\mathcal{F}$-indexability. It remains (cf. Definition 3.1) to prove that, for every state $\ell^0 < i < \ell^1$, the S_i-active policy is optimal for ν-wage problem (13) if and only if $\nu \in [\nu_i^*, \nu_{i+1}^*]$.

We present below a framework for establishing indexability, based on satisfaction of *partial conservation laws* (PCLs). Our main result is Theorem 4.1, which identifies a tractable class of indexable projects, termed *PCL-indexable*. We introduced PCLs in Niño-Mora [32, 34], in a form restricted to finite-state projects, based on LP methods. The approach below pursues a different course, extending the scope to countable-state projects. Although the following results could be grounded on infinite-dimensional LP, we have chosen instead to rely only on direct, first-principles arguments.

4.1. Decomposition and conservation laws. The PCL framework concerns a generic semi-Markov restless project as in §3, whose work (g^π) and cost (f^π) measures decompose linearly in terms of *state-action frequency measures* $x_i^{a,\pi} \geq 0$. Here, $x_i^{a,\pi}$ is a measure (e.g., discounted or long-run average) of the frequency of decision epochs at which action a is taken in state i under policy π, i.e., of the frequency of (i, a)-periods. Recall that an (i, a)-period is the time interval that starts with a decision epoch where action a is taken in state i, and ends with the next decision epoch, while an i-period is a period between decision epochs starting in state i. For consistency with such interpretation, we require such measures to satisfy the following conditions.

Assumption 4.1. *For any admissible policy $\pi \in \Pi$ and state $i \in N$:*

(i) *If π takes the active action at i-periods, then $x_i^{0,\pi} = 0$.*

(ii) *If π takes the passive action at i-periods, then $x_i^{1,\pi} = 0$.*

(iii) *For $S \in \mathcal{F}$, $j_1 \in S$, $j_2 \in S^c$, it holds that $x_{j_1}^{0, S\setminus\{j_1\}}, x_{j_2}^{1, S\cup\{j_2\}} > 0$.*

We will refer to an i-period where $i \in S$, for a given $S \subseteq N^{\{0,1\}}$, as an S-*period*, and write $S^c = N^{\{0,1\}} \setminus S$.

Definition 4.1 (Partial Conservation Laws). We say that the project's work measures satisfy *PCLs relative to $\mathcal{F}$-policies*, or *PCL($\mathcal{F}$)* for short, if there exist coefficients $w_i^S > 0$ ($i \in N^{\{0,1\}}$, $S \in \mathcal{F}$) such that, for any admissible policy $\pi \in \Pi$ and feasible active-state set $S \in \mathcal{F}$, the following holds:

$$\text{(PCL1)} \quad g^\pi + \sum_{i \in S} w_i^S x_i^{0,\pi} \geq g^S, \quad \text{with “=” if } \pi \text{ is passive at } S^c\text{-periods.}$$

$$\text{(PCL2)} \quad \sum_{i \in S} w_i^S x_i^{0,\pi} \geq 0, \quad \text{with “=” if } \pi \text{ is active at } S\text{-periods.}$$

Remark 4.1.

(i) The term "partial" refers to the fact that (PCL1)–(PCL2) only hold for the family of feasible active sets $S \in \mathcal{F}$. In the *strong conservation laws* of Shanthikumar and Yao [41], and in the *generalized conservation laws* of Bertsimas and Niño-Mora [4], analogous laws hold for *all* S.

(ii) The term "conservation" refers to the equality case in (PCL1)–(PCL2), e.g., the quantity (work) $g^\pi + \sum_{i\in S} w_i^S x_i^{0,\pi}$ is *conserved* under all admissible policies π resting the project at S^c-periods.

(iii) Satisfaction of PCL($\mathcal{F}$) ensures that ν-wage problem (13) is solved by $\mathcal{F}$-policies for a particular family of linear objectives. Namely, for $f^\pi = \sum_{i\in S} w_i^S x_i^{0,\pi}$ and $\nu = 1$ or $\nu = 0$.

We will derive satisfaction of PCLs from the following primitive requirements.

Assumption 4.2. *There exist coefficients w_i^S, c_i^S ($i \in N^{\{0,1\}}$, $S \in \mathcal{F}$), such that the following holds*:

(i) $w_i^S > 0$, *for $i \in N^{\{0,1\}}$ and $S \in \mathcal{F}$.*

(ii) *Work decomposition laws: For any admissible policy $\pi \in \Pi$ and feasible active-state set $S \in \mathcal{F}$,*

$$g^\pi + \sum_{i\in S} w_i^S x_i^{0,\pi} = g^S + \sum_{i\in S^c} w_i^S x_i^{1,\pi}.$$

(iii) *Cost decomposition laws: For any admissible policy $\pi \in \Pi$ and feasible active-state set $S \in \mathcal{F}$,*

$$f^\pi + \sum_{i\in S^c} c_i^S x_i^{1,\pi} = f^S + \sum_{i\in S} c_i^S x_i^{0,\pi}.$$

(iv) $c_i^{S_{j-1}} - c_i^{S_j} = (c_j^{S_j}/w_j^{S_j})[w_i^{S_{j-1}} - w_i^{S_j}]$, *for any controllable states $i, j \in N^{\{0,1\}}$.*

Remark 4.2.

(i) Setting $S = N^{\{0,1\}}$ in Assumption 4.2(ii, iii) we obtain that work measure g^π can be represented in terms of the $x_i^{0,\pi}$'s as

$$g^\pi = g^{N^{\{0,1\}}} - \sum_{i\in N^{\{0,1\}}} w_i^{N^{\{0,1\}}} x_i^{0,\pi},$$

while cost measure f^π can be represented as

$$f^\pi = f^{N^{\{0,1\}}} + \sum_{i\in N^{\{0,1\}}} c_i^{N^{\{0,1\}}} x_i^{0,\pi}.$$

(ii) Setting $j = i$ in Assumption 4.2(iv) we obtain that $c_i^{S_i}/w_i^{S_i} = c_i^{S_{i-1}}/w_i^{S_{i-1}}$.

The next result follows immediately from the above.

Lemma 4.1. *Under Assumptions 4.1 and 4.2*(i, ii), *PCL($\mathcal{F}$) hold.*

We will term coefficient w_i^S the (i,S)-*marginal workload*, and c_i^S the (i,S)-*marginal cost*. The next result justifies such denominations.

Lemma 4.2. *For any feasible active-state set $S \in \mathcal{F}$, and states $j_1 \in S$ and $j_2 \in S^c$*:

(a) $g^{S\setminus\{j_1\}} + w_{j_1}^S x_{j_1}^{0,S\setminus\{j_1\}} = g^S = g^{S\cup\{j_2\}} - w_{j_2}^S x_{j_2}^{1,S\cup\{j_2\}}$.

(b) $f^{S\setminus\{j_1\}} - c_{j_1}^S x_{j_1}^{0,S\setminus\{j_1\}} = f^S = f^{S\cup\{j_2\}} + c_{j_2}^S x_{j_2}^{1,S\cup\{j_2\}}$.

Proof. The left (resp. right) identities for g^S and for f^S follow by letting $\pi = S\setminus\{j_1\}$ (resp. $\pi = S\cup\{j_2\}$) in Assumption 4.2(ii, iii), respectively. □

The following result sheds light on the interpretation of Assumption 4.2(i), which plays a fundamental role in the PCL framework. It shows that such assumption represents a stronger version of the monotonicity property of work measures given in Assumption 3.1(iv).

Lemma 4.3. *Assumption 4.2*(i), *i.e., $w_i^S > 0$ for $i \in N^{\{0,1\}}$, $S \in \mathcal{F}$, is equivalent to*

$$g^{S\setminus\{j_1\}} < g^S < g^{S\cup\{j_2\}}, \quad S \in \mathcal{F}, \quad j_1 \in S, \quad j_2 \in S^c.$$

Proof. The result follows immediately from Lemma 4.2(a) and Assumption 4.1(iii). □

We will refer in what follows to the *aggregate* marginal work and cost measures

$$W^{S,0,\pi} \triangleq \sum_{i\in S} w_i^S x_i^{0,\pi}, \qquad W^{S,1,\pi} \triangleq \sum_{i\in S^c} w_i^S x_i^{1,\pi},$$
$$C^{S,0,\pi} \triangleq \sum_{i\in S} c_i^S x_i^{0,\pi}, \qquad C^{S,1,\pi} \triangleq \sum_{i\in S^c} c_i^S x_i^{1,\pi}. \tag{24}$$

Remark 4.3. By Assumption 4.2(ii), (PCL1) in Definition 4.1 can be equivalently reformulated as

$$\text{(PCL1)} \quad W^{S,1,\pi} \ge 0, \quad \text{with “=” if } \pi \text{ is passive at } S^c\text{-periods.}$$

4.2. PCL($\mathscr{F}$)**-indexable projects.** Let us introduce coefficients

$$\nu_i^S \triangleq \frac{c_i^S}{w_i^S}, \quad i \in N^{\{0,1\}}, \quad S \in \mathscr{F}, \tag{25}$$

from which we define *index* ν_i^* by

$$\nu_i^* \triangleq \nu_i^{S_i} = \nu_i^{S_{i-1}}, \quad i \in N^{\{0,1\}}. \tag{26}$$

We will term ν_i^S the (i, S)-*marginal productivity rate*. Thus, index ν_i^* is the (i, S_i)-marginal, or (i, S_{i-1})-marginal, productivity rate (cf. Remark 4.2(ii) for the equality $\nu_i^{S_i} = \nu_i^{S_{i-1}}$).

The following result gives an alternative representation for index ν_i^* in (26), which is consistent with that for the MPI in Lemma 3.4(b).

Lemma 4.4.

$$\nu_i^* = -\frac{\Delta f^{S_i}}{\Delta g^{S_i}}, \quad i \in N^{\{0,1\}}.$$

Proof. The result follows from Lemma 4.2, using Assumption 4.1(iii). □

Thus, Lemma 4.4 shows that, if the project is $\mathscr{F}$-indexable, then index ν_i^* in (26) is indeed the MPI. We introduce next a tractable project class based on the above.

Definition 4.2 (PCL($\mathscr{F}$)-Indexability). We say that the project is *PCL*($\mathscr{F}$)-*indexable* if:

(i) Assumptions 4.1 and 4.2 hold, and, hence, (cf. Lemma 4.1) PCL($\mathscr{F}$) hold.

(ii) Index ν_i^* is nondecreasing.

Our main result in this setting is that a PCL($\mathscr{F}$)-indexable project is, indeed, $\mathscr{F}$-indexable.

Theorem 4.1 (Sufficient $\mathscr{F}$-Indexability Condition). *If the project is PCL*($\mathscr{F}$)*-indexable, then it is* $\mathscr{F}$*-indexable, and* ν_i^* *is its MPI.*

To prove Theorem 4.1 we need several preliminary results. We start by relating marginal productivity rates to index values, which yields an index recursion in part (b) of the following result.

Lemma 4.5. *For any controllable states* $i, j \in N^{\{0,1\}}$:

(a) $\nu_j^{S_i} - \nu_i^* = (w_j^{S_{i-1}}/w_j^{S_i})[\nu_j^{S_{i-1}} - \nu_i^*]$.

(b) $\nu_{i+1}^* = \nu_i^* + (w_{i+1}^{S_{i-1}}/w_{i+1}^{S_i})[\nu_{i+1}^{S_{i-1}} - \nu_i^*]$.

Proof. (a) Using Assumption 4.2(iv) we obtain that, for $i, j \in N^{\{0,1\}}$,

$$\nu_j^{S_i} - \nu_i^* = \frac{c_j^{S_i}}{w_j^{S_i}} - \nu_i^* = \frac{c_j^{S_{i-1}} - \nu_i^*\big[w_j^{S_{i-1}} - w_j^{S_i}\big]}{w_j^{S_i}} - \nu_i^* = \frac{w_j^{S_{i-1}}}{w_j^{S_i}}\big[\nu_j^{S_{i-1}} - \nu_i^*\big].$$

(b) This part follows by letting $j = i + 1$ in part (a), noting that $\nu_{i+1}^{S_i} = \nu_{i+1}^*$ (cf. Remark 4.2(ii)). □

The next result represents marginal costs as finite linear combinations of marginal workloads, involving coefficients ν_i^* and $\Delta\nu_i^* = \nu_i^* - \nu_{i-1}^*$. It further reformulates such identities in terms of marginal productivity rates.

Lemma 4.6. *For any controllable states* $i, j \in N^{\{0,1\}}$:

(a) *If* $i \le j$, $c_j^{S_{i-1}} = \nu_i^* w_j^{S_{i-1}} + \sum_{k=i+1}^j w_j^{S_{k-1}} \Delta\nu_k^*$ *or, equivalently,* $\nu_j^{S_{i-1}} = \nu_i^* + \sum_{k=i+1}^j (w_j^{S_{k-1}}/w_j^{S_{i-1}})\Delta\nu_k^*$.

(b) *If* $i \ge j$, $c_j^{S_i} = \nu_i^* w_j^{S_i} - \sum_{k=j}^{i-1} w_j^{S_k} \Delta\nu_{k+1}^*$ *or, equivalently,* $\nu_j^{S_i} = \nu_i^* - \sum_{k=j}^{i-1} (w_j^{S_k}/w_j^{S_i})\Delta\nu_{k+1}^*$.

Proof. (a) For $i < j$, Assumption 4.2(iv) and *summation by parts* gives

$$\begin{aligned} c_j^{S_j} &= c_j^{S_{i-1}} + \sum_{k=i}^{j}\big[c_j^{S_k} - c_j^{S_{k-1}}\big] = c_j^{S_{i-1}} + \sum_{k=i}^{j}\nu_k^*\big[w_j^{S_k} - w_j^{S_{k-1}}\big] \\ &= c_j^{S_{i-1}} + \big[\nu_j^* w_j^{S_j} - \nu_i^* w_j^{S_{i-1}}\big] - \sum_{k=i+1}^{j} w_j^{S_{k-1}} \Delta\nu_k^*, \end{aligned}$$

whence the result follows (the case $i = j$ is trivial).

(b) For $j < i$ (again, the case $i = j$ is trivial), we have

$$c_j^{S_i} = c_j^{S_{j-1}} + \sum_{k=j}^{i} [c_j^{S_k} - c_j^{S_{k-1}}] = c_j^{S_{j-1}} + \sum_{k=j}^{i} \nu_k^* [w_j^{S_k} - w_j^{S_{k-1}}]$$
$$= c_j^{S_{j-1}} + [\nu_i^* w_j^{S_i} - \nu_j^* w_j^{S_{j-1}}] - \sum_{k=j}^{i-1} w_j^{S_k} \Delta\nu_{k+1}^*,$$

whence the result follows. This completes the proof. □

The next result is an analog of Lemma 4.6 in terms of aggregate measures.

LEMMA 4.7. *Suppose that the project is PCL($\mathscr{F}$)-indexable. Then, for any controllable state $i \in N^{\{0,1\}}$:*
(a) $C^{S_{i-1},0,\pi} = \nu_i^* W^{S_{i-1},0,\pi} + \sum_{k\in S_i} W^{S_{k-1},0,\pi} \Delta\nu_k^*$.
(b) $C^{S_i,1,\pi} = \nu_i^* W^{S_i,1,\pi} - \sum_{k\in S_{i-1}^c} W^{S_k,1,\pi} \Delta\nu_{k+1}^*$.

PROOF. (a) Using Lemma 4.6(a), we obtain

$$C^{S_{i-1},0,\pi} \triangleq \sum_{j\in S_{i-1}} c_j^{S_{i-1}} x_j^{0,\pi} = \sum_{j\in S_{i-1}} \left[\nu_i^* w_j^{S_{i-1}} + \sum_{k=i+1}^{j} w_j^{S_{k-1}} \Delta\nu_k^*\right] x_j^{0,\pi}$$
$$= \nu_i^* \sum_{j\in S_{i-1}} w_j^{S_{i-1}} x_j^{0,\pi} + \sum_{k\in S_i} \sum_{j\in S_{k-1}} w_j^{S_{k-1}} x_j^{0,\pi} \Delta\nu_k^*$$
$$= \nu_i^* W^{S_{i-1},0,\pi} + \sum_{k\in S_i} W^{S_{k-1},0,\pi} \Delta\nu_k^*,$$

where the interchange in the order of summation is justified, in the countable state case, by the nonnegativity of the terms involved.

(b) Using Lemma 4.6(b) and arguing along the same lines as in part (a) gives

$$C^{S_i,1,\pi} \triangleq \sum_{j\in S_i^c} c_j^{S_i} x_j^{1,\pi} = \sum_{j\in S_i^c} \left[\nu_i^* w_j^{S_i} - \sum_{k=j}^{i-1} w_j^{S_k} \Delta\nu_{k+1}^*\right] x_j^{1,\pi}$$
$$= \nu_i^* \sum_{j\in S_i^c} w_j^{S_i} x_j^{1,\pi} - \sum_{k\in S_{i-1}^c} \sum_{j\in S_k^c} w_j^{S_k} x_j^{1,\pi} \Delta\nu_{k+1}^*$$
$$= \nu_i^* W^{S_i,1,\pi} - \sum_{k\in S_{i-1}^c} W^{S_k,1,\pi} \Delta\nu_{k+1}^*,$$

which completes the proof. □

4.3. Workload reformulation and indexability proof. The next result is the cornerstone of our indexability proof. It formulates the difference between ν-wage problem (13)'s objective under an arbitrary policy and under an $\mathscr{F}$-policy as a linear combination of aggregate work measures.

LEMMA 4.8 (WORKLOAD REFORMULATION). *Suppose that the project is PCL($\mathscr{F}$)-indexable. Then, for any state $\ell^0 < i < \ell^1$, wage $\nu \in \mathbb{R}$ and policy $\pi \in \Pi$, ν-wage problem (13)'s objective can be written as*

$$v^\pi(\nu) = v^{S_i}(\nu) + W^{S_i,1,\pi}[\nu - \nu_i^*] + W^{S_i,0,\pi}[\nu_{i+1}^* - \nu]$$
$$+ \sum_{k\in S_{i+1}} W^{S_{k-1},0,\pi} \Delta\nu_k^* + \sum_{k\in S_{i-1}^c} W^{S_k,1,\pi} \Delta\nu_{k+1}^*.$$

PROOF. Using in turn Assumption 4.2(iii) and Lemma 4.7 gives

$$f^\pi = f^{S_i} + C^{S_i,0,\pi} - C^{S_i,1,\pi} = f^{S_i} + \nu_{i+1}^* W^{S_i,0,\pi} - \nu_i^* W^{S_i,1,\pi}$$
$$+ \sum_{k\in S_{i+1}} W^{S_{k-1},0,\pi} \Delta\nu_k^* + \sum_{k\in S_{i-1}^c} W^{S_k,1,\pi} \Delta\nu_{k+1}^*.$$

On the other hand, by Assumption 4.2(ii) we have

$$g^\pi = g^{S_i} + W^{S_i,1,\pi} - W^{S_i,0,\pi}.$$

The required expression for $v^\pi(\nu) = f^\pi + \nu g^\pi$ follows directly by substitution of the above formulae for f^π and g^π. □

We are now ready to prove the main result of this section.

Proof of Theorem 4.1. Let $\ell^0 < i < \ell^1$. We will first prove that, for any wage $\nu \in [\nu_i^*, \nu_{i+1}^*]$, the S_i-active policy is optimal for ν-wage problem (13). Consider an arbitrary policy $\pi \in \Pi$. It follows immediately from Lemma 4.8 and Definition 4.2 that, for such ν,

$$v^\pi(\nu) \geq v^{S_i}(\nu), \tag{27}$$

which establishes the required implication.

It remains to prove the reverse implication, namely, that if the S_i-active policy is optimal for the ν-wage problem, then it must be $\nu \in [\nu_i^*, \nu_{i+1}^*]$. Suppose that the S_i-active policy satisfies such optimality property, so that (27) holds for any admissible policy $\pi \in \Pi$. Then, setting $\pi = S_{i+1}$ in (27) gives

$$f^{S_{i+1}} + \nu g^{S_{i+1}} \geq f^{S_i} + \nu g^{S_i},$$

which is readily reformulated, using Assumption 3.1(iv) and Lemma 4.4 as $\nu \leq \nu_{i+1}^*$.

Further, setting $\pi = S_{i-1}$ in (27) gives

$$f^{S_{i-1}} + \nu g^{S_{i-1}} \geq f^{S_i} + \nu g^{S_i},$$

which is reformulated along the same lines as $\nu \geq \nu_i^*$. This completes the proof. □

The next result gives an insightful characterization of the MPI as a locally optimal marginal productivity rate, in a max-min relation.

Theorem 4.2. *Suppose that the project is PCL($\mathscr{F}$)-indexable. Then, for any controllable state $i \in N^{\{0,1\}}$:*

$$\max_{j \in S_i^c} \nu_j^{S_i} = \nu_i^{S_i} = \nu_i^* = \nu_i^{S_{i-1}} = \min_{j \in S_{i-1}} \nu_j^{S_{i-1}};$$

or, equivalently,

$$\max_{j \in S_i^c} \frac{f^{S_i} - f^{S_i \cup \{j\}}}{g^{S_i \cup \{j\}} - g^{S_i}} = \frac{f^{S_i} - f^{S_{i-1}}}{g^{S_{i-1}} - g^{S_i}} = \nu_i^* = \frac{f^{S_i} - f^{S_{i-1}}}{g^{S_{i-1}} - g^{S_i}} = \min_{j \in S_{i-1}} \frac{f^{S_{i-1} \setminus \{j\}} - f^{S_{i-1}}}{g^{S_{i-1}} - g^{S_{i-1} \setminus \{j\}}}.$$

Proof. From Lemma 4.6(a) we readily obtain that, for $i, j \in N^{\{0,1\}}$ with $i \leq j$,

$$\nu_j^{S_{i-1}} = \nu_i^* + \sum_{k=i+1}^{j} \frac{w_j^{S_{k-1}}}{w_j^{S_{i-1}}} \Delta\nu_k^*,$$

whence the "min" identity follows.

Further, Lemma 4.6(b) gives that, for $i, j \in N^{\{0,1\}}$ with $i \geq j$,

$$\nu_j^{S_i} = \nu_i^* - \sum_{k=j}^{i-1} \frac{w_j^{S_k}}{w_j^{S_i}} \Delta\nu_{k+1}^*,$$

whence the "max" identity follows.

The stated equivalent reformulation of the result follows from Lemma 4.2. □

Remark 4.4. In Theorem 4.2:

(i) The "max" part shows that ν_i^* is the maximum (j, S_i)-marginal productivity rate over $j \in S_i^c$.

(ii) The "min" part shows that ν_i^* is the minimum (j, S_{i-1})-marginal productivity rate over $j \in S_{i-1}$.

4.4. MPI characterization under wedge-shaped marginal workloads. We have observed that, in a variety of models, marginal workloads are *wedge shaped*, in the sense stated below. This implies the alternative, insightful characterization of the MPI given in Theorem 4.3 below, which characterizes the MPI as an *optimal marginal productivity rate relative to feasible active-state sets*. Such a result adapts to restless bandits, as shown by Gittins' [17] characterization of his index for nonrestless bandits as an *optimal average productivity rate relative to active-state sets*.

Assumption 4.3. *Marginal workload $w_i^{S_k}$ is wedge shaped as threshold state k varies, being minimized at either $k = i-1$ or $k = i$, i.e.,*

$$w_i^{S_{\ell^0}} \geq w_i^{S_{\ell^0+1}} \geq \cdots \geq w_i^{S_{i-1}}, \; w_i^{S_i} \leq w_i^{S_{i+1}} \leq \cdots, \quad i \in N^{\{0,1\}}.$$

We need the following preliminary result.

LEMMA 4.9. *Suppose that the project is PCL($\mathcal{F}$)-indexable and Assumption 4.3 holds. Then, for every controllable state $i \in N^{\{0,1\}}$, marginal productivity rate $\nu_i^{S_j}$ is nondecreasing in threshold state j:*

$$\nu_i^{S_j} \geq \nu_i^{S_{j-1}}, \quad j \in N^{\{0,1\}}, \ \textit{with "$=$" if } j = i.$$

PROOF. For the equality $\nu_i^{S_i} = \nu_i^{S_{i-1}}$ see Remark 4.2(ii). As for the inequality, reformulating the identity in Lemma 4.5(a) gives that, for $i, j \in N^{\{0,1\}}$,

$$\nu_i^{S_j} - \nu_i^{S_{j-1}} = \left(\frac{w_i^{S_{j-1}}}{w_i^{S_j}} - 1\right)\left(\nu_i^{S_{j-1}} - \nu_j^*\right). \tag{28}$$

Suppose $j \geq i+1$. Then, Assumption 4.3 ensures that the first factor in the right-hand side of (28) is nonpositive. Further, the "max" identity in Theorem 4.2 and index nondecreasingness give $\nu_i^{S_{j-1}} \leq \nu_{j-1}^* \leq \nu_j^*$, implying that the second factor is also nonpositive and, hence, the product is nonnegative.

Suppose now $j \leq i-1$. Then, the "min" identity in Theorem 4.2 gives $\nu_j^* \leq \nu_i^{S_{j-1}}$ and, hence, the second factor in the right-hand side of (28) is nonnegative. Further, Assumption 4.3 shows that the first factor is nonnegative. Therefore, so is their product, which completes the proof. □

THEOREM 4.3 (ALTERNATIVE MPI CHARACTERIZATION). *Suppose that the project is PCL($\mathcal{F}$)-indexable and Assumption 4.3 holds. Then, the MPI has the following characterization: For any controllable state $i \in N^{\{0,1\}}$,*

$$\max_{S \in \mathcal{F}: S \ni i} \nu_i^S = \nu_i^* = \min_{S \in \mathcal{F}: S^c \ni i} \nu_i^S;$$

or, equivalently,

$$\max_{S \in \mathcal{F}: S \ni i} \frac{f^{S \setminus \{i\}} - f^S}{g^S - g^{S \setminus \{i\}}} = \nu_i^* = \min_{S \in \mathcal{F}: S^c \ni i} \frac{f^S - f^{S \cup \{i\}}}{g^{S \cup \{i\}} - g^S}.$$

PROOF. By Lemma 4.9 we can write, for $j_1 \in S_{i-1}$ and $j_2 \in S_{i-1}^c$,

$$\nu_i^{S_{j_1}} \geq \nu_i^{S_i} = \nu_i^* = \nu_i^{S_{i-1}} \geq \nu_i^{S_{j_2}},$$

which implies

$$\min_{j \in N: S_j^c \ni i} \nu_i^{S_j} = \nu_i^* = \max_{j \in N: S_j \ni i} \nu_i^{S_j}, \quad i \in N^{\{0,1\}}.$$

The latter identities are readily reformulated into the stated result. The equivalent reformulation follows immediately from Lemma 4.2. □

REMARK 4.5. In Theorem 4.3:

(i) The "max" identity characterizes MPI ν_i^* as the maximum (i, S)-marginal productivity rate over $\mathcal{F}$-policies S that are active over i-periods.

(ii) The "min" identity characterizes MPI ν_i^* as the minimum (i, S)-marginal productivity rate over $\mathcal{F}$-policies S that are passive over i-periods.

5. PCL-indexability analysis of semi-Markov projects. This section deploys the PCL-indexability framework in the setting of a general semi-Markov project under three performance criteria, one of which is new. The main result is that PCL-indexability follows from satisfaction of the following simpler conditions.

ASSUMPTION 5.1.

(i) *Positive marginal workloads*: $w_i^S > 0$, *for any controllable state $i \in N^{\{0,1\}}$ and feasible active-state set $S \in \mathcal{F}$.*

(ii) *Nondecreasing index*: $\nu_i^* \leq \nu_{i+1}^*$, *for any state $\ell^0 < i < \ell^1$.*

Consider a semi-Markov project as in §3. We next describe its dynamics, along the lines in Puterman [39, Chapter 11], starting with those of its embedded process X_n. If at a decision epoch t_n where the project occupies state $X_n \triangleq X(t_n) = i$, action $a_n \triangleq a(t_n) = a$ is chosen, then the joint distribution of the length $t_{n+1} - t_n$ of the ensuing (i, a)-*period* and state X_{n+1} is given by the *transition distribution*

$$Q_{ij}^a(t) \triangleq \mathbb{P}\{t_{n+1} - t_n \leq t, \ X_{n+1} = j \mid X_n = i, \ a_n = a\},$$

with LST

$$\beta_{ij}^{a,\alpha} \triangleq \mathbb{E}\left[1_{\{X_{n+1}=j\}} e^{-\alpha(t_{n+1}-t_n)} \mid X_n = i,\ a_n = a\right] = \int_0^\infty e^{-\alpha t}\, dQ_{ij}^a(t),$$

for $\alpha > 0$. The corresponding *one-period transition probabilities* of the embedded process are

$$p_{ij}^a \triangleq \mathbb{P}\{X_{n+1} = j \mid X_n = i,\ a_n = a\} = \lim_{t\to\infty} Q_{ij}^a(t) = \lim_{\alpha \searrow 0} \beta_{ij}^{a,\alpha}.$$

From $Q_{ij}^a(t)$ we obtain the distribution of the length of an (i, a)-period,

$$F_i^a(t) \triangleq \mathbb{P}\{t_{n+1} - t_n \le t \mid X_n = i,\ a_n = a\} = \sum_{j\in N} Q_{ij}^a(t),$$

having LST

$$\beta_i^{a,\alpha} \triangleq \mathbb{E}[e^{-\alpha(t_{n+1}-t_n)} \mid X_n = i,\ a_n = a] = \sum_{j\in N} \beta_{ij}^{a,\alpha},$$

and mean

$$m_i^a \triangleq \mathbb{E}[t_{n+1} - t_n \mid X_n = i,\ a_n = a] = \int_0^\infty t\, dF_i^a(t).$$

We require, in the countable state case, that the following regularity conditions hold.

Assumption 5.2.
(i) *There exists* $\delta > 0$ *such that* $\sup_{i,a} F_i^a(\delta) < 1$.
(ii) *The expected time between decision epochs is finite, i.e.,* $m_i^a < \infty$ *for each* i, a.

Remark 5.1. Assumption 5.2(i) was introduced in Ross [40]. It implies (cf. Puterman [39, Problem 11.3]) the following:
(i) The expected number of decision epochs in any finite time interval is finite.
(ii) The row sums of matrix $\beta_{ij}^{a,\alpha}$ are uniformly bounded away from unity, i.e.,

$$\sup_{i,a} \beta_i^{a,\alpha} < 1.$$

(iii) The expected time between decision epochs is uniformly bounded away from zero, i.e.,

$$\inf_{i,a} m_i^a > 0.$$

Recall that, in general, the natural state process $X(t)$ can change state between decision epochs. The evolution of $X(t)$ *within* an (i, a)-period is characterized by

$$\hat{p}_{ij}^a(t) \triangleq \mathbb{P}\{X(t_n + t) = j \mid X_n = i,\ a_n = a,\ t_{n+1} - t_n > t\},$$

the probability that state j is occupied t time units after a decision epoch, provided the next epoch has not occurred yet.

Recall further (cf. §3) that the project continuously accrues holding costs at rate h_j^a per unit time while process $X(t)$ occupies state j and action a prevails.

5.1. Discounted criterion. We start with the (*expected total*) *discounted* criterion, where costs are continuously discounted at rate $\alpha > 0$. Letting $\mathbb{E}_i^\pi[\cdot]$ denote expectation under policy π starting at i, the discounted amount of work expended and the discounted value of costs accrued are given by

$$g_i^{\pi,\alpha} \triangleq \mathbb{E}_i^\pi\left[\int_0^\infty a(t) e^{-\alpha t}\, dt\right] \quad \text{and} \quad f_i^{\pi,\alpha} \triangleq \mathbb{E}_i^\pi\left[\int_0^\infty h_{X(t)}^{a(t)} e^{-\alpha t}\, dt\right].$$

To define appropriate work and cost measures, we will consider the initial state $X(0)$ to be drawn from a *positive probability mass function* $\mathbf{p} = (p_i)_{i\in N}$, i.e., with $p_i > 0$ for $i \in N$. Letting $\mathbb{E}^\pi[\cdot]$ be the corresponding expectation, we will use the *discounted work and cost measures* given, respectively, by

$$f^{\pi,\alpha} \triangleq \mathbb{E}^\pi\left[\int_0^\infty h_{X(t)}^{a(t)} e^{-\alpha t}\, dt\right] = \sum_{i\in N} p_i f_i^{\pi,\alpha} \quad \text{and}$$

$$g^{\pi,\alpha} \triangleq \mathbb{E}^\pi\left[\int_0^\infty a(t) e^{-\alpha t}\, dt\right] = \sum_{i\in N} p_i g_i^{\pi,\alpha}.$$

We require that such measures satisfy Assumption 3.1. Notice that our requirement that $\mathbf{p}$ be strictly positive is motivated by Assumption 3.1(iv).

To analyze the model we will reformulate it into *discrete time* in the standard fashion (cf. Puterman [39, Chapter 11]), using coefficients $\beta_{ij}^{a,\alpha}$ and $\beta_i^{a,\alpha}$ as defined above, and

$$m_i^{a,\alpha} \triangleq \mathbb{E}\left[\int_0^{t_{n+1}-t_n} e^{-\alpha t}\,dt \mid X_n = i,\ a_n = a\right] = \frac{1-\beta_i^{a,\alpha}}{\alpha},$$

$$c_i^{a,\alpha} \triangleq \mathbb{E}\left[\int_0^{t_{n+1}-t_n} h_{X(t)}^a e^{-\alpha t}\,dt \mid X_n = i,\ a_n = a\right].$$

Notice that $m_i^{a,\alpha}$ (resp. $c_i^{a,\alpha}$) is the expected discounted time (resp. cost) accrued over an (i,a)-period, $\beta_i^{a,\alpha}$ is the state- and action-dependent discrete-time discount factor, and $\beta_{ij}^{a,\alpha}$ is the discounted one-period transition probability from state i to j under action a.

We can use the above coefficients to characterize measures $g_i^{S,\alpha}$ and $f_i^{S,\alpha}$, for a given feasible active set $S \in \mathcal{F}$, as the unique solutions to linear equation systems.

Lemma 5.1 (Evaluation Equations). *For every feasible active-state set $S \in \mathcal{F}$:*

(a) *The g_i^S's are characterized by*:
$$\begin{cases} g_i^{S,\alpha} = m_i^{1,\alpha} + \sum_{j\in N} \beta_{ij}^{1,\alpha} g_j^{S,\alpha} & \text{if } i \in S \\ g_i^{S,\alpha} = \sum_{j\in N} \beta_{ij}^{0,\alpha} g_j^{S,\alpha} & \text{if } i \in N\setminus S. \end{cases}$$

(b) *The f_i^S's are characterized by*:
$$\begin{cases} f_i^{S,\alpha} = c_i^{1,\alpha} + \sum_{j\in N} \beta_{ij}^{1,\alpha} f_j^{S,\alpha} & \text{if } i \in S \\ f_i^{S,\alpha} = c_i^{0,\alpha} + \sum_{j\in N} \beta_{ij}^{0,\alpha} f_j^{S,\alpha} & \text{if } i \in N\setminus S. \end{cases}$$

We will use as *state-action frequency measure* $x_j^{a,\pi,\alpha}$ the expected total discounted number of (j,a)-periods spanned under policy π,

$$x_j^{a,\pi,\alpha} \triangleq \mathbb{E}^\pi\left[\sum_{n=0}^{\infty} 1_{\{X_n=j,\,a_n=a\}} e^{-\alpha t_n}\right] = \sum_{i\in N} p_i x_{ij}^{a,\pi,\alpha},$$

where

$$x_{ij}^{a,\pi,\alpha} \triangleq \mathbb{E}_i^\pi\left[\sum_{n=0}^{\infty} 1_{\{X_n=j,\,a_n=a\}} e^{-\alpha t_n}\right].$$

We will use vector (with row or column orientation given implicitly by context) and matrix notation, writing $\mathbf{x}^{a,\pi,\alpha} = (x_j^{a,\pi,\alpha})_{j\in N}$, $\mathbf{g}^{\pi,\alpha} = (g_i^{\pi,\alpha})_{i\in N}$, $\mathbf{f}^{\pi,\alpha} = (f_i^{\pi,\alpha})_{i\in N}$, $\mathbf{m}^{a,\alpha} = (m_j^{a,\alpha})_{j\in N}$, $\mathbf{c}^{a,\alpha} = (c_j^{a,\alpha})_{j\in N}$, and $\mathbf{B}^{a,\alpha} = (\beta_{ij}^{a,\alpha})_{i,j\in N}$; and, for $S, S' \subseteq N$, $\mathbf{B}_{SS'}^{a,\alpha} = (\beta_{ij}^{a,\alpha})_{i\in S,\,j\in S'}$, $\mathbf{f}_S^{\pi,\alpha} = (f_i^{\pi,\alpha})_{i\in S}$. We can thus represent work and cost measures $g^{\pi,\alpha}$ and $f^{\pi,\alpha}$ as linear functions of the frequency measures, by

$$\begin{aligned} g^{\pi,\alpha} &= \mathbf{x}^{1,\pi,\alpha}\mathbf{m}^{1,\alpha} = \sum_{j\in N} m_j^{1,\alpha} x_j^{1,\pi,\alpha} \\ f^{\pi,\alpha} &= \mathbf{x}^{0,\pi,\alpha}\mathbf{c}^{0,\alpha} + \mathbf{x}^{1,\pi,\alpha}\mathbf{c}^{1,\alpha} = \sum_{a\in\{0,1\}} \sum_{j\in N} c_j^{a,\alpha} x_j^{a,\pi,\alpha}, \end{aligned} \tag{29}$$

where it is understood that $x_{\ell^0}^{1,\pi,\alpha} = 0$ (recall that ℓ^0 is the single uncontrollable state).

It is well known in the theory of semi MDPs that frequency measures $x_j^{a,\pi}$ satisfy the system of linear equations given next, which formulate *detailed flow balance* identities. Let $\mathbf{I} = (\delta_{ij})_{i,j\in N}$, where δ_{ij} is Kronecker's delta, be the identity matrix indexed by the project's states.

Lemma 5.2 (Discounted Detailed Flow Balance). *For any admissible policy $\pi \in \Pi$,*

$$\mathbf{x}^{0,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{0,\alpha}) + \mathbf{x}^{1,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{1,\alpha}) = \mathbf{p}.$$

Denote by $\langle a, S\rangle$ the policy that takes action a in the initial period and follows the S-active policy thereafter. For $i \in N^{\{0,1\}}$ and $S \in \mathcal{F}$, we define the *discounted (i,S)-marginal workload* by

$$\begin{aligned} w_i^{S,\alpha} \triangleq g_i^{\langle 1,S\rangle,\alpha} - g_i^{\langle 0,S\rangle,\alpha} &= \begin{cases} g_i^{S,\alpha} - g_i^{\langle 0,S\rangle,\alpha} & \text{if } i \in S \\ g_i^{\langle 1,S\rangle,\alpha} - g_i^{S,\alpha} & \text{if } i \in S^c, \end{cases} \\ &= m_i^{1,\alpha} + \sum_{j\in N} (\beta_{ij}^{1,\alpha} - \beta_{ij}^{0,\alpha}) g_j^{S,\alpha}, \end{aligned} \tag{30}$$

the marginal increase in discounted work expended which results from using policy $\langle 1, S\rangle$ instead of $\langle 0, S\rangle$, starting at i. Recall that $S^c \triangleq N^{\{0,1\}}\setminus S$. We further define the *discounted* (i, S)*-marginal cost* by

$$c_i^{S,\alpha} \triangleq f_i^{\langle 0,S\rangle,\alpha} - f_i^{\langle 1,S\rangle,\alpha} = \begin{cases} f_i^{\langle 0,S\rangle,\alpha} - f_i^{S,\alpha} & \text{if } i \in S \\ f_i^{S,\alpha} - f_i^{\langle 1,S\rangle,\alpha} & \text{if } i \in S^c, \end{cases}$$
$$= c_i^{0,\alpha} - c_i^{1,\alpha} + \sum_{j\in N}\left(\beta_{ij}^{0,\alpha} - \beta_{ij}^{1,\alpha}\right) f_j^{S,\alpha}, \tag{31}$$

the marginal increase in cost incurred which results from adopting policy $\langle 0, S\rangle$ instead of $\langle 1, S\rangle$.

We will need the following result.

Lemma 5.3. *For every feasible active-state set* $S \in \mathcal{F}$:

(a) $\mathbf{w}_S^{S,\alpha} = (\mathbf{I}_{SN} - \mathbf{B}_{SN}^{0,\alpha})\mathbf{g}^{S,\alpha}$ *and* $\mathbf{w}_{S^c}^{S,\alpha} = \mathbf{m}_{S^c}^{1,\alpha} - (\mathbf{I}_{S^cN} - \mathbf{B}_{S^cN}^{1,\alpha})\mathbf{g}^{S,\alpha}$.

(b) $\mathbf{c}_S^{S,\alpha} = \mathbf{c}_S^{0,\alpha} - (\mathbf{I}_{SN} - \mathbf{B}_{SN}^{0,\alpha})\mathbf{f}^{S,\alpha}$ *and* $\mathbf{c}_{S^c}^{S,\alpha} = (\mathbf{I}_{S^cN} - \mathbf{B}_{S^cN}^{1,\alpha})\mathbf{f}^{S,\alpha} - \mathbf{c}_{S^c}^{1,\alpha}$.

Proof. It follows immediately from Lemma 5.1 and (30)–(31). □

Define further discounted aggregate measures $W^{S,a,\pi,\alpha}$, $C^{S,a,\pi,\alpha}$ by (24). We next establish satisfaction of work and cost decomposition laws.

Lemma 5.4. *Under any admissible policy* $\pi \in \Pi$, *the following holds*:

(a) *Workload decomposition laws*:

$$g^{\pi,\alpha} + W^{S,0,\pi,\alpha} = g^{S,\alpha} + W^{S,1,\pi,\alpha}, \quad S \in \mathcal{F}.$$

(b) *Cost decomposition laws*:

$$f^{\pi,\alpha} + C^{S,1,\pi,\alpha} = f^{S,\alpha} + C^{S,0,\pi,\alpha}, \quad S \in \mathcal{F}.$$

Proof. (a) Using Lemma 5.2, Lemma 5.1(a), Lemma 5.3(a), and $x_{\ell^0}^{1,\pi,\alpha} = 0$ gives

$$\begin{aligned} 0 &= [\mathbf{x}^{0,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{0,\alpha}) + \mathbf{x}^{1,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{1,\alpha}) - \mathbf{p}]\mathbf{g}^{S,\alpha} \\ &= \mathbf{x}^{0,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{0,\alpha})\mathbf{g}^{S,\alpha} + \mathbf{x}^{1,\pi,\alpha}[(\mathbf{I} - \mathbf{B}^{1,\alpha})\mathbf{g}^{S,\alpha} - \mathbf{m}^{1,\alpha}] - \mathbf{p}\mathbf{g}^{S,\alpha} + \mathbf{x}^{1,\pi,\alpha}\mathbf{m}^{1,\alpha} \\ &= \mathbf{x}_S^{0,\pi,\alpha}\mathbf{w}_S^{S,\alpha} - \mathbf{x}_{S^c}^{1,\pi,\alpha}\mathbf{w}_{S^c}^{S,\alpha} - g^{S,\alpha} + g^{\pi,\alpha}. \end{aligned}$$

(b) Using Lemma 5.1(b), Lemma 5.2, Lemma 5.3(b), and $x_{\ell^0}^{1,\pi,\alpha} = 0$ gives

$$\begin{aligned} 0 &= [\mathbf{x}^{0,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{0,\alpha}) + \mathbf{x}^{1,\pi,\alpha}(\mathbf{I} - \mathbf{B}^{1,\alpha}) - \mathbf{p}]\mathbf{f}^{S,\alpha} \\ &= \mathbf{x}^{0,\pi,\alpha}[(\mathbf{I} - \mathbf{B}^{0,\alpha})\mathbf{f}^{S,\alpha} - \mathbf{c}^{0,\alpha}] + \mathbf{x}^{1,\pi,\alpha}[(\mathbf{I} - \mathbf{B}^{1,\alpha})\mathbf{f}^{S,\alpha} - \mathbf{c}^{1,\alpha}] \\ &\quad - \mathbf{p}\mathbf{f}^{S,\alpha} + \mathbf{x}^{0,\pi,\alpha}\mathbf{c}^{0,\alpha} + \mathbf{x}^{1,\pi,\alpha}\mathbf{c}^{1,\alpha} \\ &= -\mathbf{x}_S^{0,\pi,\alpha}\mathbf{c}_S^{S,\alpha} + \mathbf{x}_{S^c}^{1,\pi,\alpha}\mathbf{c}_{S^c}^{S,\alpha} - f^{S,\alpha} + f^{\pi,\alpha}. \quad \square \end{aligned}$$

As in (25)–(26), and under Assumption 5.1(i), we define the *discounted* (i, S)*-marginal productivity rate* and the *discounted index* by $\nu_i^{S,\alpha} = c_i^{S,\alpha}/w_i^{S,\alpha}$ and $\nu_i^{*,\alpha} = \nu_i^{S_i,\alpha}$, respectively.

The next result establishes Assumption 4.2(iv).

Lemma 5.5. *Suppose that Assumption* 5.1(i) *holds. Then, for every state* $j \in N^{\{0,1\}}$:

(a) $f_i^{S_j,\alpha} - f_i^{S_{j-1},\alpha} = \nu_j^{*,\alpha}[g_i^{S_{j-1},\alpha} - g_i^{S_j,\alpha}]$, $i \in N$.

(b) $c_i^{S_{j-1},\alpha} - c_i^{S_j,\alpha} = \nu_j^{*,\alpha}[w_i^{S_{j-1},\alpha} - w_i^{S_j,\alpha}]$, $i \in N^{\{0,1\}}$.

Proof. (a) We have, taking $\pi = S_{j-1}$ and $S = S_j$ in Lemma 5.4(a, b):

$$g_i^{S_{j-1},\alpha} = g_i^{S_j,\alpha} + w_j^{S_j,\alpha}x_{ij}^{1,S_{j-1},\alpha}, \quad i \in N,$$
$$f_i^{S_{j-1},\alpha} + c_j^{S_j,\alpha}x_{ij}^{1,S_{j-1},\alpha} = f_i^{S_j,\alpha}, \quad i \in N.$$

Hence,

$$f_i^{S_j,\alpha} - f_i^{S_{j-1},\alpha} = c_j^{S_j,\alpha} x_{ij}^{1,S_{j-1},\alpha} = \frac{c_j^{S_j,\alpha}}{w_j^{S_j,\alpha}} w_j^{S_j,\alpha} x_{ij}^{1,S_{j-1},\alpha} = \frac{c_j^{S_j,\alpha}}{w_j^{S_j,\alpha}} \big[g_i^{S_{j-1},\alpha} - g_i^{S_j,\alpha}\big].$$

(b) Using (31) and part (a) we have that, for $i \in N^{\{0,1\}}$:

$$\begin{aligned} c_i^{S_{j-1},\alpha} - c_i^{S_j,\alpha} &= \sum_{k\in N} \big(\beta_{ik}^{1,\alpha} - \beta_{ik}^{0,\alpha}\big)\big(f_k^{S_j,\alpha} - f_k^{S_{j-1},\alpha}\big) \\ &= \nu_j^{*,\alpha} \sum_{k\in N} \big(\beta_{ik}^{1,\alpha} - \beta_{ik}^{0,\alpha}\big)\big(g_k^{S_{j-1},\alpha} - g_k^{S_j,\alpha}\big) = \nu_j^{*,\alpha}\big(w_i^{S_{j-1},\alpha} - w_i^{S_j,\alpha}\big). \end{aligned}$$

This completes the proof. □

We can now give the main result for the discounted criterion.

Theorem 5.1. *Suppose that marginal workloads $w_i^{S,\alpha}$ and index $\nu_i^{*,\alpha}$ satisfy Assumption 5.1. Then, the project is PCL($\mathscr{F}$)-indexable relative to the discounted criterion, and $\nu_i^{*,\alpha}$ is its discounted MPI.*

Proof. Assumption 5.1(i) and Lemmas 5.4–5.5 ensure satisfaction of Assumption 4.2 and, hence, of part (i) of Definition 4.2, while its part (ii) follows by Assumption 5.1(ii). The proof is completed by invoking Theorem 4.1. □

5.2. Average criterion. We turn next to the (*long-run*) *average* criterion, which we address by drawing on the above results, using a *vanishing discount approach*. In addition to the requirements stated before, we require the following *ergodicity* conditions to hold.

Assumption 5.3.

(i) *For every state $i \in N$, the S_i-active policy induces on the embedded Markov chain X_n the single positive recurrent class $S_i \cup \{i\}$.*

(ii) *Policies in Π are stable, in that there exist finite measures $x_j^{a,\pi}$, f^π, and g^π, independent of the initial state i, given for any admissible policy $\pi \in \Pi$ by*

$$\begin{aligned} x_j^{a,\pi} &\triangleq \lim_{\alpha\searrow 0} \alpha x_j^{a,\pi,\alpha} = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_i^\pi\left[\sum_{k=0}^{n} 1_{\{X_k=j,\,a_k=a\}}\right], \\ f^\pi &\triangleq \lim_{\alpha\searrow 0} \alpha f_i^{\pi,\alpha} = \lim_{t\to\infty} \frac{1}{t} \mathbb{E}_i^\pi\left[\int_0^t h_{X(s)}^{a(s)}\, ds\right], \\ g^\pi &\triangleq \lim_{\alpha\searrow 0} \alpha g_i^{\pi,\alpha} = \lim_{t\to\infty} \frac{1}{t} \mathbb{E}_i^\pi\left[\int_0^t a(s)\, ds\right]. \end{aligned}$$

(iii) *For every controllable state $i \in N^{\{0,1\}}$ and feasible active set $S \in \mathscr{F}$, there exist finite quantities w_i^S and c_i^S given by*

$$\begin{aligned} w_i^S &\triangleq \lim_{\alpha\searrow 0} w_i^{S,\alpha} = \lim_{t\to\infty} \left\{\mathbb{E}_i^{\langle 1,S\rangle}\left[\int_0^t a(s)\, ds\right] - \mathbb{E}_i^{\langle 0,S\rangle}\left[\int_0^t a(s)\, ds\right]\right\}, \\ c_i^S &\triangleq \lim_{\alpha\searrow 0} c_i^{S,\alpha} = \lim_{t\to\infty} \left\{\mathbb{E}_i^{\langle 0,S\rangle}\left[\int_0^t h_{X(s)}^{a(s)}\, ds\right] - \mathbb{E}_i^{\langle 1,S\rangle}\left[\int_0^t h_{X(s)}^{a(s)}\, ds\right]\right\}. \end{aligned}$$

As the notation suggests, we will use f^π, g^π, and $x_j^{a,\pi}$ as the average cost, work, and state-action frequency measures, respectively. We further define the *average* (i,S)*-marginal workload* and the *average* (i,S)*-marginal cost* as the quantities w_i^S and c_i^S in Assumption 5.3(iii), respectively. Thus, w_i^S (resp. c_i^S) represents the expected long-run *cumulative* marginal increase (resp. decrease) in work expended (resp. in holding cost accrued), which results from using policy $\langle 1,S\rangle$ instead of $\langle 0,S\rangle$, starting at state i.

Provided Assumption 5.1(i) holds, we define the *average* (i,S)*-marginal productivity rate* and the *average index* by $\nu_i^S = c_i^S / w_i^S$ and $\nu_i^* = \nu_i^{S_i}$, respectively.

The main result for the average criterion follows.

Theorem 5.2. *Under Assumption 5.1, the project is PCL($\mathscr{F}$)-indexable relative to the average criterion, and ν_i^* is its average MPI.*

Proof. It is readily verified that Assumptions 4.1–4.2 hold. For example, Lemmas 5.4–5.5 immediately yield average counterparts, by taking appropriate limits as α vanishes. Hence, Definition 4.2 holds and, by Theorem 4.1, the result follows. □

5.3. Mixed average-bias criterion. In a variety of models, such as that in §2, the average MPI discussed above does not exist. Such is the case in countable-state models where the average fraction of time that the project must be active is constant, as often occurs in queueing systems. We propose here to overcome such problems by introducing a new, *mixed average-bias criterion*, where cost measures are evaluated under the average criterion, while work measures correspond to Blackwell's [6] more sensitive *bias* criterion. For a review of work on mixed criteria in MDPs (which does not mention the average-bias criterion) see Feinberg and Shwartz [16]. See also Lewis and Puterman [26] for a review of research on the bias criterion. The application of the bias given here is, to the best of our knowledge, new.

In addition to Assumption 5.3, we require the following conditions to hold.

Assumption 5.4.

(i) *There exists a constant $\rho \in (0,1)$ such that, for any policy $\pi \in \Pi$ and initial state $i \in N$,*

$$\rho = \lim_{\alpha \searrow 0} \alpha g_i^{\pi,\alpha} = \lim_{t\to\infty} \frac{1}{t} \mathbb{E}_i^\pi\left[\int_0^t a(s)\,ds\right].$$

(ii) *For any $\pi \in \Pi$ and $i \in N$, there exists a finite g_i^π given by*

$$g_i^\pi \triangleq \lim_{\alpha \searrow 0}\left\{g_i^{\pi,\alpha} - \frac{\rho}{\alpha}\right\} = \mathbb{E}_i^\pi\left[\int_0^\infty (a(t)-\rho)\,dt\right].$$

(iii) *For every $i \in N^{\{0,1\}}$ and $S \in \mathscr{F}$, there exists a finite w_i^S given by*

$$w_i^S \triangleq \lim_{\alpha\searrow 0} \frac{w_i^{S,\alpha}}{\alpha} = \lim_{t\to\infty} t\left\{\mathbb{E}_i^{\langle 1,S\rangle}\left[\int_0^t a(s)\,ds\right] - \mathbb{E}_i^{\langle 0,S\rangle}\left[\int_0^t a(s)\,ds\right]\right\}.$$

We interpret *bias work measure* g_i^π as the expected total cumulative *excess work* expended over the average nominal allocation ρ under policy π, starting at i. As with the discounted criterion in §5.1, we will consider that the initial state $X(0)$ is drawn from a strictly positive probability mass function $\mathbf{p}$, and consider the work measure

$$g^\pi \triangleq \lim_{\alpha\searrow 0}\left\{g^{\pi,\alpha} - \frac{\rho}{\alpha}\right\} = \mathbb{E}^\pi\left[\int_0^\infty (a(t)-\rho)\,dt\right] = \sum_{j\in N} p_j g_j^\pi.$$

Notice that, under Assumption 5.4(iii), the w_i^S's defined for the average criterion in Assumption 5.3(iii) are zero being, hence, of no use in the PCL framework. Instead, we will use the *bias (i,S)-marginal workload w_i^S* defined in Assumption 5.4(iii), which represents the long-run limiting time-scaled marginal increase in expected work expended resulting from using policy $\langle 1,S\rangle$ instead of $\langle 0,S\rangle$, starting at i.

Measures f^π and $x_j^{a,\pi}$ and marginal costs c_i^S are defined according to the average criterion of §5.2. Provided Assumption 5.1(i) holds, we define the *average-bias (i,S)-marginal productivity rate* and the *average-bias index* by $\nu_i^S = c_i^S/w_i^S$ and $\nu_i^* = \nu_i^{S_i} = \nu_i^{S_{i-1}}$.

The main result for the mixed average-bias criterion follows.

Theorem 5.3. *Under Assumption 5.1, the project is PCL($\mathscr{F}$)-indexable relative to the average-bias criterion, and ν_i^* is its average-bias MPI.*

Proof. The result follows along the lines of Theorem 5.2 by taking appropriate limits as the discount factor vanishes in the appropriate results in §5.1 for the discounted criterion. □

6. MPI policies for optimal service control in an MTO queue. This section deploys the above framework to carry out a *PCL-indexability analysis* in the pure MTO case of §2's model. We will thus view the model as a semi-Markov restless bandit project, whose *natural state* $X(t)$ represents the number of customers in the system. The state space is thus $N = \{0,1,\dots\}$, which is partitioned into the controllable and the uncontrollable state spaces $N^{\{0,1\}} = \{1,2,\dots\}$ and $N^{\{0\}} = \{0\}$. Notice that in the uncontrollable state $\ell^0 = 0$ only the passive action $a = 0$ (server idle) is available. The sequence t_n of decision epochs at which we observe the *embedded state* $X_n = X(t_n)$ and take the *embedded action* $a_n = a(t_n)$ consists of all departure epochs *and* all arrival epochs to an idle system. Notice that this differs from the set of epochs used in the conventional *method of the embedded Markov chain*, where the state is only observed at departure instants. We will refer to a *period* between successive decision epochs $[t_n, t_{n+1})$ where $X_n = i$ and $a_n = a$ as an *(i,a)-period*. We will further use the terms *active period* and *passive period*, with the obvious meaning. The discrete-time reformulation discussed in §5, which reduces the original semi MDPs to the *embedded MDP X_n, a_n* will play a central role in our analyses.

In the analyses below we use only elementary properties of the $M/G/1$ queue. See, e.g., Kleinrock [23, Chapter 5] and the references therein. Our main concern will be to establish $\mathscr{F}$-indexability under the average-bias criterion of §5.3, for which we will derive and apply preliminary results for the discounted criterion of §5.1.

6.1. Preliminary results: Discounted criterion. We start with the discounted criterion of §5.1, with discount factor $\alpha > 0$. To reduce notational clutter, we will slightly modify the notation used in §5.1 by making α implicit wherever possible so that, e.g., it will not be included as a superscript. Further, since the active ($\beta_i^{1,\alpha}$) and the passive ($\beta_i^{0,\alpha}$) one-period discount factors do not vary with the state i, we will denote them instead, respectively, by

$$\beta_1 = \psi(\alpha) \qquad \text{and} \qquad \beta_0 = \frac{\lambda}{\alpha+\lambda}.$$

Notice that $\beta_1 = \mathbb{E}[e^{-\alpha\xi}]$ is the LST for a random variable ξ having the distribution $F^1(t)$ of an active period's duration (i.e., service time), while β_0 is the LST for the distribution $F^0(t)$ of a passive period's duration (i.e., interarrival time), which is exponentially distributed with rate λ. Similarly, we will denote the expected discounted durations of active ($m_i^{1,\alpha}$) and passive ($m_i^{0,\alpha}$) periods by

$$m_1 \triangleq \mathbb{E}\left[\int_0^{\xi} e^{-\alpha t}\,dt\right] = \frac{1-\beta_1}{\alpha} \qquad \text{and} \qquad m_0 = \frac{1-\beta_0}{\alpha} = \frac{1}{\alpha+\lambda}.$$

Let (A, ξ) be a random variable pair having the joint distribution of the number of arrivals during a service (A) and the service duration (ξ). Let

$$a_j \triangleq \mathbb{E}[1_{\{A=j\}} e^{-\alpha\xi}]$$

be the discounted probability that j customers arrive during a service, for $j \geq 0$. The a_j's are characterized by their *z-transform* $a^*(z) \triangleq \sum_{j=0}^{\infty} a_j z^j$, as shown next.

Lemma 6.1.

$$a^*(z) = \psi(\alpha+\lambda-\lambda z).$$

Proof. By conditioning on the duration of the service time we can write

$$\begin{aligned} a_j &= \int_0^{\infty} \mathbb{E}[1_{\{A=j\}} e^{-\alpha t} \mid \xi = t]\,dF^1(t) = \int_0^{\infty} \mathbb{P}\{A=j \mid \xi=t\} e^{-\alpha t}\,dF^1(t) \\ &= \int_0^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} e^{-\alpha t}\,dF^1(t) = \mathbb{E}\left[\frac{(\lambda\xi)^j}{j!} e^{-(\alpha+\lambda)\xi}\right]. \end{aligned}$$

Hence, the z-transform of the a_j's is given by

$$a^*(z) = \sum_{j=0}^{\infty} \mathbb{E}\left[\frac{(\lambda z\xi)^j}{j!} e^{-(\alpha+\lambda)\xi}\right] = \mathbb{E}[e^{-(\alpha+\lambda-\lambda z)\xi}] = \psi(\alpha+\lambda-\lambda z). \quad \square$$

Notice that

$$a^*(1) = \beta_1. \tag{32}$$

From the a_j's one can readily obtain the discounted transition probabilities β_{ij}^a (see §5).

We will further use the distribution of the length of a busy period starting with one customer, under the standard, S_0-active policy. Letting

$$T_0 \triangleq \inf\{t > 0\colon X(t^-) \neq 0,\ X(t) = 0\}$$

be the *hitting time to state* 0, the LST of the busy period's duration, $\phi = \phi(\alpha) \triangleq \mathbb{E}_1^{S_0}[e^{-\alpha T_0}]$, is well known to be characterized as the unique root satisfying $\phi \in (0,1)$ of the fixed-point equation

$$a^*(\phi) = \phi. \tag{33}$$

For example, in the exponential-service case, the latter equation becomes

$$\frac{\mu}{\alpha+\mu+\lambda(1-\phi)} = \phi, \quad \text{i.e.,} \quad \lambda\phi^2 - (\alpha+\lambda+\mu)\phi + \mu = 0.$$

Denoting the discriminant by $d = \sqrt{(\alpha+\lambda+\mu)^2 - 4\lambda\mu}$, its two roots are

$$\phi_1 = \frac{\alpha+\lambda+\mu-d}{2\lambda} \qquad \text{and} \qquad \phi_2 = \frac{\alpha+\lambda+\mu+d}{2\lambda}. \tag{34}$$

Since $0 < \phi_1 < 1 < \phi_2$, the appropriate root is $\phi = \phi_1$.

Work and marginal work measures. We address next the calculation and analysis of discounted work and marginal work measures g_i^S and w_i^S, for $S \in \mathscr{F}$. We will draw on the fact that, for each state $i \geq 0$, $X(t)$ is a *regenerative process* under the S_i-active policy, having as renewal epochs the instants at which the least recurrent state i is hit, which mark the completion of an *i-cycle*. We will denote the *hitting time to state i* by

$$T_i \triangleq \inf\{t > 0\colon X(t^-) \neq i,\ X(t) = i\}, \quad i \geq 0.$$

The next result states elementary properties of the $M/G/1$ queue, based on application of the *strong Markov property* to appropriate stopping times, or on translation-invariance identities, on which we will draw for the ensuing analyses.

Lemma 6.2. *For $s, t \geq 0$:*
(a) $\mathbb{P}_{i+k}^{S_0}\{T_k \leq s, T_0 - T_k \leq t\} = \mathbb{P}_i^{S_0}\{T_0 \leq s\}\mathbb{P}_k^{S_0}\{T_0 \leq t\}$, $i, k \geq 1$.
(b) $\mathbb{E}_i^{S_0}[e^{-\alpha T_0}] = \phi^i$, $i \geq 1$.
(c) $\mathbb{P}_0^{S_0}\{T_1 \leq s, T_0 - T_1 \leq t\} = \mathbb{P}_0^{S_0}\{T_1 \leq s\}\mathbb{P}_1^{S_0}\{T_0 \leq t\} = (1 - e^{-\lambda s})\mathbb{P}_1^{S_0}\{T_0 \leq t\}$.
(d) $\mathbb{E}_i^{S_k}[e^{-\alpha T_0}] = \beta_0^{k-i}$, $0 \leq i < k$.
(e) *The following processes have the same distribution, for $i \geq k \geq 0$:*

$$\{X(t) - k\colon t \geq 0\} \text{ under the } S_k\text{-active policy starting at } X(0) = i$$

and

$$\{X(t)\colon t \geq 0\} \text{ under the } S_0\text{-active policy starting at } X(0) = i - k.$$

We will further use the *expected discounted hitting time to* 0 starting at i under the standard policy,

$$M_{i0} \triangleq \mathbb{E}_i^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right], \quad i \geq 0,$$

and the *expected discounted busy time during a* 0*-cycle*,

$$B_{00} \triangleq \mathbb{E}_0^{S_0}\left[\int_{T_1}^{T_0} e^{-\alpha t}\,dt\right].$$

The next result gives useful recursions that characterize the M_{i0}'s and B_{00}.

Lemma 6.3.
(a) *The M_{i0}'s, for $i \geq 1$, are characterized by the recursion*

$$M_{i0} = \begin{cases} \dfrac{1-\phi}{\alpha} & \text{if } i = 1 \\[2ex] M_{10} + \phi M_{i-1,0} & \text{if } i \geq 2, \end{cases}$$

whose solution is $M_{i0} = (1 - \phi^i)/\alpha$. They further satisfy the recursion

$$M_{i+1,0} = M_{i0} + \phi^i M_{10}, \quad i \geq 1.$$

(b) $M_{00} = \dfrac{1 - \beta_0\phi}{\alpha} = \dfrac{\alpha + \lambda - \lambda\phi}{\alpha(\alpha+\lambda)}$.
(c) $B_{00} = \beta_0 M_{10}$.

Proof. (a) We have, by definition of ϕ,

$$M_{10} \triangleq \mathbb{E}_1^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] = \frac{1-\phi}{\alpha}.$$

Further, using Lemma 6.2(a, b), we obtain that, for $i \geq 1$,

$$\begin{aligned} M_{i0} &\triangleq \mathbb{E}_i^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] = \mathbb{E}_i^{S_0}\left[\int_0^{T_{i-1}} e^{-\alpha t}\,dt + e^{-\alpha T_{i-1}}\int_0^{T_0 - T_{i-1}} e^{-\alpha t}\,dt\right] \\ &= \mathbb{E}_1^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] + \mathbb{E}_1^{S_0}[e^{-\alpha T_0}]\mathbb{E}_{i-1}^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] \\ &= M_{10} + \phi M_{i-1,0}, \end{aligned}$$

as required. The solution of such recursion is easily seen to be as stated.

We further have, for $i \geq 1$,

$$\begin{aligned} M_{i+1,0} &= \mathbb{E}_{i+1}^{S_0}\left[\int_0^{T_1} e^{-\alpha t}\,dt + e^{-\alpha T_1}\int_0^{T_0-T_1} e^{-\alpha t}\,dt\right] \\ &= \mathbb{E}_i^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] + \mathbb{E}_i^{S_0}[e^{-\alpha T_0}]\mathbb{E}_1^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] \\ &= M_{i0} + \phi^i M_{10}. \end{aligned}$$

(b) By Lemma 6.2(a, c), we can write

$$\begin{aligned} M_{00} &\triangleq \mathbb{E}_0^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] = \mathbb{E}_0^{S_0}\left[\int_0^{T_1} e^{-\alpha t}\,dt + e^{-\alpha T_1}\int_0^{T_0-T_1} e^{-\alpha t}\,dt\right] \\ &= \mathbb{E}_0^{S_0}\left[\int_0^{T_1} e^{-\alpha t}\,dt\right] + \mathbb{E}_0^{S_0}[e^{-\alpha T_1}]\mathbb{E}_1^{S_0}\left[\int_0^{T_0} e^{-\alpha t}\,dt\right] \\ &= \frac{1-\beta_0}{\alpha} + \beta_0\frac{1-\phi}{\alpha} = \frac{1-\beta_0\phi}{\alpha} = \frac{\alpha+\lambda-\lambda\phi}{\alpha(\alpha+\lambda)}. \end{aligned}$$

Part (c) follows along the same lines using Lemma 6.2(c). □

The next result characterizes discounted work measures $g_i^{S_k}$.

Lemma 6.4. *For* $i \geq 0$:

(a) $g_i^{S_0} = \dfrac{1}{\alpha} - \dfrac{(1-\beta_0)\phi^i}{\alpha(1-\beta_0\phi)} = \dfrac{1}{\alpha} - \dfrac{\phi^i}{\alpha+\lambda-\lambda\phi} = \begin{cases} \dfrac{B_{00}}{\alpha M_{00}} & \text{if } i=0 \\ M_{i0} + \phi^i g_0^{S_0} & \text{if } i \geq 1. \end{cases}$

(b) *For* $k \geq 1$, $g_i^{S_k} = \begin{cases} \beta_0^{k-i} g_0^{S_0} & \text{if } 0 \leq i < k \\ g_{i-k}^{S_0} & \text{if } i \geq k. \end{cases}$

Proof. (a) Using Lemma 6.2(c) and the strong Markov property with respect to T_0, we can write

$$\begin{aligned} g_0^{S_0} &= \mathbb{E}_0^{S_0}\left[\int_0^{\infty} 1_{\{X(t)\geq 1\}}\,dt\right] \\ &= \mathbb{E}_0^{S_0}\left[\int_{T_1}^{T_0} e^{-\alpha t}\,dt + e^{-\alpha T_0}\int_0^{\infty} 1_{\{X(T_0+t)\geq 1\}} e^{-\alpha t}\,dt\right] \\ &= B_{00} + \beta_0\phi g_0^{S_0}. \end{aligned}$$

The latter identity, via Lemma 6.3, gives the stated expressions for $g_0^{S_0}$. Application of the strong Markov property with respect to T_0 and Lemma 6.2(b) further gives the identity $g_i^{S_0} = M_{i0} + \phi^i g_0^{S_0}$ for $i \geq 1$, which is readily reformulated via Lemma 6.3 to yield the stated expressions.

(b) The identity $g_i^{S_k} = \beta_0^{k-i} g_0^{S_0}$, for $0 \leq i < k$, follows from Lemma 6.2(d) and the strong Markov property with respect to T_0. The identity $g_i^{S_k} = g_{i-k}^{S_0}$ for $i \geq k \geq 1$ follows from Lemma 6.2(e). □

The next result gives recursions that characterize discounted marginal work measures w_i^S.

Lemma 6.5. *For* $i, k \geq 1$:

(a) $w_i^{S_0} = (1-\beta_0)M_{i0} = \dfrac{1-\phi^i}{\alpha+\lambda} = \begin{cases} \dfrac{1-\phi}{\alpha+\lambda} & \text{if } i=1 \\ w_1^{S_0} + \phi w_{i-1}^{S_0} & \text{if } i \geq 2. \end{cases}$

(b) $w_i^{S_k} = \begin{cases} w_1^{S_{k-i+1}} & \text{if } 2 \leq i \leq k \\ w_{i-k}^{S_0} & \text{if } i > k. \end{cases}$

(c) $w_1^{S_k} = \begin{cases} \dfrac{a_0}{\phi} w_1^{S_0} & \text{if } k=1 \\ (1-\beta_0)m_1 + \beta_0 w_1^{S_{k-1}} + \displaystyle\sum_{j=k+1}^{\infty} a_j w_{j-k}^{S_0} & \text{if } k \geq 2. \end{cases}$

Proof. (a) Using (30), Lemma 6.4(a) and Lemma 6.3(a) gives, for $i \geq 1$:

$$\begin{aligned} w_i^{S_0} &\triangleq g_i^{\langle 1, S_0\rangle} - g_i^{\langle 0, S_0\rangle} = g_i^{S_0} - \beta_0 g_{i+1}^{S_0} = \big[M_{i0} + \phi^i g_0^{S_0}\big] - \beta_0\big[M_{i+1,0} + \phi^{i+1} g_0^{S_0}\big] \\ &= \big[M_{i0} + \phi^i g_0^{S_0}\big] - \beta_0\big[M_{i0} + \phi^i M_{10} + \phi^{i+1} g_0^{S_0}\big] \\ &= (1-\beta_0) M_{i0} + \phi^i\big[(1-\beta_0\phi) g_0^{S_0} - \beta_0 M_{10}\big] = (1-\beta_0) M_{i0}, \end{aligned}$$

from which the remaining identities follow.

(b) This part follows from (30) and Lemma 6.4(b).

(c) For $k = 1$, we can write

$$\begin{aligned} w_1^{S_1} &\triangleq g_1^{\langle 1, S_1\rangle} - g_1^{\langle 0, S_1\rangle} = m_1 + \sum_{j=0}^{\infty} a_j g_j^{S_1} - g_1^{S_1} = m_1 + a_0\beta_0 g_0^{S_0} + \sum_{j=1}^{\infty} a_j g_{j-1}^{S_0} - g_0^{S_0} \\ &= m_1 + a_0\beta_0 g_0^{S_0} + \sum_{j=1}^{\infty} a_j\left[\frac{1}{\alpha} - \left(\frac{1}{\alpha} - g_0^{S_0}\right)\phi^{j-1}\right] - g_0^{S_0} \\ &= m_1 + a_0\beta_0 g_0^{S_0} + \frac{1}{\alpha}\sum_{j=1}^{\infty} a_j - \frac{1}{\phi}\left(\frac{1}{\alpha} - g_0^{S_0}\right)\sum_{j=1}^{\infty} a_j\phi^j - g_0^{S_0} \\ &= m_1 + a_0\beta_0 g_0^{S_0} + \frac{1}{\alpha}[\beta_1 - a_0] - \frac{1}{\phi}\left(\frac{1}{\alpha} - g_0^{S_0}\right)(\phi - a_0) - g_0^{S_0} \\ &= \frac{a_0}{\phi}\left[\frac{1-\phi}{\alpha} - (1-\beta_0\phi) g_0^{S_0}\right] = \frac{a_0}{\phi} w_1^{S_0}, \end{aligned}$$

where we have used Lemma 5.1, Lemma 6.4, Lemma 6.3(a), (32), (33), and $m_1 = (1-\beta_1)/\alpha$.

Let now $k \geq 2$. We can write

$$\begin{aligned} w_1^{S_k} &\triangleq g_1^{\langle 1, S_k\rangle} - g_1^{\langle 0, S_k\rangle} = m_1 + \sum_{j=0}^{\infty} a_j g_j^{S_k} - g_1^{S_k} \\ &= m_1 + \sum_{j=0}^{k-1} a_j g_j^{S_k} + \sum_{j=k}^{\infty} a_j g_j^{S_k} - g_1^{S_k} = m_1 + \beta_0 \sum_{j=0}^{k-1} a_j g_j^{S_{k-1}} + \sum_{j=k}^{\infty} a_j g_{j-1}^{S_{k-1}} - g_0^{S_{k-1}} \\ &= m_1 + \beta_0\left[w_1^{S_{k-1}} - m_1 - \sum_{j=k}^{\infty} a_j g_j^{S_{k-1}} + g_1^{S_{k-1}}\right] + \sum_{j=k}^{\infty} a_j g_{j-1}^{S_{k-1}} - g_0^{S_{k-1}} \\ &= (1-\beta_0) m_1 + \beta_0 w_1^{S_{k-1}} + \sum_{j=k}^{\infty} a_j\big[g_{j-1}^{S_{k-1}} - \beta_0 g_j^{S_{k-1}}\big] \\ &= (1-\beta_0) m_1 + \beta_0 w_1^{S_{k-1}} + \sum_{j=k+1}^{\infty} a_j w_{j-k}^{S_0}, \end{aligned}$$

where we have used Lemma 5.1, Lemma 6.4(b), and part (b). This completes the proof. □

Lemma 6.5 immediately yields a complete recursion for calculating all discounted marginal workloads $w_i^{S_k}$, for $i \geq 1$, $k \geq 0$. The calculation's backbone consists of *pivot terms* $w_1^{S_k}$, for $k \geq 0$. Calculations proceed in the order indicated by the arrows in Figure 4.

The next result shows that discounted marginal workloads are strictly positive, as required.

Proposition 6.1. *Discounted marginal workloads $w_i^{S_k}$ satisfy Assumption* 5.1(a).

Proof. The result follows immediately by induction on $k \geq 0$ using the recursions in Lemma 6.5. □

Cost and marginal cost measures. We proceed to calculate the required discounted cost and marginal cost measures f_i^S and c_i^S. In the analyses below, we will include in the notation the *shifted* holding cost rate sequence $\mathbf{h}^l$, for $l \geq 0$, relative to which cost measures are defined, with $\mathbf{h} = \mathbf{h}^0 \triangleq (h_0, h_1, \ldots)$. See §2 for definitions of $\mathbf{h}^l$, $\Delta\mathbf{h}^l$, and $\Delta^2\mathbf{h}^l$. We will thus write, e.g., $f_i^S(\mathbf{h}^l)$, which is obtained by replacing the original cost sequence $\mathbf{h}$ by $\mathbf{h}^l$, i.e., by considering that the cost rate incurred in state $X(t) = j$ is h_{j+l}.

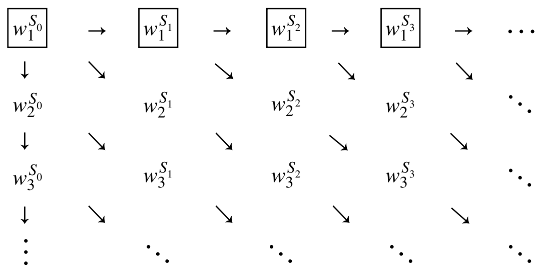

FIGURE 4. Recursive calculation of marginal workloads $w_i^{S_k}$ via Lemma 6.5.

Analogously as in the preceding analyses, we will make use of the *expected discounted cost accrued to hit state* 0 *starting at* $i \geq 0$ under the standard, S_0-active policy, defined by

$$V_{i0}(\mathbf{h}^l) \triangleq \mathbb{E}_i^{S_0}\left[\int_0^{T_0} h_{X(t)+l} e^{-\alpha t}\, dt\right]$$

Notice that $V_{00}(\mathbf{h}^l)$ is the *expected discounted cost accrued over a* 0*-cycle*.

The next result gives recursions that characterize $V_{i0}(\cdot)$ in terms of $V_{10}(\cdot)$.

LEMMA 6.6.

(a) *The* $V_{i0}(\mathbf{h}^l)$*'s, for* $i \geq 2$, *are characterized by the recursion*

$$V_{i0}(\mathbf{h}^l) = V_{10}(\mathbf{h}^{l+i-1}) + \phi V_{i-1,0}(\mathbf{h}^l),$$

whose solution is

$$V_{i0}(\mathbf{h}^l) = V_{10}(\mathbf{h}^{l+i-1} + \phi \mathbf{h}^{l+i-2} + \cdots + \phi^{i-1}\mathbf{h}^l).$$

They further satisfy the recursion

$$V_{i+1,0}(\mathbf{h}^l) = V_{i0}(\mathbf{h}^{l+1}) + \phi^i V_{10}(\mathbf{h}^l), \quad i \geq 1.$$

(b) $V_{00}(\mathbf{h}^l) = h_l m_0 + \beta_0 V_{10}(\mathbf{h}^l)$.

PROOF. (a) Arguing along the same lines as in the proof of Lemma 6.3(a), we can write, for $i \geq 2$,

$$\begin{aligned}
V_{i0}(\mathbf{h}^l) &\triangleq \mathbb{E}_i^{S_0}\left[\int_0^{T_0} h_{X(t)+l} e^{-\alpha t}\, dt\right] \\
&= \mathbb{E}_i^{S_0}\left[\int_0^{T_{i-1}} h_{X(t)+l} e^{-\alpha t}\, dt + e^{-\alpha T_{i-1}} \int_0^{T_0 - T_{i-1}} h_{X(T_{i-1}+t)+l} e^{-\alpha t}\, dt\right] \\
&= \mathbb{E}_1^{S_0}[h_{X(t)+i+l-1} e^{-\alpha t} dt] + \mathbb{E}_1^{S_0}[e^{-\alpha T_0}] \mathbb{E}_{i-1}^{S_0}[h_{X(t)+l} e^{-\alpha t} dt] \\
&= V_{10}(\mathbf{h}^{l+i-1}) + \phi V_{i-1,0}(\mathbf{h}^l)
\end{aligned}$$

as required. The solution of such recursion is easily seen to be as stated.

We further have, for $i \geq 1$,

$$\begin{aligned}
V_{i+1,0}(\mathbf{h}^l) &= \mathbb{E}_{i+1}^{S_0}\left[\int_0^{T_1} h_{X(t)+l} e^{-\alpha t}\, dt + e^{-\alpha T_1} \int_0^{T_0 - T_1} h_{X(T_1+t)+l} e^{-\alpha t}\, dt\right] \\
&= \mathbb{E}_{i+1}^{S_0}\left[\int_0^{T_1} h_{X(t)+l} e^{-\alpha t}\, dt\right] + \mathbb{E}_{i+1}^{S_0}[e^{-\alpha T_1}] \mathbb{E}_{i+1}^{S_0}\left[\int_0^{T_0 - T_1} h_{X(T_1+t)+l} e^{-\alpha t}\, dt\right] \\
&= \mathbb{E}_i^{S_0}\left[\int_0^{T_0} h_{X(t)+l+1} e^{-\alpha t}\, dt\right] + \mathbb{E}_i^{S_0}[e^{-\alpha T_0}] \mathbb{E}_1^{S_0}\left[\int_0^{T_0} h_{X(t)+l} e^{-\alpha t}\, dt\right] \\
&= V_{i0}(\mathbf{h}^{l+1}) + \phi^i V_{10}(\mathbf{h}^l).
\end{aligned}$$

(b) Arguing along the same lines as in the proof of Lemma 6.3(a), we can write

$$\begin{aligned} V_{00}(\mathbf{h}^l) &\triangleq \mathbb{E}_0^{S_0}\left[\int_0^{T_0} h_{X(t)+l} e^{-\alpha t}\,dt\right] = \mathbb{E}_0^{S_0}\left[\int_0^{T_1} h_{X(t)+l} e^{-\alpha t}\,dt + e^{-\alpha T_1}\int_0^{T_0-T_1} h_{X(t)+l} e^{-\alpha t}\,dt\right] \\ &= \mathbb{E}_0^{S_0}\left[\int_0^{T_1} h_{X(t)+l} e^{-\alpha t}\,dt\right] + \mathbb{E}_0^{S_0}[e^{-\alpha T_1}]\mathbb{E}_1^{S_0}\left[\int_0^{T_0} h_{X(T_1+t)+l} e^{-\alpha t}\,dt\right] \\ &= h_l m_0 + \beta_0 V_{10}(\mathbf{h}^l). \quad \square \end{aligned}$$

The following result gives recursions that characterize discounted cost measures $f_i^S(\mathbf{h}^l)$.

Lemma 6.7.

(a) $f_i^{S_0}(\mathbf{h}^l) = \begin{cases} \dfrac{V_{00}(\mathbf{h}^l)}{\alpha M_{00}} & \textit{if } i = 0 \\ V_{i0}(\mathbf{h}^l) + \phi^i f_0^{S_0}(\mathbf{h}^l) & \textit{if } i \geq 1. \end{cases}$

(b) *For* $k \geq 1$, $f_i^{S_k}(\mathbf{h}^l) = \begin{cases} (h_{i+l} + \cdots + \beta_0^{k-i-1} h_{k+l-1}) m_0 + \beta_0^{k-i} f_0^{S_0}(\mathbf{h}^{k+l}) & \textit{if } 0 \leq i < k \\ f_{i-k}^{S_0}(\mathbf{h}^{k+l}) & \textit{if } i \geq k. \end{cases}$

Proof. (a) Arguing analogously as in the proof of Lemma 6.4(a), we can write

$$\begin{aligned} f_0^{S_0}(\mathbf{h}^l) &= h_l m_0 + \beta_0 f_1^{S_0}(\mathbf{h}^l) \\ f_i^{S_0}(\mathbf{h}^l) &= V_{i0}(\mathbf{h}^l) + \phi^i f_0^{S_0}(\mathbf{h}^l), \quad i \geq 1. \end{aligned}$$

Solving for $f_0^{S_0}(\mathbf{h}^l)$ and $f_1^{S_0}(\mathbf{h}^l)$ gives

$$\begin{aligned} f_0^{S_0}(\mathbf{h}^l) &= h_l \frac{m_0}{1-\beta_0\phi} + \beta_0 \frac{V_{10}^{S_0}(\mathbf{h}^l)}{1-\beta_0\phi} \\ f_1^{S_0}(\mathbf{h}^l) &= h_l \frac{\phi m_0}{1-\beta_0\phi} + \frac{V_{10}^{S_0}(\mathbf{h}^l)}{1-\beta_0\phi}, \end{aligned}$$

whence the result follows, using Lemmas 6.3(b) and 6.6(b).

Part (b) follows along the same lines as Lemma 6.4(b). $\square$

The next result characterizes the required discounted marginal costs $c_i^S(\mathbf{h}^l)$.

Lemma 6.8.

(a) $c_i^{S_0}(\mathbf{h}^l) = (h_{i+l} - \phi^i h_l) m_0 + \beta_0 V_{i0}(\Delta\mathbf{h}^{l+1}) - (1-\beta_0) V_{i0}(\mathbf{h}^l)$, *for* $i \geq 1$.

(b) $c_i^{S_k}(\mathbf{h}^l) = c_{i-k}^{S_0}(\mathbf{h}^{k+l})$, *for* $i > k$.

Proof. (a) We have, for $i \geq 1$,

$$\begin{aligned} c_i^{S_0}(\mathbf{h}^l) &\triangleq f_i^{\langle 0, S_0\rangle}(\mathbf{h}^l) - f_i^{\langle 1, S_0\rangle}(\mathbf{h}^l) = h_{i+l} m_0 + \beta_0 f_{i+1}^{S_0}(\mathbf{h}^l) - f_i^{S_0}(\mathbf{h}^l) \\ &= h_{i+l} m_0 + \beta_0\big[V_{i+1,0}(\mathbf{h}^l) + \phi^{i+1} f_0^{S_0}(\mathbf{h}^l)\big] - \big[V_{i0}(\mathbf{h}^l) + \phi^i f_0^{S_0}(\mathbf{h}^l)\big] \\ &= (h_{i+l} - \phi^i h_l) m_0 + \beta_0 V_{i0}(\Delta\mathbf{h}^{l+1}) - (1-\beta_0) V_{i0}(\mathbf{h}^l), \end{aligned}$$

where we have used (31), Lemma 6.7(a), identity $V_{i+1,0}(\mathbf{h}^l) = V_{i0}(\mathbf{h}^{l+1}) + \phi^i V_{10}(\mathbf{h}^l)$ in Lemma 6.6(a), and the identities (cf. the proof of Lemma 6.7(a))

$$f_0^{S_0}(\mathbf{h}^l) = h_l m_0 + \beta_0 f_1^{S_0}(\mathbf{h}^l) = h_l m_0 + \beta_0\big[V_{10}(\mathbf{h}^l) + \phi f_0^{S_0}(\mathbf{h}^l)\big].$$

Part (b) follows from (31) and Lemma 6.7(b). $\square$

The next result gives probabilistic expressions for discounted index $\nu_i^*(\mathbf{h}) \triangleq c_i^{S_{i-1}}(\mathbf{h})/w_i^{S_{i-1}}$.

Proposition 6.2. *For* $i \geq 1$:

$$\begin{aligned} \nu_i^*(\mathbf{h}) &= \nu_1^*(\mathbf{h}^{i-1}) \\ &= \left(\lambda + \frac{\alpha}{1-\phi}\right)\mathbb{E}_0^{S_0}\left[\int_0^\infty e^{-\alpha t}\left\{\Delta h_{X(t)+i} - \frac{\alpha}{\lambda}(h_{X(t)+i-1} - h_{i-1})\right\} dt\right] \\ &= \left(\lambda + \frac{\alpha}{1-\phi}\right)\left\{\mathbb{E}_0^{S_0}\left[\int_0^\infty e^{-\alpha t} h_{X(t)+i}\,dt\right] - \mathbb{E}_1^{S_0}\left[\int_0^\infty e^{-\alpha t} h_{X(t)+i-1}\,dt\right]\right\}. \end{aligned}$$

PROOF. The first identity follows from Lemmas 6.5(b) and 6.8(b). The second one follows from

$$\begin{aligned}\nu_1^*(\mathbf{h}^{i-1}) &\triangleq \frac{c_1^{S_0}(\mathbf{h}^{i-1})}{w_1^{S_0}} = \frac{(h_i-\phi h_{i-1})m_0+\beta_0 V_{10}(\Delta\mathbf{h}^i)-(1-\beta_0)V_{10}(\mathbf{h}^{i-1})}{(1-\beta_0)M_{10}}\\
&= \frac{V_{00}(\Delta\mathbf{h}^i)+h_{i-1}(1-\phi)m_0-(1-\beta_0)V_{10}(\mathbf{h}^{i-1})}{(1-\beta_0)M_{10}}\\
&= \frac{M_{00}}{m_0M_{10}}\frac{V_{00}(\Delta\mathbf{h}^i)+h_{i-1}(1-\phi)m_0-\alpha m_0V_{10}(\mathbf{h}^{i-1})}{\alpha M_{00}}\\
&= \frac{M_{00}}{m_0M_{10}}\frac{V_{00}(\Delta\mathbf{h}^i)-\alpha m_0V_{10}(\mathbf{h}^{i-1}-h_{i-1}\mathbf{1})}{\alpha M_{00}}\\
&= \frac{M_{00}}{m_0M_{10}}\left\{\frac{V_{00}(\Delta\mathbf{h}^i)}{\alpha M_{00}}-\frac{1-\beta_0}{\beta_0}\frac{V_{00}(\mathbf{h}^{i-1}-h_{i-1}\mathbf{1})}{\alpha M_{00}}\right\}\\
&= \frac{M_{00}}{m_0M_{10}}f_0^{S_0}\left(\Delta\mathbf{h}^i-\frac{\alpha}{\lambda}(\mathbf{h}^{i-1}-h_{i-1}\mathbf{1})\right)\\
&= \frac{\alpha+\lambda(1-\phi)}{1-\phi}\mathbb{E}_0^{S_0}\left[\int_0^\infty e^{-\alpha t}\left\{\Delta h_{X(t)+i}-\frac{\alpha}{\lambda}(h_{X(t)+i-1}-h_{i-1})\right\}dt\right]\end{aligned}$$

where $\mathbf{1}$ denotes a sequence of 1's, and we have used Lemmas 6.5(a), 6.8(a), 6.6(b), 6.7(a), and 6.3(a, b), and the identities $m_0=(1-\beta_0)/\alpha$, $\beta_0=\lambda/(\alpha+\lambda)$,

$$V_{10}(h_{i-1}\mathbf{1})=h_{i-1}M_{10}=h_{i-1}\frac{1-\phi}{\alpha} \qquad \text{and} \qquad V_{00}(\mathbf{h}^{i-1}-h_{i-1}\mathbf{1})=\beta_0V_{10}(\mathbf{h}^{i-1}-h_{i-1}\mathbf{1}).$$

The third identity follows from

$$\nu_i^*(\mathbf{h})=\frac{f_i^{S_i}(\mathbf{h})-f_i^{S_{i-1}}(\mathbf{h})}{g_i^{S_{i-1}}-g_i^{S_i}}=\frac{f_0^{S_0}(\mathbf{h}^i)-f_1^{S_0}(\mathbf{h}^{i-1})}{g_1^{S_0}-g_0^{S_0}}=\left(\lambda+\frac{\alpha}{1-\phi}\right)\left[f_0^{S_0}(\mathbf{h}^i)-f_1^{S_0}(\mathbf{h}^{i-1})\right],$$

where we have used Lemmas 5.5(a), 6.4(b), 6.7(b), and 6.4(a). □

REMARK 6.1.

(i) If one can show that index $\nu_i^*(\mathbf{h})$ is nondecreasing in i under Assumption 2.1, then Theorem 5.1 (using Proposition 6.1) shows that it is indeed the discounted MPI. We emphasize that such result is *not needed* for our analysis in §6.2 of indexability under the average-bias criterion.

(ii) In particular, the second identity in Proposition 6.2 gives that, if the condition

$$\Delta^2\mathbf{h}^{i+1}\ge\frac{\alpha}{\lambda}[\Delta^2\mathbf{h}^i+\Delta h_i\mathbf{1}], \quad i\ge 2,$$

holds, then index $\nu_i^*(\mathbf{h})$ is nondecreasing in i and is thus the discounted MPI.

(iii) In the *linear* holding cost case $h_j=cj$, $j\ge 0$, i.e., $\mathbf{h}=c\mathbf{e}$, we can use Bell's [2] classic accounting argument to write

$$\alpha f_i^{S_0}(c\mathbf{e})=c\left[i+\frac{\lambda}{\alpha}-\alpha\frac{\beta_1}{1-\beta_1}g_i^{S_0}\right], \quad i\ge 0.$$

Substitution in the third identity in Proposition 6.2 gives the *constant* MPI

$$\nu_i^*(c\mathbf{e})=c\frac{\beta_1}{1-\beta_1}, \quad i\ge 1, \tag{35}$$

which agrees with the discount-optimal index rule for scheduling a multiclass $M/G/1$ queue.

(iv) We can use Proposition 6.2 to define the following vanishing-discount limiting index

$$\nu_i^{\text{average}}(\mathbf{h})\triangleq\lim_{\alpha\searrow 0}\alpha\nu_i^*(\mathbf{h})=\mu\mathbb{E}^{S_0}[\Delta h_{X+i}],$$

where X has the steady-state distribution of process $X(t)$ under policy S_0. It seems reasonable to consider $\nu_i^{\text{average}}(\mathbf{h})$ as an index appropriate to the average criterion. In §6.2 we will show that $\nu_i^{\text{average}}(\mathbf{h})$ is indeed the MPI corresponding to an appropriate (average-bias) criterion.

In the exponential-service case, we obtain simpler evaluations of discounted index $\nu_i^*(\mathbf{h})$, whence its monotonicity follows. Let $\tau \sim \text{Exp}(\alpha)$ be a random stopping time, independent of the state process $X(t)$. We draw on a classic result on the $M/M/1$ queue stating that, for $X(0) = 0$, $X(\tau)$ (under policy S_0) has a geometric distribution with success probability $1 - \rho\phi$.

THEOREM 6.1. *Suppose the service time distribution is exponential and Assumption 2.1 holds. Then, the MTO model is PCL($\mathscr{F}$)-indexable under the discounted criterion, and its discounted MPI is given by*

$$\nu_i^*(\mathbf{h}) = \mu \mathbb{E}_0^{S_0}\left[\int_0^\infty e^{-\alpha t} \Delta h_{X(t)+i}\, dt\right] = \frac{\mu}{\alpha}\mathbb{E}_0^{S_0}[\Delta h_{X(\tau)+i}] = \frac{\mu}{\alpha}(1-\rho\phi)\sum_{j=0}^\infty \Delta h_{j+i}(\rho\phi)^j, \quad i \geq 1.$$

PROOF. Let T_0 be the hitting time to 0 as before. Drawing on elementary properties of the $M/M/1$ queue, including the strong Markov property with respect to T_0, we can write, for $i \geq 1$,

$$\begin{aligned}\mathbb{E}_1^{S_0}[h_{X(\tau)+i-1}] &= \mathbb{E}_1^{S_0}[h_{X(\tau)+i-1} \mid T_0 > \tau]\mathbb{P}_1^{S_0}\{T_0 > \tau\} + \mathbb{E}_1^{S_0}[h_{X(\tau)+i-1} \mid T_0 \leq \tau]\mathbb{P}_1^{S_0}\{T_0 \leq \tau\} \\ &= \mathbb{E}_0^{S_0}[h_{X(\tau)+i}]\mathbb{P}_1^{S_0}\{T_0 > \tau\} + \mathbb{E}_0^{S_0}[h_{X(\tau)+i-1}]\mathbb{P}_1^{S_0}\{T_0 \leq \tau\}.\end{aligned}$$

It follows that

$$\mathbb{E}_0^{S_0}[h_{X(\tau)+i}] - \mathbb{E}_1^{S_0}[h_{X(\tau)+i-1}] = \mathbb{P}_1^{S_0}\{T_0 \leq \tau\}\mathbb{E}_0^{S_0}[\Delta h_{X(\tau)+i}] = \phi\mathbb{E}_0^{S_0}[\Delta h_{X(\tau)+i}].$$

Now, substitution of the above into the last identity for $\nu_i^*(\mathbf{h})$ in Proposition 6.2, gives, after straightforward simplification, the stated identities. It is now immediate that the index is nondecreasing in i under Assumption 2.1. Therefore, by Theorem 5.1, $\nu_i^*(\mathbf{h})$ is the discounted MPI, which completes the proof. □

REMARK 6.2. In the exponential-service case:

(i) It is easily seen that the α-scaled MPI $\alpha\nu_i^*(\mathbf{h})$ is decreasing in α. We can thus define, in addition to index $\nu_i^{\text{average}}(\mathbf{h})$ in Remark 6.1(iv), the limiting *myopic index*

$$\nu_i^{\text{myopic}}(\mathbf{h}) \triangleq \lim_{\alpha \nearrow \infty} \alpha\nu_i^*(\mathbf{h}) = \mu\Delta h_i, \quad i \geq 1.$$

(ii) Under a quadratic cost rate $h_j = cj^2$ we obtain the discounted MPI

$$\nu_i^*(\mathbf{h}) = \frac{c\mu}{\alpha}\left[2i - 1 + \frac{2\rho\phi}{1-\rho\phi}\right], \quad i \geq 1. \tag{36}$$

The corresponding average and myopic indices are

$$\nu_i^{\text{average}}(\mathbf{h}) = c\mu\left[2i - 1 + \frac{2\rho}{1-\rho}\right] \quad \text{and} \quad \nu_i^{\text{myopic}}(\mathbf{h}) = c\mu[2i-1].$$

6.2. Average-bias criterion. We proceed to address indexability under the average-bias criterion of §5.3, drawing on the analyses above under the discounted criterion. For clarity, when we use below the discounted measures in §6.1, we will incorporate discount factor α in the notation.

The next result characterizes *bias work measures* $g_i^{S_k}$ and $g^{S_k} \triangleq \sum_{i=0}^\infty p_i g_i^{S_k}$. We will make use of the 2nd-order Maclaurin series expansion of the length of a busy period's LST as $\alpha \searrow 0$, given by

$$\phi(\alpha) \triangleq \mathbb{E}_1^{S_0}[e^{-\alpha T_0}] = 1 - \alpha\mathbb{E}_1^{S_0}[T_0] + \frac{\alpha^2}{2}\mathbb{E}_1^{S_0}[T_0^2] + o(\alpha^2). \tag{37}$$

We readily obtain from (37) and the moments of the $M/G/1$ busy period (cf. Kleinrock [23, Chapter 5]) that

$$\phi'(0) = -\mathbb{E}_1^{S_0}[T_0] = -\frac{1/\mu}{1-\rho} \quad \text{and} \quad \phi''(0) = \mathbb{E}_1^{S_0}[T_0^2] = \frac{1/\mu^2 + \sigma^2}{(1-\rho)^3}. \tag{38}$$

LEMMA 6.9.

(a) $g_i^{S_0} = -\dfrac{\lambda}{2}\dfrac{1/\mu^2 + \sigma^2}{1-\rho} + \dfrac{i}{\mu}$, *for* $i \geq 0$.

(b) $g_i^{S_k} = g_0^{S_0} + \dfrac{i-k}{\mu}$, *for* $k \geq 1$, $i \geq 0$.

(c) $g^{S_k} = g_0^{S_0} + \dfrac{\mathbb{E}[X(0)] - k}{\mu}$, *for* $k \geq 0$.

Proof. (a) We can write

$$\begin{aligned} g_i^{S_0} &\triangleq \lim_{\alpha\searrow 0} g_i^{S_0,\alpha} - \frac{\rho}{\alpha} = \lim_{\alpha\searrow 0} \frac{1-\rho}{\alpha} - \frac{\phi^i(\alpha)}{\alpha+\lambda-\lambda\phi(\alpha)} = \lim_{\alpha\searrow 0} \frac{(1-\rho)[\alpha+\lambda-\lambda\phi(\alpha)]-\alpha\phi^i(\alpha)}{\alpha[\alpha+\lambda-\lambda\phi(\alpha)]} \\ &= \lim_{\alpha\searrow 0} \frac{(1-\rho)[1-\lambda\phi'(\alpha)]-\phi^i(\alpha)-i\alpha\phi^{i-1}(\alpha)\phi'(\alpha)}{\alpha+\lambda-\lambda\phi(\alpha)+\alpha[1-\lambda\phi'(\alpha)]} \\ &= \lim_{\alpha\searrow 0} \frac{-(1-\rho)\lambda\phi''(\alpha)-2i\phi^{i-1}(\alpha)\phi'(\alpha)-i\alpha\frac{d}{d\alpha}[\phi^{i-1}(\alpha)\phi'(\alpha)]}{1-\lambda\phi'(\alpha)+1-\lambda\phi'(\alpha)-\lambda\alpha\phi''(\alpha)} \\ &= \frac{-(1-\rho)\lambda\phi''(0)-2i\phi'(0)}{2[1-\lambda\phi'(0)]} = -\frac{\lambda}{2}\frac{1/\mu^2+\sigma^2}{1-\rho}+\frac{i}{\mu}, \end{aligned}$$

where we have used the definition of $g_i^{S_0}$ in Assumption 5.4(ii), Lemma 6.4(a), and have further applied twice L'Hôpital's rule using (38) to simplify the resulting expressions.

(b) The stated identities follow from the corresponding identities in Lemma 6.4(b), arguing along the lines of part (a)'s proof.

(c) This part follows immediately from parts (a, b) and $g^{S_k} \triangleq \sum_{i=0}^{\infty} p_i g_i^{S_k}$. □

The next result characterizes *bias marginal workloads* $w_i^{S_k}$. Notice that below a_j denotes the *undiscounted* probability that j customers arrive during a service.

Lemma 6.10. *For $i, k \geq 1$:*

(a) $$w_i^{S_0} = \frac{M_{i0}}{\lambda} = \frac{1/\lambda}{1-\rho}\frac{i}{\mu} = \begin{cases} \dfrac{1/\lambda}{1-\rho}\dfrac{1}{\mu} + w_{i-1}^{S_0} & \text{if } i \geq 2 \\ \dfrac{1/\lambda}{1-\rho}\dfrac{1}{\mu} & \text{if } i = 1. \end{cases}$$

(b) $$w_i^{S_k} = \begin{cases} w_{i-k}^{S_0} & \text{if } i > k \\ w_1^{S_{k-i+1}} & \text{if } 2 \leq i \leq k. \end{cases}$$

(c) $$w_1^{S_k} = \begin{cases} a_0 w_1^{S_0} & \text{if } k = 1 \\ \dfrac{1}{\lambda\mu} + w_1^{S_{k-1}} + \displaystyle\sum_{j=k+1}^{\infty} a_j w_{j-k}^{S_0} & \text{if } k \geq 2. \end{cases}$$

Proof. To obtain the stated identities it suffices to divide by discount factor $\alpha > 0$ the corresponding identities in Lemma 6.5, and take limits as α vanishes. □

Remark 6.3. As in the discounted case, Lemma 6.5 immediately yields a complete recursion for calculating all bias marginal workloads $w_i^{S_k}$, which proceeds as indicated in Figure 4.

The next result establishes the required properties of bias marginal workloads.

Proposition 6.3. *Bias marginal workloads $w_i^{S_k}$ satisfy the following properties*:

(a) *They are positive, i.e., Assumption 5.1 holds.*

(b) *They are wedge-shaped, satisfying Assumption 4.3 with strict inequalities, i.e., for $i \geq 1$,*

$$w_i^{S_0} > \cdots > w_i^{S_{i-1}} > w_i^{S_i} < w_i^{S_{i+1}} < \cdots.$$

Proof. (a) The result follows by induction using Lemma 6.10. See Remark 6.3.

(b) We have, using Lemma 6.10(a, b) that, for $1 \leq k \leq i-1$:

$$w_i^{S_k} - w_i^{S_{k-1}} = w_{i-k}^{S_0} - w_{i-k+1}^{S_0} = -\frac{1}{\lambda\mu(1-\rho)} < 0.$$

Further, using Lemma 6.10(b, c) gives that, for $k \geq 1$:

$$w_k^{S_k} - w_k^{S_{k-1}} = w_1^{S_1} - w_1^{S_0} = -(1-a_0)w_1^{S_0} < 0.$$

Finally, part (a) and Lemma 6.10(b, c) give that, for $k \geq i+1$,

$$w_i^{S_k} - w_i^{S_{k-1}} = w_1^{S_{k-i+1}} - w_1^{S_{k-i}} = \frac{1}{\lambda\mu} + \sum_{j=k-i+2} a_j w_{j-k+i-1}^{S_0} > 0.$$

This completes the proof. □

We next address calculation of average cost measures $f^S(\mathbf{h}^l)$, for $l \geq 0$ (see §5.2). In what follows we will denote by X a random variable having the *equilibrium distribution* of the number in system process $X(t)$ for the $M/G/1$ queue under the prevailing policy.

Lemma 6.11.

$$f^{S_k}(\mathbf{h}^l) = f^{S_0}(\mathbf{h}^{k+l}) = \mathbb{E}^{S_0}[h_{X+k+l}], \quad k \geq 0.$$

Proof. The result follows from the chain of identities

$$f^{S_k}(\mathbf{h}^l) \triangleq \lim_{\alpha \searrow 0} \alpha f_k^{S_k,\alpha}(\mathbf{h}^l) = \lim_{\alpha \searrow 0} \alpha f_0^{S_0,\alpha}(\mathbf{h}^{k+l}) = f^{S_0}(\mathbf{h}^{k+l}) = \mathbb{E}^{S_0}[h_{X+k+l}],$$

where we have used Assumption 5.3(ii) and Lemma 6.7(b). □

The next result characterizes the required average marginal costs $c_i^S(\mathbf{h}^l)$. The terms $V_{10}(\cdot)$ below are the undiscounted counterparts of corresponding namesake terms in §6.1.

Lemma 6.12.
(a) $c_i^{S_0}(\mathbf{h}^l) = (h_{i+l} - h_l)/\lambda + V_{i0}(\Delta \mathbf{h}^{l+1})$, $i \geq 1$.
(b) $c_i^{S_k}(\mathbf{h}^l) = c_{i-k}^{S_0}(\mathbf{h}^{k+l})$, $i > k$.

Proof. The result follows by taking limits as α vanishes in Lemma 6.8's identities. □

The following result gives representations for average-bias index $\nu_i^*(\mathbf{h}) \triangleq c_i^{S_{i-1}}(\mathbf{h})/w_i^{S_{i-1}}$.

Proposition 6.4.

$$\nu_i^*(\mathbf{h}) = \nu_1^*(\mathbf{h}^{i-1}) = \frac{\Delta h_i/\lambda + V_{10}(\Delta \mathbf{h}^i)}{w_1^{S_0}} = \mu \mathbb{E}^{S_0}[\Delta h_{X+i}], \quad i \geq 1.$$

Proof. Scale by α the identities in Proposition 6.2 and let α vanish to get the result. □

We can finally present the main result of this section.

Theorem 6.2. *If holding cost rates satisfy Assumption 2.1, then*:
(a) *The MTO model is PCL($\mathcal{F}$)-indexable under the average-bias criterion, having MPI $\nu_i^*(\mathbf{h})$.*
(b) *MPI $\nu_i^*(\mathbf{h})$ has the characterization in Theorem 4.3, i.e.,*

$$\max_{0 \leq k < i} \nu_i^{S_k}(\mathbf{h}) = \nu_i^*(\mathbf{h}) = \min_{k \geq i} \nu_i^{S_k}(\mathbf{h}) \quad i \geq 1.$$

Proof. (a) This part follows from Theorem 5.3, using Proposition 6.3(a) to ensure positivity of bias marginal workloads, and identity $\nu_i^*(\mathbf{h}) = \mu \mathbb{E}^{S_0}[\Delta h_{X+i}]$ in Proposition 6.4 to ensure nondecreasingness of the index $\nu_i^*(\mathbf{h})$.
(b) This part follows from Theorem 4.3 using Proposition 6.3(b). □

Remark 6.4.
(i) For a *linear cost rate* $h_j = cj$, the average-bias MPI is indeed $\nu_i^*(\mathbf{h}) = c\mu$, consistently with the average-optimal *$c\mu$-rule* for scheduling a multiclass $M/G/1$ queue.
(ii) For a *quadratic cost rate* $h_j = cj^2$, the average-bias MPI has the evaluation

$$\nu_i^*(\mathbf{h}) = c\mu\left[2i - 1 + \frac{2\rho + \lambda^2(\sigma^2 - 1/\mu^2)}{1-\rho}\right], \quad i \geq 1.$$

(iii) Denoting by $\nu_i^{*,m}(\mathbf{h})$ the average-bias MPI under cost rates $h_j = j^m$, $j \geq 0$, where $m \geq 1$ is an integer, we readily obtain the evaluation

$$\nu_i^{*,m}(\mathbf{h}) = \mu \sum_{k=0}^{m-1} \binom{m}{k} \{i^{m-k} - (i-1)^{m-k}\} \mathbb{E}^{S_0}[X^k], \quad i \geq 1.$$

Notice that if, e.g., $h_j = c_0 + c_1 j + c_2 j^2$, then $\nu_i^*(\mathbf{h}) = c_1 \nu_i^{*,1}(\{j\}) + c_2 \nu_i^{*,2}(\{j^2\})$.

(iv) Suppose costs are given customerwise, by a polynomial cost rate $h(T)$ on the current delay T accrued by a customer. To use the above results, we can obtain an *equivalent holding cost rate* h_j by drawing on, e.g., Marshall and Wolff's [28] higher-order extensions of Little's law.

7. MPI policies for optimal production control in an MTS queue. We next address the MTS case with backorders of §2's model, where there is a finished goods stock capable of holding up to and including $s \geq 1$ units, by reducing its analysis to the pure MTO case discussed above (along the lines of Morse's [29, Chapter 10] classic analysis; see also Buzacott and Shanthikumar [7, Chapter 4]). The key observation is that if $X(t)$ is the *net backorder process* defined in §2, then the shifted state process $L(t) \triangleq X(t) + s \geq 0$ evolves as the number-in-system in a corresponding MTO $M/G/1$ queue. Thus, analysis of process $X(t)$ starting at $X(0) = i$ under the S_k-active policy (work when $X(t) > k$) reduces to that of process $L(t)$ starting at $L(0) = i + s$ under the S_{k+s}-active policy (work when $L(t) > k + s$), for $i, k \geq -s$. Hence, each result in §6 for the pure MTO model translates into a corresponding result for the MTS model.

The next result, which follows immediately from such observation, states the resulting relations between bias work measures and between average cost measures for the pure MTO and the MTS cases (indicated by superscripts). The same relations hold for discounted measures. Notice that below $\mathbf{h} = \mathbf{h}^0 = (h_{-s}, h_{-s+1}, \ldots)$. To clarify, notice further that, e.g., the notation $f_i^{\mathrm{MTS}, S_k}(\mathbf{h}^l)$ refers to an MTS queue where the holding cost rate incurred in state $X(t) = j \geq -s$ is h_{j+l}, whereas $f_{i+s}^{\mathrm{MTO}, S_{k+s}}(\mathbf{h}^l)$ refers to an MTO queue where the holding cost rate incurred in state $L(t) = j + s \geq 0$ is h_{j+l}, for $l \geq 0$.

Lemma 7.1.
(a) $g_i^{\mathrm{MTS}, S_k} = g_{i+s}^{\mathrm{MTO}, S_{k+s}}$, *for* $i, k \geq -s$.
(b) $f_i^{\mathrm{MTS}, S_k}(\mathbf{h}^l) = f_{i+s}^{\mathrm{MTO}, S_{k+s}}(\mathbf{h}^l)$, *for* $i, k \geq -s,\ l \geq 0$.
(c) $w_i^{\mathrm{MTS}, S_k} = w_{i+s}^{\mathrm{MTO}, S_{k+s}}$, *for* $i \geq -s+1,\ k \geq -s$.
(d) $c_i^{\mathrm{MTS}, S_k}(\mathbf{h}^l) = c_{i+s}^{\mathrm{MTO}, S_{k+s}}(\mathbf{h}^l)$, *for* $i \geq -s+1,\ k \geq -s,\ l \geq 0$.

The next result is the MTS counterpart of Theorem 6.2. It shows that calculation of MTS MPIs reduces to that of MTO MPIs. Notice that $\mathbf{h}^s \triangleq (h_0, h_1, \ldots)$.

Theorem 7.1. *If holding cost rates satisfy Assumption* 2.1, *then*:
(a) *The MTS model is PCL*($\mathscr{F}$)*-indexable under the average-bias criterion, having MPI*

$$\begin{aligned}\nu_i^{\mathrm{MTS},*}(\mathbf{h}) &= \nu_{i+s}^{\mathrm{MTO},*}(\mathbf{h}) = \nu_1^{\mathrm{MTO},*}(\mathbf{h}^{i+s-1}) = \mu \mathbb{E}^{S_0}[\Delta h_{X+i}] \\ &= \begin{cases} \nu_i^{\mathrm{MTO},*}(\mathbf{h}^s) & \text{if } i \geq 1 \\ \nu_1^{\mathrm{MTO},*}(\mathbf{h}^s)\mathbb{P}^{S_0}\{X > -i\} + \mu \sum_{j=0}^{-i} \Delta h_{i+j} \mathbb{P}^{S_0}\{X = j\} & \text{if } -s < i \leq 0. \end{cases}\end{aligned} \tag{39}$$

(b) *MPI* $\nu_i^{\mathrm{MTS},*}(\mathbf{h})$ *has the characterization in Theorem* 4.3, *i.e.*,

$$\max_{-s \leq k < i} \nu_i^{\mathrm{MTS}, S_k}(\mathbf{h}) = \nu_i^{\mathrm{MTS},*}(\mathbf{h}) = \min_{k \geq i} \nu_i^{\mathrm{MTS}, S_k}(\mathbf{h}), \quad i \geq -s+1.$$

Proof. (a) This part follows from Theorem 6.2(a), noting first that we can write, for every controllable state $i \geq -s+1$,

$$\begin{aligned}\nu_i^{\mathrm{MTS},*}(\mathbf{h}) &\triangleq \frac{c_i^{\mathrm{MTS}, S_{i-1}}(\mathbf{h})}{w_i^{\mathrm{MTS}, S_{i-1}}} = \frac{c_{i+s}^{\mathrm{MTO}, S_{i+s-1}}(\mathbf{h})}{w_{i+s}^{\mathrm{MTO}, S_{i+s-1}}} = \nu_{i+s}^{\mathrm{MTO}, S_{i+s-1}}(\mathbf{h}) \triangleq \nu_{i+s}^{\mathrm{MTO},*}(\mathbf{h}) \\ &= \nu_1^{\mathrm{MTO},*}(\mathbf{h}^{i+s-1}) = \mu \mathbb{E}^{S_0}[\Delta h_{X+i}],\end{aligned}$$

where we have used the definition of $\nu_i^{\mathrm{MTS},*}(\mathbf{h})$, Lemma 7.1(c, d), the definition of $\nu_{i+s}^{\mathrm{MTO},*}(\mathbf{h})$, $\mathbf{h}^{i+s-1} \triangleq (h_{i-1}, h_i, \ldots)$ and Proposition 6.4. We now observe that, if $i \geq 1$, we can write, by Proposition 6.4,

$$\mu \mathbb{E}^{S_0}[\Delta h_{X+i}] = \nu_i^{\mathrm{MTO},*}(\mathbf{h}^s).$$

The second case in (39) follows by noting that, for $i \leq 0$, we can write

$$\begin{aligned}\mathbb{E}^{S_0}[\Delta h_{X+i}] &= \mathbb{E}^{S_0}[\Delta h_{X+i} \mid X > -i]\mathbb{P}^{S_0}\{X > -i\} + \sum_{j=0}^{-i} \mathbb{E}^{S_0}[\Delta h_{X+i} \mid X = j]\mathbb{P}^{S_0}\{X = j\} \\ &= \mathbb{E}^{S_0}[\Delta h_{X+1}]\mathbb{P}^{S_0}\{X > -i\} + \sum_{j=0}^{-i} \Delta h_{i+j} \mathbb{P}^{S_0}\{X = j\},\end{aligned}$$

where we have drawn on the following intuitive result: Under the S_0-active policy (for MTO process $X(t)$ under the S_0-active policy), it is clear that over any open time interval (t_1, t_2) of the sample path where $X(t) > -i$ for $t_1 < t < t_2$, the shifted process $Q(t) \triangleq X(t) + i - 1 \geq 0$ evolves as the number-in-system of an MTO $M/G/1$ queue. Hence, we obtain that in steady state:

$$\mathbb{E}^{S_0}[\Delta h_{X+i} \mid X > -i] = \mathbb{E}^{S_0}[\Delta h_{Q+1}] = \mathbb{E}^{S_0}[\Delta h_{X+1}].$$

Part (b) follows directly from Theorem 6.2(b), which completes the proof. □

While our focus is on the average-bias criterion, it is insightful to consider indexability under the discounted criterion in the exponential service case. From the above discussion and Theorem 6.1 we readily obtain the next result (where we include discount factor $\alpha > 0$ in the notation). Let ϕ and τ be as in Theorem 6.1.

THEOREM 7.2. *Suppose that service times are exponential and holding cost rates satisfy Assumption 2.1. Then, the MTS model is PCL($\mathscr{F}$)-indexable under the discounted criterion, having MPI*

$$\begin{aligned} \nu_i^{\mathrm{MTS},*,\alpha}(\mathbf{h}) &= \nu_{i+s}^{\mathrm{MTO},*,\alpha}(\mathbf{h}) = \nu_1^{\mathrm{MTO},*,\alpha}(\mathbf{h}^{i+s-1}) = \mu \mathbb{E}_0^{S_0}\left[\int_0^\infty e^{-\alpha t} h_{X(t)+i}\, dt\right] = \frac{\mu}{\alpha}\mathbb{E}_0^{S_0}[\Delta h_{X(\tau)+i}] \\ &= \begin{cases} \nu_i^{\mathrm{MTO},*,\alpha}(\mathbf{h}^s) & \text{if } i \geq 1 \\ \nu_1^{\mathrm{MTO},*,\alpha}(\mathbf{h}^s)(\rho\phi)^{-i+1} + \dfrac{\mu(1-\rho\phi)}{\alpha}\displaystyle\sum_{j=0}^{-i} \Delta h_{i+j}(\rho\phi)^j & \text{if } -s < i \leq 0. \end{cases} \end{aligned} \tag{40}$$

REMARK 7.1.

(i) For $i \geq 1$, $\nu_i^{\mathrm{MTO},*}(\mathbf{h}^s) = \nu_i^{\mathrm{MTO},*}(h_0, h_1, \ldots)$ is the bias-average MPI for the pure MTO model obtained in the previous section. Thus, the first case in (39) and in (40) shows that *the MTS MPI extends the MTO MPI*. This agrees with a result of Ha [20] on the structure of optimal policies in the linear cost exponential service case.

(ii) In the case of *linear stock holding costs* $h_j = -c^{\mathrm{F}} j$, for $j \leq -1$, we obtain the MPI evaluation

$$\nu_i^{\mathrm{MTS},*}(\mathbf{h}) = \begin{cases} \nu_i^{\mathrm{MTO},*}(\mathbf{h}^s) & \text{if } i \geq 1 \\ \left[\nu_1^{\mathrm{MTO},*}(\mathbf{h}^s) + c^{\mathrm{F}}\mu\right]\mathbb{P}^{S_0}\{X \geq -i+1\} - c^{\mathrm{F}}\mu & \text{if } -s < i \leq 0. \end{cases} \tag{41}$$

(iii) If backorder holding costs are also linear, so that $h_j = c^{\mathrm{B}} j$ for $j \geq 0$, and $h_j = -c^{\mathrm{F}} j$ for $j \leq -1$, where $c^{\mathrm{B}}, c^{\mathrm{F}} > 0$, we obtain the MPI evaluation

$$\nu_i^{\mathrm{MTS},*}(\mathbf{h}) = \begin{cases} c^{\mathrm{B}}\mu & \text{if } i \geq 1 \\ [(c^{\mathrm{B}} + c^{\mathrm{F}})\mathbb{P}^{S_0}\{X \geq -i+1\} - c^{\mathrm{F}}]\mu & \text{if } -s < i \leq 0. \end{cases} \tag{42}$$

(iv) In the $M/M/1$ case, substituting $\mathbb{P}^{S_0}\{X \geq j\} = \rho^j$ in (42) gives

$$\nu_i^{\mathrm{MTS},*}(\mathbf{h}) = \begin{cases} c^{\mathrm{B}}\mu & \text{if } i \geq 1 \\ [(c^{\mathrm{B}} + c^{\mathrm{F}})\rho^{-i+1} - c^{\mathrm{F}}]\mu & \text{if } -s < i \leq 0. \end{cases}$$

The latter turns out to be precisely the *myopic*(T) *index* of Peña-Pérez and Zipkin [38, p. 926], which they obtained through a look-ahead argument. It has been shown by De Véricourt et al. [12] that such index partly characterizes the optimal policy for scheduling a multiclass $M/M/1$ MTS queue. It is insightful to further evaluate the myopic index in Remark 6.2(i):

$$\nu_i^{\mathrm{myopic}}(\mathbf{h}) \triangleq \lim_{\alpha \nearrow \infty} \alpha \nu_i^{\mathrm{MTS},*,\alpha}(\mathbf{h}) = \mu \Delta h_i = \begin{cases} c^{\mathrm{B}}\mu & \text{if } i \geq 1 \\ -c^{\mathrm{F}}\mu & \text{if } -s < i \leq 0. \end{cases}$$

Hence, the myopic index is precisely the index obtained by Wein [44] via a heavy-traffic analysis.

Acknowledgments. The author thanks the anonymous referees for their careful reading and constructive comments that helped improve the paper's exposition. This research has been supported in part by the Spanish Ministry of Science and Technology/Education and Science under Grants BEC2000-1027 and MTM2004-02334, and a Ramón y Cajal Investigator Award; a NATO Collaborative Linkage Grant PST.CLG.976568; and the Joint Spanish-USA (Fulbright) Commission for Scientific and Technological Cooperation Grant 2000-20132.

References

[1] Ansell, P. S., K. D. Glazebrook, J. Niño-Mora, M. O'Keeffe. 2003. Whittle's index policy for a multi-class queueing system with convex holding costs. *Math. Methods Oper. Res.* **57** 21–39.
[2] Bell, C. E. 1971. Characterization and computation of optimal policies for operating an $M/G/1$ queue with removable server. *Oper. Res.* **19** 208–218.
[3] Bertsekas, D. P. 1987. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, Englewood Cliffs, NJ.
[4] Bertsimas, D., J. Niño-Mora. 1996. Conservation laws, extended polymatroids and multiarmed bandit problems; A polyhedral approach to indexable systems. *Math. Oper. Res.* **21** 257–306.
[5] Bertsimas, D., J. Niño-Mora. 2000. Restless bandits, linear programming relaxations, and a primal-dual index heuristic. *Oper. Res.* **48** 80–90.
[6] Blackwell, D. 1962. Discrete dynamic programming. *Ann. Math. Statist.* **33** 719–726.
[7] Buzacott, J. A., J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Prentice-Hall, Englewood Cliffs, NJ.
[8] Coffman, E. G., Jr., I. Mitrani. 1980. A characterization of waiting time performance realizable by single-server queues. *Oper. Res.* **28** 810–821.
[9] Cox, D. R., W. L. Smith. 1961. *Queues*. Methuen, London, UK.
[10] Dacre, M., K. Glazebrook, J. Niño-Mora. 1999. The achievable region approach to the optimal control of stochastic systems. *J. Roy. Statist. Soc., Ser. B, Statist. Methodology* **61** 747–791.
[11] D'Epénoux, F. 1960. Sur un problème de production et de stockage dans l'aléatoire. *RAIRO Rech. Opér.* **14** 3–16 [English translation. 1963. A probabilistic production and inventory problem. *Management Sci.* **10** 98–108].
[12] De Véricourt, F., F. Karaesmen, Y. Dallery. 2000. Dynamic scheduling in a make-to-stock system: A partial characterization of optimal policies. *Oper. Res.* **48** 811–819.
[13] Dusonchet, F., M.-O. Hongler. 2003. Continuous-time restless bandit and dynamic scheduling for make-to-stock production. *IEEE Trans. Robotics Automation* **19** 977–990.
[14] Edmonds, J. 1971. Matroids and the greedy algorithm. *Math. Programming* **1** 127–136.
[15] Federgruen, A., H. Groenevelt. 1988. Characterization and optimization of achievable performance in general queueing systems. *Oper. Res.* **36** 733–741.
[16] Feinberg, E. A., A. Shwartz. 2002. Mixed criteria. E. A. Feinberg, A. Shwartz, eds. *Handbook of Markov Decision Processes: Methods and Applications*. Kluwer, Boston, MA, 209–230.
[17] Gittins, J. C. 1979. Bandit processes and dynamic allocation indices. *J. Roy. Statist. Soc., Ser. B* **41** 148–177.
[18] Gittins, J. C., D. M. Jones. 1974. A dynamic allocation index for the sequential design of experiments. *Progress in Statistics, European Meeting of Statisticians, Budapest, 1972*. North-Holland, Amsterdam, The Netherlands, 241–266; *Colloquia Math. Soc. János Bolyai*, Vol. 9.
[19] Glazebrook, K. D., J. Niño-Mora. 2001. Parallel scheduling of multiclass $M/M/m$ queues: Approximate and heavy-traffic optimization of achievable performance. *Oper. Res.* **49** 609–623.
[20] Ha, A. Y. 1997. Optimal dynamic scheduling policy for a make-to-stock production system. *Oper. Res.* **45** 42–53.
[21] Kantorovich, L. V. 1959. *Ekonomicheskii Raschet Nailuchshego Ispolzovania Resursov*. Academy of Sciences USSR [English translation. 1965. *The Best Use of Economic Resources*. Harvard University Press, Cambridge, MA].
[22] Kleinrock, L. 1965. A conservation law for a wide class of queueing disciplines. *Naval Res. Logist. Quart.* **12** 181–192.
[23] Kleinrock, L. 1975. *Queueing Systems*, Vol. I: Theory. Wiley, New York.
[24] Klimov, G. P. 1974. Time-sharing service systems. I. *Theory Probab. Appl.* **19** 532–551.
[25] Koopmans, T. C. 1957. *Three Essays on the State of Economic Science*, Chapter on Allocation of resources and the price system. McGraw-Hill, New York.
[26] Lewis, M. E., M. L. Puterman. 2002. Bias optimality. E. Feinberg, A. Schwartz, eds. *Handbook of Markov Decision Processes*. Kluwer, Boston, MA, 89–111.
[27] Manne, A. S. 1960. Linear programming and sequential decisions. *Management Sci.* **6** 259–267.
[28] Marshall, K. T., R. W. Wolff. 1971. Customer average and time average queue lengths and waiting times. *J. Appl. Probab.* **8** 535–542.
[29] Morse, P. M. 1958. *Queues, Inventories and Maintenance: The Analysis of Operational Systems with Variable Demand and Supply*. Wiley, New York.
[30] Niño-Mora, J. 2000. Restless bandit index heuristics for multiclass queueing network control. Presentation *Conf. Stochastic Networks*, University of Wisconsin-Madison, Madison, WI, June 19–24.
[31] Niño-Mora, J. 2001. Restless bandits: PCL-indexability conditions and queueing control applications. Presentation, *27th Conf. Stochastic Processes and Their Applications*, Cambridge, UK, July 9–13.
[32] Niño-Mora, J. 2001. Restless bandits, partial conservation laws and indexability. *Adv. Appl. Probab.* **33** 76–98.
[33] Niño-Mora, J. 2001. Stochastic scheduling. C. A. Floudas, P. M. Pardalos, eds. *Encyclopedia of Optimization*, Vol. 5. Kluwer, Dordrecht, The Netherlands, 367–372.
[34] Niño-Mora, J. 2002. Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Math. Programming* **93** 361–413.
[35] Niño-Mora, J. 2003. Restless bandit marginal productivity indices, diminishing returns and scheduling a multiclass make-to-order/-stock queue. R. Srikant, V. V. Veeravalli, eds. *Proc. 41st Annual Allerton Conf. Comm., Control Comput.*, Monticello, IL, 100–109.
[36] Niño-Mora, J. 2004. Restless bandit dynamic allocation indices II: Multi-project case and dynamic scheduling of a multiclass make-to-order/-stock $M/G/1$ queue. Working Paper 04-09, Department of Statistics, Universidad Carlos III de Madrid, Spain, February.
[37] Papadimitriou, C. H., J. N. Tsitsiklis. 1999. The complexity of optimal queuing network control. *Math. Oper. Res.* **24** 293–305.
[38] Peña-Pérez, A., P. Zipkin. 1997. Dynamic scheduling rules for a multiproduct make-to-stock queue. *Oper. Res.* **45** 919–930.
[39] Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, New York.
[40] Ross, S. M. 1970. Average cost semi-Markov decision processes. *J. Appl. Probab.* **7** 649–656.
[41] Shanthikumar, J. G., D. D. Yao. 1992. Multiclass queueing systems: Polymatroidal structure and optimal scheduling control. *Oper. Res.* **40** S293–S299.

[42] Tsoucas, P. 1991. The region of achievable performance in a model of Klimov. Tech. Report RC16543, IBM T. J. Watson Research Center, Yorktown Heights, New York.

[43] Veatch, M. H., L. M. Wein. 1996. Scheduling a multiclass make-to-stock queue: Index policies and hedging points. *Oper. Res.* **44** 634–647.

[44] Wein, L. M. 1992. Dynamic scheduling of a multiclass make-to-stock queue. *Oper. Res.* **40** 724–735.

[45] Weiss, G. 1988. Branching bandit processes. *Probab. Engrg. Inform. Sci.* **2** 269–278.

[46] Weitzman, M. L. 2000. An "economics proof" of the supporting hyperplane theorem. *Econom. Lett.* **68** 1–6.

[47] Whittle, P. 1988. Restless bandits: Activity allocation in a changing world. J. Gani, ed. *A Celebration of Applied Probability, J. Appl. Probab. Special*, Vol. 25A. Applied Probability Trust, Sheffield, UK, 287–298.

[48] Whittle, P. 1996. *Optimal Control: Basics and Beyond*. Wiley, Chichester, UK.