Wharton

Financial Institutions Center

Managing Learning and Turnover in Employee Staffing

by
Noah Gans
Yong-Pin Zhou

99-39

The Wharton School
University of Pennsylvania

THE WHARTON FINANCIAL INSTITUTIONS CENTER

The Wharton Financial Institutions Center provides a multi-disciplinary research approach to the problems and opportunities facing the financial services industry in its search for competitive excellence. The Center's research focuses on the issues related to managing risk at the firm level as well as ways to improve productivity and performance.

The Center fosters the development of a community of faculty, visiting scholars and Ph.D. candidates whose research interests complement and support the mission of the Center. The Center works closely with industry executives and practitioners to ensure that its research is informed by the operating realities and competitive demands facing industry participants as they pursue competitive excellence.

A variety of working papers related to this subject are also available online at:

http://fic.wharton.upenn.edu/fic

If you would like to learn more about the Center or become a member of our research community, please let us know of your interest.

Anthony M. Santomero
Director

Managing Learning and Turnover in Employee Staffing[*]

Noah Gans   Yong-Pin Zhou

Operations and Information Management Department

The Wharton School, University of Pennsylvania

Philadelphia, PA 19104-6366

July 27, 1999

Abstract

We study the employee staffing problem in a service organization that uses employee service capacities to meet random, non-stationary service requirements. The employees experience learning and turnover on the job, and we develop a Markov Decision Process (MDP) model that explicitly represents the stochastic nature of these effects. Theoretical results are developed that show the optimal hiring policy is of a state-dependent "hire-up-to" type, similar to the inventory "order-up-to" policy. This holds for discounted-cost MDP's under both finite and infinite planning horizons.

We also develop structural properties of the optimal policy to facilitate computation of the optimal hiring numbers. For two important special cases of the general model, we prove the optimality of a myopic policy under both stationary and stochastically increasing service requirements. Moreover, we show that in these two cases, when service requirements are k-periodic, it is sufficient to solve a k-period MDP problem with appropriate end-of-horizon cost function. When general, non-stationary service requirements are present, we prove the existence of a one-sided "smoothing effect" of the optimal hire-up-to levels.

Numerical results show that the use of state-dependent hire-up-to policies may offer significant cost savings over simpler hiring policies. In particular, our results show that when employee capacity increase due to learning is substantial and flexible incremental capacity (overtime) is tight, a fully state-dependent policy out-performs a policy that hires only on the basis of the total number of employees in the system.

Our problem formulation and results suggest natural connections to the classic results in inventory literature. We also discuss many of the connections and distinctions in the paper.

1   Introduction

We consider the employee staffing problem at a service organization. Suppose there are service requirements the organization must meet, and a forecast of these requirements is available for several periods of time in the future. The forecast may be based on historical data and a projection of the organization's business path. The organization's objective is to use the capacity provided by its employees to meet the service requirements in a least cost fashion.

The essential factor in the problem that differentiates it from more traditional capacity planning problems is that the people providing these services are not homogeneous. Different employees will have different service capacities (or skill levels), and these service capacities change over time. When people learn on the job, their service capacities increase; when they turn over their service capacities are lost. Learning can take place because of either initial training or on-the-job training. Turnover can happen because of a mismatch with the job or a better career opportunity elsewhere.

Learning and turnover are typically random elements. Therefore the numbers of employees in the organization – and the need to hire additional workers – can be difficult to predict. In addition, the initial training period may be long. These factors make hiring to meet future service requirements a difficult problem to solve effectively.

1.1   Our Approach

We take a hierarchical approach to this problem that mirrors the hierarchical planning strategies used in a manufacturing environment. A long-term, high-level staffing problem corresponds to long-term factory capacity planning. A medium-term, mid-level work-force scheduling problem corresponds to the mid-level planning for which many companies use Materials Requirements Planning (MRP). Finally there exists a moment-by-moment, low-level work-assignment problem that is the analogue of real-time shop-floor control in a manufacturing firm.

Example: a Telephone Call Center   Figure 1 is a hypothetical example of a telephone call center call volume forecast. We have about six months' data. The service standards in these call centers are something like: "on average a phone call is put on hold for 20 seconds or less" or "85% of the time a phone call is put on hold for 20 seconds or less."

Typically, the problem is disaggregated into short-term, medium-term, and long-term planning
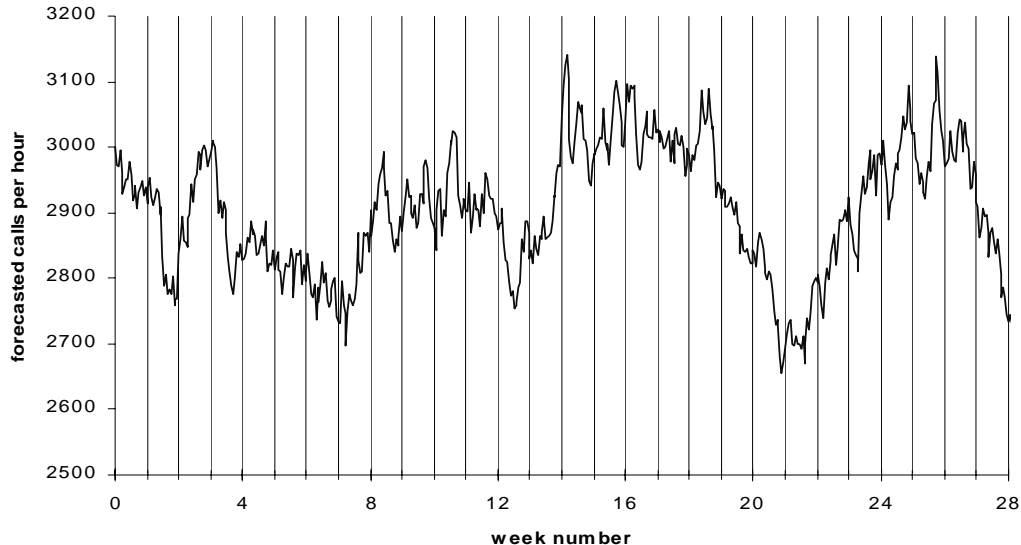
Figure 1: Call Center Example

components. In the short run, the call center must solve a real-time control problem, assigning incoming calls to available Customer Service Representatives (CSR's). The medium-term problem, staff scheduling, is typically solved on a weekly basis. In any given week, the numbers of employees of different types is fixed, and a schedule is developed for those employees that minimizes overtime and outsourcing costs, subject to the call center's service level requirements. ("Outsourcing" refers to a common practice in which one call center diverts incoming calls to another on a contract basis. In addition to paying a long-term retainer, the diverting call center typically pays a per-call fee for this service.) The long-term component seeks to hire the right numbers of people and train them the right way so that, when the sequence of weekly scheduling problems is solved, they produce a low-cost – if not least-cost – solution to the global problem.

In this paper, we will focus on the long-range hiring problem. It is formulated as a discrete-time, continuous-state-space Markov Decision Process (MDP), where the state variable vector represents the numbers of people at different levels on the learning curve. This approach allows us to model naturally the randomness in the system and to prove the existence of desirable structural properties

2

of the optimal hiring policy.

To suppress the other two lower-level problems, we will assume that there exists an "operating cost" function, $O(\cdot)$, which, given the numbers of employees at different skill levels and a forecast for service requirements, will tell us how much the total operating costs will be. This operating cost function should be based on an efficient, if not optimal, solution to the scheduling problem.

For example, in call centers, managers usually use commercial software to do workforce scheduling each week. This software typically solves a large-scale mixed integer program, and it returns a schedule which shows the amount of overtime and outsourcing used for the week. With this software, it is not difficult to generate $O(\cdot)$ as a response function.

1.2   Overview of Results

We show that under a discounted-cost criterion, convexity of the operating cost function (along with certain other costs, such as hiring and wages) is propagated through MDP value iteration. Therefore, when $O(\cdot)$ is convex, the optimal hiring policy can be characterized as a state-dependent "hire-up-to" policy. That is, there exist state-dependent critical numbers for each period. If at the beginning of a period the number of entry-level employees exceeds the critical number then no hiring will be done. Otherwise, new employees will be hired to bring the number of entry-level employees back to the critical number. This holds in both finite and infinite horizon cases, with general, non-stationary service requirements.

We also develop results that offer a more detailed characterization of the optimal policy in important special cases. In particular, we show that, when service requirements are stationary or increasing, and when there is

- no learning, no training leadtime, and stochastic turnover, or

- no learning, positive training leadtime, and deterministic turnover,

then a "myopic" policy is optimal. A "myopic" policy is a policy under which we optimize a one-period static problem for each period, rather than the dynamic multi-period problem of the MDP.

In the second special case – with no learning, positive leadtime, and deterministic turnover – we also show that the optimal policy is a "hire-up-to staffing position" policy that is analogous

to classic results concerning "inventory position" in the inventory literature. While the inventory position is the sum of all goods on hand and on order, the staffing position in our model is a weighted sum of all employees working and in training. In our model, the presence of turnover makes the weights less than or equal to one.

In both special cases, when demands are k-periodic, then a k-period analogue of a myopic policy is optimal. That is, it is sufficient to solve a k-period MDP with appropriate end-of-horizon cost functions. When service requirements are non-stationary, we also prove there exists a one-sided "smoothing effect". Karlin [20] and Zipkin [29] have similar results in the inventory setting.

When there is learning, however, the results developed for our special cases no longer hold. Numerical results in §8 show that, when the problem data are periodic, there appears to be a two-sided smoothing effect. This is different from the one-sided smoothing effect demonstrated by Karlin [20] and Zipkin [29]. A similar two-sided smoothing effect is observed in the production-inventory setting by Kapuściński and Tayur [17].

Using numerical examples, we also study the benefit obtained through the use of our model. More specifically, we compare the cost performance of the optimal hiring policy, which depends on the number of employees at each capacity level, to a simpler policy, which bases hiring decisions only upon a single number, the total number of current employees across all types. In many cases, such as when the capacity increase that comes with learning is small or the cost of being under-staffed is small, the simpler, "single-number" policy performs quite well. When there is significant learning and the costs of being under-staffed are high, however, the optimal policy may provide a significant cost advantage.

As is clear from our MDP formulation and the structure of the optimal hiring policy, there exist intimate connections between our results and those of inventory theory. The "inventory" in our system is the population of employees, and "demand" corresponds to turnover. Conversely, our results may be seen as extensions of classic results from inventory theory, and from an inventory-theoretic viewpoint, our model may be viewed in two ways: as a single location model with multiple types of inventories and as a multi-echelon inventory system. We will discuss the similarities and differences in more detail after we present our general result, Theorem 1.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Then, in §3 and §4 we develop our model and main results. We then return in §5 to discuss further modeling considerations, in particular the structure of the operating cost function, $O(\cdot)$, and the

representation of leadtimes induced by hiring and training. Section 6 develops the optimality of myopic policies, and related structural results, for the special cases discussed above, and §7 develops analogous structural results when the data are periodic. Numerical analysis and insights are given in §8. Finally, in §9 we discuss our results and offer directions for future research. Connections to inventory theory are remarked in §4, §6 and §7, as the results are developed.

2   Literature Review

Holt, Modigliani, Muth, and Simon's seminal manpower planning model and its linear hiring rules [14] have inspired a long stream of research papers. For example, see Orrebeck, Schuette, and Thompson [23], Ebert [6], Gaimon and Thompson [7], and Khoshnevis and Wolfe [22], all of which use some mathematical programming approach. Using a partial differential equation model, Gerchak, Parlar, and Sengupta [8] solve the recruitment rate required to maintain a fixed capacity in an organization. Grinold and Stanford [10] develop a dynamic programming model based on linear costs, linear capacity (or budget) constraints and deterministic, fractional flow of employees.

In all of the above papers, the flow of employees among different types, including learning and turnover, is modeled as deterministic. In the model we propose, however, we allow learning and turnover to be stochastic and use a distinct MDP approach. This approach allows us to more carefully model the dynamics of learning and turnover among employees.

There is a separate stream of research that develops Markov-Chain models for human resource planning problems. Bartholomew, Forbes and McClean [3] provides an excellent summary of research in this area. In this paper we adopt a model closely related to the "mixed-exponential" model found in the human resource literature. Little work in this area has been devoted to the control aspect of the recruitment process, however, while control of the hiring process is the central question for us.

Two recent exceptions which consider aspects of control are papers by Bordoloi and Matsuo [5] and Pinker and Shumsky [25]. The first paper derives steady-state performance measures for (heuristic) linear control rules that are applied in a manufacturing environment that is similar to ours. It does not, however, consider the nature of optimal control policies. The second paper considers the improvement in quality which comes with worker experience. This focus on quality is complementary to the capacity analysis of this paper.

As is noted in the Introduction, there also exists a close connection between our results and those of classic inventory theory. In addition to papers by Karlin [19, 20] and Zipkin [29], our work is most closely related to work on inventory systems with spoilage by Iglehart and Jaquette [15], as well as other work on inventory-level dependent demand by Gerchak and Wang [9]. There is also a close connection between our results and those in the production-inventory literature. For example, see Aviv and Federgruen [2] and Kapuściński and Tayur [17] (and the references therein).

3   Model

In our discrete-time, continuous-state-space MDP, the number of planning periods under consideration is assumed to be T ($T = \infty$ in the infinite planning horizon case). Time periods will be indexed by t, $t = 0, 1, \ldots, T$. The length of time included in one period may be a week, a month, a quarter, or a year, depending on the application.

3.1   State Space and Control

In any period an employee may have attained one of m discrete skill levels, $i = 1, 2, \ldots, m$. In applications these skill levels may correspond to service speeds or acquired skills. In the former case, all employees do the same job but their service speeds increase with i. In the latter case, there are many job types. Employees start at level 1, with the most basic job skills, and progressively acquire more skills to become capable of handling more types of jobs. A combination of both cases is also possible: employees not only acquire new skills as they progress, they also become faster at these skills. In this paper we will use terms "service capacity" and "skill level" interchangeably, but the reader should keep in mind that they imply service speed and/or skill sets.

We also assume turnover may occur at any level i, $1 \leq i \leq m$, but hiring is made only at the entry level, level 1. So when new employees are hired they start at level 1 and progressively go through the levels until they turn over. We use a vector $(n_{1,t}, \ldots, n_{i,t}, \ldots, n_{m,t})$ to denote the numbers of employees at all levels $i = 1, \ldots, m$ before hiring at the beginning of period t. Since only type-1 employees will be hired, the vector becomes $(y_t, n_{2,t}, \ldots, n_{m,t})$ after x_t people are hired, where $y_t = n_{1,t} + x_t$. We treat $n_{i,t}$ as real variables. This relaxation of integrality makes our proofs, which rely on convexity properties, less burdensome. For large operations with many employees, the continuous approximation should be reasonable.

3.2 Learning, Turnover, and System Dynamics

Learning and turnover are accounted for only at the end of each time period. We assume that, of the $n_{i,t}$ level-i employees in period t, $\tilde{q}_{i,t}(n_{i,t})$ will quit and $\tilde{l}_{i,t}(n_{i,t})$ will learn and move to skill level $i+1$ in period $t+1$. Here "$\tilde{\ }$" is used to denote random numbers, indicating the fact that learning and turnover occur randomly. Note that the i and t subscripts mean that learning and turnover can have different patterns at different levels in different periods.

Typically, the rates at which learning and turnover occur have particular functional forms. The log-linear "learning curve" has been widely used in manufacturing (see Yelle [28]), and there is empirical evidence that it also exists in service operations. An example with directory assistance operators can be found in Gustafson [11]. Turnover typically decreases with job tenure: employees whose "job fit" is poor tend to leave after a short time; employees with longer tenure are self-selected to have better job fit (for example, see Jovanovic [16] and Parsons [24]). In our experience working with telephone call centers, we have found that the turnover rate can decline by more than 50% after an initial training and adjustment period.

Structurally, however, our model and its results do not depend on the rates at which learning and turnover take place. The "typical" learning-curve and turnover behaviors described above affect only the problem data used in the model.

Nevertheless, our model formulation does implicitly make some assumptions concerning learning and turnover. More specifically, it assumes that an employee's probability of learning or turning over in any period is independent of that person's learning and turnover probability in other periods. That is, the processes are Markovian.

Because individual employees typically learn and turn over independently of each other, a natural distribution to use for modeling learning and turnover in each period is the multinomial distribution. It implies that a person's length of stay at any level is geometrically distributed with different parameters for different levels. This "mixed geometric" distribution is a discrete version of the "mixed exponential" tenure length distribution in Bartholomew, Forbes and McClean [3], which has been widely used in the human resources literature.

When the number of type i employees turnover is binomially distributed – in which case employees turn over but do not learn – then we can use well-known results showing that binomial distributions are "totally positive of order 3" (TP_3) to show that the value function is convex in

the number of employees. This result relies heavily on both the TP_3 property and the fact that, in this case, the value function is one-dimensional.

As Karlin [21] states, however, generalization of the TP_3 property to higher dimensional cases and multinomial distributions is very difficult, and to the best of our knowledge no such results exist in the literature. This analytical difficulty with multinomial distributions prompts us to approximate them with stochastic proportions.

More specifically, we assume that $\tilde{l}_{i,t}$ and $\tilde{q}_{i,t}$ are independently distributed random variables with support on $[0,1]$ that represent the "stochastic proportion" of people who learn and turn over ($\tilde{l}_{m,t} = 0$). Therefore, in period t,

$$\tilde{q}_{i,t}(\cdot) = \begin{cases} \tilde{q}_{1,t} \cdot y_t & \text{if } i = 1 \\ \tilde{q}_{i,t} \cdot n_{i,t} & \text{if } i > 1 \end{cases} \quad \text{and} \quad \tilde{l}_{i,t}(\cdot) = \begin{cases} \tilde{l}_{1,t} \cdot (1 - \tilde{q}_{1,t}) \cdot y_t & \text{if } i = 1 \\ \tilde{l}_{i,t} \cdot (1 - \tilde{q}_{i,t}) \cdot n_{i,t} & \text{if } i > 1 \end{cases}.$$

Given the numbers of employees in the system at time t, $(n_{1,t}, \ldots, n_{m,t})$, and the fractions of employees turning over and learning, $\tilde{q}_{i,t}$ and $\tilde{l}_{i,t}$, the numbers of employees at $t+1$ are straightforward to calculate. The following equations represent the system evolution:

$$y_t = n_{1,t} + x_t, \qquad\qquad\qquad x_t \geq 0 \qquad (1)$$

$$n_{1,t+1} = (1 - \tilde{l}_{1,t})(1 - \tilde{q}_{1,t})y_t, \qquad (2)$$

$$n_{2,t+1} = (1 - \tilde{l}_{2,t})(1 - \tilde{q}_{2,t})n_{2,t} + \tilde{l}_{1,t}(1 - \tilde{q}_{1,t})y_t, \qquad (3)$$

$$n_{i,t+1} = (1 - \tilde{l}_{i,t})(1 - \tilde{q}_{i,t})n_{i,t} + \tilde{l}_{i-1,t}(1 - \tilde{q}_{i-1,t})n_{i-1,t}, \qquad 2 < i \leq m. \qquad (4)$$

Remark 1 The analytical difficulties with multinomial distributions to which we refer have long been recognized in the production and inventory literature. For an inventory model with spoilage, Iglehart and Jaquette [15] use the TP_3 property to prove that an "order-up-to" policy is optimal, but this result holds only for a single product, one-stage deterioration, and zero leadtime. To obtain results for more general models, other researchers who have investigated inventory spoilage and random production yields have also resorted to stochastic proportion models. For example, see Henig and Gerchak [12] and the references therein.

Note that the nature of the error introduced by using stochastic proportions is of second order. The means of the stochastic proportions can be set to equal the multinomial probabilities that an individual turns over or learns. When the distribution is multinomial, however, the relative dispersion of the fraction of employees quitting or advancing changes with the number of employees, but it does not change with stochastic proportions.

To the extent that, in steady state, the number of employees of one type remains large and relatively stable, these second-order differences should not induce undue bias. In our experience developing numerical examples that use stochastic proportions we have not found these errors to be significant (see §8).

3.3 Service Requirements

The employee staffing problem is defined in discrete time with time periods that are long enough to capture the effects of learning and turnover: weeks, months, quarters. Demand for service, along with the quality standards with which demand must be met, are typically defined on a much smaller time scale, however. This prompts us to define the service requirements as a vector, $\overrightarrow{D_t}$. Elements of the vector may represent the service requirements for different sub-intervals within the larger, discrete time period.

For example, in a telephone call center, hiring may take place on a weekly or monthly basis, while call volume forecasts and service standards are specified in 15-minute intervals. Therefore, when the planning period is one week and there is only one job type, the service requirement vector $\overrightarrow{D_t}$ may have a dimension of 672, corresponding to the call volume forecasts for every 15-minute interval of a week. When there are J job types, the service requirement vector $\overrightarrow{D_t}$ would have a dimension of $J \times 672$.

Thus, the definition of $\overrightarrow{D_t}$ allows us to model service requirements that are non-stationary over time and include multiple types of service. Furthermore, the elements of $\overrightarrow{D_t}$ need not be scalars at all. For example, by letting the elements of $\overrightarrow{D_t}$ be probability distributions we may explicitly model the uncertainty inherent in demand forecasts.

3.4 Costs and the Objective Function

In any period t, three types of costs will be incurred. First, a fixed cost of h will be incurred to hire a new employee. This cost typically includes advertising for, interviewing, and testing of job applicants, when appropriate. It may also include one-time training costs that are independent of wages.

Second, we let W_i be the wage cost of a type-i employee. This is the "fully loaded" compensation cost for one period and typically includes wages, benefits, and other direct personnel costs. Note

that these costs may vary with the skill level attained by the employee.

All other costs that arise when the $(y_t, n_{2,t}, \ldots, n_{m,t})$ employees serve demand, $\overrightarrow{D_t}$, are captured in the *operating cost function*, $O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t})$. These are "variable" costs that change with the level of the service requirements. They typically include employee overtime and the variable cost of outsourcing.

Note that we intentionally let the definition of $O_t(\cdot)$ remain a bit vague at this point. In the context of our hierarchical approach, it reflects the (often difficult) work assignment and personnel scheduling problems that must be solved in period t. It must also incorporate service-level constraints that ensure that "adequate" capacity is obtained to "reasonably" serve demand. Because the nature of the work assignment and personnel scheduling problems – and the resulting form of $O_t(\cdot)$ – may vary substantially from one setting to another, we leave the function undefined. Our only technical requirement is that $O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t})$ be jointly convex in $(y_t, n_{2,t}, \ldots, n_{m,t})$. In §5.1 we offer simple, general examples of $O_t(\cdot)$ in which this convexity property holds.

Finally, we use α to denote the one-period discount factor in the MDP and minimize the expected total discounted cost. Therefore the objective is to minimize the function

$$\min_{x_0, y_0, \ldots, x_T, y_T} \mathbf{E} \left\{ \sum_{t=0}^{T} \left[\alpha^t \left(h x_t + W_1 y_t + \sum_{i=2}^{m} W_i n_{i,t} + O_t(y_t, n_{2,t}, \ldots, n_{m,t}) \right) \right] \right\}, \qquad (5)$$

subject to the system dynamics (1), (2), (3), and (4).

Note that for simplicity we only consider costs incurred during the planning horizon in (5), therefore the "end-of-horizon" cost function is zero. Other types of "end-of-horizon" cost functions can also be incorporated.

4 Optimality of the Hire-up-to Policies

We will not solve the mathematical program (5) directly. Instead we solve the following equivalent Markov Decision Process (MDP) problem. (See §4.2 and §5.1 in Puterman [26] for this equivalence.)

Let $V_t(n_{1,t}, \ldots, n_{m,t} | T)$ denote the total discounted future cost at the beginning of period t when the numbers of employees on hand are $(n_{1,t}, \ldots, n_{m,t})$ and the planning horizon is T time periods. Then we have the following recursive expression:

$$V_t(n_{1,t}, \ldots, n_{m,t} | T) \;=\; \min_{x_t \geq 0} \left\{ h x_t + W_1(n_{1,t} + x_t) + \sum_{i=2}^{m} W_i n_{i,t} + O_t(n_{1,t} + x_t, n_{2,t}, \ldots, n_{m,t}) \right.$$

$$+\alpha \mathbf{E}_{\{\tilde{q}_{1,t},\dots,\tilde{l}_{1,t},\dots\}} \left\{ V_{t+1}(n_{1,t+1},\dots,n_{m,t+1}|T) \right\} \Big\}$$

$$= \min_{y_t \geq n_{1,t}} \left\{ J_t(y_t, n_{2,t}, \dots, n_{m,t}|T) \right\} - hn_{1,t} + \sum_{i=2}^{m} W_i n_{i,t}, \tag{6}$$

where

$$J_t(y_t, n_{2,t}, \dots, n_{m,t}|T) \stackrel{\text{def}}{=} H_t(y_t, n_{2,t}, \dots, n_{m,t}) + \alpha \mathbf{E}_{\{\tilde{q}_{1,t},\dots,\tilde{l}_{1,t},\dots\}} \left\{ V_{t+1}(n_{1,t+1},\dots,n_{m,t+1}|T) \right\},$$

$$H_t(y_t, n_{2,t}, \dots, n_{m,t}) \stackrel{\text{def}}{=} (h + W_1)y_t + O_t(y_t, n_{2,t}, \dots, n_{m,t}),$$

and (2), (3), and (4) hold.

Remark 2 Note that $O_t(y_t, \dots, n_{m,t})$, and therefore $J_t(y_t, \dots, n_{m,t}|T)$, is not defined for $y_t < 0$. Later we will show the convexity of these functions, but for now, suppose convexity holds for J_t. When $\frac{\partial J_t}{\partial y_t}(0, \dots, n_{m,t}|T) \leq 0$, the minimum of $J_t(y_t, \dots, n_{m,t}|T)$ is achieved at $y_t^*(n_{2,t}, \dots, n_{m,t}|T) \geq 0$, and there is no need for this definition. However, when $\frac{\partial J_t}{\partial y_t}(0, \dots, n_{m,t}|T) > 0$, it is optimal to not hire: $y_t^*(n_{2,t}, \dots, n_{m,t}|T) = 0$. In this case, we extend the definition of J_t so that $J_t(y_t, \dots, n_{m,t}|T) = J_t(0, \dots, n_{m,t}|T)$ for $y_t < 0$. This extension makes the first-order-condition, that $\lim_{y_t \to y_t^{*-}(n_{2,t},\dots,n_{m,t}|T)} \frac{\partial J_t}{\partial y_t}(y_t(n_{2,t}, \dots, n_{m,t}|T), n_{2,t}, \dots, n_{m,t}|T) = 0$, valid, and it preserves the convexity of J_t.

In this section, our main result shows that the optimal control is a state-dependent hire-up-to policy:

Definition 1 *A policy is of the state-dependent "hire-up-to" type if, for any* $(n_{1,t}, n_{2,t}, \dots, n_{m,t})$, *there exists* $y_t^*(n_{2,t}, \dots, n_{m,t})$ *such that the optimal hiring number* x_t^* *is*

$$x_t^* = \begin{cases} y_t^*(n_{2,t}, \dots, n_{m,t}) - n_{1,t} & \text{if } n_{1,t} < y_t^*(n_{2,t}, \dots, n_{m,t}) \\ 0 & \text{otherwise.} \end{cases}$$

For the finite-horizon problem, this result holds with no additional assumptions. For the infinite-horizon case, however, we need to make the following two technical assumptions about the state space and cost. In particular, we assume that for any planning horizon T there exists large constants, M and K, such that:

Assumption 1 $y_t^* + \sum_{i=2}^{m} n_{i,t} \leq M \qquad \forall i, t.$

Assumption 2 $H_t(y_t, n_{2,t}, \dots, n_{m,t}) \leq K, \ \forall t \ and \ \forall (y_t, n_{2,t}, \dots, n_{m,t}) \ where \ y_t + \sum_{i=2}^{m} n_{i,t} \leq M.$

Assumption 1 states that at any time the total number of employees after hiring does not exceed M. Because learning and turnover will not increase the total number of employees in the organization, this ensures that if we start with fewer than M people in total, then we will never exceed that number. For example, M could be the total population on earth. As a result, the state space can be reduced to a bounded and closed, and therefore compact, subset of $\mathbf{R}^m$.

Assumption 2 states that whenever we have finite number of people in the organization, the total one-period cost is bounded. This is straightforward for hiring and wage costs. It also holds for the operating costs. While we have not explicitly defined $O_t(\cdot)$, we know that as long as there exist finite bounds on all service requirements and cost parameters, the total operating cost in any time period is bounded.

Now we are ready to state our main result:

Theorem 1

(i) For $T < \infty$, $J_t(y_t, n_{2,t}, ... n_{m,t}|T)$ is jointly convex in $(y_t, n_{2,t}, ... n_{m,t})$ and $V_t(n_{1,t}, \ldots, n_{m,t}|T)$ is jointly convex in $(n_{1,t}, \ldots, n_{m,t})$ for all t.

(ii) For $T = \infty$, under Assumptions 1 and 2, $J_t(y_t, n_{2,t}, ... n_{m,t}|\infty)$ is jointly convex in $(y_t, n_{2,t}, ... n_{m,t})$ and $V_t(n_{1,t}, \ldots, n_{m,t}|\infty)$ is jointly convex in $(n_{1,t}, \ldots, n_{m,t})$ for all t.

Therefore, for both cases, in any time period the optimal policy is of the "hire-up-to" type.

Sketch of Proof   The proof of part (i) is similar to classic proofs of convexity results in the inventory literature and will be presented in Appendix A. It differs from that for more traditional inventory models, however, because of the dynamics in which "inventory" changes type. In particular, we use the following lemmas to show that convexity of the MDP value function is preserved in the presence of learning and turnover and that convexity is propagated through the minimization that is central to MDP recursion. They follow from Theorems 5.7 and 5.3 in Rockafellar [27] respectively.

Lemma 1 *If $g(y_1, \ldots, y_n)$ is jointly convex in $(y_1, \ldots, y_n)$ and $f_i(x_1, \ldots, x_m), i = 1, \ldots, n$, are linear functions, then $h(x_1, x_2, \ldots, x_m) = g(f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m))$ is also jointly convex in $(x_1, \ldots, x_m)$.*
Lemma 2 *Let $f(n_1, n_2, \ldots, n_m) = \inf_{y \geq n_1} \{h(y, n_2, \ldots, n_m)\}$. If $h(y, n_2, \ldots, n_m)$ is jointly convex in $(y, n_2, \ldots, n_m)$, then $f(n_1, n_2, \ldots, n_m)$ is jointly convex in $(n_1, n_2, \ldots, n_m)$.*

Standard MDP results provide that with stationary $\overrightarrow{D_t}$, the infinite horizon cost functions are the limits of finite horizon cost functions. That is, $\lim_{T\to\infty} V_t(n_{1,t}, \ldots, n_{m,t}|T) = V_t(n_{1,t}, \ldots, n_{m,t}|\infty)$ and $\lim_{T\to\infty} J_t(y_t, \ldots, n_{m,t}|T) = V_t(y_t, \ldots, n_{m,t}|\infty)$. As a result, convexity of cost function is preserved and the optimal policy is again of hire-up-to type. For non-stationary $\overrightarrow{D_t}$ (or, non-stationary costs in the inventory models), Karlin [18] states that similar results can be shown by using limit results but does not provide a proof. Our proof of part (ii), also presented in Appendix A, confirms this claim.

$$\triangle$$

Remark 3 There are two ways to view our system as an inventory system. First, our model may be viewed as a single location model with multiple types of inventories. Here turnover corresponds to "spoilage" or some other special type of inventory-level-dependent demand, and learning corresponds to the change of type among different types of inventories.

Second, our model may be viewed as a multi-echelon inventory system. Here, different levels on the learning curve correspond to different system echelons. In turn, learning is the analogue of transfer between echelons, and turnover represents the echelons' inventory-level dependent demand. The fact that learning and turnover are random and that costs are not separable makes these inventory models difficult to solve.

Note also that the optimality of "up-to" type of policy has also been established for other capacitated production systems with stationary or periodic demand. See Aviv and Federgruen [2] and Kapuściński and Tayur [17] for recent results, as well as references to other work.

5 Modeling Considerations

5.1 Operating Cost Function

Now we will discuss in detail the operating cost function $O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t})$, and describe how it can be modeled to address the uncertainty in the service requirement forecast, as well as to capture the essence of the assignment and work-force scheduling problems.

We will start with a simple example. Suppose there is only one job type. For the work-force scheduling problem within any time period t, we assume that there are in total S sub-intervals in

the period, indexed by $s = 1, 2, \ldots, S$. Within this time period, there may exist WS feasible work schedules, indexed by $w = 1, 2, \ldots, \mathit{WS}$. These work schedules may or may not include overtime, but all of them satisfy workplace rules and regulations regarding breaks, overtime, and so on.

Each type-i employee is assumed to have a processing capacity of μ_i. For now, we will assume that service requirements are specified as requirements for total processing capacities. We use $\overrightarrow{D_t} = \{D_t^s\}_{s=1,\ldots,S}$ to denote these requirements for each sub-interval s of period t.

Let

$$I(w, s) = \begin{cases} 1 & \text{if work schedule } w \text{ requires working in sub-interval } s \\ 0 & \text{otherwise} \end{cases},$$

and let x_{iw} denote the number of type-i employees on work schedule w. For simplicity, we will assume that a fixed cost of C_{iw} is incurred to assign a type-i employee to work schedule w.

When the service requirements cannot be met with available regular time and overtime of all employees, outsourcing will be used to satisfy the residual service requirements. More specifically, we let variables z_s denote the amount of work to be outsourced in sub-interval s. It will again be assumed that a fixed cost of OS_s is incurred for each unit of outsourced work in sub-interval s.

Therefore, the operating cost function is the minimum cost obtained by solving the following Linear Program (LP) for each $(y_t, n_{2,t}, \ldots, n_{m,t})$:

$$
\begin{aligned}
O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t}) \quad = \quad \min_{x_{iw}, z_s} \quad & \sum_{i=1}^{m} \sum_{w=1}^{\mathit{WS}} C_{iw} x_{iw} + \sum_{s=1}^{S} OS_s z_s \\
\text{s.t.} \quad & \sum_{i=1}^{m} \sum_{w: I(w,s)=1} \mu_i x_{iw} + z_s \geq D_t^s \qquad \forall s \\
& \sum_{w=1}^{\mathit{WS}} x_{1w} \leq y_t \\
& \sum_{w=1}^{\mathit{WS}} x_{iw} \leq n_{i,t} \qquad \forall 1 < i \leq m \\
& x_{iw}, \ z_s \geq 0 \qquad \forall i, w, s
\end{aligned}
\tag{7}
$$

Note that this LP is a continuous approximation to the real scheduling problem in which all x_{iw} assume integer values. Even though this LP may give fractional optimal solutions of x_{iw} that are not directly implementable, for large organizations with many employees, the minimum objective function obtained from this LP, $O_t(\cdot)$, is likely to be close to the real operating cost function. Berman, Larson, and Pinker [4] use a similar LP formulation to solve the workforce and workflow scheduling problem at high volume factories such as USPS Mail Processing Centers.

Proposition 1 *For any fixed $\overrightarrow{D_t}$, $O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t})$ is jointly convex in $(y_t, n_{2,t}, \ldots, n_{m,t})$.*

Proof   Suppose for $(y_t, n_{2,t}, \ldots, n_{m,t})$, $\{x_{iw}, z_s\}$ achieves the minimum in $O_t(y_t, n_{2,t}, \ldots, n_{m,t})$, and for $(y_t', n_{2,t}', \ldots, n_{m,t}')$, $\{x_{iw}', z_s'\}$ achieves the minimum. Then for $\lambda(y_t, n_{2,t}, \ldots, n_{m,t}) + (1 - \lambda)(y_t', n_{2,t}', \ldots, n_{m',t}')$, the combination $\lambda\{x_{iw}, z_s\} + (1 - \lambda)\{x_{iw}', z_s'\}$ is feasible. Moreover,

$$
\begin{aligned}
&\lambda O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t}) + (1 - \lambda)O_t(y_t', n_{2,t}', \ldots, n_{m,t}'; \overrightarrow{D_t}) \\
&= \quad \lambda\left(\sum_{i,w} C_{iw} x_{iw} + \sum_s OS_s z_s\right) + (1 - \lambda)\left(\sum_{i,w} C_{iw} x_{iw}' + \sum_s OS_s z_s'\right) \\
&\stackrel{(*)}{\geq} \quad \sum_{i,w} C_{iw}\left(\lambda x_{iw} + (1 - \lambda)x_{iw}'\right) + \sum_s OS_s\left(\lambda z_s + (1 - \lambda)z_s'\right) \\
&\geq \quad O_t(\lambda(y_t, n_{2,t}, \ldots, n_{m,t}) + (1 - \lambda)(y_t', n_{2,t}', \ldots, n_{m,t}'); \overrightarrow{D_t}) \qquad (8)
\end{aligned}
$$

Note that for the linear cost functions, as is the case in this example, (*) is actually an equality. But in general $\geq$ is sufficient. When the cost functions are convex, $\geq$ clearly holds. Therefore, Proposition 1 holds for general convex cost functions $C_{iw}(\cdot)$ and $OS_s(\cdot)$ as well (see Proposition 2 below).

$\triangle$

Indeed, in more general cases the scheduling problem is more difficult. For example, both the demand forecast and the set of employee skills may include multiple types of service. Nevertheless, we can adapt the original LP to take these into account.

Suppose that there are J types of jobs, indexed by $j = 1, \ldots, J$, and that each employee type, i, specifies a group of employees that are capable of processing the same types of jobs and work at the same speed. The service requirement forecast $\overrightarrow{D_t}$ now includes forecasts for each job type as well : $\overrightarrow{D_t} = \left\{D_t^{sj}\right\}_{s=1,\ldots,S}^{j=1,\ldots,J}$. These service requirements may be modeled as distributions to account for the forecast uncertainty. The cost of assigning x_{iw} type-i employees to work schedule w is $C_{iw}(x_{iw})$, and the cost of outsourcing z_{sj} type-j work in sub-interval s is $OS_{sj}(z_{sj})$. Both cost functions can be general convex functions. This includes the previous linear costs as special cases.

The following mathematical program therefore generalizes the original LP:

$$
O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t}) \quad = \quad \min_{x_{iw}, z_{sj}} \sum_{i=1}^{m}\sum_{w=1}^{WS} C_{iw}(x_{iw}) + \sum_{s=1}^{S}\sum_{j=1}^{J} OS_{sj}(z_{sj})
$$

15

$$\text{s.t.} \quad f(x_{11}, \ldots, x_{mWS}, z_{11}, \ldots, z_{SJ}; \overrightarrow{D_t}) \geq 0 \qquad (9)$$

$$\sum_{w=1}^{WS} x_{1w} \leq y_t$$

$$\sum_{w=1}^{WS} x_{iw} \leq n_{i,t} \qquad \forall 1 < i \leq m$$

$$x_{iw}, \; z_{sj} \geq 0 \qquad \forall i, w, s, j$$

Here, (9) is a service-level constraint that generalizes (7). Because forecasts can be uncertain, the service constraints may also be specified in distributional form. For example, it may say "the choice of $(x_{11}, \ldots, x_{mWS}, z_{11}, \ldots, z_{SJ})$ should be such that the probability of meeting a certain service standard is more than a certain percentage, $a\%$". Clearly the evaluation of equation (9) is not transparent. Indeed, it includes the low-level work assignment aspect of the problem. Nevertheless, by imposing a mild restriction on (9), we can guarantee the convexity of the function $O_t(\cdot)$.

Proposition 2 *For a fixed $\overrightarrow{D_t}$, if*

(i) the set $\{(x_{11}, \ldots, x_{mWS}, z_{11}, \ldots, z_{SJ}) : f(x_{11}, \ldots, x_{mWS}, z_{11}, \ldots, z_{SJ}; \overrightarrow{D_t}) \geq 0\}$ is convex,

(ii) the cost functions $C_{iw_i}(\cdot)$, and $OS_{sj}(\cdot)$ are convex,

then given the distributions of $\overrightarrow{D_t} = \left\{ D_t^{sj} \right\}_{s=1,\ldots,S}^{j=1,\ldots,J}$, $O_t(y_t, n_{2,t}, \ldots, n_{m,t}; \overrightarrow{D_t})$ is jointly convex.

Proof The proof is similar to that of Proposition 1. We only need to note that condition (i) is used to ensure the convexity of the feasible region, and condition (ii) guarantees (*) in (8) still holds.

$$\triangle$$

Conditions (i) and (ii) appear to be reasonable. For example, for condition (i), if two solutions to the math program, $(x_{11}, \ldots, x_{mWS}, z_{11}, \ldots, z_{SJ})$, provide acceptable service more than $a\%$ of the time, it is reasonable to assume that a linear combination of them will also do so. This assumption is true, for example, when service requirements, $\overrightarrow{D_t}$, are specified only as requirements for total processing capacity. Since the marginal cost of each extra unit of outsourced work is typically increasing, it is also reasonable to assume $OS_{sj}(\cdot)$ to be convex. $C_{iw_i}(\cdot)$ is usually linear.

5.2 Training/Hiring Leadtime

In organizations in which employees need to acquire a certain level of service proficiency before they are put "on-line", training is an important component of the hiring process. For example, in some call centers, initial training for a CSR may take more than 2 months to complete. In this case, hiring leadtime (of which training is a major component) becomes very important in capacity planning for future periods.

Because our model already includes multiple employee types, it is straightforward to incorporate this leadtime. If the hiring/training leadtime is λ periods, then we add λ more employee types (thus $m > \lambda$) and use employee types 1 to λ to denote the trainees. For organizations in which training mainly takes place on the job, $\mu_i > 0, \forall i = 1, \ldots, \lambda$ denotes the service output these trainees may generate. For other organizations, in which employees must receive a minimum amount of initial training before working, $\mu_i = 0, \forall i = 1, \ldots, \lambda$. Note that W_i for $1 \leq i \leq \lambda$ are the trainee wage and benefit rates. If the hiring process itself takes some time, then the first few W_i's may be 0. It is clear that our formulation offers flexibility in modeling all kinds of hiring leadtimes and their associated costs.

6 Stationary or Increasing Service Requirements: Cases in Which Myopic Policies are Optimal

Even though we have characterized the optimal hiring policy, it may be difficult to implement. The high dimensionality of the state space makes the computational task required formidable, particularly in organizations with many employees.

When the service requirements are stationary or (stochastically) increasing, however, myopic policies may be optimal. Under myopic policies we optimize one-period static problems for each period, rather than the MDP's dynamic, multi-period problem, thus greatly reducing the computational requirement.

Definition 2 *A policy is myopic if, in each period t, an action $x_t^G(n_{1,t}, \ldots, n_{m,t})$ is taken where $x_t^G(n_{1,t}, \ldots, n_{m,t})$ is the minimizer of some one-period cost function $G_t(n_{1,t} + x_t, \ldots, n_{m,t}; \overrightarrow{D}_t)$. A myopic policy is optimal if $x_t^G(n_{1,t}, \ldots, n_{m,t})$ also minimizes $V_t(n_{1,t}, \ldots, n_{m,t}; \overrightarrow{D}_t, \overrightarrow{D}_{t+1}, \ldots)$ for any $(n_{1,t}, \ldots, n_{m,t})$.*

In this section we examine two cases for which myopic policies are optimal. While myopic policies are not optimal in general, they may also prove to be effective, computationally practical heuristics for a broader class of problem instances than the special cases in which they are optimal.

6.1 No Learning, No Hiring Leadtime

In some organizations very little initial training is used, and learning on the job is so fast or so little that we can model all the employees as one type. The prototypical example of this type of operation is a "fast-food" restaurant. Even though learning is not a factor in these organizations, turnover remains an important problem.

Since there is only one employee type, when hiring leadtime is zero, the state space becomes one-dimensional. We will drop the subscript i and keep the meaning of variables and parameters such as n, x, y and W. If we let $\tilde{r}_t = 1 - \tilde{q}_t$ and

$$G_t(y_t; \overrightarrow{D_t}) \stackrel{\text{def}}{=} \mathbf{E}_{\{\tilde{q}_t\}} \left\{ ((1 - \alpha\tilde{r}_t)h + W)y_t + O_t(y_t; \overrightarrow{D_t}) \right\} = ((1 - \alpha\mathbf{E}\{\tilde{r}_t\})h + W)y_t + O_t(y_t; \overrightarrow{D_t}), \quad (10)$$

then given any hiring policy, π,

$$
\begin{aligned}
V(n_0|\pi; \overrightarrow{D_t}) &= \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^t \left[h(y_t - n_t) + Wy_t + O_t(y_t; \overrightarrow{D_t}) \right] \right\} \\
&= \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^t \left[(h + W)y_t + O_t(y_t; \overrightarrow{D_t}) \right] \right\} - \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=1}^{\infty} \alpha^t hn_t \right\} - hn_0 \\
&\stackrel{(**)}{=} -hn_0 + \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^t \left[(h + W)y_t + O_t(y_t; \overrightarrow{D_t}) \right] \right\} - \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^{t+1} h\tilde{r}_t y_t \right\} \\
&= -hn_0 + \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^t [((1 - \alpha\tilde{r}_t)h + W)y_t + O_t(y_t)] \right\} \\
&= -hn_0 + \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \left\{ \sum_{t=0}^{\infty} \alpha^t G_t(y_t; \overrightarrow{D_t}) \right\}, \quad\quad (11)
\end{aligned}
$$

where $y_t \geq n_t$ and $n_{t+1} = \tilde{r}_t y_t$.

Equality (**) holds because $n_{\tau+1}$ is independent of $\{\tilde{r}_t\}_{t=\tau+1}^{\infty}$ and y_τ is independent of $\{\tilde{r}_t\}_{t=\tau}^{\infty}$, which implies that for any τ,

$$\mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \{n_{\tau+1}\} = \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\tau}\}} \{n_{\tau+1}\} = r_\tau \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\tau-1}\}} \{y_\tau\} = r_\tau \mathbf{E}_{\{\{\tilde{r}_t\}_{t=0}^{\infty}\}} \{y_\tau\}.$$

Note that the above transformation, (11), also holds in the finite horizon case, if we assume an "end-of-horizon" cost function of $V_{T+1}(n_{T+1}) = -\alpha hn_{T+1}$.

18

Thus (10) defines the one-period cost function $G(\cdot)$ that is the basis of a myopic policy. When the demand distribution $\{\overrightarrow{D_t}\}$ is stationary it is not difficult to show that repeatedly optimizing $G(\cdot)$ minimizes $V(\cdot)$ as well. Furthermore, when the sequences of demands is "stochastically increasing", myopic policies continue to be optimal, given the following assumption.

Assumption 3 *For two forecast distributions, if $\overrightarrow{D_1} \preceq \overrightarrow{D_2}$, then $dO_t(y_t; \overrightarrow{D_1})/dy_t \geq dO_t(y_t; \overrightarrow{D_2})/dy_t$ for all t and y_t.*

Recall that a random variable X is said to be stochastically less than or equal to another random variable Y, $X \preceq Y$, if their corresponding cumulative distribution functions F_X and F_Y satisfy $F_X(x) \geq F_Y(x)$ for all x. Then Assumption 3 states that the marginal value of an extra employee in reducing the operating cost is higher when the service requirement is stochastically higher.

Theorem 2 *Suppose system parameters, except the service requirements, are stationary. Let $y_t^G = \operatorname{argmin}_{y_t} G_t(y_t)$. Choose the smallest when multiple minimizers exist. Then, given Assumption 3, when there is no learning and no hiring leadtime,*

    (i) for both the infinite-horizon case and the finite-horizon case with end-of-horizon cost function $V_{T+1}(n_{T+1}) = -\alpha h n_{T+1}$, if the service requirements are stationary or stochastically increasing, $y_t^G \leq y_{t+1}^G$ for all t;

    (ii) if $y_t^G \leq y_{t+1}^G$ for all t, then the myopic policy of minimizing $G_t(y_t)$ in each period t is optimal. That is, the "hire-up-to" numbers in the optimal policy are $(y_0^*, y_1^*, \ldots) = (y_0^G, y_1^G, \ldots)$.

Therefore, when service requirements are stationary or stochastically increasing, the myopic policy is optimal. In steady-state, with stationary requirements, $x_{t+1}^* = \tilde{q}_t y_t^G$, and with increasing requirements, $x_{t+1}^* = (y_{t+1}^G - y_t^G) + \tilde{q}_t y_t^G$.

Proof Part (i). When service requirements are stationary, $G_t(y_t)$ is the same for all t, therefore $y_t^G = y_{t+1}^G$ for all t. When service requirements are stochastically increasing, $\overrightarrow{D_t} \preceq \overrightarrow{D}_{t+1}$, then, by Assumption 3, $O_t'(y_t; \overrightarrow{D_t}) \geq O_{t+1}'(y_t; \overrightarrow{D}_{t+1})$ and $G_t'(y) \geq G_{t+1}'(y)$. First order conditions state that $\lim_{y \to y_t^{G-}} G_t'(y) = \lim_{y \to y_{t+1}^{G-}} G_{t+1}'(y) = 0$. Hence the convexity of both $G_t(\cdot)$ and $G_{t+1}(\cdot)$ implies $y_t^G \leq y_{t+1}^G$.

Part(ii). We note that even though the costs underlying $G_t(\cdot)$ in the staffing problem differ from the costs underlying one-period costs in inventory problems, the convexity and separability of the $G_t(\cdot)$'s, along with the fact that the sequence of solutions, $\{y_t^G\}$, is increasing, are sufficient to ensure that classic arguments from inventory theory hold. For example, see Proposition 3-2 and Theorem 3-1 in Heyman and Sobel [13].

As a result, in steady state, when service requirements are stationary, $y_t = y_t^* = y_t^G$, then $n_{t+1} = \tilde{r}_t y_t = \tilde{r}_t y_t^G$, and $x_{t+1}^* = y_{t+1}^* - n_{t+1} = (1 - \tilde{r}_t) y_t^* = \tilde{q}_t y_t^G$. When the service requirements are increasing, so are y_t^G. And $x_{t+1}^* = y_{t+1}^* - n_{t+1} = (y_{t+1}^G - y_t^G) + \tilde{q}_t y_t^G$.

$\triangle$

The implication here is that if service requirements are stationary, we are simply hiring to replace those who have just left ("x-for-x"). If the service requirements are increasing, we hire to replace the turnover, as well as to expand to meet increasing demand. Therefore, not only is the optimal policy myopic, the optimal hiring action in steady state is also very easy to understand and simple to exercise.

Remark 4 Iglehart and Jaquette [15] have shown that, for an inventory model with inventory-level-dependent deterioration, an order-up-to policy is optimal. While the nature of the cost function in their model is different from that in our staffing model, adaptation of their proof to a discrete-state-space version of our problem shows that, in this special case, a myopic hire-up-to policy is optimal even when turnover is binomially distributed. The proof hinges on the TP_3 property of binomial distributions. (See Proposition 3.2 in Karlin [21].)

6.2   Positive Hiring Leadtime, Deterministic Turnover, and No Learning

In this section we consider positive hiring leadtime cases and let λ be the leadtime. As before, we will assume no learning on the job so there are $(\lambda + 1)$ types of employees. The first λ types are in the process of being hired or trained, and only type $(\lambda + 1)$ has productive capacity, $\mu_{\lambda+1} > 0$. We will also assume that the turnover rates are deterministic and stationary. That is, $\tilde{q}_{i,t} = q_i, \tilde{r}_{i,t} = r_i = 1 - q_i$ with probability one for constants $q_i, i = 1, \ldots, \lambda + 1$.

Definition 3 *The staffing position in period t before hiring is*

$$SP_t^- = n_t r_{\lambda+1}^\lambda + \sum_{\tau=1}^{\lambda-1} \left[x_{t+\tau-\lambda} \left(\prod_{s=1}^{\lambda} r_s \right) (r_{\lambda+1})^{\lambda-\tau} \right] \tag{12}$$

and after hiring is $SP_t^+ = SP_t^- + x_t \prod_{s=1}^{\lambda} r_s$.

Remark 5 Staffing position follows in spirit the "inventory position" concept in the inventory literature (e.g. see Arrow, Karlin and Scarf [1]). Because turnover reduces the number of people in each period, however, the staffing position in our setting is a weighted sum of all the people in the pipeline. For example, suppose $x_{t+1-\lambda}$ trainees are hired at the beginning of period $t+1-\lambda$. Then, due to trainee turnover, $x_{t+1-\lambda} \left(\prod_{s=1}^{\lambda} r_s \right)$ will be available to work at the beginning of period $t+1$, but out of these people, only $x_{t+1-\lambda} \left(\prod_{s=1}^{\lambda} r_s \right) (r_{\lambda+1})^{\lambda-1}$ will remain at the beginning of period $t + \lambda$, due to on-the-job turnover. Trainees at different levels will start working in different periods so the proportions of them that will "survive" until period $t + \lambda$ are different.

Just as $SP_t^+ = y_{t+\lambda}$, we can transform hiring costs and numbers hired over one leadtime to develop results that are analogous to Theorem 2. Specifically, if we are in period $t - \lambda$ then the future cost of hiring now for one person in period t becomes

$$\hat{h} = \frac{\alpha^{-\lambda} (h + W_1) + \sum_{i=1}^{\lambda-1} \left[\alpha^{-\lambda+i} \left(\prod_{s=1}^{i} r_s \right) W_{i+1} \right]}{\prod_{s=1}^{\lambda} r_s}. \tag{13}$$

Here $\{W_1, \ldots, W_\lambda\}$ represent wages that may be paid to employees in different stages of hiring and training. That is, if the leadtime includes an initial training period, then wages paid during training may be included here. Otherwise, they may be set to zero.

Then using $\{SP_t^+\}$ and $\hat{h}$ in place of $\{y_t\}$ and h, we may myopically solve the hiring problem, as in Theorem 2, to generate the optimal hiring numbers $\{\hat{x}_t\}$. Note, however, that $\hat{x}_t$ corresponds to the net number of new hires remaining in period t. The actual number hired in period $t - \lambda$ is

$$x_{t-\lambda}^* = \frac{\hat{x}_t}{\prod_{s=1}^{\lambda} r_s}. \tag{14}$$

Thus, we have the following analogue to Theorem 2:

Corollary 1 *If there is no learning, turnover is deterministic and stationary, the hiring leadtime is λ and service requirements are stationary or stochastically increasing, then a myopic policy is optimal. Furthermore, the policy uses (12) as the system state and (13) as hiring cost, and it determines net hiring numbers $\{\hat{x}_t\}$ using the algorithm of Theorem 2. It then calculates optimal hiring numbers $\{x_{t-\lambda}^*\}$ from $\{\hat{x}_t\}$ using (14).*

7 Non-stationary Requirements: Periodicity and Smoothing

In many service organizations, customer demand for service is highly non-stationary. For example, retail stores and catalog vendors experience holiday-season spikes in demand. Similarly, for call centers in retail financial services, tax season means high call volumes. In this section, we will derive structural properties of the optimal hiring policy when the service requirements are non-stationary. First, in §7.1 and §7.2, we will study structural properties of the optimal policy when service requirements are non-stationary. Then in §7.3 we will study more general non-stationary service requirements. In all three sections, we focus on the special cases of §6.

7.1 No Learning, No Hiring Leadtime

We assume that service requirements are periodic with a fixed cycle length. Similarly we assume that other model parameters, such as costs and turnover rates, are also either stationary or periodic, though not necessarily of the same cycle lengths. This allows us to treat the whole process as periodic with a cycle length that is the least common multiplier of the lengths of all the cycles embedded in the system. We will denote this system cycle length by k.

We only consider the infinite planning horizon case. Note that the optimality of hire-up-to policy is shown in Theorem 1. As the optimal decision depends only on the current state and cost data and service requirements that are k-periodic, the optimal hire-up-to levels are also k-periodic. We denote them by $y_0^*, y_1^*, \ldots, y_{k-1}^*$.

Theorem 3 *Suppose Assumptions 1 and 2 hold. When there is no learning, no hiring leadtime, and the service requirements are k-periodic,*

(i) *there always exist periods $\underline{t}$ (mod k) in which a myopic policy is optimal;*

(ii) *the optimal policy is "almost" myopic: the optimal hire-up-to levels can be found through solution of a k-period MDP with $\underline{t}$ as the last period and $-\alpha h n_{\underline{t}+1}$ as the end-of-horizon cost function.*

Proof   For part (i), we first prove that if $y_t^* \leq y_{t+1}^*$, then the myopic policy is optimal in period t, i.e., $y_t^* = y_t^G = \operatorname{argmin}_{y_t} G_t(y_t)$. Then if we let $y_{\underline{t}}^*$ be the smallest hire-up-to number among $y_0^*, y_1^*, \ldots, y_{k-1}^*$, the myopic policy is optimal in period $\underline{t}$.

To prove this, we note that since y_t^G is the smallest minimizer of $G_t(y_t)$, it is also the smallest value such that its left derivative $\lim_{y_t \to y_t^{G-}} G_t'(y_t) = 0$. Similarly, y_t^* is the smallest value such that $\lim_{y_t \to y_t^{*-}} J_t'(y_t) = 0$.

Because $y_t^* \le y_{t+1}^*$, when $y_t \le y_t^*$, $J_t(y_t) = (h + W)y_t + O_t(y_t) + \alpha(J_{t+1}(y_{t+1}^*) - hr_t y_t) = G_t(y_t) + \alpha J_{t+1}(y_{t+1}^*)$. Therefore, when $y_t \le y_t^*$, $G_t'(y_t) = J_t'(y_t)$. Hence $\lim_{y_t \to y_t^{*-}} G_t'(y_t) = \lim_{y_t \to y_t^{*-}} J_t'(y_t) = 0$, and y_t^{*-} is the smallest such number. Therefore $y_t^G = y_t^*$.

Part (ii) then follows directly.

$\triangle$

Remark 6 Note that Theorem 3 only establishes the existence of $\underline{t}$, a procedure still has to be developed to find $\underline{t}$. In the inventory literature, Karlin [19, 20] and Zipkin [29] have analyzed inventory systems with seasonal demands. [20] has similar results to Theorem 3, and both [20] and [29] give k-stage algorithms to find $\underline{t}$ and calculate the optimal order-up-to levels. Interested readers are referred to [20] and [29].

7.2   Positive Hiring Leadtime, Deterministic Turnover, and No Learning

Now turnover is deterministic and there is a positive hiring leadtime, λ. In this case, we can use versions of the staffing position, $\{SP_t^-\}$, and hiring cost, $\hat{h}$, of §6.2 to transform the staffing problem into an analogous zero leadtime system. In this periodic setting, however, the data need not be stationary, and we may modify the transformations (12) and (13) to include additional time indices.

With these transformations, the arguments that lead to Theorem 3 can be used directly to prove the following analogue:

Corollary 2 *Suppose Assumptions 1 and 2 hold. When there is no learning, turnover is deterministic, the hiring leadtime is λ, and service requirements are k-periodic, then the results of Theorem 3 hold.*

Again, the optimal "net" staffing numbers obtained in the zero leadtime analogue, $\{\hat{x}_t\}$, must be transformed using a modified version of (14) to obtain the optimal hiring numbers $\{x_t^*\}$.

23

7.3 Monotonicity Properties and Smoothing Effect

In this part, we will study the effect of non-stationary demand on the optimal "hire-up-to" numbers. The service requirements can be any sequence of non-stationary distributions. While the ideas and proofs presented here are similar to those in Karlin's inventory model in [19] (Zipkin [29] has similar results), our setting is different. In inventory models demand affects the system dynamics, but in our model demand for service affects only the costs. The "demand" that drives the system dynamics in our model is, in fact, the learning and turnover of employees.

In both special cases in which a myopic policy is optimal, the state space can be collapsed into one dimension – either the actual staffing level or the weighted staffing position. The optimal hire-up-to staffing position numbers (which reduce to the actual staffing level numbers when $\lambda = 0$) have the following structural properties. (Proof of the following theorem can be found in the Appendix.)

To stress the dependence of hire-up-to numbers on the service requirement forecast, we will use $y_t^*(D_{t+\lambda}, D_{t+\lambda+1}, \ldots)$ to denote the optimal staffing position in period t when the demand stream from period $t + \lambda$ on is $D_{t+\lambda}, D_{t+\lambda+1}, \ldots$

Theorem 4 *Suppose there is no learning, no hiring leadtime, and Assumption 3 holds. For any sequence of service requirements $D_{t+\lambda}, D_{t+\lambda+1}, \ldots$,*

(i) *if $D_{t+\lambda} \succeq D_{t+\lambda+1}$, then $y_t^*(D_{t+\lambda}, D_{t+\lambda+1}, \ldots) \geq y_{t+1}^*(D_{t+\lambda+1}, D_{t+\lambda+2}, \ldots)$;*

(ii) *if $D_{t+\lambda} \preceq D_{t+\lambda+1}$ but $y_t^*(D_{t+\lambda}, D_{t+\lambda+1}, \ldots) > y_{t+1}^*(D_{t+\lambda+1}, D_{t+\lambda+2}, \ldots)$, then*
$$y_{t+1}^*(D_{t+\lambda+1}, D_{t+\lambda+2}, \ldots) > y_{t+2}^*(D_{t+\lambda+2}, D_{t+\lambda+3}, \ldots).$$

This theorem shows a one-sided smoothing effect. Part (i) show that when the service requirement decreases, so does the hire-up-to number. Part (ii) shows, however, that but when the service requirement increases, the hire-up-to number may go down instead. This happens if the optimal hire-up-to number after that period also decreases. The one-sided nature of the smoothing effect is driven by the fact that we are free to hire, but are not free to fire.

In our general model, however, this one-sided effect may not hold. In particular, the presence of learning guarantees that there is a positive probability that, even without hiring, service capacity will grow to exceed what is needed in future periods. In turn, the optimality of myopic policies, which is based on the impossibility of future over-capacity, no longer exists. The approach we take here to prove the one-sided smoothing effect does not hold in the general case precisely because it

relies heavily on the optimality of myopic policies in some periods. The numerical analysis presented in the next section suggests that there exist a two-sided, rather than one-sided, smoothing effect.

8 Numerical Analysis

In this section, we perform two sets of numerical experiments. The first identifies conditions under which hiring policies that account for employee learning may significantly out-perform those that do not. The second indicates that the presence of learning may induce a "two-sided smoothing effect" into periodic problems which differs from the one-sided effect demonstrated in the inventory literature. These numerical analyses are based on average-cost models.

8.1 Average Cost as the Method of Comparison

When the planning horizon is finite, the average-cost problem is a special case of the discounted-cost problem (with $\alpha = 1$). When planning horizon is infinite, the optimal average-cost policy generates a constant average one-period cost (for all the states that are in the same chain of the MDP), rather than the discounted costs that may depend on the system's starting state. The state-independent nature of the average cost makes the comparison of alternative policies and MDP systems clearer.

To use average costs, we must first develop a set of properties for the optimal hiring policy under the average-cost criterion. The proposition below assumes that there exist an average-cost optimal policy and its corresponding value function.

Proposition 3 *Let $V_0^\alpha(\cdot)$ be the α-discounted cost function, and suppose the average-cost function, $\lim_{T\to\infty} \sum_{t=0}^{T-1} \left[hx_t + W_1 y_t + \sum_{i=2}^m W_i n_{i,t} + O_t(y_t, n_{2,t}, \ldots, n_{m,t}) \right] / T$ exists. Then*

(i) $\lim_{\alpha\to 1}(1-\alpha)V_0^\alpha(n_{1,0}, \ldots, n_{m,0})$ exists, and

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{1}{T} \left[hx_t + W_1 y_t + \sum_{i=2}^m W_i n_{i,t} + O_t(y_t, n_{2,t}, \ldots, n_{m,t}) \right] = \lim_{\alpha\to 1}(1-\alpha)V_0^\alpha(n_{1,0}, \ldots, n_{m,0});$$

(ii) the average-cost function is convex and the optimal hiring policy is of the hire-up-to type.

Proof Part *(i)* of the proposition follows from Proposition 4-7 in Heyman & Sobel [13]. Part *(ii)* follows naturally from *(i)* since the convexity of discounted-cost functions has already been established in Theorem 1.

△

Part (i) of the proposition establishes the relationship between the discounted-cost and average-cost criteria: when α is close to 1, the average cost function divided by $(1 - \alpha)$ can be used to approximate the discounted cost function. Therefore comparisons based on average-cost criterion can be used to infer comparisons based on discounted-cost as well.

Note that we do not prove the existence of an average-cost optimal policy for our system, so technically we should regard our numerical results with some caution. Nevertheless, in all of our numerical examples, MDP value iteration has converged without problems. Furthermore, we find the clarity of the numerical comparisons afforded by the average-cost model to outweigh this technical limitation.

8.2   Learning Effect

We have established the optimality of state-dependent hire-up-to policies. In large organizations with many employees, the MDP state space grows quickly, and for these policies, MDP value iteration becomes computationally intensive. From §6 and §7 we see, however, that when learning is ignored, there exist specially-structured policies that are easy to analyze, understand, and implement. So what do we gain by explicitly modeling the learning effect?

To answer this question we perform numerical tests in which we compare the performance of hiring policies that account for learning with those which do not. Results presented below suggest that, in certain cases, a more careful modeling of learning may help to significantly reduce operating costs. In particular, when the learning curve is steep and the cost of buying incremental, flexible capacity is high, then more sophisticated policies can provide significant cost savings.

In the following numerical example, each time period corresponds to a three-month quarter, and service requirements are assumed to be stationary and specified as total capacity requirements (e.g. number of calls per quarter).

We assume that there are two types of employees, slow and fast. A state (S, F) then means that there are S slow and F fast employees available. The more experienced, fast employees have a fixed capacity increase over newer, slow employees. In our numerical experiments we systematically vary the percentage increase in capacity that comes with learning from 20% to 80%.

Other parameters related to system dynamics are as follows. New (slow) employees have a 48% annual turnover rate (15% quarterly), and experienced (fast) employees turn over at 34% annually (10% per quarter). In all examples 100% of the new employees that do not quit become fast. Thus, after an initial three-month break-in period, learning is deterministic. In the examples, we also restrict employee numbers – hired, learning, and turning over – to be integral.

The example's cost parameters are constructed to be consistent with data from actual call center operations. On average the fully loaded annual wage is $25,978.48 for an employee with annual processing capacity of 20,000. Employees with other capacities are paid proportionally. Overtime is paid at a 50% premium, and outsourcing cost is set to be extremely high so that it will not be a viable option. There is also a $1000 fixed cost for each employee hired.

In our numerical experiments, we also systematically vary the maximum amount of overtime available to the organization as percentage of the total regular time available. Typically work rules nominally allow this limit to be 50%. If, for example, an employee normally works a 40-hour week, then the maximum overtime available would be 20 additional hours. In practice, however, the actual percentage of overtime available is limited by other constraints imposed by the weekly scheduling problem. In particular, daily or weekly peaks in demand for service often make the use of overtime, which is available at off-peak times, an ineffective means of adding marginal capacity. In our examples, we run three scenarios for the maximum available overtime: 10%, 30% and 50%.

We consider two alternative hiring policies. In one learning is explicitly considered, and the optimal hiring number is a function of the number of each of the two types of employees. In the other learning is ignored, and the optimal hiring number depends only on the total number of employees. We call this a "single-number" policy. Note that for both policies the actual system dynamics remain the same: different types of employees still turn over differently and provide capacities accordingly. The difference is that the information about different employee numbers is tracked in one policy, while it is ignored in the other policy.

We develop MDP solution procedures for both of these policies and compare their optimal costs, varying systematically the percentage of capacity increase from slow to fast, and the maximum overtime percentage. A cost comparison is shown in Figure 2.

It is clear from Figure 2 that when there is plenty of overtime (30% or 50%), no matter how much the capacity increase is from slow to fast, the single-number hiring policy, which treats all different types of employees as the same, performs almost as well as the optimal policy. When
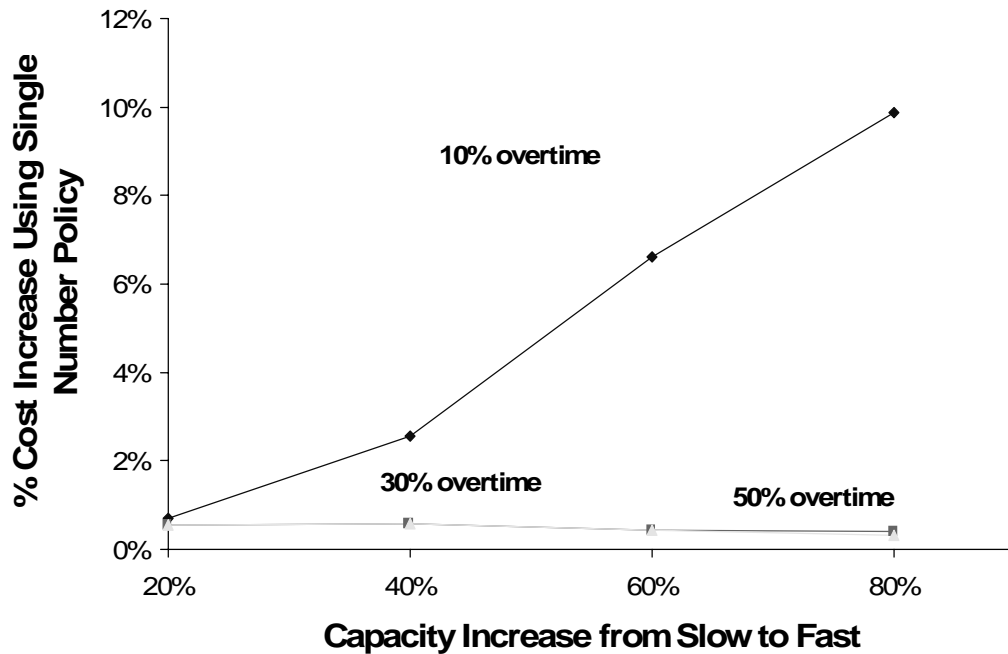
Figure 2: Difference in Average Cost between Optimal and Single-Number Policies

overtime availability is tight (10%), however, the difference becomes significant as the capacity increase gets larger.

Figure 3 shows that, for the case of 80% capacity increase (with learning) and a 10% overtime maximum, the bulk of the cost increase is due to increased staff levels under the single-number hiring policy. Indeed, the single-number policy systematically over-staffs relative to the staffing levels of the optimal plan.

Because outsourcing costs are prohibitively high, both the optimal and the single-number policy seek to ensure, with probability close to one, that there exist enough people so that the service requirement can be met from regular capacity and overtime. While the optimal policy can recognize differences between fast and slow employees, the single-number policy does not. To ensure that there exists enough capacity on hand, the single-number policy must hire as if (nearly) all of the workers on hand were slow.

Figures 4 and 5 detail the equilibrium distribution of capacity for both policies under two scenarios, 80% capacity increase and 50% overtime maximum and 80% capacity increase and 10% overtime maximum. Both figures show that the left tails of both policies' capacity distributions
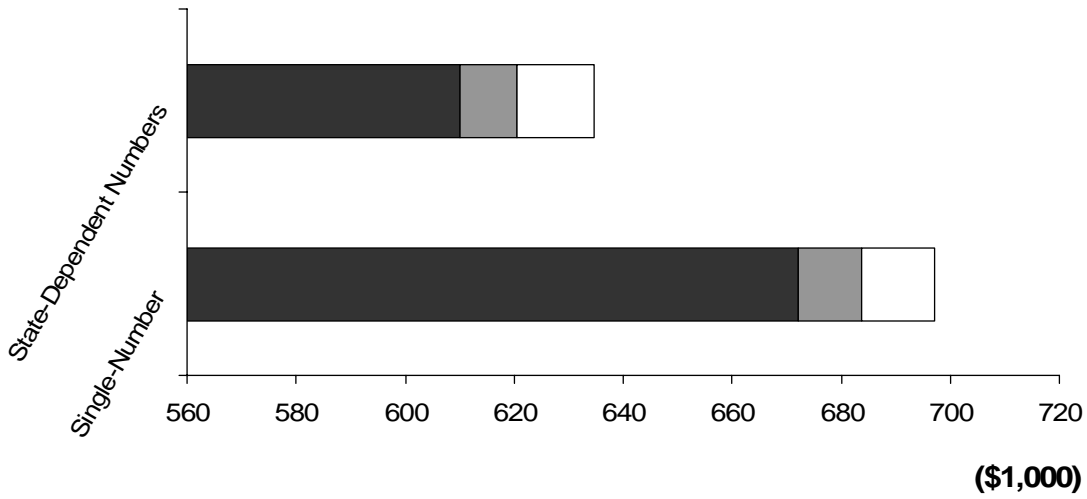
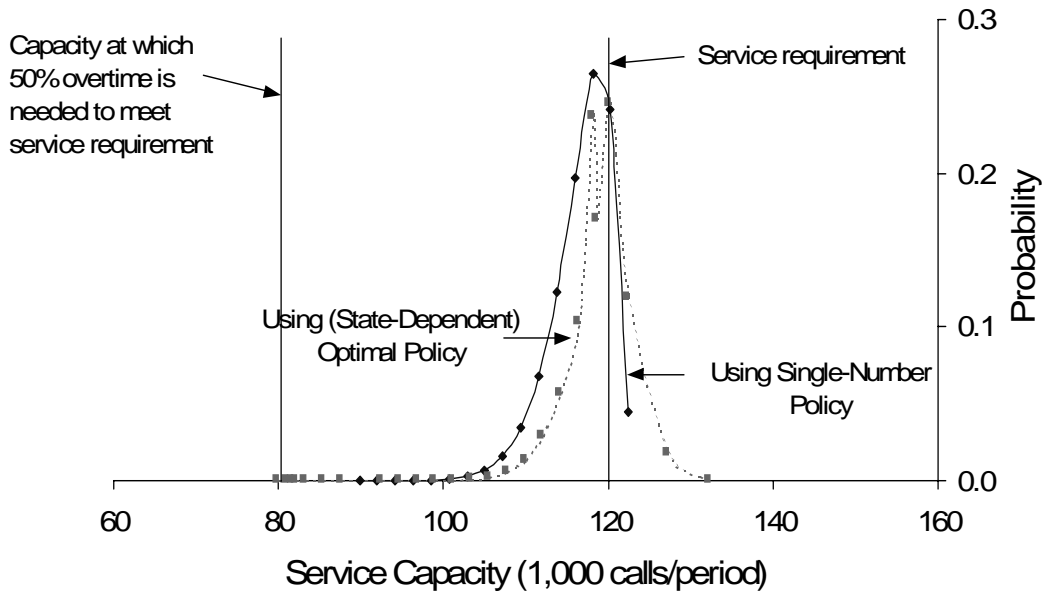Figure 3: Cost Breakdown of Optimal and Single-number Policies

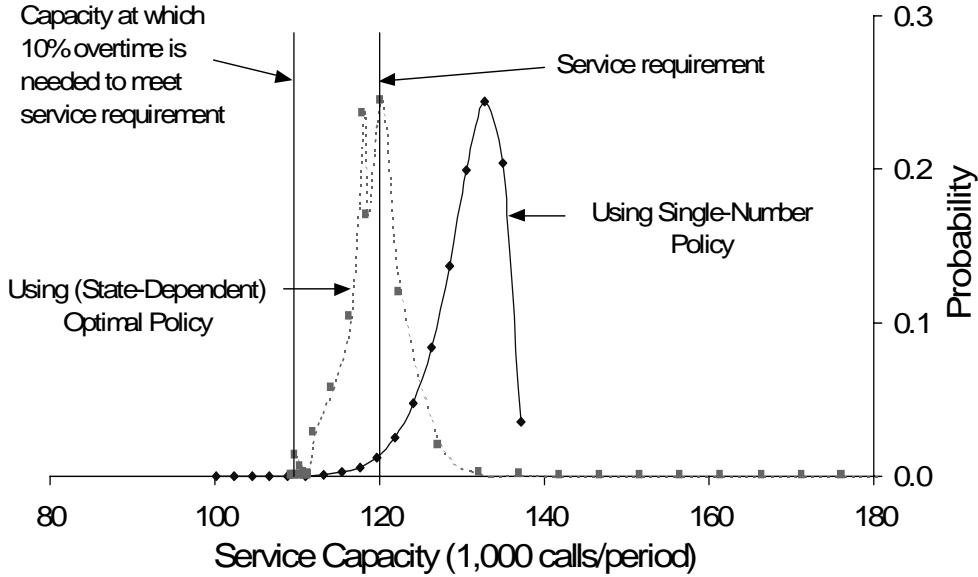Figure 4: Distribution of Capacity when Maximum Overtime is 50%

29

Figure 5: Distribution of Capacity when Maximum Overtime is 10%

trail off at the maximum overtime percentage allowed.

Figure 4 also shows that, in the 50% overtime case, both the optimum policy and the single-number policy consistently use overtime. Here, the probability that the single-number policy has excess capacity is low, and the relatively flat tradeoff between hiring more workers and using more overtime evidently allows the single-number policy to perform well.

In Figure 5, however, the single-number policy consistently maintains costly excess capacity. Here the optimal policy is able to maintain it's staffing level roughly at the service requirement, where the operating cost function appears to be flat.

In conclusion, our numerical studies imply that the cost increase incurred by using a single-number policy instead of the optimal, state-dependent policy increases with the percentage increase in service capacity from slow to fast. But more importantly, the availability of flexible capacity (in the form of overtime or part-time labor) in the organization is the best indicator of how well the single-number policy will perform.

Remark 7 We have performed all of the analyses in this section twice. First we have used multi-nomial learning and turnover probabilities (see §3.2 for discussion). Then we have developed

30

a stochastic proportion approximation, which samples from beta distributions to determine the turnover proportions. In each simulation, the parameters of the beta distribution were chosen to match the mean and variance of the aggregate turnover rate (across all states in steady-state) obtained using multinomial distribution. In all the cases we have considered, the cost difference between using the multinomial distribution and its stochastic proportion approximation has been within 0.5%. All the numerical results presented in this paper were generated using stochastic proportions.

8.3 Two-sided Smoothing Effect

In §7, we prove that in two special cases, when the state space can be collapsed into one dimension and the service requirements are periodic, there exists a one-sided smoothing effect. In these cases, when the optimal myopic hire-up-to number goes down from one time period to the next, so does the true optimal number. But when the myopic number goes up, the true optimal number may still go down if the myopic number further goes down in the following period. In the presence of learning, however, this one-sided ("upside") smoothing need not hold.

In Figure 6 we give an example in which there exists a "downside" smoothing effect. In this example, capacity improvement from slow to fast is 40% and maximum overtime is 10%. The periodic service requirements are {20,000, 30,000, 20,000, 120,000}, and the figure compares the myopic hire-up-to number for state (0,5) to the optimal hire-up-to number for the same state. We see that from period 2 to period 3, even though the myopic hire-up-to levels goes down (reflecting the decrease in service requirement), the true optimal number goes up, in anticipation of a further increase of service requirement in the following period. Kapuściński and Tayur [17] observe a similar two-sided smoothing effect in a capacitated production-inventory system. In our system, the downside smoothing effect is mostly caused by employee learning and consequent capacity increase, while in their setting, it is caused by the limit on production capacity.

9 Conclusion

The MDP approach to the employee staffing problem allows for direct modeling of the stochastic nature of learning and turnover, and it yields structural results that provide insight into system dynamics and control. Our numerical results show that, when the learning curve is steep and
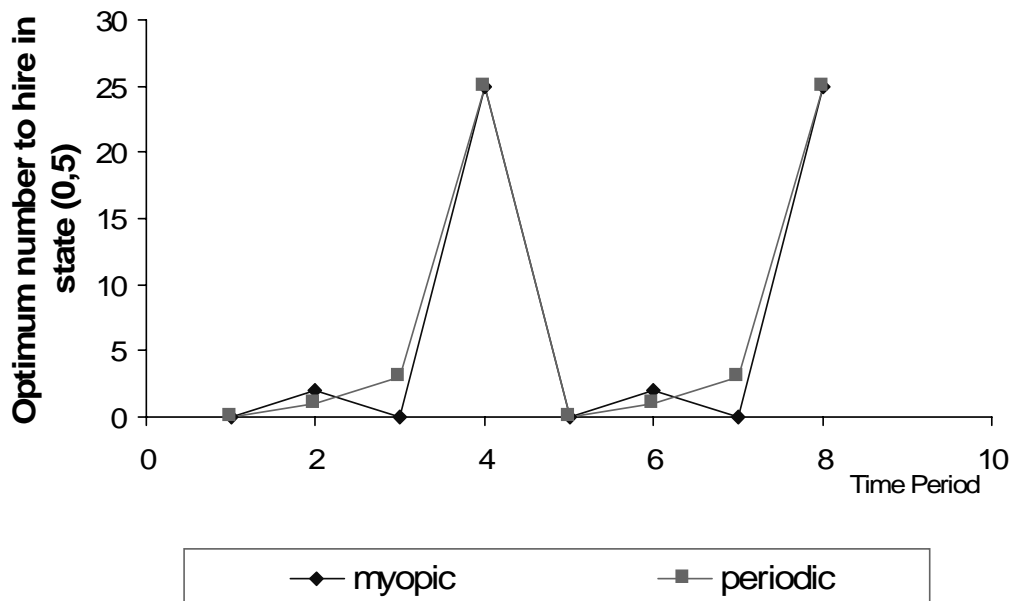
Figure 6: Downside Smoothing Effect: optimal numbers to hire in state (0,5) when the service requirements are 4-periodic: 20000, 30000, 20000, 120000.

capacity flexibility is limited, the more careful control afforded by this approach may provide significant cost benefits.

The optimal hiring policy does not yield to easy or fast computation, however. The MDP's state space explodes with the size of the work force and the number of employee types (speeds, skills). A natural extension of our current work would be to identify and evaluate effective heuristics that are more computationally efficient. For example, when learning is not dramatic, variants of the myopic policies of §6 may work well.

Another approach would be to search for optimal hire-up-to numbers using simulation-based optimization techniques. For example, Kapuściński and Tayur [17] have shown these techniques to be effective for finding optimal "up-to" levels in capacitated, multi-echelon production problems. Additional work must be done in our case, however, to validate and evaluate the effectiveness of the technique.

A   Proof of Theorem 1

Proof of part (i) First, let T be finite and fixed. We use induction.

For $t = T + 1$, formulation (5) implies $V_{T+1}(n_{1,T+1}, \ldots, n_{m,T+1}|T) \equiv 0$. Then $V_{T+1}(\cdot|T)$ is (trivially) jointly convex in $(n_{1,T+1}, \ldots, n_{m,T+1})$. Next, suppose $V_{t+1}(n_{1,t+1}, n_{2,t+1}, \ldots, n_{m,t+1}|T)$ is jointly convex in $(n_{1,t+1}, n_{2,t+1}, \ldots, n_{m,t+1})$. Then in two steps we prove that $J_t(y_t, \ldots, n_{m,t}|T)$ and $V_t(n_{1,t}, \ldots, n_{m,t}|T)$ are, in turn, jointly convex in $(y_t, \ldots, n_{m,t})$ and $(n_{1,t}, \ldots, n_{m,t})$ respectively.

Step 1. We first prove that $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$ is jointly convex in $(y_t, n_{2,t}, \ldots n_{m,t})$. Due to the assumption that $O_t(y_t, n_{2,t}, \ldots, n_{m,t})$ is jointly convex in $(y_t, n_{2,t}, \ldots, n_{m,t})$, $H_t(y_t, n_{2,t}, \ldots, n_{m,t})$ is also jointly convex in $(y_t, n_{2,t}, \ldots, n_{m,t})$. Because the integral of a convex function is again convex, we only need to prove that $V_{t+1}(n_{1,t+1}, n_{2,t+1}, \ldots, n_{m,t+1}|T)$ is jointly convex in $(y_t, n_{2,t}, \ldots, n_{m,t})$ for any realization of $(\tilde{q}_{1,t}, \ldots, \tilde{q}_{m,t}, \tilde{l}_{1,t}, \ldots \tilde{l}_{m-1,t})$. This is true due to Lemma 1 and the fact – from (2), (3), and (4) – that for fixed $(\tilde{q}_{1,t}, \ldots, \tilde{q}_{m,t}, \tilde{l}_{1,t}, \ldots, \tilde{l}_{m-1,t})$, $n_{1,t+1}, n_{2,t+1}, \ldots, n_{m,t+1}$ are all linear in $(y_t, n_{2,t}, \ldots, n_{m,t})$.

Step 2. Now, given that $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$ is jointly convex in $(y_t, n_{2,t}, \ldots n_{m,t})$, we prove that $V_t(n_{1,t}, \ldots, n_{m,t}|T)$ is jointly convex in $(n_{1,t}, \ldots, n_{m,t})$. This follows from a direct application of Lemma 2. Note that $-h n_{1,t} + \sum_{i=2}^{m} W_i n_{i,t}$ is jointly convex in $(n_{1,t}, \ldots, n_{m,t})$.

Therefore, by repeating steps 1 and 2, we know that $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$ is jointly convex in $(y_t, n_{2,t}, \ldots n_{m,t})$ and $V_t(n_{1,t}, \ldots, n_{m,t}|T)$ is jointly convex in $(n_{1,t}, \ldots, n_{m,t})$ for all t.

As a result of $J_t(n_{1,t}, \ldots, n_{m,t}|T)$ being jointly convex in $(n_{1,t}, \ldots, n_{m,t})$ for all t we know that there exists $y_t^*(n_{2,t}, \ldots, n_{m,t}|T)$ such that $y_t^*(n_{2,t}, \ldots, n_{m,t}|T)$ minimizes $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$ without the constraint $y_t \geq n_{1,t}$. When there are multiple minimizers to $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$, we will take the smallest. Furthermore, given any $(n_{1,t}, n_{2,t}, \ldots, n_{m,t})$, because $J_t(y_t, n_{2,t}, \ldots n_{m,t}|T)$ is convex, the optimal hiring policy with the constraint will be to hire up to $y_t^*(n_{2,t}, \ldots, n_{m,t}|T)$ if $n_{1,t} < y_t^*(n_{2,t}, \ldots, n_{m,t}|T)$, and not hire (and thus remain $n_{1,t}$) otherwise. This is exactly the "hire-up-to" policy of Definition 1.

$\triangle$

Proof of part (ii) We will need the following lemma.

Lemma 3 $V_t(\cdot|T)$ *is increasing in* T.

Proof of Lemma 3   Pick any two cost functions $V_{t+1}^1(\cdot)$ and $V_{t+1}^2(\cdot)$ such that $V_{t+1}^1(n_{1,t+1}, \ldots, n_{m,t+1}|T) \leq$
$V_{t+1}^2(n_{1,t+1}, \ldots, n_{m,t+1}|T)$ for any $(n_{1,t+1}, \ldots, n_{m,t+1})$. Moreover, let the minima of $V_t^1(\cdot)$ and $V_t^2(\cdot)$
be achieved at y_t^{*1} and y_t^{*2} respectively. Then

$$V_t^1(n_{1,t}, \ldots, n_{m,t}|T)$$
$$= \quad H_t(y_t^{*1}, n_{2,t}, \ldots, n_{m,t}) + \alpha \mathbf{E}_{\{\tilde{q}_{1,t}, \ldots, \tilde{l}_{1,t}, \ldots\}} \left\{ V_{t+1}^1(n_{1,t+1}^1, \ldots, n_{m,t+1}^1|T) \right\} - hn_{1,t} + \sum_{i=2}^m W_i n_{i,t}$$
$$\leq \quad H_t(y_t^{*2}, n_{2,t}, \ldots, n_{m,t}) + \alpha \mathbf{E}_{\{\tilde{q}_{1,t}, \ldots, \tilde{l}_{1,t}, \ldots\}} \left\{ V_{t+1}^1(n_{1,t+1}^2, \ldots, n_{m,t+1}^2|T) \right\} - hn_{1,t} + \sum_{i=2}^m W_i n_{i,t}$$
$$\leq \quad H_t(y_t^{*2}, n_{2,t}, \ldots, n_{m,t}) + \alpha \mathbf{E}_{\{\tilde{q}_{1,t}, \ldots, \tilde{l}_{1,t}, \ldots\}} \left\{ V_{t+1}^2(n_{1,t+1}^2, \ldots, n_{m,t+1}^2|T) \right\} - hn_{1,t} + \sum_{i=2}^m W_i n_{i,t}$$
$$= \quad V_t^2(n_{1,t}, \ldots, n_{m,t}|T).$$

Therefore, the MDP value iteration is an increasing mapping. When T is increased to $T + 1$,
the values of $V_t(\cdot|T)$ change as if the end-of-horizon cost function had been increased from 0 to a
positive function. Therefore, by applying the monotonicity of the mapping repeatedly, we find that
the function $V_t(\cdot|T)$ also increases.

$\triangle$

It is not difficult to see that, because total one-period cost is bounded by K (due to Assumption
2), $V_t(\cdot|T) \leq \frac{K}{1-\alpha}$ for any T and $t \leq T$.

Thus $\{V_t(\cdot|T) : T = 0, 1, \ldots\}$ is a monotone, bounded sequence. Because monotone bounded
sequences always converge, if we let $T \to \infty$, then

$$\lim_{T \to \infty} V_t(n_{1,t}, \ldots, n_{m,t}|T) = U_t(n_{1,t}, \ldots, n_{m,t}) \tag{15}$$

for some finite function $U_t(n_{1,t}, \ldots, n_{m,t})$.   Moreover, from Theorem 10.8 in Rockafellar [27],
$V_t(n_{1,t}, \ldots, n_{m,t}|T)$ converges to $U_t(n_{1,t}, \ldots, n_{m,t})$ uniformly, and $U_t(n_{1,t}, \ldots, n_{m,t})$ is jointly convex
in its arguments. By applying Lebesgue's Dominated Convergence Theorem, we obtain

$$U_t(n_{1,t}, \ldots, n_{m,t})$$
$$= \quad \min_{y_t \geq n_{1,t}} \left\{ H_t(y_t, n_{2,t}, \ldots, n_{m,t}) + \alpha \mathbf{E}_{\{\tilde{q}_{1,t}, \ldots, \tilde{l}_{1,t}, \ldots\}} \left\{ U_{t+1}(n_{1,t+1}, \ldots, n_{m,t+1}) \right\} \right\} - hn_{1,t} + \sum_{i=2}^m W_i n_{i,t}.$$

Hence these $U_t(\cdot)$'s satisfies the MDP optimality equation. As a result, they are the infinite-
horizon cost functions. That is,

$$V_t(n_{1,t}, \ldots, n_{m,t}|\infty) = U_t(n_{1,t}, \ldots, n_{m,t}) = \lim_{T \to \infty} V_t(n_{1,t}, \ldots, n_{m,t}|T).$$

Therefore, the infinite-horizon cost functions $V_t(\cdot|\infty)$ are jointly convex, and the "hire-order-to" policy is, again, optimal in the infinite-horizon case.

$\triangle$

B   Proof of Theorem 4

In the following proof we will assume $\lambda = 0$. When $\lambda > 0$, we need to modify the hiring cost to be $\alpha^{-\lambda} h$, but the proof is similar.

Let $V_t(n_t; D_t, D_{t+1}, \ldots)$ denote the total discounted cost if the current staffing level is n_t and the demand stream from now on is $D_t, D_{t+1}, \ldots$ Then

$$
\begin{aligned}
V_t(n_t; D_t, D_{t+1}, \ldots) &= \min_{y_t \geq n_t} \left\{ (h + W)y_t + O_t(y_t; D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}} \left\{ V_{t+1}(\tilde{r}_t y_t; D_{t+1}, D_{t+2}, \ldots) \right\} \right\} - h n_t \\
&= \min_{y_t \geq n_t} \left\{ J_t(y_t; D_t, D_{t+1}, \ldots) \right\} - h n_t.
\end{aligned}
$$

To prove parts (i) and (ii), we need the following lemmas. Lemma 4 shows that when the service requirements are higher, the hire-up-to levels are higher. Lemma 5 shows that the myopic policies always hire at least as many employees as the optimal policy. This is natural because the myopic policy assumes a higher end-of-period "salvage value": it assumes that every employee that stays will be needed in the next period. Lemma 6 shows that if the optimal hiring number goes up in the next period, then the myopic policy is optimal in this period.

Lemma 4 *Let* $\phi, \phi, \ldots$ *and* $\chi, \chi, \ldots$ *be two sequences of stationary service requirements. If* $\phi \preceq \chi$, *then* $y_t^*(\phi, \phi, \ldots) \leq y_t^*(\chi, \chi, \ldots) \ \forall t$;

Proof   For stationary service requirements, Theorem 2 states that the myopic policy of minimizing $G(\cdot)$ in each period is optimal. As a result, $\lim_{y_t \to y_t^{*-}(\phi, \phi, \ldots)} G_t'(y_t; \phi) = 0$ and $\lim_{y_t \to y_t^{*-}(\chi, \chi, \ldots)} G_t'(y_t; \chi) = 0$. According to Assumption 3, $\phi \preceq \chi$ implies $O_t'(\cdot; \phi) \geq O_t'(\cdot; \chi)$, and hence $G_t'(\cdot; \phi) \geq G_t'(\cdot; \chi)$. The convexity of $G(\cdot)$ then immediately implies $y_t^*(\phi, \phi, \ldots) \leq y_t^*(\chi, \chi, \ldots) \ \forall t$.

$\triangle$

Lemma 5 *For any* $\{D_t, D_{t+1}, \ldots\}$, $y_t^*(D_t, D_{t+1}, \ldots) \leq y_t^*(D_t, D_t, \ldots)$;

Proof   We first note that

$$V'_t(n_t; D_t, D_{t+1}, \ldots) = \begin{cases} -h & \text{if } n_t \le y^*_t(D_t, D_{t+1}, \ldots) \\ -h + J'_t(n_t; D_t, D_{t+1}, \ldots) & \text{if } n_t > y^*_t(D_t, D_{t+1}, \ldots). \end{cases}$$

Therefore $V'_t(n_t; D_t, D_{t+1}, \ldots) \ge -h$ for all t, n_t, and $\{D_t, D_{t+1}, \ldots\}$.

We know that $y^*_t(D_t, D_t, \ldots)$ satisfies

$$(h + W) + O'_t(y^*_t(D_t, D_t, \ldots); D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\{-h\tilde{r}_t\} = 0, \tag{16}$$

and $y^*_t(D_t, D_{t+1}, \ldots)$ satisfies

$$\begin{aligned} 0 &= (h + W) + O'_t(y^*_t(D_t, D_t+1, \ldots); D_t) \\ &\quad + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\left\{V'_{t+1}(\tilde{r}_t y^*_t(D_t, D_{t+1}, \ldots); D_{t+1}, D_{t+2}, \ldots)\tilde{r}_t\right\} \\ &\ge (h + W) + O'_t(y^*_t(D_t, D_{t+1}, \ldots); D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\{-h\tilde{r}_t\}, \end{aligned} \tag{17}$$

because $V'_{t+1}(n_t; D_{t+1}, D_{t+2}, \ldots) \ge -h$ for all n_t.

Now it is obvious from (16) and (17) that $O'_t(y^*_t(D_t, D_{t+1}, \ldots); D_t) \le$ $O'_t(y^*_t(D_t, D_t, \ldots); D_t)$. Hence $y^*_t(D_t, D_{t+1}, \ldots) \le y^*_t(D_t, D_t, \ldots)$.

$\triangle$

Lemma 6 If $y^*_t(D_t, D_{t+1}, \ldots) \le y^*_{t+1}(D_t+1, D_{t+2}, \ldots)$, then $y^*_t(D_t, D_{t+1}, \ldots) = y^*_t(D_t, D_t, \ldots)$.

Proof   We know that $J'_t(y^*_t(D_t, D_{t+1}, \ldots); D_t, D_{t+1}, \ldots) = 0$, which is

$$(h + W) + O'_t(y^*_t(D_t, D_t+1, \ldots); D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\{V'_{t+1}(\tilde{r}_t y^*_t(D_t, D_{t+1}, \ldots); D_{t+1}, D_{t+2}, \ldots)\tilde{r}_t\} = 0. \tag{18}$$

Since $y^*_t(D_t, D_{t+1}, \ldots) \le y^*_{t+1}(D_{t+1}, D_{t+2}, \ldots)$, $\tilde{r}_t y^*_t(D_t, D_{t+1}, \ldots) \le y^*_{t+1}(D_{t+1}, D_{t+2}, \ldots)$ with probability 1. Moreover, we know $V'_{t+1}(x; D_{t+1}, D_{t+2}, \ldots) = -h$ if $x \le y^*_{t+1}(D_{t+1}, D_{t+2}, \ldots)$, therefore $(h + W) + O'_t(y^*_t(D_t, D_{t+1}, \ldots); D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\{-h\tilde{r}_t\} = 0$. Similarly, we also know that $(h + W) + O'_t(y^*_t(D_t, D_t, \ldots); D_t) + \alpha \mathbf{E}_{\{\tilde{r}_t\}}\{-h\tilde{r}_t\} = 0$. So $y^*_t(D_t, D_{t+1}, \ldots) = y^*_t(D_t, D_t, \ldots)$, because $O_t(\cdot)$ is convex. (If there are multiple roots we choose the smallest one so these two should still be equal).

$\triangle$

Proof of Theorem 4   Now suppose, by contradiction, that part (i) is false, i.e., $D_t \succeq D_{t+1}$ and $y_t^*(D_t, D_{t+1}, \ldots) < y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots)$, then by Lemmas 6, 4, and 5, $y_t^*(D_t, D_{t+1}, \ldots) = y_t^*(D_t, D_t, \ldots) \geq y_{t+1}^*(D_{t+1}, D_{t+1}, \ldots) \geq y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots)$. A contradiction. Similarly, suppose, by contradiction, that part (ii) is false, i.e. $D_t \preceq D_{t+1}$, $y_t^*(D_t, D_{t+1}, \ldots) > y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots)$, and $y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots) \leq y_{t+2}^*(D_{t+2}, D_{t+3}, \ldots)$. Then, again by Lemma 6, $y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots) = y_{t+1}^*(D_{t+1}, D_{t+1}, \ldots)$, and by Lemmas 5 and 4, $y_t^*(D_t, D_{t+1}, \ldots) \leq y_t^*(D_t, D_t, \ldots) \leq y_{t+1}^*(D_{t+1}, D_{t+1}, \ldots) = y_{t+1}^*(D_{t+1}, D_{t+2}, \ldots)$. A contradiction.

$\triangle$

References

[1] Kenneth J. Arrow, Samuel Karlin, and Herbert Scarf. *Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, Stanford, CA, 1958.

[2] Yossi Aviv and Awi Federgruen. Stochastic inventory models with limited production capacity and periodically varying parameters. *Probability in the Engineering and Informational Sciences*, 11(1):107–135, 1997.

[3] D.J. Bartholomew, A.F. Forbes, and S.I. McClean. *Statistical Techniques for Manpower Planning.* John Wiley and Sons, New York, 2 edition, 1991.

[4] Oded Berman, Richard C. Larson, and Edieal Pinker. Scheduling workforce and workflow in a high volume factory. *Management Science*, 43(2):158–172, February 1997.

[5] Sanjeev Bordoloi and Hirofumi Matsuo. Human resource planning in knowledge-intensive operations: A model for learning with stochastic turnover. Working paper, University of Texas at Austin, September 1998.

[6] Ronald J. Ebert. Aggregate planning with learning curve productivity. *Management Science*, 23(2), October 1976.

[7] Cheryl Gaimon and Gerald L. Thompson. A distributed parameter cohort personnel planning model that uses cross-sectional data. *Management Science*, 30(6):750–764, June 1984.

[8] Yigal Gerchak, Mahmut Parlar, and S. Sankar Sengupta. On manpower planning in the presence of learning. *Engineering Costs and Production economics*, 20:295–303, 1990.

[9] Yigal Gerchak and Yunzeng Wang. Periodic-review inventory models with inventory-level-dependent demand. *Naval Research Logistics*, 41:99–116, 1994.

[10] Richard C. Grinold and Robert E. Stanford. Optimal control of a graded manpower system. *Management Science*, 20(8):1201–1216, April 1974.

[11] H. W. Gustafson. Force-loss cost analysis. In W. H. Mobley, *Employee Turnover: Causes, Consequences, and Control.* Addison-Wesley, Reading, Massachusetts, 1982.

[12] Mordechai Henig and Yigal Gerchak. The structure of periodic review policies in the presence of random yield. *Operations Research*, 38(4):634–643, 1990.

[13] Daniel P. Heyman and Matthew J. Sobel. *Stochastic Models in Operations Research Volume II: Stochastic Optimization.* McGraw-Hill Book Company, 1984.

[14] C.C. Holt, F. Modigliani, J. Muth, and H.A. Simon. *Planning Production, Inventory and Work Force.* Prentice-Hall, Englewood Cliffs, N.J., 1960.

[15] Donald L. Iglehart and Stratton C. Jaquett. Multi-class inventory models with demand a function of inventory level. *Naval Research Logistics Quarterly*, 16:495–502, 1969.

[16] B. Jovanovic. Job matching and the theory of turnover. *Journal of Political Economy*, 87:972–990, 1979.

[17] Roman Kapuściński and Sridhar Tayur. A capacitated production-inventory model with periodic demand. *Operations Research*, 46(6):899–911, November-December 1998.

[18] Samuel Karlin. The structure of dynamic programming models. *Naval Research Logistics Quarterly*, 2:285–294, 1955.

[19] Samuel Karlin. Dynamic inventory policy with varying stochastic demands. *Management Science*, 6(3):231–258, 1960.

[20] Samuel Karlin. Optimal policy for dynamic inventory process with stochastic demands subject to seasonal variations. *Journal of SIAM*, 8(4):611–629, 1960.

[21] Samuel Karlin. *Total Positivity*, volume 1. Stanford University Press, Stanford, CA, 1968.

[22] Behrokh Khoshnevis and Philip M. Wolfe. A short-cycle product aggregate planning model incorporating improvement curve productivity. *Engineering Costs and Production Economics*, 10:217–233, 1986.

[23] M.G. Orrbeck, D.R. Schuette, and H.E. Thompson. The effect of worker productivity on production smoothing. *Management Science*, 14(6), February 1968.

[24] D. O. Parsons. Quit rates over time: a search and information approach. *American Economic Review*, 63:390–401, 1973.

[25] Edieal Pinker and Robert Shumsky. The efficiency-quality tradeoff of crosstrained workers. To appear in *Manufacturing and Service Operations Management*, University of Rochester, 1999.

[26] Martin L. Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, NY, 1994.

[27] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.

[28] Louis E. Yelle. The learning curve: historical review and comprehensive survey. *Decision Sciences*, 10:302–328, 1979.

[29] Paul Zipkin. Critical number policies for inventory models with periodic data. *Management Science*, 35(1):71–80, 1989.