Math. Meth. Oper. Res. (2005) 62: 499–514
DOI 10.1007/s00186-005-0033-5

ORIGINAL ARTICLE

Hyun-Soo Ahn · Rhonda Righter
J. George Shanthikumar

Staffing decisions for heterogeneous workers with turnover

Received: January 2005 / Revised version: March 2005 / Published online: 16 November 2005

Abstract In this paper we consider a firm that employs heterogeneous workers to meet demand for its product or service. Workers differ in their skills, speed, and/or quality, and they randomly leave, or turn over. Each period the firm must decide how many workers of each type to hire or fire in order to meet randomly changing demand forecasts at minimal expense. When the number of workers of each type can by continuously varied, the operational cost is jointly convex in the number of workers of each type, hiring and firing costs are linear, and a random fraction of workers of each type leave in each period, the optimal policy has a simple hire-up-to/fire-down-to structure. However, under the more realistic assumption that the number of workers of each type is discrete, the optimal policy is much more difficult to characterize, and depends on the particular notion of discrete convexity used for the cost function. We explore several different notions of discrete convexity and their impact on structural results for the optimal policy.

Keywords optimal staffing · discrete optimization · dynamic programming · multimodularity

---

H-S. Ahn
Operations and Management Science,
University of Michigan Business School,
701 Tappan St.
Ann Arbor, MI 48109-1234, USA
E-mail: hsahn@umich.edu

R. Righter (✉) · J.G. Shanthikumar
Department of Industrial Engineering and Operations Research,
University of California,
Berkeley, CA 94720, USA
E-mail: rrighter@ieor.berkeley.edu, shanthikumar@ieor.berkeley.edu

1 Introduction

It is a great privilege to participate in this workshop and special volume in honor of Arie Hordijk. He is a wonderful mathematician, with great insight and creativity, and he has had a huge impact on the applied probability community. Especially for the second author, Arie has been a role model, mentor, and friend.

We consider a firm that employs heterogeneous wokers to meet demand for its product or service. Workers differ in their skills, speed, and/or quality, and they randomly leave, or turn over. Each period the firm must decide how many workers of each type to hire or fire in order to meet randomly changing demand forecasts at minimal expense. When the number of workers of each type can be continuously varied, the operational cost is jointly convex in the number of workers of each type, hiring and firing costs are linear, and a random fraction of workers of each type leave in each period, the optimal policy has a simple hire-up-to/fire-down-to structure. However, under the more realistic assumption that the number of workers of each type is discrete, the optimal policy is much more difficult to characterize, and depends on the particular notion of discrete convexity used for the cost function. We explore several different notions of discrete convexity and their impact on structural results for the optimal staffing policy.

We model our system as a discrete-time Markov decision process (MDP). Let $\mathbf{n}_t = (n_{1t}, \ldots, n_{mt})$ be a non-negative m-dimensional vector that represents the current (at time t) number of workers of each of m types before hiring decisions are made. Let θ_t represent the current state of the environment. The environment may affect the distribution of the demand during the period, the pool of available employees from which we may hire, the probabilities that employees will leave, and the costs we incur. At the beginning of the period, based on $\mathbf{n}_t$ and θ_t, the firm must decide how many workers of each type to hire (or fire). Let $\mathbf{d}_t = (d_{1t}, \ldots, d_{mt})$ represent our hiring (firing if $d_{it} < 0$) decisions at time t, where $\mathbf{d}_t$ is a function of $\mathbf{n}_t$ and θ_t though we suppress the dependence notationally. Let $\mathbf{N}_{t+1}$ be the worker vector at the end of the period (beginning of the next period), which is a random function of $\mathbf{n}_t + \mathbf{d}_t$ and θ_t. Using α as the one-period discount factor, our objective is to minimize the total expected discounted cost,

$$\begin{aligned} V_t(\mathbf{n}_t, \theta_t) &= \min_{\mathbf{d}_t \geq -\mathbf{n}_{tt}} E \sum_{j=t}^{T} \alpha^j E\left[c_t(\mathbf{N}_j, \mathbf{d}_j, \theta_j) + C_t(\mathbf{N}_j + \mathbf{d}_j, \theta_j) | \mathbf{n}_t, \mathbf{d}_t, \theta_t\right] \\ &= \min_{\mathbf{d}_t \geq -\mathbf{n}_t} [c_t(\mathbf{n}_t, \mathbf{d}_t, \theta_t) + C_t(\mathbf{n}_t + \mathbf{d}_t, \theta_t) \\ &\quad + \alpha E[V_{t+1}(\mathbf{N_{t+1}}(\mathbf{n}_t + \mathbf{d}_t, \theta_t), \theta_{t+1}) | \mathbf{n}_t, \mathbf{d}_t, \theta_t]], \end{aligned}$$

where c represents our hiring and firing costs, and C represents our expected operational costs under the optimal (or possibly heuristic) operational policy for meeting demand during the period, including wages for workers. We will assume linear hiring and firing costs, so

$$c_t(\mathbf{n}_t, \mathbf{d}_t, \theta_t) = \sum_i [h_{it}(\theta_t) d_i^+ + f_{it}(\theta_t) d_i^-]$$

for all $\mathbf{n}_t$, where $h_{it}(\theta_t) > 0$ and $f_{it}(\theta_t) > 0$, and $d_i^+ = \max\{d_i, 0\}$ and $d_i^- = \max\{-d_i, 0\}$. Then, using the arguments of Dixit (1997), Eberly and van Mieghem

(1997), Gans and Zhou (2002), and Schmidt and Nahmias (2003, preprint), we can show the following, where $\mathbf{n}^i = (n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_m)$.

Theorem 1 *Suppose n_i and d_i, and hence $y_i = n_i + d_i$ can take on continuous values, i.e. $\mathbf{n}, \mathbf{d}, \mathbf{y} \in \mathbf{R}_+^m$, and $C_t(y_1, \ldots, y_m, \theta)$ is jointly convex in y for all θ. Also suppose that $N_{it}(y_{it}, \theta_t) = R_{it}(\theta_t) y_{it}$, where $R_{it}(\theta_t)$, $i = 1, 2, \ldots, m$, are independent random fractions with support on* [0, 1]. *Then the optimal policy has the following hire-up-to/fire-down-to structure. For each i there exist two functions $U_{it}(\mathbf{n}^i, \theta) \leq D_{it}(\mathbf{n}^i, \theta)$, such that for a given starting state $\mathbf{n} = \mathbf{n}_t$, for each i, if $n_i < U_{it}(\mathbf{n}^i)$ hire up to $U_{it}(\mathbf{n}^i)$ type i workers, i.e., hire $U_{it}(\mathbf{n}^i) - n_i$ type i workers, if $n_i > D_{it}(\mathbf{n}^i)$ fire down to $D_{it}(\mathbf{n}^i)$ workers, and otherwise do not hire or fire type i workers.*

Dixit (1997) and Eberly and van Mieghem (1997) study a dynamic investment problem where the (continuous) decision in each period is how much to invest or disinvest in each of multiple resources (labor and capital in the case of Dixit 1997), and in which there is no turnover of the resources. They show the optimality of the "Invest/Stay put/Disinvest (ISD)" policy, which invests up to a critical number, or disinvests (fires) down to a critical number, or makes no changes. We will henceforth call the hire-up-to/fire-down-to policy of Theorem 1 the ISD policy. Gans and Zhou (2002), and Schmidt and Nahmias (2003, preprint) consider staffing models in which the resources are workers that may leave, and where the number of workers can take on a continuous range of values. Their models also permit learning, or shifts of workers from one type to another. For Gans and Zhou (2002), only type 1 workers may be hired and none may be fired, and they show the optimality of a hire-up-to policy for type 1 workers. Schmidt and Nahmias (2003, preprint) show the optimality of an ISD policy when there is only one type of worker, and in a special case for two types of workers.

Other research in capacity investment is based on the single-period multi-product newsvendor model, and ignores issues of turnover. See Fine and Freund (1990), Shumsky and Zhang (2003, preprint), van Mieghem and Rudi (2002), van Mieghem (1998, 2002, preprint), Harrison and van Mieghem (1999), Netessine et al. (2002), and the references therein.

The book of Bartholomew et al. (1991) gives an overview of the use of Markov and deterministic models for managing human resources with learning and turnover. Pinker and Shumsky (2000) include experience-based learning, in which the quality of the work depends on the amount of similar work servers have done before. This tends to make flexibility look relatively less attractive because workers who do many different things may not be as good at any particular one. Their model includes a tenure model in which the retention rate depends on the stage of the model, which in turn depends on the worker's tenure. See also Misra et al. (2003, preprint) for a learning model where the problem is to simultaneously determine salesforce size and pricing of products sold. Bordoloi and Matsuo (2001) consider a two-station tandem production model in which new workers work at stage 1, then they learn to do stage 2 work, and then they are considered flexible and can do both. Their model assumes flexible workers train stage 1 and stage 2 workers, and that some workers may leave. They use control theory to determine the number of new workers to hire.

Pinker and Larson (2003) consider both the strategic problem of hiring regular workers and contracting part-time workers at the beginning of the planning horizon along with the operational problem of scheduling part-time workers over time depending on regular worker absenses and current workload. See also Berman and Larson (1994) and Larson and Pinker (2000).

Another approach to staffing problems with heterogeneous workers is based on queueing models, and much of this work has been in the call center context. Refer to Mandelbaum (2003) for an excellent annotated research bibliography on call centers, and see Gans et al. (2003), as well as Koole and Mandelbaum (2002) and Pinedo et al. (2000), for an overview. Perry and Nilsson (1994) use simple queueing approximations to study staffing for servers with heterogeneous skill sets. Koole et al. (2003) consider staffing levels of generalists and specialists assuming overflow routing and no losses. Using stochastic fluid models, Harrison and Zeevi (2003, preprint) reduce the staffing problem to a multi-dimensional newsvendor problem. Borst and Seri (1999) consider heuristics for the combined staffing and routing problem for skill-based routing, using target delays and service levels in terms of probability of delay exceeding a certain threshold. The skill matrix and number of available servers are given. They obtain conditions characterizing the range of reasonable server configurations, and propose two simple credit schemes for assigning calls to servers. Armony and Maglaris (2003) also consider both staffing and routing for a queueing model with a single class of customer, so server "skills" correspond to "speeds" of serving the common customer type. They show that always routing customers to the fastest available server is asymptotically optimal, and then show how to determine staffing levels to minimize costs subject to the constraint that steady-state waiting probabilities cannot exceed a pre-specified level.

We consider discrete-space models, in which the number of workers at any time, and the number to be hired or fired, and the random number that leave, must take integer values. This makes the problem much harder than in the continuous case of Gans and Zhou (2002) and Schmidt and Nahmias (2003, preprint), and we can obtain only partial results.

Theorem 1 follows from the following facts that for continuously valued $\mathbf{n}, \mathbf{y}$, i.e., $\mathbf{n}, \mathbf{y} \in \mathbf{R}_+^m$.

(1) If $f(\mathbf{n})$ is jointly convex in $\mathbf{n}$, then $Ef(\mathbf{N}(\mathbf{n}))$ is jointly convex in $\mathbf{n}$ where $N_i(n_i)$ is an independent random fraction of n_i.
(2) If $f(\mathbf{n}, \mathbf{y})$ is jointly convex in $(\mathbf{n}, \mathbf{y})$, then $\inf_{\mathbf{y} \in \mathbf{A}} f(\mathbf{n}, \mathbf{y})$ is jointly convex in $\mathbf{n}$ where $\mathbf{A} \subset \mathbf{R}_+^m$ is a convex set.

Property (1) says that continuous joint convexity is preserved under random fractional transformations; we will call it a *preservation property*. Property (2) says that continuous joint convexity propagates, after optimization, from one period to the next; we will call it a *propagation property*. We will study various notions of discrete convexity to see what structural results for the optimal policy can be obtained. A key issue for any notion of discrete convexity is whether it has a preservation and a propagation property. Of course, for a discrete model the random turnover transformation cannot be a random fraction. The obvious discrete analogue is a binomial model, $N_i(n_i) \sim Binomial(n_i, p_i)$ for some, possibly random, p_i, and we will explore the implications of this model.

We first see what structural results we can obtain for our discrete-space model without any convexity assumptions. We then study structural, preservation, and propagation properties of componentwise convexity, supermodularity, multimodularity, and directional convexity. Our work was inspired by the elegant work of

Altman, Gaujal, and Hordijk (2000) on multimodularity. We initially hoped that we would be able to develop a discrete analogue of Theorem 1 under the assumption of multimodular costs, but we were able to obtain only partial results. Indeed, for none of the notions of discrete convexity we considered were we able to do all three of: (1) completely characterize the optimal policy for (2) random, binomial, turnover for (3) a multistage problem. That is, with heterogeneous workers, no definition of discrete convexity gave us all three of: full characterization, preservation, and propagation.

We note here that if there were only one type of worker, or if the cost function were separable (effectively reducing the problem to separate single-worker-type problems), then a discrete analogue of Theorem 1 would hold. In particular, the notion of convexity is well-defined and unproblematic, and there is no problem with its propagation over stages or its preservation under binomial transformations (see, e.g., Karlin (1968) or the proof of Lemma 10). However, we are interested in situations where workers are flexible, i.e., some types may partially substitute for others, so there will be interaction terms in any appropriate cost function.

2 Preliminary results for the discrete-space model

We start by developing a partial characterization of the optimal policy without convexity assumptions. In this case, we need not worry about preservation or propagation, so to ease the notational burden, we assume a single-period deterministic model. That is, with the obvious simplifications of our earlier notation,

$$V(\mathbf{n}) = \min_{\mathbf{d} \geq -\mathbf{n}} \left[\sum_i [h_i d_i^+ + f_i d_i^-] + C(\mathbf{n} + \mathbf{d}) \right].$$

We first note that when each n_i is sufficiently large or small, it is reasonable to expect that there is a target value, $\mathbf{y}^*$, such that it is optimal to hire/fire to the target. That is, we are trying to characterize the "corners" of the space, where it will be optimal to hire or fire for each worker type. Each corner can be characterized by a subset of worker types, $\mathcal{H}$, where workers types in $\mathcal{H}$ will be hired, and those types in $\mathcal{H}^c$ will be fired. More precisely, let $\mathcal{S} = \{1, 2, \dots, n\}$ be the set of worker types, let $\mathcal{H} \subset \mathcal{S}$ be the set of worker types that potentially will be hired, and let

$$V_{\mathcal{H}}(\mathbf{n}) := \min_{\{\mathbf{y}: y_i \geq n_i, i \in \mathcal{H}, y_i \leq n_i, i \in \mathcal{H}^c\}} W_{\mathcal{H}}(\mathbf{y}) - \left[\sum_{i \in \mathcal{H}} h_i n_i - \sum_{i \in \mathcal{H}^c} f_i n_i \right],$$

where

$$W_{\mathcal{H}}(\mathbf{y}) := \sum_{i \in \mathcal{H}} h_i y_i - \sum_{i \in \mathcal{H}^c} f_i y_i + C(\mathbf{y}), \quad \mathbf{y} \in \mathbf{Z}_+^m,$$

is a function independent of $\mathbf{n}$. We assume that for each corner characterized by $\mathcal{H}$, there exists at least one target ($k \geq 1$) defined by $\mathbf{y}^*_{(k)}(\mathcal{H}) := \arg\min\{W_{\mathcal{H}}(\mathbf{y}) : \mathbf{y} \in \mathbf{Z}_+^m\}$, where k indexes the targets in case of multiple targets. A sufficient (and reasonable) condition for the existence of $\mathbf{y}^*_{(k)}(\mathcal{H})$ for all $\mathcal{H}$ is that $C(\mathbf{y})$ has a finite lower bound, and we will henceforth make this assumption. Note that if two minima $\mathbf{y}^*_{(k)}(\mathcal{H})$ and $\mathbf{y}^*_{(l)}(\mathcal{H})$ are such that $y^*_{(k)i}(\mathcal{H}) \geq y^*_{(l)i}(\mathcal{H})_i$ for all $i \in \mathcal{H}$ and $y^*_{(k)i}(\mathcal{H}) \leq y^*_{(l)i}(\mathcal{H})_i$ for all $i \in \mathcal{H}^c$, then we can ignore $\mathbf{y}^*_{(l)}(\mathcal{H})$, and we will no longer consider $\mathbf{y}^*_{(l)}(\mathcal{H})$ a minimum. Let $\mathcal{C}(\mathcal{H}) = \cup_k\{\mathbf{n}\colon y^*_{(k)i}(\mathcal{H}) \geq n_i$ for $i \in \mathcal{H}$ and $y^*_{(k)i}(\mathcal{H}) \leq n_i$ for $i \in \mathcal{H}^c\}$ be the "corner" defined by $\mathcal{H}$ and $\mathbf{y}^*_{(k)}(\mathcal{H})$, $k = 1, 2, \ldots$. It is not hard to show that $\max_k\{\mathbf{y}^*_{(k),i}(\mathcal{H})\} \leq \min_k\{\mathbf{y}^*_{(k),i}(\mathcal{H} \setminus i)\}$ (see Lemma 2 below), so that "corners" for distinct $\mathcal{H}$ are non-overlapping. We can partially characterize the optimal policy as follows. If it is possible to move to $\mathbf{y}^*_{(k)}(\mathcal{H})$ by hiring workers of types in $\mathcal{H}$ and firing those of types in $\mathcal{H}^c$, then it is optimal to do so. The proof is straightforward and is omitted.

Lemma 1 *The optimal policy for all* $\mathbf{n} \in \mathcal{C}(\mathcal{H})$ *is to hire/fire to one of the target values* $\mathbf{y}^*_{(k)}(\mathcal{H})$*, i.e.,* $\mathbf{d}^* = \mathbf{y}^*_{(k)}(\mathcal{H}) - \mathbf{n}$.

So, for example, given a staffing level that minimizes costs if we only hire workers, $y^*_k(\mathcal{S})$, if we start with fewer workers of each type, $n_i \leq y^*_{(k)i}(\mathcal{S})$, then it is optimal to hire up to $y^*_k(\mathcal{S})$.

Let $H_i(\mathbf{n}^i) := \arg\min_{n_i \geq 0} C(\mathbf{n}) + h_i n_i$ and $F_i(\mathbf{n}^i) := \arg\min_{n_i \geq 0} C(\mathbf{n}) - f_i n_i$, where $\mathbf{n}^i = (n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_m)$. Also, if there are multiple minima, H_i^k and F_i^k, we choose the largest for H_i and the smallest for F_i.

Lemma 2 $\max_k H_i^k(\mathbf{n}^i) \leq \min_k F_i^k(\mathbf{n}^i)$.

Proof Fix i and $\mathbf{n}^i$ and suppose $H := \max_k H_i^k(\mathbf{n}^i) > F := \min_k F_i^k(\mathbf{n}^i)$, and let $(\mathbf{n}^i, H) = (n_1, \ldots, n_{i-1}, H, n_{i+1}, \ldots, n_m)$ with $(\mathbf{n}^i, F)$ similarly defined. By definition we have $C(\mathbf{n}^i, H) + hH \leq C(\mathbf{n}^i, F) + hF$ and $C(\mathbf{n}^i, F) - fF \leq C(\mathbf{n}^i, H) - fH$, i.e., $f(H - F) \leq C(\mathbf{n}^i, H) - C(\mathbf{n}^i, F) \leq h(F - H)$. But this with $H > F$ gives us $0 < 0$, a contradiction.

Note that a global optimum for $\mathcal{H}$, $\mathbf{y}^*_{(k)}(\mathcal{H})$, must occur where the functions H_i, $i \in \mathcal{H}$, and F_i, $i \in \mathcal{H}^c$ intersect, because we ignore $\mathbf{y}^*_{(l)}(\mathcal{H})$ if there exists $y^*_{(k)i}(\mathcal{H})$ such that $y^*_{(k)i}(\mathcal{H}) \geq y^*_{(l)i}(\mathcal{H})_i$ for all $i \in \mathcal{H}$ and $y^*_{(k)i}(\mathcal{H}) \leq y^*_{(l)i}(\mathcal{H})_i$ for all $i \in \mathcal{H}^c$, and we take the largest for H_i and the smallest for F_i in the case of multiple H_i and F_i.

Now we can divide the space $\mathbf{Z}^n_+$ into a set of interior regions, $\mathcal{I} = \cup\mathcal{I}_k$, and an exterior region, $\mathcal{E}$, where $\mathbf{n} \in \mathcal{I}$ if $H_i(\mathbf{n}^i) \leq n_i \leq F_i(\mathbf{n}^i)$ for all i, and $\mathbf{n} \in \mathcal{E}$ otherwise. Note that $\mathcal{C}(\mathcal{H})\backslash\{\mathbf{y}^*_{(k)}(\mathcal{H})\}_k \subset \mathcal{E}$ for all $\mathcal{H}$. We let $\mathcal{B} \subset \mathcal{I}$ be the set of boundary points, so $\mathbf{n} \in \mathcal{B}$ if for all $i \in \{1, \ldots, m\}$, $H_i(\mathbf{n}^i) \leq n_i \leq F_i(\mathbf{n}^i)$, and for some $j \in \{1, \ldots, m\}$, $n_j = H_j(\mathbf{n}^j)$ or $n_j = F_j(\mathbf{n}^j)$. Since a global optimum for $\mathcal{H}$, $\mathbf{y}^*_{(k)}(\mathcal{H})$, must occur where the functions H_i, $i \in \mathcal{H}$, and F_i, $i \in \mathcal{H}^c$ intersect, we have $\mathbf{y}^*_{(k)}(\mathcal{H}) \in \mathcal{B}$ for all k and $\mathcal{H}$. Let us further define contiguous regions of $\mathcal{I}$, $\mathcal{I}_k$, as follows. First define the neighbors of a point $\mathbf{n}$ as all the points that can be reached from $\mathbf{n}$ by hiring or firing at most one worker of each type, (the hyper-cube with $\mathbf{n}$ at its center and with length 2 in each dimension), i.e., $N(\mathbf{n}) = \{\mathbf{n} + \sum_{i=1}^m k_i\mathbf{e}_i, k_i = -1, 0, 1, i = 1, \ldots, m\}$, where $\mathbf{e}_i$ is the vector with 1 for its ith component and 0 for all other components. Let $\mathcal{I}_k$ form a partition of $\mathcal{I}$ such that if $\mathbf{n} \in \mathcal{I}_k$, $\mathbf{r} \in N(\mathbf{n})$, and $\mathbf{r} \in \mathcal{I}$ then $\mathbf{r} \in \mathcal{I}_k$ and if $\mathbf{n} \in \mathcal{I}_k$ either $\mathcal{I}_k = \{\mathbf{n}\}$ or $\mathbf{r} \in \mathcal{I}_k$ for some $\mathbf{r} \in N(\mathbf{n})$. Let $\mathcal{B}_k$ be the corresponding partition of $\mathcal{B}$. See Figure 1 for a two-dimensional example, where we illustrate continuous functions F and H for simplicity. The boundaries, $\mathcal{B}_1$ and $\mathcal{B}_2$ are given by the heavy curves

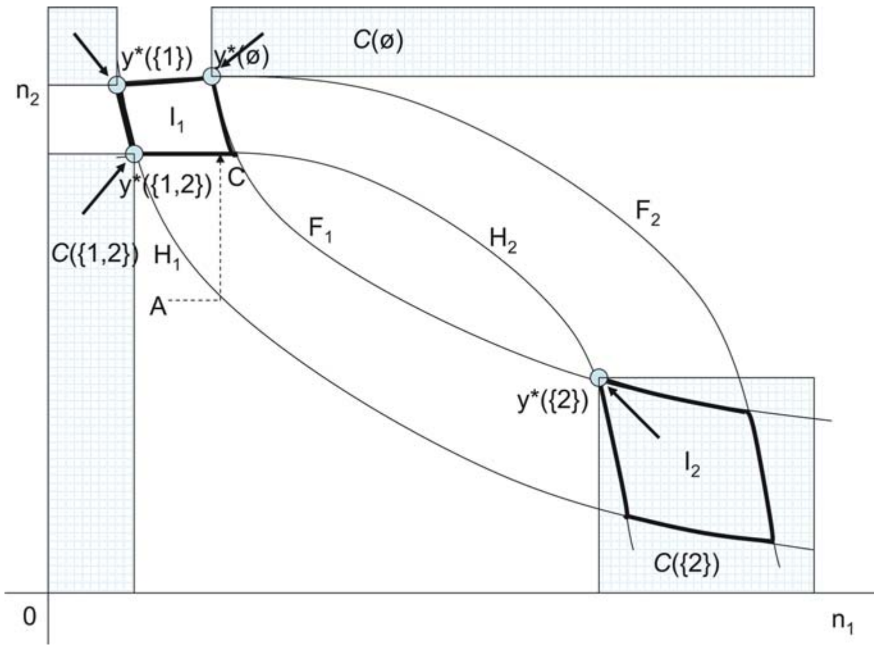

Fig. 1 Characterization of the optimal policy

surrounding $\mathcal{I}_1$ and $\mathcal{I}_2$, and in this example $\mathcal{I}_2$ can be ignored because it is a subset of $\mathcal{C}(\emptyset)$.

For our next result we must also assume (as is reasonable in practice) that there is an upper bound on the number of workers of each type that can be hired. Let M_i be the upper bound for type i, and let $\mathbf{Z}^m_{+,\mathbf{M}} = \{\mathbf{n} : 0 \leq n_i \leq M_i, n_i \in \mathbf{Z}, i = 1, \ldots, m\}$. Then it will never be optimal to hire/fire to a point $\mathbf{n} \in \mathcal{E}$.

Lemma 3 *Suppose the domain of C is $\mathbf{Z}^m_{+,\mathbf{M}}$. Then for $\mathbf{n} \in \mathcal{E}$ it is optimal to hire/fire to some point $\mathbf{y} \in \mathcal{I}$, and for $\mathbf{n} \in \mathcal{C}(\mathcal{H})$ it is optimal to hire/fire to some $\mathbf{y}^*_{(k)}(\mathcal{H})$.*

Proof Choose any $\mathbf{n} \in \mathcal{E}$, and suppose $n_i < H_i(\mathbf{n}^i)$, say. Then we have, by definition of $H_i(\mathbf{n}^i)$, that

$$C(\mathbf{y}) + h_i(y_i - n_i) < C(\mathbf{n}),$$

where $\mathbf{y} = (n_1, \ldots, n_{i-1}, H_i(\mathbf{n}^i), n_{i+1}, \ldots, n_m)$. If $\mathbf{y} \notin \mathcal{B}$ but $\mathbf{y} \in \mathcal{C}(\mathcal{H})$ for some $\mathcal{H}$, we can minimize the cost by moving to $\mathbf{y}^*_{(k)}(\mathcal{H}) \in \mathcal{B}$, from Lemma 1. Otherwise we can repeat the argument to find $\mathbf{y}'$ that equals $\mathbf{y}$ except for replacing n_j with $H_j(\mathbf{n}^j)$ or $F_j(\mathbf{n}^j)$ for some j, and so that the cost will be lower to move to $\mathbf{y}'$ rather than staying at $\mathbf{n}$ or $\mathbf{y}$. Continuing to repeat the argument we will finally have a point $\hat{\mathbf{y}} \in \mathcal{B}$ such that the cost to hire/fire to $\hat{\mathbf{y}}$ will be less than the cost of staying at $\mathbf{n} \in \mathcal{E}$. Note that we never return to a point previously visited because each move strictly decreases the cost, and $\mathbf{Z}^m_{+,\mathbf{M}}$ is finite, so the process converges. The idea is illustrated in Figure 1, where the cost of staying at point $A \in \mathcal{E}$ is greater than the cost of hiring type 1 workers to a point on H_1 and then hiring type 2 workers to a point on $\mathcal{B}_1$.

We need the upper bounds M_i when $m > 2$, because without them hiring an infinite number of workers may be optimal, even with the assumption of finite target values $\mathbf{y}^*(\mathcal{H})$. Suppose $m = 3$, $H_3(n_1, n_2) = 5$ and $F_3(n_1, n_2) = 10$ for all n_1, n_2, $H_1(n_2, n_3) = 3n_2$ for all n_3, and $H_2(n_1, n_3) = 3n_1$ for all n_3, and we start at the point $(7, 7, 7)$. Then following the argument in the proof above we will continue to decrease the cost as we move back and forth from $(n_1, 3n_1, 7)$ to $(3n_2, n_2, 7)$ with ever higher values of n_1 and n_2.

We can extend the idea of Lemma 1 of having global "target values," to a notion of local target values. Suppose that only hiring is permitted, i.e., we restrict ourselves to $\mathcal{H} = \mathcal{S}$. Then the optimal (hiring) policy can be characterized by a partition of $\mathbf{Z}^m_{+,\mathbf{M}}$ into subregions $R^{(i)}$ defined by a sequence of target values, $\mathbf{y}^{(i)} \in \mathbf{Z}^m_{+,\mathbf{M}}$ as follows. Let $\mathbf{y}^{(0)} = \mathbf{y}^*(\mathcal{S}) := \arg\min\{W_{\mathcal{S}}(\mathbf{y}) : \mathbf{y} \in \mathbf{Z}^m_+\}$ (where, in the case of mulitple values, one is chosen arbitrarily) and let $R^{(0)} = \{\mathbf{n} : \mathbf{n} \le \mathbf{y}^{(0)}\}$. Then let $\mathbf{y}^{(i)} := \arg\min\{W_{\mathcal{S}}(\mathbf{y}) : \mathbf{y} \in \mathbf{Z}^m_+ \backslash \cup_{j=0}^{i-1} R^{(j)}\}$ and $R^{(i)} = \{\mathbf{n} : \mathbf{n} \le \mathbf{y}^{(i)}, \mathbf{n} \in \mathbf{Z}^m_+ \backslash \cup_{j=0}^{i-1} R^{(j)}\}$. This is because given $\mathbf{n}$, the optimal policy depends on $\mathbf{n}$ only through the constraint that $\mathbf{y} \ge \mathbf{n}$. Thus, if $\mathbf{y}^*$ is optimal for $\mathbf{n}$, it is also optimal for any $\mathbf{m}$ such that $\mathbf{n} \le \mathbf{m} \le \mathbf{y}$. We say that the optimal policy has a "target-box" structure.

Lemma 4 *If only hiring is permitted, i.e., $\mathcal{H} = \mathcal{S}$, then the optimal policy has the target-box structure defined above.*

If firing is permitted as well, we obtain a set of partitions, $R^{(i)}_{\mathcal{H}}$ and corresponding targets $\mathbf{y}^{(i)}_{\mathcal{H}}$ for each $\mathcal{H}$, such that the optimal policy starting in state $\mathbf{n}$ is to hire/fire to $\mathbf{y}^{(i)}_{\mathcal{H}^*}$ where $\mathbf{y}^{(i)}_{\mathcal{H}_*}$ is such that $W_{\mathcal{H}^*}(\mathbf{y}^{(i)}_{\mathcal{H}^*}) = \min_{\mathcal{H}} W_{\mathcal{H}}(\mathbf{y}^{(i)}_{\mathcal{H}})$.

Lemmas 1 and 3 give us some structure on the optimal policy, but to refine it we will need to make some sort of convexity assumptions.

3 Notions of discrete convexity

3.1 Component-wise convexity

The simplest notion of discrete convexity is component-wise convexity. A function on the integers, $C(\mathbf{n})$, is component-wise convex (cwcx) if it is convex in n_i for all $\mathbf{n}^i$. In this case we can further characterize the single-stage optimal policy: for points in the exterior it is optimal to move to a point on the boundary, and for points in some interior region $\mathcal{I}_k$, it is optimal to do nothing or to move to a point on the boundary of a different interior region. Let $C_{\mathcal{H}}(\mathbf{y}) = \sum_{i\in\mathcal{H}} h_i y_i - \sum_{i\in\mathcal{H}^c} f_i y_i + C(\mathbf{y})$, so $\mathbf{y}^*(\mathcal{H})$ minimizes $C_{\mathcal{H}}$. Note that if C is cwcx, then so is $C_{\mathcal{H}}$. We have the following corollary to Lemma 3.

Corollary 1 *If C is a component-wise convex function on $\mathbf{Z}^m_{+,\mathbf{M}}$, then for $\mathbf{n} \in \mathcal{I}_k$ it is optimal to do nothing or to hire/fire to a point $\mathbf{y} \in \mathcal{B}\backslash\mathcal{B}_k$, for $\mathbf{n} \in \mathcal{E}$ it is optimal to hire/fire to some point $\mathbf{y} \in \mathcal{B}$, and for $\mathbf{n} \in \mathcal{C}(\mathcal{H})$ it is optimal to hire/fire to $\mathbf{y}^*(\mathcal{H})$.*

Proof Now choose any k and $\mathbf{n} \in \mathcal{I}_k$ (note that $\mathcal{B}_k \subset \mathcal{I}$). From the last lemma, it will not be optimal to move to a point $\mathbf{y} \in \mathcal{E}$, so first choose some point $\mathbf{y} \in \mathcal{I}_k$, say $\mathbf{y} \geq \mathbf{n}$. (The argument for other $\mathbf{y} \in \mathcal{I}_k$ is similar). Then, letting $(\mathbf{n}_i, \mathbf{y}) = (n_1, \ldots, n_i, y_{i+1}, \ldots, y_m)$, the cost to move from $\mathbf{n}$ to $\mathbf{y}$ will be

$$\begin{aligned}\Delta &:= C(\mathbf{y}) - C(\mathbf{n}) + \sum_{i=1}^{m} h_i(y_i - n_i) \\ &= C(\mathbf{y}) - C(\mathbf{n}_1, \mathbf{y}) + h_1(y_1 - n_1) + C(\mathbf{n}_1, \mathbf{y}) - C(\mathbf{n}_2, \mathbf{y}) + h_2(y_2 - n_2) \\ &\quad + \cdots + C(\mathbf{n}_{m-1}, \mathbf{y}) - C(\mathbf{n}) + h_m(y_m - n_m) =: \sum_{i=1}^{m} \Delta_i \geq 0.\end{aligned}$$

The last inequality follows because, for all i, $H_i((\mathbf{n}_i, \mathbf{y})^i) \leq n_i \leq y_i$, so $\Delta_i \geq 0$ from the convexity of C in direction i. Therefore, from any point in $\mathcal{I}_k$, it is not optimal to move to another point in $\mathcal{I}_k$. Now choose some point $\mathbf{y} \in \mathcal{I}\backslash\mathcal{I}_k$, say $\mathbf{y} \geq \mathbf{n}$. (Again the argument for other $\mathbf{y} \in \mathcal{I}\backslash\mathcal{I}_k$ is similar). Let $\mathbf{y}'$ be a point such that $\mathbf{y} \geq \mathbf{y}' \geq \mathbf{n}$ and $\mathbf{y}' \in \mathcal{B}_k$. An argument similar to the one above shows that from $\mathbf{n}$ it will cost less to move to $\mathbf{y}'$ than to move to $\mathbf{y}$. The rest of the result follows from Lemma 3.

We can extend Lemma 4, using the idea of the proof of Corollary 1, to further characterize the structure of the optimal hiring policy, when firing is not permitted for any worker type.

Lemma 5 *If only hiring is permitted, i.e., $\mathcal{H} = \mathcal{S}$, and C is component-wise convex, then the optimal policy has the target-box structure of Lemma 4 for* $\mathbf{n}$ *such that $n_i < H_i(\mathbf{n}^i)$ for at least one i. For all other* $\mathbf{n}$, *the optimal policy is to do nothing.*

We now consider whether component-wise convexity propagates and is preserved under binomial transformations.

We say that a family of random vectors $\{N(\mathbf{n}) = (N_1(\mathbf{n}), \ldots, N_n(\mathbf{n})), \mathbf{n} \in \mathbf{Z}^m\}$ is stochastically component-wise convex, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SCWCX$, if $Ef(N(\mathbf{n}))$ is cwcx for any cwcx function f. We say that it is stochastically cwcx in the sample path sense, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SCWCX(sp)$, if for any $\mathbf{n} \in \mathbf{Z}^m$, for any $i = 1, \ldots, m$, and for any cwcx function f we can construct N_1, N_2, N_3, N_4, on a common probability space such that

$$N_1 =_{st} N(\mathbf{n}), N_2 =_{st} N(\mathbf{n} + \mathbf{e}_i), N_3 =_{st} N(\mathbf{n} + \mathbf{e}_i), N_4 =_{st} N(\mathbf{n} + \mathbf{2e}_i),$$

and such that, with probability 1,

$$f(N_1) - f(N_2) \geq f(N_3) - f(N_4).$$

Of course $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SCWCX(sp) \Longrightarrow \{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SCWCX$.

Let $B_i(\mathbf{n})$, $i = 1, \ldots, m$ be binomially distributed random variables with parameters n_i and p_i for some p_i, and let $B(\mathbf{n}) = (B_1(\mathbf{n}), \ldots, B_n(\mathbf{n}))$.

Lemma 6 $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SCWCX(sp)$.

Proof Fix i. Given $\mathbf{n} \in \mathbf{Z}^m$, first generate $N_1 =_{st} B(\mathbf{n})$. With probability p_i^2 let $N_2 = N_3 = N_1 + \mathbf{e}_i$ and $N_4 = N_1 + 2\mathbf{e}_i$,

with probability $(1 - p_i)^2$ let $N_2 = N_3 = N_4 = N_1$,

with probability $p_i(1 - p_i)$ let $N_2 = N_1$, and $N_3 = N_4 = N_1 + \mathbf{e}_i$,

and with probability $p_i(1 - p_i)$ let $N_3 = N_1$, and $N_2 = N_4 = N_1 + \mathbf{e}_i$. The result follows.

Note that the lemma also holds for random, not necessarily independent p_i's. The following example shows that component-wise convexity is not propagated in the dynamic programming recursion, and a local minimum need not be a global minimum.

Example 1 Consider the cwcx function: $f(n_1, n_2) = 2 - n_1 - n_2 + n_1 n_2, \mathbf{n} \in \mathbf{Z}_+^2$. There are local minima at (0,2) and (2,0) of value 0, and a local minimum at (1,1) of value 1. Also $g(n_1) = \min_{n_2} f(n_1, n_2)$ takes on the values 0, 1, and 0 for $n_1 = 0$, 1, and 2 respectively, so is not convex. Thus the structure of Corollary 1 holds only for a single stage problem with random turnover.

We will next consider a class of discrete functions that has the preservation and propagation properties, and allows us to further characterize the optimal policy, though it still does not guarantee that an ISD policy is optimal.

3.2 Supermodularity

We say that C is supermodular (submodular) if $C(\mathbf{n} + \mathbf{e}_i) - C(\mathbf{n})$ is increasing (decreasing) in $\mathbf{n}^i$, for all i. A supermodular cost function indicates that workers of different training levels are substitutes, i.e., the advantage of additional workers of one type is decreasing in the number of workers of other types (so the cost is increasing). This may the case, for example, when workers with more training may be able to replace workers with less training. Alternatively, worker types may be complements, e.g., when they work together, in which case a submodular cost function is appropriate.

Lemma 7 *If C is supermodular (submodular) then $H_i(\mathbf{n}^i)$ and $F_i(\mathbf{n}^i)$ are decreasing (increasing).*

Proof We show the result for supermodular functions; the submodular case is similar. Pick some point $\mathbf{n}^i$ for some i, and let

$$n_i = H_i(\mathbf{n}^i) := \arg\min_{n_i}(C(\mathbf{n}) + h_i n_i).$$

Then $C(\mathbf{n} + ke_i) - C(\mathbf{n}) \geq 0$ for all $k > 0$, and from supermodularity, $C(\mathbf{n} + ke_i + e_j) - C(\mathbf{n} + e_j) \geq 0$, so $H_i(\mathbf{n} + e_j) \leq H_i(\mathbf{n})$.

We can show that for our model, supermodularity is propagated in the multistage problem when only hiring is permitted, so the optimal policy has the target-box structure of Lemma 4, and when we have only $m = 2$ worker types. Let $\hat{C}(y_1, y_2) = C(y_1, y_2) + h_1 y_1 + h_2 y_2$. We want to show that V_t is supermodular given that C and V_{t+1} are supermodular, where $V_t(\mathbf{n}) = \min_{\mathbf{y} \geq \mathbf{n}}\{\hat{C}(\mathbf{y}) - h_1 n_1 - h_2 n_2 + V_{t+1}(\mathbf{y})\} = \min_{\mathbf{y} \geq \mathbf{n}}\{\hat{C}(\mathbf{y}) + V_{t+1}(\mathbf{y})\} - h_1 n_1 - h_2 n_2$. First note that it is easy to show that $\hat{C} + V_{t+1}$ is supermodular given that C and V_{t+1} are supermodular.

Lemma 8 *If* $f(\mathbf{y}) := \hat{C}(\mathbf{y}) + V_{t+1}(\mathbf{y})$ *is supermodular, then*

$$g(\mathbf{n}) := \min_{\mathbf{y} \geq \mathbf{n}} f(\mathbf{y})$$

is supermodular.

Proof We need to show that for four points, $\mathbf{n}_i$, $i = 1, 2, 3, 4$, such that $\mathbf{n}_2 = \mathbf{n}_1 + k_1\mathbf{e}_1$, $\mathbf{n}_3 = \mathbf{n}_1 + k_2\mathbf{e}_2$, $\mathbf{n}_4 = \mathbf{n}_1 + k_1\mathbf{e}_1 + k_2\mathbf{e}_2$, we have $f(\mathbf{y}_1) + f(\mathbf{y}_4) \geq f(\mathbf{y}_2) + f(\mathbf{y}_3)$, where $\mathbf{y}_i = \arg\min_{\mathbf{y} \geq \mathbf{n}_i} f(\mathbf{y})$. We consider the following cases:

(1) $\mathbf{y}_1 \geq \mathbf{n}_4$: In this case, because of the target-box structure of the optimal policy (Lemma 4), we must have that $\mathbf{n}_i$, $i = 1, 2, 3, 4$, are all in the same subregion, and have the same target: $\mathbf{y}_1 = \mathbf{y}_2 = \mathbf{y}_3 = \mathbf{y}_4$, so $f(\mathbf{y}_1) + f(\mathbf{y}_4) \geq f(\mathbf{y}_2) + f(\mathbf{y}_3)$ trivially.

(2) $\mathbf{y}_1 \geq \mathbf{n}_3$ but $\mathbf{y}_1$ is not greater than or equal to $\mathbf{n}_4$: then $\mathbf{y}_1 = \mathbf{y}_3$, so $f(\mathbf{y}_1) = f(\mathbf{y}_3)$, and $f(\mathbf{y}_4) \geq f(\mathbf{y}_3)$ and $f(\mathbf{y}_4) \geq f(\mathbf{y}_2)$, so $f(\mathbf{y}_1) + f(\mathbf{y}_4) \geq f(\mathbf{y}_2) + f(\mathbf{y}_3)$.

(3) $\mathbf{y}_1 \geq \mathbf{n}_2$ but $\mathbf{y}_1$ is not greater than or equal to $\mathbf{n}_4$: same argument as case (2).

(4) $\mathbf{n}_1 \leq \mathbf{y}_1 \leq \mathbf{n}_4$: Let $\hat{\mathbf{y}}_2 = (y_{41}, y_{12})$ and let $\hat{\mathbf{y}}_3 = (y_{11}, y_{42})$, where $\mathbf{y}_i = (y_{i1}, y_{i2})$, $i = 1, 4$. Then $f(\mathbf{y}_1) + f(\mathbf{y}_4) \geq f(\hat{\mathbf{y}}_2) + f(\hat{\mathbf{y}}_3)$, and $f(\hat{\mathbf{y}}_i) \geq f(\mathbf{y}_i)$, $i = 2, 3$.

Furthermore, supermodularity is preserved under state transformation due to binomial turnover. We say that a family of random vectors $\{N(\mathbf{n}) = (N_1(\mathbf{n}), \ldots, N_n(\mathbf{n})), \mathbf{n} \in \mathbf{Z}^m\}$ is stochastically supermodular, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SSM$, if $Ef(N(\mathbf{n}))$ is supermodular for any supermodular function f. We say that it is stochastically supermodular in the sample path sense, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SSM(sp)$, if for any $\mathbf{n} \in \mathbf{Z}^m$, for any $i = 1, \ldots, m$, and for any supermodular function f we can construct N_1, N_2, N_3, N_4, on a common probability space such that

$$N_1 =_{st} N(\mathbf{n}), N_2 =_{st} N(\mathbf{n} + \mathbf{e}_i), N_3 =_{st} N(\mathbf{n} + \mathbf{e}_j), N_4^2 =_{st} N(\mathbf{n} + \mathbf{e}_i + \mathbf{e}_j),$$

and such that, with probability 1,

$$f(N_1) - f(N_2) \geq f(N_3) - f(N_4).$$

Of course $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SSM(sp) \Longrightarrow \{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SSM$. The proof of the following is similar to that of Lemma 6.

Lemma 9 $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in SSM(sp)$.

While supermodularity provides a refinement of the policy structure close to that of "hire-up-to/fire-down-to," further refinement can be achieved if additional convexity properties are satisfied.

3.3 Directional convexity

We say that C is directionally convex (dcx) if C is cwcx and supermodular, and from our earlier results, we can get some structure on the optimal policy, though not a simple ISD structure. Assuming we have only $m = 2$ types of workers and only hiring is permitted, from Lemmas 5 and 7, we can construct the optimal policy with the following algorithm. It successively characterizes target-boxes and unknown boxes. Here each box is characterized by two points, $\mathbf{x}$ and $\mathbf{y}$, such that $\mathbf{n}$ is in the box if $\mathbf{x}^{(i)} < \mathbf{n} \leq \mathbf{y}^{(i)}$, and the optimal policy for such $\mathbf{n}$ is to hire-up-to $\mathbf{y}^{(i)}$.

1. Set $i = 0$ and let the first unknown box be $U^{(0)} = \mathbf{Z}^2_{+,\mathbf{M}}$, and let $u^{(0)}_{1a} = u^{(0)}_{2a} = 0$, $u^{(0)}_{1b} = u^{(0)}_{2b} = M$, so $U^{(0)} = \{(n_1, n_2) : u^{(0)}_{1a} \le n_1 \le u^{(0)}_{1b}, u^{(0)}_{2a} \le n_2 \le u^{(0)}_{2b}\}$.
2. For $n_1 = u^{(i)}_{1a}, \dots, u^{(i)}_{1b}$ find $H^{(i)}_2(n_1) = \arg\min_{n_2 = u^{(i)}_{2a}, \dots, u^{(i)}_{2b}} \hat{C}(n_1, n_2)$.
3. Find $y^{(i)}_1 = \arg\min_{n_1 = u^{(i)}_{1a}, \dots, u^{(i)}_{1b}} \hat{C}(n_1, H^{(i)}_2(n_1))$, and let $y^{(i)}_2 = H^{(i)}_2(y^{(i)}_1)$. Thus, $\mathbf{y}^{(i)}$ is the global minimum for $U^{(i)}$, and the optimal policy for $\mathbf{u}^{(i)}_a \le \mathbf{n} \le \mathbf{y}^{(i)}$ is to hire-up-to $\mathbf{y}^{(i)}$ (this is a target-box) and the optimal policy for $\mathbf{y}^{(i)} \le \mathbf{n} \le \mathbf{u}^{(i)}_b$ is to do nothing.
4. Let $u^{(i+1)}_{1a} = u^{(i)}_{1a}$, $u^{(i+1)}_{1b} = y^{(i)}_1$, $u^{(i+1)}_{2a} = y^{(i)}_2 + 1$, $u^{(i+1)}_{2b} = u^{(i)}_{2b}$, $U^{(i+1)} = \{(n_1, n_2) : u^{(i+1)}_{1a} \le n_1 \le u^{(i+1)}_{1b}, u^{(i+1)}_{2a} \le n_2 \le u^{(i+1)}_{2b}\}$.
5. Set i to $i + 1$ and repeat 2–4 until $U^{(i)} = \emptyset$.

4′. Let $u^{(i+1')}_{1a} = y^{(i)}_{1a} + 1$, $u^{(i+1')}_{1b} = u^{(i)}_{1b}$, $u^{(i+1')}_{2a} = u^{(i)}_{2a}$, $u^{(i+1')}_{2b} = y^{(i)}_2$, $U^{(i+1')} = \{(n_1, n_2) : u^{(i+1')}_{1a} \le n_1 \le u^{(i+1')}_{1b}, u^{(i+1')}_{2a} \le n_2 \le u^{(i+1')}_{2b}\}$.

5′. Set i to $i + 1'$ and repeat 2, 3, and 4′ until $U^{(i)} = \emptyset$.

Consider the following example of a dcx function.

$$\hat{C}(\mathbf{n}) = 15 + 0.9n_1 + 1.3n_2 - \min\{13, 2n_1 + 3n_2\}.$$

The optimal policy is shown below, where we give the values of $\hat{C}$ for $0 \le n_i \le 7$, with the values of n_1 given along the bottom, and the values of n_2 along the left side. Target values are indicated with bold script, and lines indicate target-boxes. Thus, the target-boxes are $U^{(0)} = \{\mathbf{n} : 0 \le n_1 \le 2, 0 \le n_2 \le 3\}$, $U^{(1)} = \{\mathbf{n} : 0 \le n_1 \le 2, n_2 = 4\}$, $U^{(1')} = \{\mathbf{n} : 3 \le n_1 \le 5, 0 \le n_2 \le 1\}$, and $U^{(2')} = \{\mathbf{n} : 6 \le n_1 \le 7, n_2 = 0\}$. Above the target-boxes it is optimal to do nothing.

7	11.1	12.0	12.9	13.8	14.7	15.6	16.5	17.4
6	9.8	10.7	11.6	12.5	13.4	14.3	15.2	16.1
5	8.5	9.4	10.3	11.2	12.1	13.0	13.9	14.8
4	8.2	**8.1**	9.0	9.9	10.8	11.7	12.6	13.5
3	9.9	8.8	**7.7**	8.6	9.5	10.4	11.3	12.2
2	11.6	10.5	9.4	8.3	**8.2**	9.1	10.0	10.9
1	13.3	12.2	11.1	10.0	8.9	**7.8**	8.7	9.6
0	15.0	13.9	12.8	11.7	10.6	9.5	8.4	**8.3**
	0	1	2	3	4	5	6	7

Let us compare the optimal policy for this example with what it would be if we were permitted to hire fractional workers. In that case, the global minimum is at (0, 4.33), so, if we start with no workers, the optimal policy is to hire no type 1 workers and 4.33 type 2 workers (vs. hiring 2 type 1 workers and 3 type 2 workers in the discrete case). If we start with 0 type 1 workers and 4 type 2 workers, then in the continous case the optimal policy is to hire no type 1 workers and 0.33 type 2 workers (vs. hiring 1 type 1 worker and 0 type 2 workers in the discrete case).

Though directional convexity is preserved under binomial turnover from Lemmas 6 and 9, it is not propagated. Indeed, the cwcx function of Example 1 is also supermodular, and hence is dcx, and it does not satisfy the propagation property.

3.4 Multimodularity

Multimodularity strengthens the notion of directional convexity, because multimodular functions are cwcx and supermodular. We will see that it gives us a full characterization of the optimal policy as an ISD policy for deterministic costs, and it satisfies the propagation property. Unfortunately, it does not satisfy the binomial preservation property.

We suppose there are $m = 2$ types of workers. A function $f(\mathbf{n})$ for $\mathbf{n} \in \mathbf{Z}^2$ is multimodular if the following three inequalities hold, where the first two correspond to a type of midpoint convexity, and the third is supermodularity. For all $\mathbf{n}, \mathbf{m} \in \mathbf{Z}^m$,

$$f(\mathbf{n}+2\mathbf{e}_1) + f(\mathbf{n} + \mathbf{e}_2) \geq f(\mathbf{n} + \mathbf{e}_1) + f(\mathbf{n} + \mathbf{e}_1 + \mathbf{e}_2) \tag{1}$$
$$f(\mathbf{n}+2\mathbf{e}_2) + f(\mathbf{n} + \mathbf{e}_1) \geq f(\mathbf{n} + \mathbf{e}_2) + f(\mathbf{n} + \mathbf{e}_1 + \mathbf{e}_2) \tag{2}$$
$$f(\mathbf{n}) + f(\mathbf{n} + \mathbf{e_1} + \mathbf{e}_2) \geq f(\mathbf{n} + \mathbf{e}_1) + f(\mathbf{n} + \mathbf{e}_2). \tag{3}$$

See Hajek (1985) and Altman et al. (2000).

A local optimum is a global optimum for multimodular functions where a local optimum is defined as follows. We say $\mathbf{n}^*$ is a local multimodular optimum if, $f(\mathbf{n}^*) \leq f(\mathbf{n}^*+\mathbf{e}_i)$; $f(\mathbf{n}^*) \leq f(\mathbf{n}^*-\mathbf{e}_i)$, $i = 1, 2$, and $f(\mathbf{n}^*) \leq f(\mathbf{n}^*+\mathbf{e}_1-\mathbf{e}_2)$. If f is multimodular and $\mathbf{n}^*$ is a local optimum, then it is a global optimum, i.e., $f(\mathbf{n}^*) \leq f(\mathbf{m})$ for all $\mathbf{m} \in \mathbf{Z}^2$. Note that the condition for local optimum is stronger than that claimed in Altman et al. (2000). The requirement of the stronger condition was recently shown by Murota (2004).

For multimodular functions, even in higher dimensions than two, the optimal policy has the ISD structure. This can be shown from Theorem 1 and the fact that the piecewise affine interpolation of a multimodular function is jointly convex (Altman et al. 2000). See also Narongwanich et al. (2003, preprint).

Theorem 2 *If C is multimodular then the optimal policy has the following hire-up-to/fire-down-to structure. For each i there exist two functions $U_i(\mathbf{x}^i) \leq D_i(\mathbf{x}^i)$, such that for a given starting state $\mathbf{x}$, for each i, if $x_i < U_i(\mathbf{x}^i)$ hire up to $U_i(\mathbf{x}^i)$ type i workers, i.e., hire $U_i(\mathbf{x}^i)$–x_i type i workers, if $x_i > D_i(\mathbf{x}^i)$ fire down to $D_i(\mathbf{x}^i)$ workers, and otherwise do not hire or fire type i workers.*

Moreover, multimodular functions propagate under the dynamic programming recursion. That is, when f is multimodular, so is g where $g(\mathbf{n}) = \min_{\mathbf{m}} f(\mathbf{m}, \mathbf{n})$ (Narongwanich et al. 2003).

We say that a family of random vectors $\{N(\mathbf{n}) = (N_1(\mathbf{n}), \ldots, N_n(\mathbf{n})), \mathbf{n} \in \mathbf{Z}^m\}$ is stochastically multimodular, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^m\} \in$ SMM, if $Ef(N(\mathbf{n}))$ is multimodular for any multimodular function f. We say that it is stochastically multimodular in the sample path sense, $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in$ SMM(sp), if for any $\mathbf{n} \in \mathbf{Z}^2$ and for $j = 1, 2, 3$ [corresponding to inequalities (1)–(3)] and for any multimodular function f we can construct $N_1^j, N_2^j, N_3^j, N_4^j$, on a common probability space such that

$$N_1^1 =_{st} N(\mathbf{n} + \mathbf{e}_2),\ N_2^1 =_{st} N(\mathbf{n} + \mathbf{e}_1),$$
$$N_3^1 =_{st} N(\mathbf{n} + \mathbf{e}_1 + \mathbf{e}_2),\ N_4^1 =_{st} N(\mathbf{n} + 2\mathbf{e}_1),$$
$$N_1^2 =_{st} N(\mathbf{n} + \mathbf{e}_1),\ N_2^2 =_{st} N(\mathbf{n} + \mathbf{e}_2),$$

$$N_3^2 =_{st} N(\mathbf{n}+\mathbf{e}_1+\mathbf{e}_2),\ N_4^2 =_{st} N(\mathbf{n}+2\mathbf{e}_2),$$
$$N_1^2 =_{st} N(\mathbf{n}),\ N_2^2 =_{st} N(\mathbf{n}+\mathbf{e}_1),$$
$$N_3^2 =_{st} N(\mathbf{n}+\mathbf{e}_2),\ N_4^2 =_{st} N(\mathbf{n}+\mathbf{e}_1+\mathbf{e}_2),$$

and such that, for $j = 1, 2, 3$, with probability 1,

$$f(N_1^j) + f(N_4^j) \ge f(N_2^j) + f(N_3^j).$$

Of course $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in \text{SMM(sp)} \Longrightarrow \{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in \text{SMM}$.

We will show below that multimodularity is not preserved in general for binomial transformations, but first we show that a special case of a binomial random variable does indeed preserve multimodularity, in the strong, sample path, sense.

Let $B_i(\mathbf{n})$, $i = 1, 2$ be independent binomially distributed random variables with parameters n_i and p_i for some p_i , and let $B(\mathbf{n}) = (B_1(\mathbf{n}), B_2(\mathbf{n}))$.

Lemma 10 *If* $p_1 = p_2 =: p$, $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in \text{SMM(sp)}$.

Proof We show the result for the first inequality, N_i^1; the arguments for the other two are similar. Given $\mathbf{n} \in \mathbf{Z}^2$, first generate $N_0^1 =_{st} B(\mathbf{n})$. With probability p^2 let

$$N_1^1 = N_0^1 + \mathbf{e}_2,\ N_2^1 = N_0^1 + \mathbf{e}_1,\ N_3^1 = N_0^1 + \mathbf{e}_1 + \mathbf{e}_2,\ N_4^1 = N_0^1 + 2\mathbf{e}_1,$$

with probability $(1-p)^2$ let $N_1^1 = N_2^1 = N_3^1 = N_4^1 = N_0^1$,
with probability $p(1-p)$ let $N_1^1 = N_2^1 = N_0^1$, $N_3^1 = N_4^1 = N_0^1 + \mathbf{e}_1$,
and with probability $p(1-p)$ let $N_1^1 = N_3^1 = N_0^1 + \mathbf{e}_2$, $N_2^1 = N_4^1 = N_0^1 + \mathbf{e}_2$.
Then $N_1^1, N_2^1, N_3^1, N_4^1$ have the appropriate marginal distributions, and inequality (1)–(3) holds for f multimodular.

Of course the same proof shows that even for p a random variable, $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in \text{SMM(sp)}$ (by conditioning on p). The following result shows that the assumption of identical departure rates for different worker types cannot be relaxed for stochastic multimodularity to be preserved.

Proposition 1 $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \notin$ *SMM in general.*

Proof For $\{B(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in SMM$ we need, from inequality (1),

$$\begin{aligned}\Delta_1 := \ & Ef(B(\mathbf{n}+\mathbf{e}_2)) + Ef(B(\mathbf{n}+2\mathbf{e}_1)) \\ & - [Ef(B(\mathbf{n}+\mathbf{e}_1)) + Ef(B(\mathbf{n}+\mathbf{e}_1+\mathbf{e}_2))] \ge 0.\end{aligned}$$

We have

$$\begin{aligned}\Delta_1 = \ & p_2 Ef(B(\mathbf{n})+\mathbf{e}_2) + (1-p_2)Ef(B(\mathbf{n})) + p_1^2 Ef(B(\mathbf{n})+2\mathbf{e}_1) \\ & +2p_1(1-p_1)Ef(B(\mathbf{n})+\mathbf{e}_1) + (1-p_1)^2 Ef(B(\mathbf{n})) \\ & -[p_1 Ef(B(\mathbf{n})+\mathbf{e}_1) + (1-p_1)Ef(B(\mathbf{n})) + p_1 p_2 Ef(B(\mathbf{n})+\mathbf{e}_1+\mathbf{e}_2) \\ & +p_1(1-p_2)Ef(B(\mathbf{n})+\mathbf{e}_1) + p_2(1-p_1)Ef(B(\mathbf{n})+\mathbf{e}_2) \\ & +(1-p_1)(1-p_2)Ef(B(\mathbf{n}))] \\ = \ & p_1\{(p_1-p_2)\,[Ef(B(\mathbf{n})+2\mathbf{e}_1) - 2Ef(B(\mathbf{n})+\mathbf{e}_1) + Ef(B(\mathbf{n}))] \\ & +p_2[Ef(B(\mathbf{n})+\mathbf{e}_2) + Ef(B(\mathbf{n})+2\mathbf{e}_1) - Ef(B(\mathbf{n})+\mathbf{e}_1) \\ & -Ef(B(\mathbf{n})+\mathbf{e}_1+\mathbf{e}_2)]\}\end{aligned}$$

The second term is positive from inequality (1), and if $p_1 \geq p_2$ the first term is positive because multimodular functions are componentwise convex. On the other hand, if we compute Δ_2 for inequality (2), we will need $p_2 \geq p_1$ for the corresponding first term to be positive. Thus, we can construct an example so that one of them is negative for appropriately chosen f, p_1, and p_2. For example, consider $f(n, m) = n^2 + m^2 + 2mn$, which is a multimodular function. Then $\Delta_1 \geq 0$ iff $p_1 \geq p_2$, but $\Delta_2 \geq 0$ iff $p_1 \leq p_2$.

We mention here another stochastic model that preserves multimodularity. In this case, all of the workers of a given type leave with some (possibly random) probability. This is basically the stochastic assumption for the model of Narongwanich et al. (2003, preprint) in which (all of the) capacity of a given type can become obsolete. The proof is similar to that of Lemma 10.

Lemma 11 *Suppose $N_i(\mathbf{n}) = n_i$ with some, possibly random, probability p_i and 0 with probability $1 - p_i$, where the N_i's are not necessarily independent. Then* $\{N(\mathbf{n}), \mathbf{n} \in \mathbf{Z}^2\} \in \mathrm{SMM(sp)}$.

Summarizing our strongest results for multimdodular cost functions, we have the following

Theorem 3 *Suppose that for our original, multistage model, $C_t(y_1, y_2, \theta)$ is multimodular in $\mathbf{y}$ for all θ. Also suppose that $N_{it}(y_{it}, \theta_t)$, $i = 1, 2$, are independent binomial random variables with parameters y_{it} and θ_t (for all i), or $N_{it}(y_{it}, \theta_t)$, $i = 1, 2$ are all-or-nothing random variables as defined in Lemma 11 with parameters y_{it} and θ_{it}. Then the optimal policy is an ISD policy.*

Theorem 4 *Suppose that for a **single**-stage model, $C(y_1, y_2, \theta)$ is multimodular in $\mathbf{y}$ for all θ. Also suppose that $N_{it}(y_{it}, \theta_t)$ are independent binomial random variables with parameters y_{it} and θ_{it}. Then the optimal policy is an ISD policy.*

Acknowledgements We would like to thank the organizers of this workshop and special volume for their efforts, and the referee for helpful comments. We especially wish to thank Arie Hordijk for his inspiration and dedication to the field.

References

Altman E, Gaujal B, Hordijk A (2000) Multimodularity, convexity, and optimizations properties. Math Oper Res 25:324–347

Armony M, Maglaris C (2003) On customer contact centers with a call-back option: customer decisions, routing rules, and system design. Oper Res 52:271–292

Bartholomew DJ, Forbes AF, McClean SI (1991) Statistical techniques for manpower planning, 2nd edn. Wiley, New York

Berman O, Larson RC (1994) Determining optimal pool size of a temporary call-in work force. Eur J Oper Res 73:55–64

Bordoloi SK, Matsuo H (2001) Human resource planning in knowledge-intensive operations: a model for learning with stochastic turnover. Eur J Oper Res 130:169–189

Borst S, Seri P (1999) Robust algorithms for sharing agents with multiple skills. Working paper, Bell Laboratories, Lucent Technologies, Murray Hill, NJ

Dixit A (1997) Investment and employment dynamics in the short run and the long run. Oxford Econ Pap 49:1–20

Eberly JC, van Mieghem JA (1997) Multi-factor dynamic investment under uncertainty. J Econ Theory 75:345–387

Fine CH, Freund RM (1990) Optimal investment in product-flexible manufacturing capacity. Manage Sci 36:449–466
Gans N, Koole G, Mandelbaum A (2003) Telephone call centers: tutorial, review, and research prospects. Manuf Serv Oper Manage 5:79–141
Gans N, Zhou Y-P (2002) Managing learning and turnover in employee staffing. Oper Res 50:991–1006
Hajek B (1985) Extremal splittings of point processes. Math Oper Res 10:543–556
Harrison JM, van Mieghem JA (1999) Multi-resource investment strategies: operational hedging under demand uncertainty. Eur J Oper Res 113:17–29
Harrison JM, Zeevi A (2003) A method for staffing large call centers based on stochastic fluid models (preprint)
Karlin S (1968) Total positivity, vol 1. Stanford University Press, Stanford
Koole G, Mandelbaum A (2002) Queueing models of call centers: an introduction. Ann Oper Res 113:41–59
Koole G, Pot A, Talim J (2003) Routing heuristics for multi-skill call centers. In: Proceedings of the 2003 winter simulation conference
Larson RC, Pinker EJ (2000) Staffing challenges in financial services. In: Melnick EL, Nayyar PR, Pinedo ML, Seshadri S (eds). Creating value in financial services: strategies, operations, and technologies, vol 17. Kluwer, Boston, pp 327–356
Mandelbaum A (2003) Call centers: research bibliography with abstracts. Downloadable from: http://ie.technion.ac.il/serveng
Murota K (2003) Discrete convex analysis. SIAM, Philadelphia
Murota K (2004) Note on multimodularity and L-convexity, report METR 2004-31. Department of Mathematical Informatics, The University of Tokyo, Tokyo, Japan
Misra S, Pinker E, Shumsky RA (2003) Sizing and structuring a slaesforce in the presence of job flexibility and experience-based learning (preprint)
Narongwanich W, Duenyas I, Birge JR (2003) Optimal portfolio of reconfigurable and dedicated capacity under uncertainty (preprint)
Netessine S, Dobson G, Shumsky RA (2002) Flexible service capacity: optimal investment and the impact of demand correlation. Oper Res 50:375–388
Perry M, Nilsson A (1994) Performance modeling of automatic call distributors: operator services staffing with heterogeneous positions. In: Proceedings of the 14th international teletraffic congress, ITC-14. Elsevier, Amsterdam, pp 1023–1032
Pinedo M, Seshadri S, Shanthikumar JG (2000) Call centers in financial services: strategies, technologies and operations. In: Melnick EL, Nayyar PR, Pinedo ML, Seshadri S (eds). Creating value in financial services: strategies, operations, and technologies, vol 18. Kluwer, Boston, pp 357–388
Pinker EJ, Larson RC (2003) Optimizing the use of contingent labor when demand is uncertain. Eur J Oper Res 144:39–55
Pinker EJ, Shumsky RA (2000) The efficiency-quality trade-off of cross-trained workers. Manuf Serv Oper Manage 2:32–48
Schmidt CP, Nahmias S (2003) Workforce planning under uncertainty (preprint)
Shumsky RA, Zhang F (2003) Dynamic capacity management with substitution (preprint)
van Mieghem JA (1998) Investment strategies for flexible resources. Manage Sci 44:1071–1078
van Mieghem JA (2002) Capacity portfolio investment and hedging: review and new directions (preprint)
van Mieghem JA, Rudi N (2002) Newsvendor networks: inventory management and capacity investment with discretionary activities. Manuf Serv Oper Manage 4:313–335