Technical Memorandum Number 751

Mathematical Models for Explaining the Emergence of Specialization in Performing Tasks

by

Daniel Solow and Joseph Szmerekovsky

August 2001

Department of Operations
Weatherhead School of Management
Case Western Reserve University
10900 Euclid Avenue
Cleveland, Ohio 44106-7235

Mathematical Models for Explaining the Emergence of Specialization in Performing Tasks

Daniel Solow and Joseph Szmerekovsky

Department of Operations
Weatherhead School of Management
Case Western Reserve University
Cleveland, OH 44106
Fax: (216) 368–4776
e-mail: dxs8@po.cwru.edu

August 28, 2001

Abstract

In an evolving community consisting of many individuals, it is often the case that the individuals tend, over time, to become more specialized in performing the tasks necessary for survival and growth of the community as a whole. In this work, linear and nonlinear mathematical models are presented for providing insights as to when and why this functional specialization emerges.

1 Introduction

In a *community* made up of *individuals*, each of whom can perform any of a number of necessary *tasks*, it is often the case that the individuals spend virtually all of their time performing one, or at most a few, tasks. This property—hereafter referred to as **functional specialization**—appears to emerge over time and is found, for example, in living organisms (where the individuals are cells), in human societies (where the individuals are people), in business organizations (where the individuals are the employees), and the like. On considering these and other examples, it seems likely that the reason functional specialization emerges and persists is because the community as a whole derives a benefit from functional specialization. Based on this assumption, the contribution of the work here is a collection of models that provides a mathematical justification for the value of functional specialization to the community and hence the reason for the emergence of this phenomenon.

Static models are presented and analyzed in Section 2. It is first shown how a linear model can fail to explain the emergence of functional specialization. Then, conditions are provided under which functional specialization arises in certain nonlinear models. In Section 3, functional specialization is shown to emerge in a dynamic model that changes from one time period to the next. It is assumed that the reader is familiar with the linear programming and transportation problems [see Bazaraa et al. (1990)], and nonlinear programming problems [see Bazaraa et al. (1993)].

2 Static Models in which Functional Specialization Emerges

In this section, static linear and nonlinear models are proposed for studying the emergence of functional specialization. To that end, consider a community made up of N individuals, each of whom can perform any of T tasks (usually $N >> T$). Functional specialization is related to the amount of time each individual devotes to each task. To model the time commitments of the individuals, define the following variables:

$$x_{it} = \text{the fraction of time that individual } i \text{ devotes to performing task } t \quad (i = 1, \ldots, N; t = 1, \ldots, T).$$

Of course, the sum of the fractions of time an individual spends on all tasks must be 1. Thus, the variables must satisfy the following constraints:

$$\sum_{t=1}^{T} x_{it} = 1, \qquad i = 1, \ldots, N \tag{1}$$

$$x_{it} \geq 0, \qquad \text{all } i \text{ and } t. \tag{2}$$

For the community to survive, it is assumed that the values of the variables $\mathbf{x} = (x_{it})$ need to satisfy P *survival constraints*, each of the form $g_j(\mathbf{x}) \geq 0$, and from which the community then receives a total benefit of $f(\mathbf{x})$. The objective is to determine values for the variables $\mathbf{x} = (x_{it})$ that achieve the largest value of $f(\mathbf{x})$ while satisfying the survival constraints $g_j(\mathbf{x})$ and also the constraints in (1) and (2). Thus, the proposed model is:

$$\begin{array}{lrcl} \max & f(\mathbf{x}) & & \\ \text{s.t.} & g_j(\mathbf{x}) & \geq & 0, \quad j = 1, \ldots, P \\ & \sum_{t=1}^{T} x_{it} & = & 1, \quad i = 1, \ldots, N \\ & x_{it} & \geq & 0, \quad \text{all } i \text{ and } t. \end{array} \tag{3}$$

An implicit assumption in the model in (3) is that the optimal time commitments of the individuals are chosen so as to benefit the community as a whole, rather than the individuals. Such an assumption is valid in the following situations:

- A community in which the benefits of the individuals are linked closely with those of the community. For example, in early human clans, the survival of an individual depended almost entirely on the survival of the clan as a whole. Another example might be an ant colony or the organs in the human body.
- A community in which there is an internal or external pressure on the individuals to perform in the interest of the community. For example, a society with a benevolent dictator or a leader that elicits such behavior from the individuals. Another example is a social structure in which societal goals outweigh individual goals.
- A community in which incentives are such that the individuals act in the best interest of the community. For example, a company in which employees are given incentives so that individual benefit and corporate benfit coincide.

For such communities, the model in (3) is appropriate. In the simplest case, one might consider a linear objective function and linear survival constraints in (3). The drawback of such a model is presented next.

2.1 A Linear Model and Its Drawbacks

To create a linear objective function, let

c_{it} = the (known) benefit to the community when individual i devotes 100% of its time to task t $(i = 1, \ldots, N; t = 1, \ldots, T)$.

Assuming that the benefits are scaled linearly and that the total benefit to the community is the sum of the individual contributions, the objective function in (3) becomes:

$$\max \sum_{i=1}^{N} \sum_{t=1}^{T} c_{it} x_{it} \tag{4}$$

Linear survival constraints arise when, for example, it is assumed that the total amount of time devoted by all individuals to a task t must be at least a specified amount, say, l_t. In this case, the survival constraints are:

$$\sum_{i=1}^{N} x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T. \tag{5}$$

Combining the various pieces in (4), (5), (1), and (2), the goal is to determine values for the variables x_{it} so as to solve the following linear programming problem (LP):

$$\begin{array}{llcll} \max & \sum_{i=1}^{N} \sum_{t=1}^{T} c_{it} x_{it} & & & \\ \text{s.t.} & \sum_{i=1}^{N} x_{it} & \geq & l_t, & t = 1, \ldots, T \\ & \sum_{t=1}^{T} x_{it} & = & 1, & i = 1, \ldots, N \\ & x_{it} & \geq & 0, & \text{all } i, t \end{array} \tag{6}$$

Functional specialization would be indicated if each individual spends 100% of its time on a single task. That is, functional specialization emerges if, in the optimal solution to problem (6), for each individual i, there is exactly one task s such that $x_{is} = 1$ and for each other task $t \neq s$, $x_{it} = 0$. As shown in Appendix A, it is possible to convert problem (6) to a balanced transportation problem [see Bazaraa et al. (1990)] for which it is known that the optimal solution obtained by the simplex algorithm consists entirely of the integers 0 and 1, provided that all data (c_{it} and l_t) are integers. Alternatively, it is not hard to show that the constraint matrix associated with the standard-form version of (6) is totally unimodular [see Ahuja et al. (1993)] and so every basic feasible solution is integer.

Unfortunately, the LP model in (6) and the fact that the optimal solution obtained by the simplex algorithm is integer is not a compelling argument that functional specialization *necessarily* emerges. This is because, while it is true that the simplex algorithm produces an optimal integer solution—thus indicating functional specialization—there can, and generally will, be other optimal solutions that contain fractional values for the variables—meaning that the community can do just as well without functional specialization. To illustrate, consider the following specific instance of the foregoing LP in which there are $N = 3$ individuals and $T = 2$ tasks:

$$
\begin{array}{lllll}
\text{maximize} & x_{11} + x_{21} + x_{31} + x_{12} + x_{22} + x_{32} & & & \\
& & & & \\
\text{s.t.} & \textbf{Task Requirements} & & & \\
& x_{11} + x_{21} + x_{31} & \geq & 1 & \text{(Task 1)} \\
& x_{12} + x_{22} + x_{32} & \geq & 1 & \text{(Task 2)} \\
& & & & \\
& \textbf{Individual Fractions Sum to 1} & & & \\
& x_{11} + x_{12} & = & 1 & \text{(Individual 1)} \\
& x_{21} + x_{22} & = & 1 & \text{(Individual 2)} \\
& x_{31} + x_{32} & = & 1 & \text{(Individual 3)} \\
& & & & \\
& \text{all } x_{it} & \geq & 0 &
\end{array}
\tag{7}
$$

The following optimal solution, in which Individuals 1 and 2 specialize in Task 1 and Individual 3 specializes in Task 2, provides an overall benefit of 3 to the community:

$$
\begin{array}{ll}
x_{11} = 1, & x_{12} = 0 \\
x_{21} = 1, & x_{22} = 0 \\
x_{31} = 0, & x_{32} = 1
\end{array}
$$

However, the optimal solution in which each of the three individuals spends half of the time on each of the two tasks also provides the same benefit of 3 to the community and is obtained without any specialization.

2.2 Nonlinear Models that Exhibit Functional Specialization

For functional specialization necessarily to emerge, it must be the case that *every* optimal solution to (3) is integer. This would happen, for example, when the linear programming model in (6) has a unique optimal solution, that is, when the associated dual problem is nondegenerate [see Bazaraa et al. (1990)]. An alternative explanation for functional specialization to emerge is when the objective and/or survival constraints of the model in (3) are nonlinear functions that possess certain properties. Such models, together with conditions under which every optimal solution is integer, are presented now.

2.2.1 A Model with a Nonlinear Objective Function

Consider the following model, in which the goal is to find values for the variables $\mathbf{x} = (x_{it})$ so as to:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \sum_{i=1}^{N} x_{it} \geq l_t, \quad t = 1, \ldots, T \\ & \sum_{t=1}^{T} x_{it} = 1, \quad i = 1, \ldots, N \\ & x_{it} \geq 0, \quad \text{all } i, t \end{aligned} \tag{8}$$

In the event that the objective function f is convex, it is well known [Bazaraa et al. (1993)] that (8) has an optimal solution at an extreme point of the feasible region. Furthermore, because of the special structure of the constraints in the foregoing model, every extreme point of the feasible region is integer. It follows that there is an optimal integer solution at an extreme point of (8). However, as mentioned in Section 2.1, in general there can also be optimal solutions that are not integer. A class of functions that ensures that every optimal solution to (8)—and other more general models—is integer is described next. In so doing, the following notations are used:

$S = \{\mathbf{y} \in R^T : \sum_{t=1}^{T} y_t = 1 \text{ and } y_t \geq 0 \text{ for } t = 1, \ldots, T\}$.

$S^N = S \times \cdots \times S =$ the N-fold Cartesian product of S.

$x_{it} =$ the fraction of time that individual i devotes to performing task t $(i = 1, \ldots, N; t = 1, \ldots, T)$.

$\mathbf{x}_i = (x_{i1}, \ldots, x_{iT}) \in S =$ the fractions of time devoted by individual i to each of the T tasks $(i = 1, \ldots, N)$.

$\mathbf{x} =$ $N \times T$ vector of variables $(\mathbf{x}_1, \ldots, \mathbf{x}_N) \in S^N$.

2.2.2 Sublinear Functions

In this section, a property of three different types of functions is identified and subsequently used to ensure that every optimal solution to a nonlinear programming model is integer. The property is first defined as follows for a function $h : [0, 1] \to R^1$ (see Figure 1):

Definition 1 *A continuous function $h : [0, 1] \to R^1$ is* **sublinear** *if and only if the graph of h lies strictly below the line segment connecting $h(0)$ and $h(1)$, that is, for all $0 < x < 1$, $h(x) < (1 - x)h(0) + xh(1)$.*

Example 1: Any strictly convex continuous function on R^1, such as $h(x) = x^2$, is sublinear. This is because, by definition, a strictly convex function h satisfies the property that for all

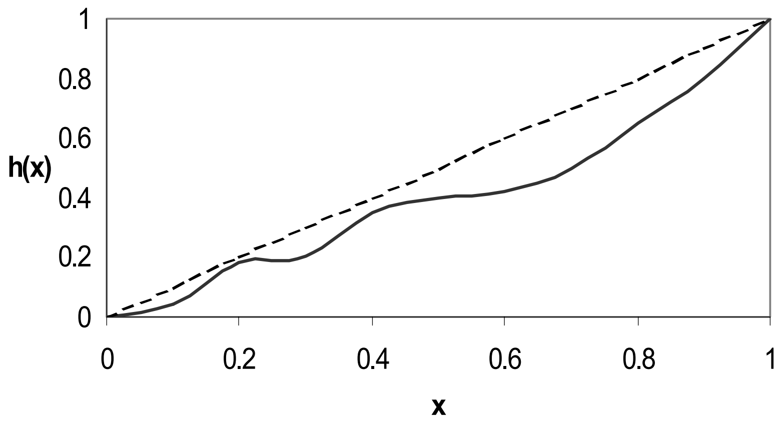

Figure 1

$y, z \in R^1$ with $y \neq z$ and for all $0 < t < 1$, $h((1-t)y + tz) < (1-t)h(y) + th(z)$ and so, for $y = 0$, $z = 1$, and $t = x$, it follows that

$$h(x) = h((1-x)0 + x(1)) < (1-x)h(0) + xh(1).$$

The importance of Definition 1 is that the maximum of a sublinear function h over the interval $[0, 1]$ is at 0 or 1. A natural generalization of this property to a function $h : S \to R^1$ follows.

Definition 2 *A continuous function $h : S \to R^1$ is* **sublinear** *if and only if h achieves its maximum value of $c = \max\{h(\mathbf{x}) : \mathbf{x} \in S\}$ only at a vertex of S, that is, for all $\mathbf{z} \in S$, $h(\mathbf{z}) = c \Rightarrow$ there is an integer t with $1 \leq t \leq T$ such that $z_t = 1$.*

Example 2: If $h_t : [0, 1] \to R^1$ are sublinear functions such that for each $t = 1, \ldots, T$, $h_t(0) = a$ and $h_t(1) = b$, then the following function $h : S \to R^1$ is sublinear:

$$h(\mathbf{z}) = \sum_{t=1}^{T} h_t(z_t).$$

This is because if $\mathbf{z} \in S$ has fractional values, then $h(\mathbf{z}) < h(\mathbf{y})$ for any integer point $\mathbf{y} \in S$ since

$$\begin{aligned}
h(\mathbf{z}) &= \sum_{t=1}^{T} h_t(z_t) \\
&= \sum_{\{t:0<z_t<1\}} h_t(z_t) + \sum_{\{t:z_t=0\}} h_t(z_t) \\
&< \sum_{\{t:0<z_t<1\}} [(1-z_t)h_t(0) + z_t h_t(1)] + \sum_{\{t:z_t=0\}} (1-z_t)h_t(0) \\
&= \sum_{t=1}^{T} (1-z_t)h_t(0) + \sum_{\{t:0<z_t<1\}} z_t h_t(1) \\
&\leq (T-1)a + b \\
&= h(\mathbf{y}).
\end{aligned}$$

In particular, the function h in Example 2 achieves its maximum value at, and only at, any integer point in S.

The final type of sublinear function considered here is described in the following two definitions.

Definition 3 *A function $h : S^N \to R^N$ is* **decomposable** *if and only if there are functions $h_i : S \to R^1$, for $i = 1, \ldots, N$, such that for all $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in S^N$, $h(\mathbf{x}) = (h_1(\mathbf{x}_1), \ldots, h_N(\mathbf{x}_N))$.*

Definition 4 *A decomposable continuous function $h : S^N \to R^N$, in which $h(\mathbf{x}) = (h_1(\mathbf{x}_1), \ldots, h_N(\mathbf{x}_N))$, is* **sublinear** *if and only if for each $i = 1, \ldots, N$, $h_i : S \to R^1$ is sublinear.*

With the concept of sublinear functions, it is now possible to establish that functional specialization necessarily emerges in certain nonlinear models, that is, that every optimal solution is integer.

2.2.3 A Model with a Nonlinear Objective Function and Constraints

In addition to a nonlinear objective function $f(\mathbf{x})$, suppose that the P survival constraints are also nonlinear and have the form $g_j(\mathbf{x}) \geq 0$, so:

$g : S^N \to R^P$ = the P survival constraints that the variables must satisfy.

$f : S^N \to R^1$ = the benefit the community derives from the time allocations of all individuals.

The objective now is to find values for the variables $\mathbf{x} = (x_{it})$ so as to solve the following nonlinear programming problem (NLP):

$$\begin{array}{lrcl}
\max & f(\mathbf{x}) & & \\
\text{s.t.} & g(\mathbf{x}) & \geq & \mathbf{0} \\
& \mathbf{x} & \in & S^N
\end{array} \qquad (9)$$

The next theorem provides conditions under which functional specialization necessarily emerges in the optimal solution to the NLP in (9), that is, conditions under which every optimal solution to the NLP is integer. That theorem uses the following definition.

Definition 5 *For vectors $\mathbf{y}, \mathbf{z} \in R^N$, by $\mathbf{y} \preceq \mathbf{z}$ it is meant that:*

(a) For each $i = 1, \ldots, N$, $y_i \leq z_i$ and

(b) There is an integer j with $1 \leq j \leq N$ such that $y_j < z_j$.

Theorem 1 *If there is a sublinear function $h : S^N \to R^N$ such that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$ and there is a feasible solution $\bar{\mathbf{x}}$ for the NLP in (9) such that for each $i = 1, \ldots, N$, $h_i(\bar{\mathbf{x}}_i) = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\}$, then any optimal solution must be integer.*

Proof. Let $\mathbf{x}^* \in S^N$ be an optimal solution for (9). It will be shown by contradiction that $\mathbf{x}^*$ is integer. So, suppose that $\mathbf{x}^*$ has fractional values. Then there are integers i and t with $1 \leq i \leq N$ and $1 \leq t \leq T$ such that $0 < x^*_{it} < 1$. A contradiction is reached by showing that $f(\mathbf{x}^*) < f(\bar{\mathbf{x}})$. To that end, define

$$\mathcal{F} = \{\ 1 \leq i \leq N : \text{there is a } t \text{ with } 1 \leq t \leq T \text{ such that } 0 < x^*_{it} < 1\} \neq \emptyset.$$

Using the sublinearity of h and letting $c_i = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\} = h_i(\bar{\mathbf{x}}_i)$ for each $i = 1, \ldots, N$, it follows that

$$\begin{aligned} h_i(\mathbf{x}^*_i) &< c_i = h_i(\bar{\mathbf{x}}_i), \quad \text{for each } i \in \mathcal{F} \\ h_i(\mathbf{x}^*_i) &\leq c_i = h_i(\bar{\mathbf{x}}_i), \quad \text{for each } i \notin \mathcal{F}. \end{aligned} \tag{10}$$

In other words, (10) says that $\mathbf{x}^*$ and $\bar{\mathbf{x}}$ are two vectors in S^N for which $h(\mathbf{x}^*) \preceq h(\bar{\mathbf{x}})$. The hypothesis now ensures that $f(\mathbf{x}^*) < f(\bar{\mathbf{x}})$ and so $\mathbf{x}^*$ is not optimal. This contradiction establishes the claim that all optimal solutions are integer. □

Theorem 1 provides conditions under which every optimal solution is integer. However, the theorem does not indicate *which* integer solution in S^N is optimal. The next theorem shows that the feasible solution in the hypotheses of Theorem 1 is optimal, provided that the objective function f is continuous.

Theorem 2 *If f is continuous on S^N and there is a sublinear function $h : S^N \to R^N$ such that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$ and there is a feasible solution $\mathbf{x}^*$ for the NLP in (9) such that for each $i = 1, \ldots, N$, $h_i(\mathbf{x}^*_i) = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\}$, then $\mathbf{x}^*$ is optimal and integer.*

Proof. Once $\mathbf{x}^*$ is shown to be optimal, Theorem 1 ensures that $\mathbf{x}^*$ is integer. So, to see that $\mathbf{x}^*$ is optimal, let $\mathbf{x}$ be any feasible solution for the NLP. It will be shown that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$.

Case 1. The feasible solution $\mathbf{x}$ has a fractional value. In this case, define

$$\mathcal{F} = \{\ 1 \leq i \leq N : \text{there is a } t \text{ with } 1 \leq t \leq T \text{ such that } 0 < x_{it} < 1\} \neq \emptyset.$$

Using the sublinearity of h and letting $c_i = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\} = h_i(\mathbf{x}_i^*)$ for each $i = 1, \ldots, N$, it follows that

$$\begin{aligned} h_i(\mathbf{x}_i) &< c_i = h_i(\mathbf{x}_i^*), \quad \text{for each } i \in \mathcal{F} \\ h_i(\mathbf{x}_i) &\leq c_i = h_i(\mathbf{x}_i^*), \quad \text{for each } i \notin \mathcal{F}. \end{aligned} \tag{11}$$

In other words, (11) says that $\mathbf{x}$ and $\mathbf{x}^*$ are two vectors in S^N for which $h(\mathbf{x}) \preceq h(\mathbf{x}^*)$. The hypothesis now ensures that $f(\mathbf{x}) < f(\mathbf{x}^*)$.

Case 2. The feasible solution $\mathbf{x}$ is integer. Note that the argument in Case 1 does not apply because $\mathcal{F} = \emptyset$. So, consider a sequence of points $(\mathbf{x}^k)$ in S^N, each of which has a fractional value, and that converges to $\mathbf{x}$. Repeating the argument in Case 1 for each $k = 1, 2, \ldots$, it follows that $h(\mathbf{x}^k) \preceq h(\mathbf{x}^*)$. Thus, by the hypothesis, for each $k = 1, 2, \ldots$,

$$f(\mathbf{x}^k) < f(\mathbf{x}^*). \tag{12}$$

The fact that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ now follows by taking the limit as $k \to \infty$ on both sides of (12) and using the continuity of f.

The proof is now complete because it has been shown that for any feasible solution $\mathbf{x}$, $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ and so $\mathbf{x}^*$ is optimal. □

Theorem 2 reduces the problem of solving the NLP in (9) to finding a feasible integer solution $\mathbf{x}^* \in S^N$ such that for each $i = 1, \ldots, N$, $h_i(\mathbf{x}_i^*) = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\}$. For each i, the point $\mathbf{x}_i^*$ can be found by evaluating the T integer points in S to determine one that provides the maximum value of h_i. Then, $\mathbf{x}^* = (\mathbf{x}_1^*, \ldots, \mathbf{x}_N^*)$ is the desired optimal solution to the NLP in (9), provided that $\mathbf{x}^*$ is feasible. The process of finding such a solution $\mathbf{x}^*$ is simplified even further when each h_i achieves the same maximum value at all integer points in S, as illustrated in the following example.

Example 3: For each i and t, let $h_{it} : [0, 1] \to R^1$ be sublinear functions (for example, $h_{it}(x_{it}) = x_{it}^2$). The continuous function

$$f(\mathbf{x}) = \sum_{i=1}^{N} \sum_{t=1}^{T} h_{it}(x_{it})$$

satisfies the hypotheses of Theorems 1 and 2. To see that this is so, consider the functions $h_i : S \to R^1$ defined by

$$h_i(\mathbf{z}) = \sum_{t=1}^{T} h_{it}(z_t), \quad \text{for } i = 1, \ldots, N.$$

Now each h_i is sublinear (see Example 2 in Section 2.2.2) and so the decomposable function $h : S^N \to R^N$ defined by $h(\mathbf{x}) = (h_1(\mathbf{x}_1), \ldots, h_N(\mathbf{x}_N))$ for each $\mathbf{x} \in S^N$ is also sublinear.

Also, h satisfies the property that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$. This is because, if $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, then, for each $i = 1, \ldots, N$,

$$\sum_{t=1}^{T} h_{it}(x_{it}) = h_i(\mathbf{x}_i) \leq h_i(\mathbf{y}_i) = \sum_{t=1}^{T} h_{it}(y_{it}), \tag{13}$$

with strict inequality holding for at least one value of i. It then follows by summing (13) over i that

$$f(\mathbf{x}) = \sum_{i=1}^{N}\sum_{t=1}^{T} h_{it}(x_{it}) < \sum_{i=1}^{N}\sum_{t=1}^{T} h_{it}(y_{it}) = f(\mathbf{y}).$$

By Theorem 2, any feasible integer solution $\mathbf{x}^*$ for the NLP in (9) such that for each $i = 1, \ldots, N$, $h_i(\mathbf{x}_i^*) = \max\{h_i(\mathbf{z}) : \mathbf{z} \in S\}$ is optimal. In the event that each h_i achieves the same maximum value of c_i at all integer points in S, any feasible integer solution $\mathbf{x} \in S^N$ is optimal for the NLP in (9). This occurs when, for example, each h_{it} satisfies $h_{it}(0) = a_i$ and $h_{it}(1) = b_i$, as happens when $h_{it}(x_{it}) = x_{it}^2$, in which $h_{it}(0) = 0$ and $h_{it}(1) = 1$.

All of the models presented so far are static and therefore remain the same over time. A dynamic model is presented now for studying the emergence of functional specialization as the system changes over time.

3 A Dynamic Model in which Functional Specialization Emerges

The proposed dynamic model incorporates the fact that the more time an individual spends on a task, the better that individual becomes at performing that task. This is analogous to a neural network in which the pathways are reinforced by repeated use [see Muller and Reinhardt (1991)]. Thus, let

A_{it}^k = the (known) productive fraction of each time unit that individual i puts into task t in time period k $(i = 1, \ldots, N;$ $t = 1, \ldots, T; k = 0, 1, \ldots)$.

$\mathcal{A}$ = the set of all $N \times T$ matrices A such that for all i and t, $0 \le A_{it} \le 1$.

The closer the value A_{it}^k is to 1, the more efficient individual i is at performing task t. Moreover, the value $A_{it}^k x_{it}$ represents the *effective* number of time units individual i devotes to task t in period k.

For the community to survive, it is assumed that the total number of effective time units devoted by all individuals to each task t in period k must be at least a specified amount, l_t. Thus, the survival constraints in period k are:

$$\sum_{i=1}^{N} A_{it}^k x_{it} \ge l_t, \quad \text{for } t = 1, \ldots, T.$$

Furthermore, let the function $f(\mathbf{x})$ be the benefit to the community from the time allocations $\mathbf{x}$ of the individuals. Putting together the pieces, the goal is to determine, for each time period k, values for the variables $\mathbf{x} = (x_{it})$ so as to solve the following nonlinear programming problem, denoted by NLP(A^k):

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \sum_{i=1}^{N} A_{it}^k x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T \quad [\text{NLP}(A^k)] \\ & \mathbf{x} \in S^N \end{aligned} \tag{14}$$

An optimal solution $\mathbf{x}^k$ to NLP(A^k) need not be integer. However, the values of $\mathbf{x}^k = (x_{it}^k)$ are used now to change the values of A_{it}^k to reflect the fact that individuals who spend time on a task become more efficient at that task. For example, if individual i spends 100% of its time on task t in period k ($x_{it}^k = 1$), then the value of A_{it}^k should increase by some fraction F ($0 < F < 1$) of the way to 1. More generally, if individual i spends the fraction x_{it}^k of its time on task t in period k, then A_{it}^k increases by the fraction $x_{it}^k F$ of the way to 1, so,

$$A_{it}^{k+1} = A_{it}^k + x_{it}^k F(1 - A_{it}^k). \tag{15}$$

On the basis of the update formula in (15), the following dynamic model is proposed:

A Dynamic Nonlinear Programming Model

Step 0: Let $A^0 \in \mathcal{A}$ be a given matrix for which NLP(A^0) is feasible and set $k = 0$.

Step 1: Let $\mathbf{x}^k$ be an optimal solution to NLP(A^k).

Step 2: Use $\mathbf{x}^k$ and the update formula in (15) to compute the matrix A^{k+1}. Set $k = k+1$ and go to Step 1.

The goal is to provide conditions under which the matrices A^k generated by this process converge to a matrix A^∞ for which functional specialization necessarily emerges—that is, every convergent subsequence of $\mathbf{x}^k$ converges to an integer optimal solution of NLP(A^∞). The first step in this direction is to ensure that it is always possible to obtain an optimal solution to NLP(A^k) in Step 1, as established in the next theorem.

Theorem 3 *If NLP(A^0) is feasible and f is continuous on S^N, then for each $k = 0, 1, \ldots$, NLP(A^k) has an optimal solution, $\mathbf{x}^k$.*

Proof. The statement is true for $k = 0$ because, by hypothesis, NLP(A^0) is feasible. Furthermore, the constraints $\mathbf{x} \in S^N$ together with the continuity of f ensure that NLP(A^0) has an optimal solution.

Assume now that the statement is true for k and let $\mathbf{x}^k$ be an optimal solution for NLP(A^k). Then $\mathbf{x}^k$ is feasible for NLP(A^{k+1}). This is because $\mathbf{x}^k \in S^N$ and, from the update formula in (15), for each i and t, $A_{it}^{k+1} \geq A_{it}^k$, so

$$\sum_{i=1}^{N} A_{it}^{k+1} x_{it}^k \geq \sum_{i=1}^{N} A_{it}^k x_{it}^k \geq l_t.$$

Thus, NLP(A^{k+1}) is feasible. The constraints that $\mathbf{x} \in S^N$ together with the continuity of f ensure that NLP(A^{k+1}) has an optimal solution. The result now follows by induction. □

From here on, it is assumed that NLP(A^0) is feasible and that $\mathbf{x}^k$ is optimal for NLP(A^k) and generates A^{k+1}. The next theorem establishes that the matrices A^k converge, element by element, to a matrix A^∞ for which the associated NLP(A^∞) has an optimal solution if f is continuous on S^N.

Theorem 4 *If NLP(A^0) is feasible and f is continuous on S^N, then there is a matrix $A^\infty \in \mathcal{A}$ such that for each i and t, $A^\infty_{it} = \lim_{k\to\infty} A^k_{it}$ and NLP(A^∞) is optimal.*

Proof. For each i and t, the update formula in (15) ensures that the sequence of real numbers (A^k_{it}) is monotonically non-decreasing and bounded above by 1. It follows from the Monotone Convergence Theorem [see Bartle and Sherbert (1992)] that for each i and t, there is a real number A^∞_{it} with $0 \leq A^\infty_{it} \leq 1$ such that $A^\infty_{it} = \lim_{k\to\infty} A^k_{it}$.

It is now shown that NLP(A^∞) is feasible. To that end, from Theorem 3, let $\mathbf{x}^k \in S^N$ be an optimal solution to NLP(A^k), so, for each k,

$$\sum_{i=1}^{N} A^\infty_{it} x^k_{it} \geq \sum_{i=1}^{N} A^k_{it} x^k_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T. \tag{16}$$

Thus, $\mathbf{x}^k$ is feasible for NLP(A^∞). Finally, the constraints $\mathbf{x} \in S^N$ and the continuity of f ensure that NLP(A^∞) has an optimal solution, thus completing the proof. □

The issue now is what is happening to the sequence $(\mathbf{x}^k)$ of optimal solutions to NLP(A^k). Because $(\mathbf{x}^k)$ belongs to the compact set S^N, there is a subsequence K and a point $\mathbf{x}^* \in S^N$ such that $(\mathbf{x}^k)$ converges to $\mathbf{x}^*$ as $k \in K$. The next theorem, whose proof is given in Appendix B, establishes that $\mathbf{x}^*$ is an optimal solution to NLP(A^∞).

Theorem 5 *If f is continuous on S^N and the sequence of matrices (A^k) converges element by element to a matrix $A^\infty \in \mathcal{A}$ and for each $k = 0, 1, \ldots$, $\mathbf{x}^k$ solves NLP(A^k), then any convergent subsequence of $(\mathbf{x}^k)$ converges to an optimal solution $\mathbf{x}^*$ of NLP(A^∞).*

For functional specialization to emerge, it is necessary to establish that $\mathbf{x}^*$ is integer. This is accomplished by showing that, under suitable conditions on f, $\mathbf{x}^*$ is an optimal extreme point of the following Nonlinear Transportation Problem, denoted by NLTP:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \sum_{i=1}^{N} u_{it} x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T \quad (NLTP) \\ & \mathbf{x} \in S^N \end{aligned} \tag{17}$$

where u_{it} is defined as follows:

$$u_{it} = \begin{cases} 1, & \text{if } x^*_{it} > 0 \\ 0, & \text{if } x^*_{it} = 0 \end{cases} \tag{18}$$

It is now possible to show that, with no conditions on f other than continuity, $\mathbf{x}^*$ is optimal for NLTP. To that end, the following lemma is proved first.

Lemma 1 *If* $\mathbf{x}_{it}^* > 0$*, then* $A_{it}^{\infty} = 1$.

Proof. From the update formula in (15), for each $k = 0, 1, \ldots,$

$$A_{it}^{k+1} = A_{it}^k + x_{it}^k F(1 - A_{it}^k) \tag{19}$$

Taking the limit on both sides of (19) over k in the subsequence for which $(\mathbf{x}^k)$ converges to $\mathbf{x}^*$ and recalling that (A_{it}^k) converges to A_{it}^{∞} yields

$$A_{it}^{\infty} = A_{it}^{\infty} + x_{it}^* F(1 - A_{it}^{\infty}).$$

The result that $A_{it}^{\infty} = 1$ follows by subtracting A_{it}^{∞} from both sides of the foregoing equation and then dividing by $x_{it}^* F > 0$. $\square$

Theorem 6 *If* f *is continuous on* S^N *and* $\mathbf{x}^*$ *is the optimal solution for NLP(*A^{∞}*) in Theorem 5, then* $\mathbf{x}^*$ *is optimal for NLTP.*

Proof. It is first shown that $\mathbf{x}^*$ is feasible for NLTP. To that end, note from Theorem 5 that $\mathbf{x}^*$ is optimal, and hence feasible, for NLP(A^{∞}), so

$$\sum_{i=1}^{N} A_{it}^{\infty} x_{it}^* \geq l_t, \quad \text{for } t = 1, \ldots, T \tag{20}$$

Consider a fixed value of t between 1 and T and define $X^+ = \{1 \leq i \leq N : x_{it}^* > 0\}$ and $X^0 = \{1 \leq i \leq N : x_{it}^* = 0\}$. Then,

$$\begin{aligned}
\sum_{i=1}^{N} u_{it} x_{it}^* &= \sum_{i \in X^+} u_{it} x_{it}^* + \sum_{i \in X^0} u_{it} x_{it}^* \\
&= \sum_{i \in X^+} (1) x_{it}^* + \sum_{i \in X^0} u_{it}(0) && \text{(definition of } u_{it}, X^+, \text{ and } X^0) \\
&= \sum_{i \in X^+} A_{it}^{\infty} x_{it}^* + \sum_{i \in X^0} A_{it}^{\infty}(0) && \text{(Lemma 1)} \\
&= \sum_{i=1}^{N} A_{it}^{\infty} x_{it}^* && \text{(definition of } X^0) \\
&\geq l_t && \text{[from (20)]}
\end{aligned}$$

To see that $\mathbf{x}^*$ is optimal for NLTP, let $\mathbf{x}$ be any feasible solution for NLTP. It will be shown that $f(\mathbf{x}^*) \geq f(\mathbf{x})$ by establishing that $\mathbf{x}$ is feasible for NLP(A^{∞}), for then, because, by hypothesis, $\mathbf{x}^*$ is optimal for NLP(A^{∞}), $f(\mathbf{x}^*) \geq f(\mathbf{x})$. Thus, it remains to show that $\mathbf{x}$

is feasible for NLP(A^∞), but,

$$\begin{aligned}
l_t &\le \sum_{i\in X^+} u_{it}x_{it} + \sum_{i\in X^0} u_{it}x_{it} && (\mathbf{x} \text{ is feasible for NLTP})\\
&= \sum_{i\in X^+} (1)x_{it} + \sum_{i\in X^0} (0)x_{it} && (\text{definition of } u_{it})\\
&= \sum_{i\in X^+} A^\infty_{it}x_{it} + \sum_{i\in X^0} (0)x_{it} && (\text{Lemma 1})\\
&\le \sum_{i\in X^+} A^\infty_{it}x_{it} + \sum_{i\in X^0} A^\infty_{it}x_{it}\\
&= \sum_{i=1}^{N} A^\infty_{it}x_{it}. \qquad \square
\end{aligned}$$

Although Theorem 6 establishes that $\mathbf{x}^*$ is optimal for NLTP, $\mathbf{x}^*$ need not be integer. For example, consider the following linear programming problems—denoted by LP(A^k)—in which the community accrues a benefit from each task t in each period that is equal to c_t times the number of time units dedicated to task t:

$$\begin{aligned}
\max \quad & \sum_{t=1}^{T} c_t x_{it}\\
\text{s.t.} \quad & \sum_{i=1}^{N} A^k_{it}x_{it} \ge l_t, \quad \text{for } t = 1,\ldots,T \quad [\text{LP}(A^k)]\\
& \mathbf{x} \in S^N
\end{aligned} \tag{21}$$

Although it can be shown that the limiting problem LP(A^∞) has an optimal integer solution, a convergent subsequence of optimal solutions to LP(A^k) need not converge to an integer optimal solution.

Nonlinearity is needed to obtain such a result. The next theorem provides a sufficient condition on f to ensure that $\mathbf{x}^*$ is integer and hence that functional specialization emerges from the solutions to the dynamic model.

Theorem 7 *If f is continuous and strictly convex on S^N, then every optimal solution to NLTP, and hence $\mathbf{x}^*$, is integer.*

Proof. Because f is strictly convex, an optimal solution of NLTP can occur only at an extreme point of NLTP. Furthermore, every extreme point of NLTP is integer because of the transportation-like constraints. Thus, every optimal solution of NLTP is integer. $\square$

Theorem 7 establishes the emergence of functional specialization from the solutions to NLP(A^k) in the dynamic model when f is strictly convex. This is because any convergent subsequence of optimal time allocations for NLP(A^k) approaches $\mathbf{x}^*$, which is an optimal integer solution to both NLTP and NLP(A^∞). Note that there could be non-integer optimal solutions to NLP(A^∞), however, because of the way in which the system evolves, under the conditions of Theorem 7, no convergent subsequence of optimal solutions to NLP(A^k)

will converge to such a non-integer solution. One might ask what would happen if the dynamic model began with $B^0 = A^\infty$ and a non-integer optimal solution, $\mathbf{y}^0$, to NLP(B^0) were obtained. In this case, the dynamic model results in a sequence of matrices, (B^k), and corresponding optimal solutions, $\mathbf{y}^k$, to NLP(B^k) for which, under the conditions of Theorem 7, any convergent subsequence of $(\mathbf{y}^k)$ converges to a solution $\mathbf{y}^*$ in which functional specialization emerges. In other words, under the conditions of Theorem 7, functional specialization is an inevitable outcome of the dynamic model, regardless of the starting conditions.

4 Extensions and Future Research

In the dynamic model in Section 3, it is assumed that individuals never get worse at tasks and thus the matrices A^k are monotonically non-decreasing, as seen in the update formula in (15). A more realistic assumption is that if an individual i spends no time on task t in period k ($x_{it}^k = 0$), then the value of A_{it}^k decreases in the next period by the fraction F ($0 < F < 1$) of the way to 0. In this case, the update formula becomes:

$$A_{it}^{k+1} = \begin{cases} A_{it}^k + x_{it} F(1 - A_{it}^k), & \text{if } x_{it}^k > 0 \\ \\ (1-F)A_{it}^k, & \text{if } x_{it}^k = 0 \end{cases} \tag{22}$$

In this case, Theorem 3 in Section 3 still applies, meaning that as long as NLP(A^0) is feasible and f is continuous on S^N, each subsequent NLP(A^k) has an optimal solution. However, the remaining theorems in Section 3 need not hold. For example, Theorem 4 is no longer valid because the sequence (A^k) of matrices is no longer monotone. It is, however, possible to prove the following theorem.

Theorem 8 *If NLP(A^0) is feasible and f is continuous on S^N, then there is a matrix $A^\infty \in \mathcal{A}$ and a subsequence K such that for each i and t, $A_{it}^\infty = \lim_{k \in K} A_{it}^k$, the sequence $(A_{it}^k)_{k \in K}$ is monotone, and NLP(A^∞) is optimal.*

Theorem 5 in Section 3 holds for the subsequence K in Theorem 8. That is, by defining, for each $j = 1, 2, \ldots$, the matrix B^j to be the j^{th} matrix in the subsequence $(A^k)_{k \in K}$, the hypotheses, and hence the conclusion, of Theorem 5 apply to the sequence of matrices (B^j). Thus, any convergent subsequence of optimal solutions for NLP(B^j) converges to an optimal solution of NLP($B^\infty = A^\infty$). However, letting $\mathbf{y}^j$, for $j = 1, 2, \ldots$ be an optimal solution for NLP(B^j), it is an open question as to whether any convergent subsequence of $(\mathbf{y}^j)$ converges to an integer optimal solution $\mathbf{y}^*$ for NLP(B^∞). In particular, no suitable version of Lemma 1 has been found as yet for this case. Nonetheless, repeated simulations with a linear objective function have resulted in the emergence of functional specialization in this modified dynamic model.

Additional mathematical difficulties manifest themselves in the dynamic model when the update formula includes a critical fraction of time, b (or b_{it}), for which an individual who spends at least the fraction b of time at a task gets proportionately better at that task and an individual who spends less than the fraction b gets proportionally worse at that task. In this case, the update formula becomes:

$$A_{it} = \begin{cases} A_{it} + \left(\frac{x_{it}-b}{1-b}\right) F(1 - A_{it}), & \text{if } x_{it} \geq b \\ \\ \left[1 - \left(\frac{b-x_{it}}{b}\right) F\right] A_{it}, & \text{if } x_{it} < b \end{cases} \tag{23}$$

In this case, even if NLP(A^0) is feasible, a subsequent NLP(A^k) can become infeasible and so the dynamic model is no longer well defined. To illustrate, consider the following numerical example in which two individuals must allocate their time among two different tasks so as to maximize a linear objective function. For the initial problem, NLP(A^0), it is assumed that each individual is 100% efficient at both tasks. The problem NLP(A^0) therefore is the following:

$$\begin{array}{lll} \text{maximize} & x_{11} + x_{21} + x_{12} + x_{22} & \\ & & \\ \text{s.t.} & x_{11} + x_{21} & \geq \ 1 \\ & x_{12} + x_{22} & \geq \ 1 \\ & & \\ & x_{11} + x_{12} & = \ 1 \\ & x_{21} + x_{22} & = \ 1 \\ & & \\ & \text{all } x_{it} & \geq \ 0 \end{array} \tag{24}$$

The following solution, in which each individual spends one half of its time on Task 1 and one half of its time on Task 2, is optimal for this problem:

$$\begin{array}{ll} x_{11} = 1/2, & x_{21}, = 1/2 \\ x_{12} = 1/2, & x_{22}, = 1/2 \end{array}$$

However, if $F = 1$ and $b = 2/3$ is the minimum time requirement to maintain performance on a particular task then, using the update formula in (23), the following problem, NLP(A^1), is easily seen to be infeasible:

$$\begin{array}{lll} \text{maximize} & x_{11} + x_{21} + x_{12} + x_{22} & \\ & & \\ \text{s.t.} & \frac{5}{6}x_{11} + \frac{5}{6}x_{21} & \geq \ 1 \\ & \frac{5}{6}x_{12} + \frac{5}{6}x_{22} & \geq \ 1 \\ & & \\ & x_{11} + x_{12} & = \ 1 \\ & x_{21} + x_{22} & = \ 1 \\ & & \\ & \text{all } x_{it} & \geq \ 0 \end{array} \tag{25}$$

Another direction for future research is to extend the dynamic model in Section 3 to allow for an increasing number of tasks, individuals, and minimum survival requirements (l_t) over time. In particular, at what rate can these parameters increase so that functional specialization still emerges?

As a final direction for future research, recall that the models developed here are based on the assumption that the optimal time allocations are made on the basis of what is best for the community. A different approach is needed when the individuals allocate their time so as to maximize their own interests rather than those of the community as a whole. These and other related questions are currently under investigation.

Conclusion

In this work, linear and nonlinear static and dynamic models are proposed for studying the emergence of functional specialization. These models apply to a community in which the individuals allocate their time in the best interest of the community. It is shown how a linear model is inadequate to account for the emergence of functional specialization. Rather, it is nonlinearity in the objective function that provides an explanation for this phenomenon. In particular, sufficient conditions are provided on a nonlinear objective function under which functional specialization necessarily emerges in a static model and also in a dynamic model that allows for the individuals to get better at tasks over time.

Appendix A

In this appendix, it is shown how to transform the linear programming model (6) in Section 2.1 to a balanced transportation problem. Assuming that $N > \sum_t l_t$, the transformation is accomplished by creating a single dummy task, $T+1$, with a requirement of $l_{T+1} = N - \sum_{t=1}^{T} l_t$. For each individual i, set the value of $c_{iT+1} = \max\{c_{it} : 1 \leq t \leq T\}$. Then, the balanced transportation problem is:

$$
\begin{array}{lrcll}
\max & \sum_{i=1}^{N} \sum_{t=1}^{T+1} c_{it} w_{it} & & & \\
\text{s.t.} & \sum_{i=1}^{N} w_{it} & = & l_t, & t = 1, \ldots, T+1 \\
& \sum_{t=1}^{T+1} w_{it} & = & 1, & i = 1, \ldots, N \\
& w_{it} & \geq & 0, & \text{all } i, t
\end{array}
\tag{26}
$$

The next two theorems establish that any optimal solution to (6) provides an optimal solution to (26) with the same objective function value and vice versa.

Theorem 9 *If* $\mathbf{x}$ *is an optimal solution for (6) with objective function value* $c(\mathbf{x})$*, then there is an optimal solution to (26) whose objective function value is at least* $c(\mathbf{x})$*.*

Proof. The key observation is that, in the optimal solution to (6), for any task $t = 1, \ldots, T$ for which $\sum_i x_{it} > l_t$ and for any individual $i = 1, \ldots, N$ with $x_{it} > 0$, any amount of x_{it} up to $l_t - \sum_i x_{it}$ can be diverted from task t to task $T+1$ in (26) without affecting the optimal objective function value, $c(\mathbf{x})$. To see that this is so, for any individual i, let j be an integer for which $c_{ij} = \max_{t=1,\ldots,T}\{c_{it}\}$. Then, for a task t for which $\sum_i x_{it} > l_t$ and any individual i with $x_{it} > 0$, it must be that $c_{it} = c_{ij}$, for otherwise, it would be possible to create a better feasible solution for (6) than $\mathbf{x}$ by reducing x_{it} by some positive amount δ and increasing the value of x_{ij} by δ.

On the basis of the foregoing observation, a feasible solution to (26) is constructed by diverting, from each task t for which $\sum_i x_{it} > l_t$ and some of the individuals i, some amount of x_{it} to task $T+1$. To specify these amounts, for any task t for which $\sum_i x_{it} > l_t$, define

$$
i(t) \quad = \quad \text{the largest integer for which } \sum_{i=1}^{i(t)} x_{it} < l_t
$$

$$
y_{it} \quad = \quad \begin{cases} x_{it}, & \text{for } i = 1, \ldots, i(t) \\ l_t - \sum_{i=1}^{i(t)} x_{it}, & \text{for } i = i(t)+1 \\ 0, & \text{for } i = i(t)+2, \ldots, T \end{cases} \quad \text{and } w_{it} = \begin{cases} y_{it}, & \text{for } t = 1, \ldots, T \\ 1 - \sum_{j=1}^{T} y_{ij}, & \text{for } t = T+1 \end{cases}
$$

The foregoing values of w_{it} are feasible for (26) and provide at least as good an objective function value as $c(\mathbf{x})$ for (6) because any amount of x_{it} that is diverted to task $T+1$ accrues a benefit of $c_{iT+1} = \max\{c_{it} : 1 \leq t \leq T\} \geq c_{it}$. Thus, the optimal objective function value of (26) is at least as good as $c(\mathbf{x})$. $\square$

Theorem 10 *If* $\mathbf{w}$ *is an optimal solution for (26) with objective function value* $c(\mathbf{w})$ *then there is an optimal solution to (6) whose objective function value is at least* $c(\mathbf{w})$.

Proof. For each $i = 1, \ldots, N$, let $j(i)$ be the first integer for which $c_{ij(i)} = \max\{c_{it} : 1 \leq i \leq N\}$. Then the following feasible solution $\mathbf{x}$ to (6) has the same objective function value as $c(\mathbf{w})$:

$$x_{it} = \begin{cases} w_{it} + w_{iT+1}, & \text{if } t = j(i) \\ \\ w_{it} & \text{if } t \neq j(i) \end{cases}$$

Thus, the optimal objective function value of (6) is at least as good as $c(\mathbf{w})$. $\square$

Appendix B

Theorem 5 is proved in this appendix. To that end, the following theorem first establishes conditions under which the optimal objective function values of NLP(A^k) converge to the optimal objective function value of NLP(A^∞). Throughout, it is assumed that f is continuous on S^N, NLP(A^0) is feasible and hence, by Theorem 4, each NLP(A^k) is optimal, as is NLP(A^∞).

Theorem 11 *If for each* $k = 0, 1, \ldots$, $\mathbf{x}^k$ *is optimal for NLP(*A^k*) and generates* A^{k+1} *and* $\mathbf{x}^*$ *is optimal for NLP(*A^∞*), then* $(f(\mathbf{x}^k)) \to f(\mathbf{x}^*)$ *as* $k \to \infty$.

Proof. By the monotonicity of the matrices A^k, each $\mathbf{x}^k$ is feasible for NLP(A^{k+1}) and so the sequence $(f(\mathbf{x}^k))$ is monotonically non-decreasing and bounded above by $f(\mathbf{x}^*)$ because each $\mathbf{x}^k$ is also feasible for NLP(A^∞) (see (16) in the proof of Theorem 4 in Section 3). Thus, the limit of $(f(\mathbf{x}^k))$ exists and it remains to show that this limit is bounded below by $f(\mathbf{x}^*)$. This is done in two parts.

Case 1. There is no integer t with $1 \leq t \leq T$ for which there is an integer $k(t)$ and a feasible solution $\mathbf{x}(t)$ for NLP($A^{k(t)}$) with $\sum_{i=1}^{N} A_{it}^{k(t)} x(t)_{it} > l_t$.

In this case, it is shown that $\mathbf{x}^*$ is feasible for every NLP(A^k) and so

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^k), \quad k = 1, 2, \ldots.$$

The desired lower-bound result then follows on taking the limit as $k \to \infty$.

To see that $\mathbf{x}^*$ is feasible for every NLP(A^k), it is shown that $A^0 = A^1 = \cdots$, for suppose not. Then there are integers k, h, and t such that $A_{ht}^k < A_{ht}^{k+1}$ (so $x_{ht}^k \neq 0$). It then follows that

$$\begin{aligned}
\sum_{i=1}^{N} A_{it}^{k+1}\mathbf{x}_{it}^{k} &= \sum_{i\neq h} A_{it}^{k+1}x_{it}^{k} + A_{ht}^{k+1}x_{ht}^{k} \\
&\geq \sum_{i\neq h} A_{it}^{k}x_{it}^{k} + A_{ht}^{k+1}x_{ht}^{k} \qquad \text{(monotonocity of } A^k\text{)} \\
&> \sum_{i\neq h} A_{it}^{k}x_{it}^{k} + A_{ht}^{k}x_{ht}^{k} \qquad \text{(choice of } k \text{ and } x_{ht}^k \neq 0\text{)} \\
&\geq l_t \qquad (\mathbf{x}^k \text{ is feasible for NLP}(A^k)).
\end{aligned}$$

Taking $k(t) = k+1$ and $\mathbf{x}(t) = \mathbf{x}^k$ contradicts the hypotheses of this case.

Case 2. There is an integer t with $1 \leq t \leq T$ for which there is an integer $k(t)$ and a feasible solution $\mathbf{x}(t)$ for NLP$(A^{k(t)})$ with $\sum_{i=1}^{N} A_{it}^{k(t)}x(t)_{it} > l_t$.

In this case, the lower-bound result is obtained by producing a sequence $(\mathbf{w}^j)$ converging to $\mathbf{x}^*$ and, for each j, an integer $k_j \geq j$ such that $\mathbf{w}^j$ is feasible for NLP(A^{k_j}), the latter yielding

$$f(\mathbf{w}^j) \leq f(\mathbf{x}^{k_j}), \quad j = 1, 2, \ldots$$

On taking limits on both sides of the foregoing inequality, it then follows by the convergence of $(\mathbf{w}^j)$ to $\mathbf{x}^*$, the continuity of f, and the convergence of $(f(\mathbf{x}^k))$ that

$$f(\mathbf{x}^*) = \lim_{j\to\infty} f(\mathbf{w}^j) \leq \lim_{j\to\infty} f(\mathbf{x}^{k_j}) = \lim_{k\to\infty} f(\mathbf{x}^k).$$

The sequence $(\mathbf{w}^j)$ is constructed by showing that for every $\epsilon > 0$, there is an integer $k(\epsilon)$ such that the ϵ-neighborhood of $\mathbf{x}^*$ and the feasible region of NLP$(A^{k(\epsilon)})$ intersect. The sequence $(\mathbf{w}^j)$ is then obtained by setting $\epsilon = 1/j$, $k_j = \max\{k(1/j), j\}$ and letting $\mathbf{w}^j$ be any point in both the ϵ-neighborhood of $\mathbf{x}^*$ and the feasible region of NLP(A^{k_j}). Thus, let $\epsilon > 0$.

Now $k(\epsilon)$ is chosen so that a point $\mathbf{z}$ on the line segment between $\mathbf{x}^*$ and a new point $\mathbf{x}$ (constructed below) is both feasible for NLP$(A^{k(\epsilon)})$ and in the ϵ-neighborhood of $\mathbf{x}^*$. To construct this new point $\mathbf{x}$, from the hypotheses of this case, the task constraints can be renumbered so that (a) for all $t = 1, \ldots, t^*$, there is an integer $k(t)$ and a feasible solution $\mathbf{x}(t)$ for NLP$(A^{k(t)})$ with $\sum_{i=1}^{N} A_{it}^{k(t)}x(t)_{it} > l_t$ and (b) for all $t = t^*+1, \ldots, T$, there does not exist an integer $k(t)$ and a feasible solution $\mathbf{x}(t)$ for NLP$(A^{k(t)})$ with $\sum_{i=1}^{N} A_{it}^{k(t)}x(t)_{it} > l_t$. Now let $m = \max\{k(t) : t = 1, \ldots, t^*\}$ and

$$\mathbf{x} = \sum_{t=1}^{t^*} x(t)/t^*.$$

Note that $\mathbf{x}$ is a convex combination of feasible points for NLP(A^m) and, as such, is itself feasible for NLP(A^m).

Consider now a point $\mathbf{z} = \lambda\mathbf{x} + (1-\lambda)\mathbf{x}^*$, where $\lambda \in (0, 1)$ is chosen sufficiently close to 0 so that $\mathbf{z}$ is in the ϵ-neighborhood of $\mathbf{x}^*$. It is now shown that there is an integer $K(\lambda)$, which is in fact the desired $k(\epsilon)$, such that $\mathbf{z}$ is feasible for NLP$(A^{K(\lambda)})$. Specifically, using the fact that the matrices (A^k) converge to A, let $K(\lambda)$ be such that

$$A_{it}^{\infty} - A_{it}^{K(\lambda)} < \frac{1}{N}\left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t\right), \quad t = 1, \ldots, t^*;\ i = 1, \ldots, N. \tag{27}$$

Such a value for $K(\lambda)$ can be chosen because the right side of (27) satisfies, for each $t = 1, \ldots, t^*$,

$$\begin{aligned}
\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} &= \lambda \sum_{j=1}^{N} A_{jt}^{\infty} x_{jt} + (1-\lambda) \sum_{j=1}^{N} A_{jt}^{\infty} x_{jt}^* && \text{(definition of } \mathbf{z}\text{)} \\
&\geq \lambda \sum_{j=1}^{N} A_{jt}^{m} x_{jt} + (1-\lambda) l_t && \text{(monotonicity and feasibility of } \mathbf{x}^*\text{)} \\
&> \lambda l_t + (1-\lambda) l_t = l_t && \text{[definition of } m,\ t^*, \text{ and } \lambda \in (0,1)].
\end{aligned}$$

It remains to show that $\mathbf{z}$ is feasible for NLP$(A^{K(\lambda)})$. To that end, consider first a task constraint t with $1 \leq t \leq t^*$, then

$$\begin{aligned}
\sum_{i=1}^{N} A_{it}^{K(\lambda)} z_{it} &> \sum_{i=1}^{N} A_{it}^{\infty} z_{it} - \tfrac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t\right) z_{it} && \text{[from (27)]} \\
&\geq \sum_{i=1}^{N} A_{it}^{\infty} z_{it} - \left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t\right) && (\mathbf{z} \in S^N) \\
&= l_t.
\end{aligned}$$

Finally, for the task constraints $t > t^*$, as in Case 1, it is shown that for all $i = 1, \ldots, N$, $A_{it}^0 = A_{it}^1 = \cdots$, for suppose not. Then there are integers k and h such that $A_{ht}^k < A_{ht}^{k+1}$ (so $x_{ht}^k \neq 0$). It then follows that

$$\begin{aligned}
\sum_{i=1}^{N} A_{it}^{k+1} \mathbf{x}_{it}^k &= \sum_{i \neq h} A_{it}^{k+1} x_{it}^k + A_{ht}^{k+1} x_{ht}^k \\
&\geq \sum_{i \neq h} A_{it}^{k} x_{it}^k + A_{ht}^{k+1} x_{ht}^k && \text{(monotonocity of } A^k\text{)} \\
&> \sum_{i \neq h} A_{it}^{k} x_{it}^k + A_{ht}^{k} x_{ht}^k && \text{(choice of } k \text{ and } h \text{ and } x_{ht}^k \neq 0\text{)} \\
&\geq l_t.
\end{aligned}$$

But this contradicts the definition of t^*. In other words, the task constraints $t > t^*$ are the same for every NLP(A^k) and so $\mathbf{z}$, being a convex combination of $\mathbf{x}$ and $\mathbf{x}^*$, both of which satisfy all of these task constraints, must itself satisfy these constraints.

The proof of Theorem 11 is now complete. □

With the aid of Theorem 11, it is possible to prove Theorem 5. Specifically, consider any subsequence of $(\mathbf{x}^k)$ that converges to a point $\mathbf{x}$. It is clear that $\mathbf{x}$ is feasible for NLP(A^{∞}). It remains to show that $\mathbf{x}$ is optimal for NLP(A^{∞}). To that end, let $\mathbf{x}^*$ be optimal for NLP(A^{∞}). From Theorem 11, it follows that $(f(\mathbf{x}^k)) \to f(\mathbf{x}^*)$ as $k \to \infty$. The desired conclusion that $\mathbf{x}$ is optimal for NLP(A^{∞}) follows using the continuity of f and the subsequence $(\mathbf{x}^{k_j})$ because

$$f(\mathbf{x}) = \lim_{j \to \infty} f(\mathbf{x}^{k_j}) = \lim_{k \to \infty} f(\mathbf{x}^k) = f(\mathbf{x}^*).$$

The proof of Theorem 5 is now complete. □

References

Ahuja, K. R., Magnanti, L. T. and Orlin, B. J. (1993), *Network Flows, Theory, Algorithms, and Applications*, Prentice Hall, New Jersey.

Bartle, R. G., and Sherbert, D. R. (1992), *Introduction to Real Analysis*, 2nd ed., John Wiley and Sons, New York.

Bazaraa, M. S., Jarvis, J. J., and Sherali, H. D., (1990), *Linear Programming and Network Flows*, 2nd ed., John Wiley and Sons, New York.

Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. (1993), *Nonlinear Programming: Theory and Algorithms*, 2nd ed., John Wiley and Sons, New York.

Muller, B., and Reinhardt, J. (1991), *Neural Networks: An Introduction*, Springer-Verlag, Berlin.