Mathematical Models for Explaining the Emergence of Specialization in Performing Tasks

Daniel Solow and Joesph Szmerekovsky[†]

Dept. of Operations
Case Western Reserve University
Cleveland, OH 44106
e-mail: `dxs8@po.cwru.edu`
Fax: (216) 368–6250

[†]Dept. of Mgmt., Marketing and Finance
North Dakota State University
Fargo, ND 58105

February 1, 2004

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. The contribution in this work is a collection of linear and nonlinear mathematical models that provide insights as to when and why functional specialization emerges in general, rather than specific, settings. The results from these models, which are based on an evolutionary approach, apply to communities in which individuals allocate their time in the best interest of the community as a whole.

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 in each of the following settings:

- Living organisms, where the individuals are cells [see, for example, Hawthorne and Via (2001), Rushworth (2000), and Turken and Swick (1999)].

- Human societies, where the individuals are people [see, for example, Abdel-Rahman (1996) and Proudman and Redding (2000)].

- Business organizations, where the individuals are the employees [see, for example, Ng and Yang (1997)].

One example in which specialization is assumed arises in mathematical models for solving the task-assignment problem in business. For example, given a group of tasks and a

group of people, each of whom can perform one or more of the tasks, the maximum matching algorithm developed by Edmonds (1965) is designed to determine the maximum number of people that can be assigned to tasks in such a way that no person is assigned to more than one task. That is, it is assumed that each person specializes in performing at most one task. A generalization of this problem is the assignment problem [see surveys by Derigs (1985), Martello and Toth (1987), and Bertsekas (1991)], in which the objective is to assign qualified people to at most one task so as to incur the least total cost.

On considering these and other examples, one is led naturally to the question of how functional specialization arose and why it persists. From an evolutionary point of view, it could be argued that, through adaptation, functional specialization has been selected and endures because it has desirable properties that lead to a "fitter" community. That is, the reason functional specialization emerges and persists is because the community, as a whole, derives a survival 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—aimed at capturing the concept of *fitness* of the community—are presented and analyzed in Section 2. It is first shown how a linear model can fail to explain the emergence of functional specialization, thus leading to the need for nonlinear models. Conditions are then provided under which functional specialization arises in certain nonlinear models. In Section 3, functional specialization is shown to emerge in a dynamic model that captures the fact that the more time an individual spends on a task, the more efficient that individual becomes at performing the task. 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. These models are designed to reflect a general, rather than a specific, community so as to have the broadest possible applicability. 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:

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

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:

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

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

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 $f(\mathbf{x})$, which is a measure of the fitness of the community. From an evolutionary point of view, the objective is to determine values for the variables $\mathbf{x} = (x_{it})$ that achieve the fittest community, that is, 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:

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

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, while not valid in all communities, is known to occur in some communities, such as the following:

- A community in which the benefits and survival 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 human body and its organs.

- 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 or a social structure in which societal goals outweigh individual goals. In a supply chain, Lamb et al. (1996) suggest that "channel captains" use their dominant positions to alter behavior of individual members of the supply chain (see Gaski (1984) for an appropriate literature review and Munson et al. (1999) for a discussion of power and conflict in supply chains).

- A community in which incentives are such that the individuals act in the best interest of the community. Such examples of biological systems are given in Buss (1987). As another example, consider a company in which employees are given incentives so that individual benefit and corporate benefit coincide. More recently, much research has been done on the use of incentives to coordinate the supply chain, that is, to get the manufacturers, wholesalers, and retailers to act in such a way as to maximize profits for the entire supply chain, rather than for each individual agent [see, for instance, Munson and Rosenblatt (2001), Mantrala and Raman (1999) and Emmons and Gilbert (1998)].

For such communities, the general model in (3) is applicable. In contrast, this model is not appropriate for communities in which the individuals act in their own best interest

rather than in the interest of the community as a whole. Much literature is devoted to such communities. For example, the concept of a Nash equilibrium in game theory [see Yuan et al. (1998) and also Borgers (2001) for a survey of books on evolutionary games] is predicated on the assumption that each player chooses a strategy that maximizes their individual profit. Also, in many supply chain settings, the manufacturers, wholesalers, and retailers are assumed to act in their own self interest [see, for example, Lee (2001) and Fulkerson and Staffend (1997)]. Another example occurs in the economic theory of international trade where comparative advantages cause different countries to specialize in the export of certain goods [see, for example, Yi (2003) and Evenett and Keller (2002)]. Even when individuals act in their own interest, it could be argued that unless such actions are in the long-term interest of the community, the community will not survive. Nevertheless, to study how functional specialization emerges in communities where the individuals act in their own interest might best be done with models different from the ones used here and is outside the scope of this research.

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

2.1   A Linear Model and Its Drawbacks

To create a linear objective function, let

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

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. Assuming that each individual has K time units available , the survival constraints are:

$$K \left(\sum_{i=1}^{N} x_{it} \right) \geq L_t \quad \text{or} \quad \sum_{i=1}^{N} x_{it} \geq l_t = L_t/K \quad \text{for } t = 1, \ldots, T. \tag{5}$$

The constraints in (5) are the simplest possible linear constraints and are used here as a starting point. More general linear constraints of the form $\sum_{i=1}^{N} a_{it} x_{it} \geq l_t$ are considered in the dynamic model in Section 3.

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):

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

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

$$\sum_{t=1}^{T} x_{it} = 1, \quad i = 1, \ldots, N \tag{6}$$

$$x_{it} \geq 0, \quad \text{all } i, t$$

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 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:

maximize $\quad x_{11} + x_{21} + x_{31} + x_{12} + x_{22} + x_{32}$

s.t. **Task Requirements**

$$\begin{array}{llll} x_{11} + x_{21} + x_{31} & \geq & 1 & \text{(Task 1)} \\ x_{12} + x_{22} + x_{32} & \geq & 1 & \text{(Task 2)} \end{array}$$

Individual Fractions Sum to 1 $\tag{7}$

$$\begin{array}{llll} x_{11} + x_{12} & = & 1 & \text{(Individual 1)} \\ x_{21} + x_{22} & = & 1 & \text{(Individual 2)} \\ x_{31} + x_{32} & = & 1 & \text{(Individual 3)} \end{array}$$

$$\text{all } x_{it} \quad \geq \quad 0$$

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:

$$x_{11} = 1, \quad x_{12} = 0$$
$$x_{21} = 1, \quad x_{22} = 0$$
$$x_{31} = 0, \quad x_{32} = 1$$

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:

$$\max \quad f(\mathbf{x})$$

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

$$\sum_{t=1}^{T} x_{it} = 1, \quad i = 1, \ldots, N \tag{8}$$

$$x_{it} \geq 0, \quad \text{all } i, t$$

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.

Figure 1
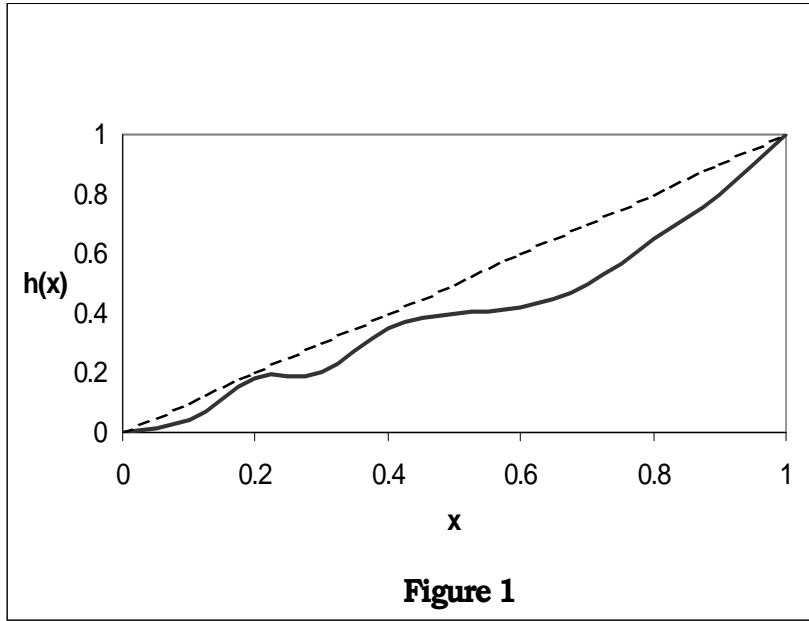
In so doing, the following notations are used:

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

$$\mathbf{e}^t = \text{the vertex of } S \text{ with all 0's except for a 1 in position } t.$$

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

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

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

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

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

7

$y, z \in R^1$ 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 $h(x) = h((1-x)0 + x(1)) < (1-x)h(0) + xh(1)$.

A natural generalization of the sublinear property to a function $h : S \to R^1$ is that the graph of the function on the interior of S lies strictly below the plane through the points $(\mathbf{e}^t, h(\mathbf{e}^t))$, $t = 1, \ldots, T$. Noting that a point $\mathbf{z} \in S$ can be expressed as a convex combination of the extreme points $\mathbf{e}^t$ of S leads to the following definition.

Definition 2 *A continuous function* $h : S \to R^1$ *is* **sublinear** *if and only if for all* $\mathbf{z} = (z_1, \ldots, z_T) \in S$ *with at least one* z_t *with* $0 < z_t < 1$, *it follows that*

$$h(\mathbf{z}) = h\left(\sum_{t=1}^{T} z_t \mathbf{e}^t\right) < \sum_{t=1}^{T} z_t h(\mathbf{e}^t).$$

Any strictly convex continuous function on R^T when restricted to S is sublinear, as is the following function $h : S \to R^1$.

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 at least one z_t with $0 < z_t < 1$, then, noting that each $h(\mathbf{e}^t) = (T-1)a + b$, it follows that

$$
\begin{aligned}
h(\mathbf{z}) &= \sum_{t=1}^{T} h_t(z_t) \\[2mm]
&= \sum_{\{t:0<z_t<1\}} h_t(z_t) + \sum_{\{t:z_t=0\}} h_t(z_t) \\[2mm]
&< \sum_{\{t:0<z_t<1\}} [(1-z_t)h_t(0) + z_t h_t(1)] + \sum_{\{t:z_t=0\}} (1-z_t)h_t(0) \\[2mm]
&= \sum_{t=1}^{T} (1-z_t)h_t(0) + \sum_{\{t:0<z_t<1\}} z_t h_t(1) \\[2mm]
&\leq (T-1)a + b \\[2mm]
&= \sum_{t=1}^{T} z_t h(\mathbf{e}^t).
\end{aligned}
$$

If you think of the value of $h(\mathbf{z})$ as the benefit an individual derives from making time allocations $\mathbf{z} \in S$, as is done from here on, sublinearity provides the basis for establishing functional specialization because a sublinear function ensures that the individual achieves maximum benefit by, and only by, specializing in exactly one task, as shown in the following lemma whose proof (along with proofs of all other lemmas and theorems) is given in Appendix B.

8

Lemma 1 *If $h : S \to R^1$ is sublinear, then h achieves its maximum value of c at, and only at, a vertex of S. That is, for all $\mathbf{z} \in S$ with $h(\mathbf{z}) = c$, there is a component t such that $z_t = 1$, or equivalently, if $\mathbf{z} \in S$ has fractional values, then $h(\mathbf{z}) < c$.*

However, recall that the domain of the objective function f of the community is S^N rather than S, thus giving rise to the need for the following final type of sublinear function.

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.*

The component functions $h_1, \ldots, h_N$ in Definition 4 should be thought of as the benefit functions of the N individuals in the community.

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 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):

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

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 definitions.

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 $y_j < z_j$.*

Definition 6 *The function $f : S^N \to R^1$ is **specializable** if and only if there is a sublinear function $h : S^N \to R^N$ such that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$.*

A specializable objective function f for a community means that if all individuals with time allocations $\mathbf{y}$ receive at least as much personal benefit as they do with time allocations $\mathbf{x}$, and at least one individual is strictly better off with $\mathbf{y}$, then the community as a whole is better off with $\mathbf{y}$ rather than $\mathbf{x}$.

The next theorem establishes that functional specialization necessarily emerges when the objective function of the community is specializable, provided that the time allocations that maximize each individual's benefit satisfy the survivability constraints.

Theorem 1 *If the objective function f of the NLP in (9) is specializable and the time allocations $\bar{\mathbf{x}}$ in which each individual achieves the maximum personal benefit satisfies the survivability constraints of the NLP, then any optimal solution results in specialization.*

Theorem 1 provides conditions under which functional specialization emerges and, in fact, requires the existence of a particular feasible solution, namely, $\bar{\mathbf{x}}$. If such a feasible solution does not exist, it may be necessary for individuals to work at more than one task to satisfy the survivability constraints. Even if the conditions of Theorem 1 hold, the theorem does not indicate *which* tasks the individuals specialize in. The next theorem shows that the time allocations in the feasible solution in the hypotheses of Theorem 1 indicate the tasks in which the individuals specialize, provided that the objective function f of the community is continuous.

Theorem 2 *If f is specializable and continuous on the extreme points of S^N and the time allocations $\mathbf{x}^*$ in which each individual i achieves the maximum personal benefit satisfies the survivability constraints of the NLP, then an optimal solution to the NLP is for each individual i to specialize in that task that maximizes the personal benefit of individual i.*

Theorem 2 means that the solution of the NLP in (9) is for each individual to specialize in a task that provides that individual with the maximum personal benefit, provided that doing so collectively satisfies the survivability constraints. These results are illustrated in the following example.

Example 3: For each i and t, let $h_{it} : [0,1] \to R^1$ be sublinear functions that represent the benefit individual i derives from allocating the fraction x_{it} to task t (for example, $h_{it}(x_{it}) = x_{it}^2$). The continuous community-benefit 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 individual benefit functions $h_i : S \to R^1$ defined by

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

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 $h(\mathbf{x}) = (h_1(\mathbf{x}_1), \dots, h_N(\mathbf{x}_N))$ for each $\mathbf{x} \in S^N$ is also sublinear.

Also, f is specializable, that is, 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 individual $i = 1, \ldots, N$,

$$\sum_{t=1}^{T} h_{it}(x_{it}) = h_i(\mathbf{x}_i) \leq h_i(\mathbf{y}_i) = \sum_{t=1}^{T} h_{it}(y_{it}), \qquad (10)$$

with strict inequality holding for at least one individual i. It then follows by summing (10) over i that

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

Assuming that the time allocations $\mathbf{x}^*$ in which each individual i achieves maximum personal benefit satisfies the survivability constraints, by Theorem 1, any optimal solution exhibits functional specialization. Then, by Theorem 2, the specialization indicated by $\mathbf{x}^*$ is optimal for the NLP.

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, $0 \leq A_{it} \leq 1$. (Any such matrix is called an *efficiency matrix*).

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, in period k, the survival constraints—which are a generalization of those in (5)—are:

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

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):

$$\max \quad f(\mathbf{x})$$

$$\text{s.t.} \quad \sum_{i=1}^{N} A_{it}^k x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T \quad [\text{NLP}(A^k)] \tag{11}$$

$$\mathbf{x} \in S^N$$

An optimal solution $\mathbf{x}^k$ to NLP(A^k) need not result in specialization. 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. As a starting point that allows appropriate analysis, we begin by assuming that 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 increases 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,

$$A_{it}^{k+1} = A_{it}^k + x_{it}^k F(1 - A_{it}^k). \tag{12}$$

Generalizations and extensions to this learning formula are discussed in Section 4. On the basis of the update formula in (12), the following dynamic model is proposed:

A Dynamic Nonlinear Programming Model

Step 0: Let $A^0 \in \mathcal{A}$ be a given efficiency 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 (12) 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 efficiency matrices A^k generated by this process converge to an efficiency 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, $\mathbf{x}^k$.*

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} according to (12). The next theorem establishes that the efficiency matrices A^k converge, element by element, to an efficiency 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, $A_{it}^\infty = \lim_{k \to \infty} A_{it}^k$ and NLP(A^∞) is optimal.*

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 establishes that $\mathbf{x}^*$ is an optimal solution to NLP(A^∞).

Theorem 5 *If f is continuous on S^N and the sequence of efficiency matrices (A^k) converges element by element to an efficiency 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:

$$\max \quad f(\mathbf{x})$$

$$\text{s.t.} \quad \sum_{i=1}^{N} u_{it} x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T \quad (NLTP) \tag{13}$$

$$\mathbf{x} \in S^N$$

where u_{it} is defined as follows:

$$u_{it} = \begin{cases} 1, & \text{if } x_{it}^* > 0 \\ 0, & \text{if } x_{it}^* = 0 \end{cases} \tag{14}$$

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 used.

Lemma 2 *If, in the optimal solution $\mathbf{x}^*$, an individual i spends any positive fraction of time on a task t (that is, if $\mathbf{x}_{it}^* > 0$), then, in the limit, that individual is 100% efficient at task t (that is, $A_{it}^\infty = 1$).*

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.*

Although Theorem 6 establishes that $\mathbf{x}^*$ is optimal for NLTP, $\mathbf{x}^*$ need not result in specialization. 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:

$$\max \quad \sum_{t=1}^{T} c_t \sum_{i=1}^{N} x_{it}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} A_{it}^k x_{it} \geq l_t, \quad \text{for } t = 1, \ldots, T \quad [\text{LP}(A^k)] \tag{15}$$

$$\mathbf{x} \in S^N$$

13

Although it can be shown that the limiting problem LP(A^∞) has an optimal solution that results in specialization, a convergent subsequence of optimal solutions to LP(A^k) need not converge to such an 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 $\mathbf{x}^*$, results in specialization.*

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, a special form of learning is used, namely, the update formula in (12). A natural question to ask is whether functional specialization emerges when other forms of learning apply (see Chase et al. (1998) for a discussion of learning curves and a bibliography). To that end, define $L : R^{n^2} \times \mathcal{A} \to \mathcal{A}$ as a *learning function* that associates to each time-allocation vector $\mathbf{x} \in R^{n^2}$ and efficiency matrix $A \in \mathcal{A}$, the updated efficiency matrix $L(\mathbf{x}, A) \in \mathcal{A}$. Thus, if $A^k \in \mathcal{A}$ is the efficiency matrix in period k and $\mathbf{x}^k$ is an optimal solution to NLP(A^k), then the efficiency matrix in period $k + 1$ is $A^{k+1} = L(\mathbf{x}^k, A^k)$. It is not hard to show that all of the results in Section 3—including the emergence of functional specialization—hold when the learning function satisfies the following properties [as does the learning function defined in (12)]:

(a) (Monotonicity) For all time allocations $\mathbf{x} \in R^{n^2}$ and efficiency matrices $A \in \mathcal{A}$, $L(\mathbf{x}, A) \geq A$.

(b) (Continuity) L is continuous in its arguments.

(c) (No Improvement) If the time allocations $\mathbf{x} \in R^{n^2}$ and the efficiency matrix $A \in \mathcal{A}$ satisfy $A = L(\mathbf{x}, A)$, then for all individuals i and for all tasks t, either individual i spends no time on task t (that is, $x_{it} = 0$) or individual i is 100% efficient at task t (that is, $A_{it} = 1$).

In the update formula in (12) as well as in the foregoing conditions on the learning function, it is assumed that individuals never get worse at tasks and thus the matrices A^k are monotonically non-decreasing. 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 a fraction F of the way to 0. In this case, for example, the update formula becomes:

$$A_{it}^{k+1} = \begin{cases} A_{it}^k + x_{it}F(1 - A_{it}^k), & \text{if } x_{it}^k > 0 \\ \\ (1 - F)A_{it}^k, & \text{if } x_{it}^k = 0 \end{cases} \qquad (16)$$

With (16), 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 efficiency 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 an efficiency matrix $A^\infty \in \mathcal{A}$ and a subsequence K such that for each individual i and task 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 jth 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 optimal solution $\mathbf{y}^*$ for NLP(B^∞) that results in specialization. In particular, no suitable version of Lemma 2 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:

$$A_{it} = \begin{cases} A_{it} + \left(\frac{x_{it}-b}{1-b}\right) F(1 - A_{it}), & \text{if } x_{it} \geq b \\ \\ \left[1 - \left(\frac{b-x_{it}}{b}\right) F\right] A_{it}, & \text{if } x_{it} < b \end{cases} \qquad (17)$$

With (17), 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:

$$\text{maximize} \quad x_{11} + x_{21} + x_{12} + x_{22}$$

$$
\begin{array}{llcl}
\text{s.t.} & x_{11} + x_{21} & \geq & 1 \\
& x_{12} + x_{22} & \geq & 1 \\
& & & \\
& x_{11} + x_{12} & = & 1 \\
& x_{21} + x_{22} & = & 1 \\
& & & \\
& \text{all } x_{it} & \geq & 0
\end{array}
\qquad (18)
$$

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 (17), the following problem, $\text{NLP}(A^1)$, is easily seen to be infeasible:

$$\text{maximize} \quad x_{11} + x_{21} + x_{12} + x_{22}$$

$$
\begin{array}{llcl}
\text{s.t.} & \frac{5}{6}x_{11} + \frac{5}{6}x_{21} & \geq & 1 \\
& \frac{5}{6}x_{12} + \frac{5}{6}x_{22} & \geq & 1 \\
& & & \\
& x_{11} + x_{12} & = & 1 \\
& x_{21} + x_{22} & = & 1 \\
& & & \\
& \text{all } x_{it} & \geq & 0
\end{array}
\qquad (19)
$$

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:

$$\max \quad \sum_{i=1}^{N} \sum_{t=1}^{T+1} c_{it} w_{it}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} w_{it} = l_t, \quad t = 1, \ldots, T+1 \tag{20}$$

$$\sum_{t=1}^{T+1} w_{it} = 1, \quad i = 1, \ldots, N$$

$$w_{it} \geq 0, \quad \text{all } i, t$$

The next two theorems establish that any optimal solution to (6) provides an optimal solution to (20) 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 (20) 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 (20) 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 (20) 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

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

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

18

The foregoing values of w_{it} are feasible for (20) 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 $c_{iT+1} = \max\{c_{it} : 1 \le t \le T\} \ge c_{it}$. Thus, the optimal objective function value of (20) is at least as good as $c(\mathbf{x})$. □

Theorem 10 *If* $\mathbf{w}$ *is an optimal solution for (20) 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 $c_{ij(i)} = \max\{c_{it} : 1 \le i \le N\}$. Then the following feasible solution $\mathbf{x}$ to (6) has the same objective function value as $c(\mathbf{w})$:

$$
x_{it} = \begin{cases} w_{it} + w_{iT+1}, & \text{if } t = j(i) \\[2mm] w_{it} & \text{if } t \neq j(i) \end{cases}
$$

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

Appendix B

All lemmas and theorems in this work are proved here.

Proof of Lemma 1 in Section 2.2.3. Because h is a continuous function on the nonempty compact set S, h achieves its maximum value at a point in S. So, let $\mathbf{z} \in S$ be such that $h(\mathbf{z}) = c = \max\{h(\mathbf{x}) : \mathbf{x} \in S\}$. To see that $\mathbf{z}$ must be an extreme point of S, suppose not. Then there is a component t such that $0 < z_t < 1$. But then by the fact that h is sublinear, it follows that

$$
c = h(\mathbf{z}) = h\left(\sum_{t=1}^{T} z_t \mathbf{e}^t\right) < \sum_{t=1}^{T} z_t h(\mathbf{e}^t) \le \sum_{t=1}^{T} z_t c = c.
$$

This contradiction proves the lemma. □

Proof of Theorem 1 in Section 2.2.3. 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 $1 \le i \le N$ and $1 \le t \le T$ such that $0 < x_{it}^* < 1$. A contradiction is reached by showing that $f(\mathbf{x}^*) < f(\bar{\mathbf{x}})$. To that end, define

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

Because f is specializable, there is a sublinear function $h : S^N \to R^N$ such that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$. 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 from Lemma 1 applied to each h_i that

$$
h_i(\mathbf{x}_i^*) < c_i = h_i(\bar{\mathbf{x}}_i), \quad \text{for each } i \in \mathcal{F}
$$

(21)

$$
h_i(\mathbf{x}_i^*) \le c_i = h_i(\bar{\mathbf{x}}_i), \quad \text{for each } i \notin \mathcal{F}.
$$

In other words, (21) 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. □

Proof of Theorem 2 in Section 2.2.3. 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

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

Because f is specializable, there is a sublinear function $h : S^N \to R^N$ such that for all $\mathbf{x}, \mathbf{y} \in S^N$ with $h(\mathbf{x}) \preceq h(\mathbf{y})$, $f(\mathbf{x}) < f(\mathbf{y})$. 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 from Lemma 1 applied to each h_i that

$$h_i(\mathbf{x}_i) < c_i = h_i(\mathbf{x}_i^*), \quad \text{for each } i \in \mathcal{F}$$

$$h_i(\mathbf{x}_i) \leq c_i = h_i(\mathbf{x}_i^*), \quad \text{for each } i \notin \mathcal{F}. \tag{22}$$

In other words, (22) 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 for which the sequence 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 $k = 1, 2, \ldots$,

$$f(\mathbf{x}^k) < f(\mathbf{x}^*). \tag{23}$$

The fact that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ now follows by taking the limit as $k \to \infty$ on both sides of (23) and using the continuity of f at the extreme point $\mathbf{x}$.

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. $\square$

Proof of Theorem 3 in Section 3. The statement is true for $k = 0$ because, by hypothesis, $\text{NLP}(A^0)$ is feasible. Furthermore, the constraints $\mathbf{x} \in S^N$ together with the continuity of f ensure that $\text{NLP}(A^0)$ has an optimal solution.

Assume now that the statement is true for k and let $\mathbf{x}^k$ be an optimal solution for $\text{NLP}(A^k)$. Then $\mathbf{x}^k$ is feasible for $\text{NLP}(A^{k+1})$. This is because $\mathbf{x}^k \in S^N$ and, from the update formula in (12), 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, $\text{NLP}(A^{k+1})$ is feasible. The constraints that $\mathbf{x} \in S^N$ together with the continuity of f ensure that $\text{NLP}(A^{k+1})$ has an optimal solution. The result now follows by induction. $\square$

Proof of Theorem 4 in Section 3. For each i and t, the update formula in (12) ensures that the sequence of real numbers (A_{it}^k) 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 $0 \leq A_{it}^\infty \leq 1$ such that $A_{it}^\infty = \lim_{k \to \infty} A_{it}^k$.

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,

$$\sum_{i=1}^{N} A_{it}^\infty x_{it}^k \geq \sum_{i=1}^{N} A_{it}^k x_{it}^k \geq l_t, \quad \text{for } t = 1, \ldots, T. \tag{24}$$

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. $\qquad\square$

To prove Theorem 5 in Section 3, the following lemma 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^∞).

Lemma 3 *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 of Lemma 3. 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 (24) 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 $\sum\limits_{i=1}^{N} A_{it}^{k(t)} x(t)_{it} > l_t$.

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

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

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 $\sum\limits_{i=1}^{N} A_{it}^{k(t)} x(t)_{it} > l_t$.

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 $\mathrm{NLP}(A^{k(\epsilon)})$ 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 $\mathrm{NLP}(A^{k_j})$. 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 $\mathrm{NLP}(A^{k(\epsilon)})$ 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 $\mathrm{NLP}(A^{k(t)})$ 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 $\mathrm{NLP}(A^{k(t)})$ 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 $\mathrm{NLP}(A^m)$ and, as such, is itself feasible for $\mathrm{NLP}(A^m)$.

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 $\mathrm{NLP}(A^{K(\lambda)})$. Specifically, using the fact that the matrices (A^k) converge to A, let $K(\lambda)$ be such that

$$A_{it}^{\infty} - A_{it}^{K(\lambda)} < \frac{1}{N}\left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t\right), \quad t = 1, \ldots, t^*; \; i = 1, \ldots, N. \tag{25}$$

Such a value for $K(\lambda)$ can be chosen because the right side of (25) satisfies, for each $t = 1, \ldots, t^*$,

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

It remains to show that $\mathbf{z}$ is feasible for $\mathrm{NLP}(A^{K(\lambda)})$. To that end, consider first a task

constraint t with $1 \leq t \leq t^*$, then

$$\sum_{i=1}^{N} A_{it}^{K(\lambda)} z_{it} > \sum_{i=1}^{N} A_{it}^{\infty} z_{it} - \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t \right) z_{it} \quad \text{[from (25)]}$$

$$\geq \sum_{i=1}^{N} A_{it}^{\infty} z_{it} - \left(\sum_{j=1}^{N} A_{jt}^{\infty} z_{jt} - l_t \right) \qquad (\mathbf{z} \in S^N)$$

$$= l_t.$$

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

$$\sum_{i=1}^{N} A_{it}^{k+1} \mathbf{x}_{it}^k = \sum_{i \neq h} A_{it}^{k+1} x_{it}^k + A_{ht}^{k+1} x_{ht}^k$$

$$\geq \sum_{i \neq h} A_{it}^k x_{it}^k + A_{ht}^{k+1} x_{ht}^k \qquad \text{(monotonocity of } A^k)$$

$$> \sum_{i \neq h} A_{it}^k x_{it}^k + A_{ht}^k x_{ht}^k \qquad \text{(choice of } k \text{ and } h \text{ and } x_{ht}^k \neq 0)$$

$$\geq l_t.$$

But this contradicts the definition of t^*. In other words, the task constraints $t > t^*$ are the same for every $\text{NLP}(A^k)$ 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 Lemma 3 is now complete. □

Proof of Theorem 5 in Section 3. Consider any subsequence of $(\mathbf{x}^k)$ that converges to a point $\mathbf{x}$. It is clear that $\mathbf{x}$ is feasible for $\text{NLP}(A^\infty)$. It remains to show that $\mathbf{x}$ is optimal for $\text{NLP}(A^\infty)$. To that end, let $\mathbf{x}^*$ be optimal for $\text{NLP}(A^\infty)$. From Theorem 3, it follows that $(f(\mathbf{x}^k)) \to f(\mathbf{x}^*)$ as $k \to \infty$. The desired conclusion that $\mathbf{x}$ is optimal for $\text{NLP}(A^\infty)$ 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. □

Proof of Lemma 2 in Section 3. From the update formula in (12), for each $k = 0, 1, \ldots$,

$$A_{it}^{k+1} = A_{it}^k + x_{it}^k F(1 - A_{it}^k) \tag{26}$$

Taking the limit on both sides of (26) 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 $x_{it}^* F > 0$. □

Proof of Theorem 6 in Section 3. 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 $\text{NLP}(A^\infty)$, so

$$\sum_{i=1}^{N} A_{it}^\infty x_{it}^* \geq l_t, \quad \text{for } t = 1, \ldots, T \tag{27}$$

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,

$$
\begin{aligned}
\sum_{i=1}^{N} u_{it} x_{it}^* &= \sum_{i \in X^+} u_{it} x_{it}^* + \sum_{i \in X^0} u_{it} x_{it}^* \\
&= \sum_{i \in X^+} (1) x_{it}^* + \sum_{i \in X^0} u_{it}(0) \quad \text{(definition of } u_{it}, X^+, \text{ and } X^0) \\
&= \sum_{i \in X^+} A_{it}^\infty x_{it}^* + \sum_{i \in X^0} A_{it}^\infty (0) \quad \text{(Lemma 2)} \\
&= \sum_{i=1}^{N} A_{it}^\infty x_{it}^* \quad \text{(definition of } X^0) \\
&\geq l_t \quad \text{[from (27)]}
\end{aligned}
$$

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,

$$
\begin{aligned}
l_t &\leq \sum_{i \in X^+} u_{it} x_{it} + \sum_{i \in X^0} u_{it} x_{it} \quad (\mathbf{x} \text{ is feasible for NLTP}) \\
&= \sum_{i \in X^+} (1) x_{it} + \sum_{i \in X^0} (0) x_{it} \quad \text{(definition of } u_{it}) \\
&= \sum_{i \in X^+} A_{it}^\infty x_{it} + \sum_{i \in X^0} (0) x_{it} \quad \text{(Lemma 2)} \\
&\leq \sum_{i \in X^+} A_{it}^\infty x_{it} + \sum_{i \in X^0} A_{it}^\infty x_{it} \\
&= \sum_{i=1}^{N} A_{it}^\infty x_{it}. \quad \square
\end{aligned}
$$

Proof of Theorem 7 in Section 3. 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. $\square$

References

Abdel-Rahman, H. M. (1996), 'When do Cities Specialize in Production', *Regional Science and Urban Economics*, **26**, 1-22.

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.

Bertsekas, D. P. (1991), *Linear Network Optimization: Algorithms and Codes*, The MIT Press, MA.

Borgers, T. (2001), 'Recent Books on Evolutionary Game Theory', *Journal of Economic Surveys*, **152**, 237-250.

Buss, L. W. (1987), *The Evolution of Individuality*, Princeton University Press, New Jersey.

Chase, R., Aquilano, N., and Jacobs, R. (1998), *Production and Operations Management: Manufacturing and Service*, 8th edition, Irwin/Mc-Graw Hill, New York.

Derigs, U. (1985), 'The shortest Augmenting Path Method for Solving Assignment Problems - motivation and computational experience', in: Monma, C. L. (1985), *Algorithms and Software for Optimization - Part I*, *Annals of Operations Research*, **4**, 57-102.

Edmonds, J. (1965), 'Pths, Trees and Flowers', *Canadian Journal of Mathematics*, **17**, 449-467.

Emmons, H. and Gilbert, S. M. (1998), 'Returns Policies in Pricing and Inventory Decisions for Catalogue Goods', *Managment Science*, **442**, 276-283.

Evenett, S. and Keller, W. (2002), 'On Theories Explaining the Success of the Gravity Equation', *The Journal of Political Economy*, **11042**, 281-316.

Fulkerson, B. and Staffend, G. (1997), 'Decentralized Control in the Customer Focused Enterprise', *Annals of Operations Research*, **75**, 325-333.

Gaski, J. F. (1984), 'The Theory of Power and Conflict in Channels of Distribution', *Journal of Marketing*, **483**, 9-29.

Hawthorne, D. J. and Via, S. (2001), 'Genetic Linkage of Ecological Specialization and Reproductive Isolation in Pea Aphids', *Nature (London)*, **4126850**, 904-907.

Lamb Jr., C. W., Hair Jr., J. F., and McDaniel, C. (1996), *Marketing*, 3rd ed., South-Western College Publishing, Ohio.

Lee, C. H. (2001), 'Coordinated Stocking, Clearance Sales, and Return Policies for a Supply Chain', *European Journal of Operational Research*, **131**, 491-513.

Mantrala, M. K. and Raman, K. (1999), 'Demand Uncertainty and Supplier's Return Policies for a Multi-Store Style-Good Retailer', *European Journal of Operational Research*, **115**, 270-284.

Martello, S. and Toth, P. (1987), 'Linear Assignment Problems', in: Martello, S., Laporte, G., Minoux, M., Ribeiro, C. (1987), *Surveys in Combinatorial Optimization*, *Annals of Discrete Mathematics*, **31**, 259-282.

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

Munson, C. L. and Rosenblatt, M. J. (2001), 'Coordinating a Three-Level Supply Chain with Quantity Discounts', *IIE Transactions*, **33**, 371-384.

Munson, C. L., Rosenblatt, M. J., and Rosenblatt, Z. (1999), 'The Use and Abuse of Power in Supply Chains', *Business Horizons*, **42**1, 55-65.

Ng, Y.-K. and Yang, X. (1997), 'Specialization, Information, and Growth: A Sequential Equilibrium Analysis', *Review of Development Economics*, **13**, 257-274.

Proudman, J., and Redding, S. (2000), 'Evolving Patterns of International Trade', *Review of International Economics*, **83**, 373-396.

Rushworth, M. F. S. (2000), 'Anatomical and Functional Subdivision within the Primate Lateral Prefrontal Cortex', *Psychobiology*, **28**2, 187-196.

Turken, A. U. and Swick, D. (1999), 'Response Selection in the Human Anterior Cingulate Cortex', *Nature Neuroscience*, **2**10, 920-924.

Yi, K.-M. (2003), 'Can Vertical Specialization Explain the Growth of World Trade', *The Journal of Political Economy*, **111**1, 52-102.

Yuan, G. X.-Z., Isac, G., Tan, K.-K. and Yu, J. (1998), 'The Study of Minimax Inequalities, Abtract Economics and Applications to Variational Inequalities and Nash Equilibria', *Acta Applicandae Mathematicae*, **54**, 135-166.