1-2005

The Asymptotic Optimal Partition and Extensions of the Nonsubstitution Theorem

Julio R. Hasfura-Buenaga
Trinity University, jhasfura@trinity.edu

Allen G. Holder
Trinity University, aholder@trinity.edu

Jeffrey Stuart

The Asymptotic Optimal Partition and Extensions of The Nonsubstitution Theorem

Julio-Roberto Hasfura-Buenaga[†], Allen Holder[††*], and Jeffrey Stuart[†††]

March 13, 2002

Abstract

The data describing an asymptotic linear program rely on a single parameter, usually referred to as time, and unlike parametric linear programming, asymptotic linear programming is concerned with the steady state behavior as time increases to infinity. The fundamental result of this work shows that the optimal partition for an asymptotic linear program attains a steady state for a large class of functions. Consequently, this allows us to define an asymptotic center solution. We show that this solution inherits the analytic properties of the functions used to describe the feasible region. Moreover, our results allow significant extensions of an economics result known as the Nonsubstitution Theorem.

Key Words: Asymptotic Linear Programming, Analytic Matrix Theory, Optimal Partition, Mathematical Economics, Nonsubstitution Theorem

---

[†] Department of Mathematics, Trinity University, San Antonio, TX, USA.

[††] Hearin Center for Enterprise Science, School of Business Administration, The University of Mississippi, University, MS, USA.

[†††] Department of Mathematics, Pacific Lutheran University, Tacoma, WA, USA.

[*] Research supported by ONR grant N00014-01-1-0917. Research conducted at Trinity University.

1 Introduction

The data describing many business and economic linear programs depend on a single parameter t, usually viewed as time. As such, understanding the dynamics of a solution as time progresses is important, and *steady-state* properties are often desired. A property *stabilizes* if it attains a steady-state for all sufficiently large t, (typical properties are feasibility and boundedness).

The foundational work on asymptotic linear programming was done by Jeroslow in [15] and [16], where the author assumes that the data functions are rational. In [15], the author shows that an optimal basis becomes stable for sufficiently large t, and that the number of basic optimal solutions stabilizes. This article also shows how to use the simplex method to produce a steady-state optimal basis. The continuity properties of a basic optimal solution near its poles are investigated in [16]. Bernard [3, 4] has studied the complexity of updating a basis in the special case of the data being linear in t. Economic models are developed and analyzed in [2] and [4].

Throughout, we are concerned with the asymptotic linear program

$$LP(t) \quad \min\{c^T(t)x : A(t)x = b(t), x \geq 0\},$$

and it associated dual

$$LD(t) \quad \max\{b^T(t)y : A^T(t)y + s = c(t), s \geq 0\},$$

where $A(t) : \mathbb{R} \to \mathbb{R}^{m \times n}$, $b(t) : \mathbb{R} \to \mathbb{R}^m$, and $c(t) : \mathbb{R} \to \mathbb{R}^n$. For any $t \in \mathbb{R}$, the *data instance* defining $LP(t)$ is $(A(t), b(t), c(t))$. The feasible region for $LP(t)$ is denoted by $\mathcal{P}(t)$, and the strict interior is $\mathcal{P}^o(t) = \{x \in \mathcal{P}(t) : x > 0\}$. Similarly, the dual feasible region is $\mathcal{D}(t)$, and its strict interior is $\mathcal{D}^o(t) = \{(y, s) \in \mathcal{D}(t) : s > 0\}$. The primal and dual optimal sets are denoted by $\mathcal{P}^*(t)$ and $\mathcal{D}^*(t)$, respectively. The necessary and sufficient optimality conditions for $LP(t)$ and $LD(t)$ are

$$\begin{aligned} A(t)x &= b(t), \quad x \geq 0, & (1) \\ A^T(t)y + s &= c(t), \quad s \geq 0, \text{ and} & (2) \\ x^T s &= 0. & (3) \end{aligned}$$

The theoretical elegance and robust computational behavior of the simplex method dominated the linear programming literature until the 1980s. However, the lack of a polynomial time simplex algorithm lead researchers to investigate other solution techniques, and in 1979 Khachiyan [18] developed an interior point algorithm that showed that the class of linear programs is solvable in polynomial time. While Khachiyan's result substantially added to the theory of linear programming, the practical performance of this algorithm was disappointing. As such, the mathematical programming community's focus remained on the simplex algorithm. This changed in 1986 when Karmarkar [17] claimed to have an interior point algorithm that out performed the simplex algorithm. This claim was heavily scrutinized by the academic community, and we now understand that interior point algorithms are not just viable alternatives to the simplex algorithm, but that they do indeed out perform simplex based procedures on large problems.

The most prevalent interior point algorithms are called path-following interior point algorithms, and these algorithms follow an infinitely smooth curve, called the *central path*, towards optimality. Our succinct development of the central path is adequate for our purposes, but interested readers are directed to the three texts of Roos, Telaky, and Vial [23], Wright [27], and Ye [28] for a complete development. The central path is constructed by replacing the complementarity constraint in (3) with

$$Xs = \mu e, \tag{4}$$

where X is the diagonal matrix of x, μ is positive, and e is the vector of ones. Notice that this constraint requires an x and a (y, s) such that $x > 0$ and $s > 0$, and hence, it requires that the primal and dual strict interiors be nonempty —i.e. $\mathcal{P}^o(t) \neq \emptyset$ and $\mathcal{D}^o(t) \neq \emptyset$. Because we are interested in the solutions provided by path-following interior point algorithms, we make the following assumption.

Assumption 1 *For sufficiently large $t \in \mathbb{R}$, the strict interiors of the primal and dual feasible regions are nonempty.*

Assumption 1 is equivalent to assuming that the primal and dual optimal sets are bounded for large t [27], and without loss in generality we assume throughout that t is large enough to satisfy this assumption. The x and s components of a solution to the system (1), (2), and (4) are unique and are denoted by $x(\mu, t)$ and $s(\mu, t))$ (see any of [21, 23, 27, 28]). The reason that y is not guaranteed to be unique is that y and s are not guaranteed to be related in a one-to-one fashion —i.e. $A(t)$ is not guaranteed to have full row rank. To overcome this difficulty, we set $y(\mu, t) = (A^T(t))^+(c(t) - s(\mu, t))$, where $(A^T(t))^+$ is the Moore-Penrose pseudo inverse of $A^T(t)$. We make the following naming conventions for a fixed t.

$$\begin{aligned}
\text{The central path at time t} &: \quad \{(x(\mu, t), y(\mu, t), s(\mu, t)) : \mu > 0\} \\
\text{The primal central path at time t} &: \quad \{x(\mu, t) : \mu > 0\} \\
\text{The dual central path at time t} &: \quad \{(y(\mu, t), s(\mu, t) : \mu > 0\}
\end{aligned}$$

The central path has a unique limit, called the *center solution*, which is in the strict interior of the optimal set. Denoting this limit by $(x^*(t), y^*(t), s^*(t))$, we have for sufficiently large t that

$$\begin{aligned}\lim_{\mu \downarrow 0} x(\mu,t) &= x^*(t) \in \mathcal{P}^*(t), \text{ and} \\ \lim_{\mu \downarrow 0} (y(\mu,t), s(\mu,t)) &= (y^*(t), s^*(t)) \in \mathcal{D}^*(t).\end{aligned}$$

Unlike a basic optimal solution, the analytic center solution is always *strictly complementary*, meaning that $(x^*(t))^T s^*(t) = 0$ and $x^*(t) + s^*(t) > 0$. (An early result due to Goldman and Tucker guarantees that every solvable linear program has such a solution [7].) Any strictly complementary solution induces the *optimal partition*, which for sufficiently large t is defined by

$$\begin{aligned}B(t) &= \{i : x_i^*(t) > 0\}, \text{ and} \\ N(t) &= \{1, 2, 3, \ldots, n\} \backslash B(t).\end{aligned}$$

The set $B(t)$ indicates the collection of primal variables allowed to be positive at optimality, and the set $N(t)$ indicates the collection of primal variables that are zero in every optimal solution. The roles of $B(t)$ and $N(t)$ are reversed for the dual problem, so $N(t)$ indexes the dual slack variables allowed to be positive at optimality, and $B(t)$ indicates the collection of dual slack variables forced to be zero at optimality. Allowing a set subscript on a vector (matrix) to be the subvector (submatrix) corresponding with the components (columns) indexed by the set, we have that the optimal partition characterizes the optimal sets as follows,

$$\begin{aligned}\mathcal{P}^*(t) &= \{x \in \mathcal{P}(t) : x_{N(t)} = 0\} \\ &= \{x : A_{B(t)}(t) x_{B(t)} = b(t), x_{B(t)} \geq 0, x_{N(t)} = 0\}\end{aligned} \tag{5}$$

and

$$\begin{aligned}&\mathcal{D}^*(t) = \{(y,s) \in \mathcal{D}(t) : s_{B(t)} = 0\} = \\ &\quad \{(y,s) : A_{B(t)}^T(t) y = c_{B(t)}(t), A_{N(t)}^T(t) y + s_{N(t)} = c_{N(t)}^T(t), s_{N(t)} \geq 0\}.\end{aligned} \tag{6}$$

The strict interiors of the optimal sets are

$$\begin{aligned}(\mathcal{P}^*(t))^o &= \{x \in \mathcal{P}^*(t) : x_{B(t)} > 0\}, \text{ and} \\ (\mathcal{D}^*(t))^o &= \{(y,s) \in \mathcal{D}^*(t) : s_{N(t)} > 0\}.\end{aligned}$$

The primal center solution is the analytic center of $\mathcal{P}^*(t)$, and the dual center solution is the analytic center of $\mathcal{D}^*(t)$. This means that $x^*(t)$ is the unique solution to

$$\max \left\{ \sum_{i \in B(t)} \ln(x_i) : A_{B(t)}(t) x_{B(t)} = b(t), x_{B(t)} > 0, x_{N(t)} = 0 \right\}. \tag{7}$$

Similarly, $(y^*(t), s^*(t))$ solves

$$\max \left\{ \sum_{i \in N(t)} \ln(s_i) : A^T_{B(t)}(t) y = c_{B(t)}(t), A^T_{N(t)}(t) y + s_{N(t)} = c^T_{N(t)}(t), s_{N(t)} > 0 \right\}.$$

The necessary and sufficient Lagrange conditions for the mathematical program in (7) are the existence of a ρ and a γ such that

$$\left.\begin{array}{rcl} A_{B(t)}(t) x_{B(t)} &=& b(t),\ x_{B(t)} > 0, \\ A^T_{B(t)}(t)\rho + \gamma &=& 0,\ \gamma > 0, \text{ and} \\ X_{B(t)}\gamma &=& e. \end{array}\right\} \tag{8}$$

The dual multipliers ρ and γ are not $y^*(t)$ and $s^*(t)$. Since the mathematical program in (7) is strictly convex, $x^*_{B(t)}(t)$ uniquely satisfies these equations. Also, since $X^*_{B(t)}(t)$ is invertible, the third equation implies that γ is also unique. However, if $A_{B(t)}(t)$ does not have full row rank, the linear relationship between ρ and γ is not one-to-one. Subsequently, ρ is unique only if $A_{B(t)}(t)$ has full row rank. We later use the fact that $A_{B(t)}(t)$ and $b(t)$ could have been replaced in (7) by a submatrix of $A_{B(t)}(t)$ having full row rank and a corresponding subvector of $b(t)$ —i.e. via row reduction. If such a substitution were undertaken, we have that the solution to (8) is unique and that $x^*_{B(t)}(t)$ remains uniquely optimal (but γ and ρ are different). Similar conditions are available for the dual center solution.

Our goal is to revisit the topics first investigated by Jeroslow, but instead of dealing with basic optimal solutions, we deal with the optimal partition and the center solution. We note that our approach is more general for the following two reasons. First, if *LP*(t) and *LD*(t) have unique solutions for sufficiently large t, the center solution is basic. Since we show in Section 2 that the center solution stabilizes, our results include the case of unique optimal basis —i.e. our results reduce to Jeroslow's results when the optimal solution is unique for all sufficiently large t. Second, our analysis is more general because it does not require that the data be rational in t (asymptotic linear programs in the literature have been built with rational functions [15, 16] and linear functions [2, 3, 4, 29]). In fact, the only restriction made on $A(t)$, $b(t)$, and $c(t)$ is that they adhere to Assumption 2.

Assumption 2 *We assume that the triple $(A(t), b(t), c(t))$ is well-behaved, meaning that there exists a time T, such that for $t \geq T$, the functions $A(t)$, $b(t)$, and $c(t)$ are continuous and have the property that the determinants of all square submatrices of*

$$\begin{bmatrix} A(t) & 0 & b(t) \\ 0 & A^T(t) & c(t) \end{bmatrix}$$

are either constant or have no roots.

For example, if $(A(t), b(t), c(t))$ is rational, the determinants of the square submatrices are rational and Assumption 2 is satisfied. However, the class of functions with which we deal is substantially larger than the set of rational functions.

We are interested in properties that reach a *steady state* or *stabilize* as time attains sufficiently large values. One of the main results of this paper shows that there exists a time T, such that for all $t \geq T$, the optimal partition stabilizes. In other words, we show that there exists a time T, such that the components of an optimal solution required to be zero at T are precisely the decision variables that must be zero for each $t \geq T$. Hence, the collection of variables that must be zero in an optimal solution stabilizes.

The paper proceeds as follows. In Section 2 we present a simple argument showing that the optimal partition stabilizes. Using this result, we develop some analytic properties in Section 3. In Section 4 we show that the results of Section 2 have economic implications by extending a famous economics result called The Nonsubstitution Theorem. Conclusions and directions for future research are located in Section 5.

Some brief notes on notation are warranted before we begin our development. A superscript $^+$ on a matrix indicates the Moore-Penrose pseudo inverse (a good reference is Campbell and Meyer [5]). Capitalizing a vector variable forms a diagonal matrix whose main diagonal is comprised of the elements of the vector. So, if x and γ are vectors, $X = \text{diag}(x_1, x_2, \ldots, x_n)$ and $\Gamma = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n)$. The rank, column space, and null space of a matrix A are denoted $\text{rank}(A)$, $\text{col}(A)$, and $\text{null}(A)$, respectively. The determinant of the matrix A is $\det(A)$. The collection of real valued functions having n continuous derivatives is denoted $\mathcal{C}^n$, and we use the standard notation that $\mathcal{C}^0$ is the set of continuous functions. For notational ease, we say that the matrix function $M(t)$ is in $\mathcal{C}^n$ if every component function of $M(t)$ is in $\mathcal{C}^n$. Other notation is standard within the mathematical programming community and may be found in the *Mathematical Programming Glossary* [8].

2 The Asymptotic Optimal Partition

The main objective of this section is to establish that the optimal partition stabilizes, and we define the *asymptotic optimal partition* to be the optimal partition that attains a steady-state. The following example clarifies our objectives.

Example 1 *Consider*

$$A(t) = \left[1, 1 + \frac{1}{e^t}\right] \ , \ b(t) = \left(\frac{1+t}{t}\right) \ , \ \textit{and} \ \ c(t) = \begin{pmatrix} 1/t \\ \tan^{-1}(t) \end{pmatrix}.$$

Let $\hat{x}(t)$ be an optimal solution at time t. Then,

$$\begin{array}{rcl} A_{(1,1)}(t)c_2(t) < A_{(1,2)}(t)c_1(t) & \Rightarrow & \hat{x}(t)_1 = 0, \\ A_{(1,1)}(t)c_2(t) > A_{(1,2)}(t)c_1(t) & \Rightarrow & \hat{x}(t)_2 = 0, \text{ and} \\ A_{(1,1)}(t)c_2(t) = A_{(1,2)}(t)c_1(t) & \Rightarrow & \text{neither } \hat{x}(t)_1 \text{ or } \hat{x}(t)_2 \text{ must be zero.} \end{array}$$

These conditions imply that

$$\begin{array}{l} \text{for } 0 < t < 1.34961, \text{ we have } (B(t)|N(t)) = (\{2\}|\{1\}), \\ \text{for } t > 1.34961, \text{ we have } (B(t)|N(t)) = (\{1\}|\{2\}), \text{ and} \\ \text{for } t = 1.34961, \text{ we have } (B(t)|N(t)) = (\{1,2\}|\emptyset). \end{array}$$

So, the collection of indices of the decision variables that must be zero in an optimal solution stabilizes after 1.34961. *However, if we replace c with*

$$c(t) = \frac{1}{t}\begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

we have that the components forced to be zero at optimality change with every solution to $\tan(t) = 1 + 1/e^t$. *Since this equation has an unbounded sequence of solutions, the desired stability does not exist. Notice that for this c, we have $\|c(t)\| = 1/t$, which is monotonically decreasing. Hence, component functions that provide monotonic norms are not sufficient. We also point out that the optimal partition exists for $t = \infty$ (assuming t is in $\mathbb{R}^* = \mathbb{R} \cup \{\infty\}$). In this case we have that $A(\infty) = [1, 1]$, $b(\infty) = (1)$, and $c(\infty) = (0, 0)^T$, which implies that $(B(\infty)|N(\infty)) = (\{1, 2\}|\emptyset)$. We mention this to distinguish the difference between behavior at ∞, which we are not investigating, and asymptotic behavior, which we are investigating. In this last situation we have that the optimal partition does not stabilize because for every t^1 we can find a larger t^2 such that the optimal partitions are different. However, the partition does exist for $t = \infty$.*

Let $\{(B^1|N^1), (B^2|N^2), \ldots, (B^{2^n}|N^{2^n})\}$ be all possible two set partitions of $\{1, 2, \ldots, n\}$. For any fixed time, one of these partitions is the *optimal partition* for $LP(t)$. We relate t to a partition by defining $\phi(t) : \mathbb{R} \to \{1, 2, \ldots, 2^n\}$, such that the optimal partition of $LP(t)$ is $(B^{\phi(t)}, N^{\phi(t)})$. We note that ϕ is well defined because the optimal partition is unique. The goal of this section may now be stated as showing that there exists T such that $\phi(t)$ is constant for $t \geq T$.

For $j = 1, 2, \ldots, 2^n$, let $v_j = (v_1^T, v_2^T, v_3^T)^T$ be partitioned as $\left(x_{B^j}^T, y^T, s_{N^j}^T\right)^T$. Define

$$H_j(t) = \begin{bmatrix} A_{B^j}(t) & 0 & 0 \\ 0 & A_{B^j}^T(t) & 0 \\ 0 & A_{N^j}^T(t) & I \end{bmatrix} \text{ and } h_j(t) = \begin{pmatrix} b(t) \\ c_{B^j}(t) \\ c_{N^j}(t) \end{pmatrix}.$$

We say that v_j is sufficiently positive, written $v_j \gg 0$, if $v_1 > 0$ and $v_3 > 0$. Observe that $\hat{v}_{\phi(t)} = (\hat{v}_1^T, \hat{v}_2^T, \hat{v}_3^T)^T$ relates to

$$((x^T_{B\phi(t)}, x^T_{N\phi(t)}), y^T, (s^T_{B\phi(t)}, s^T_{N\phi(t)})) = ((x^T_{B\phi(t)}, 0), y^T, (0, s^T_{N\phi(t)}))$$

in a one-to-one fashion —i.e. $\hat{v}_1 \leftrightarrow x_{B\phi(t)}$, $\hat{v}_2 \leftrightarrow y$, and $\hat{v}_3 \leftrightarrow s_{N\phi(t)}$. From (5) and (6) we now see that the set of sufficiently positive solutions to $H_{\phi(t)}(t)v_{\phi(t)} = h_{\phi(t)}(t)$ is isomorphic to $(\mathcal{P}^*(t))^o \times (\mathcal{D}^*(t))^o$. Also, from the fact that the optimal partition is unique, we have the important property that the equation $H_j(t)v_j = h_j(t)$ has a sufficiently positive solution if, and only if, $j = \phi(t)$.

The proof that the optimal partition stabilizes depends on the following three lemmas. The first of these lemmas shows that the rank of a matrix attains a steady-state under Assumption 2.

Lemma 1 *Let $M(t)$ be a matrix function whose component functions have the property that there exists a time T, such that for all $t \geq T$, the determinants of all square submatrices are either constant or have no roots. Then, the* $\text{rank}(M(t))$ *stabilizes.*

Proof: Let T be such that for all $t \geq T$, the determinants of all square submatrices of $M(t)$ have either become constant or have no roots. Let $S(T)$ be a maximal submatrix of $M(T)$ with nonzero determinant. Then, all larger square submatrices have a determinant of zero for $t \geq T$. Since $\det(S(t)) \neq 0$ for $t \geq T$, we have that $\text{rank}(M(t)) = \text{rank}(S(t))$ for $t \geq T$. ■

The second lemma shows that the optimal partition remains constant over a neighborhood provided that $h_j(t)$ remains in the column space of $H_j(t)$, and that the Moore-Penrose pseudo inverse of $H_j(t)$ is continuous. The continuity of $H_j^+(t)$ might appear self serving, but as we shall see, this condition is tied closely to the rank of $H_j(t)$, which is easier to deal with.

Lemma 2 *Let t_0 be large enough to satisfy Assumption 1, and set $j = \phi(t_0)$. Let $\mathcal{N}$ be a neighborhood of t_0 such that $H_j^+(t)$ is continuous over $\mathcal{N}$ and that $h_j(t) \in \text{col}(H_j(t))$ for $t \in \mathcal{N}$. Then, the optimal partition is constant over some neighborhood about t_0.*

Proof: Let $v_j(t_0)$ be a sufficiently positive solution to $H_j(t_0)v_j = h_j(t_0)$. Then, $v_j(t_0) = H_j^+(t_0)h_j(t_0) + q(t_0)$, where $q(t_0) \in \text{null}(H_j(t_0))$. Let

$$v_j(t) = H_j^+(t)h_j(t) + (I - H_j^+(t)H_j(t))(q(t_0) + H_j^+(t_0)h_j(t_o) - H_j^+(t)h_j(t)).$$

The proof follows once we show that for t sufficiently close to t_0, $v_j(t)$ is a sufficiently positive solution to $H_j^+(t_0)v_j = h_j(t_0)$. First, since

$$(I - H_j^+(t)H_j(t))(q(t_0) + H_j^+(t_0)h_j(t_o) - H_j^+(t)h_j(t)) \in \text{Null}(H_j(t))$$

we have

$$
\begin{aligned}
H_j(t)v_j(t) &= H_j(t)H_j^+(t)h_j(t) \\
&= h_j(t),
\end{aligned}
$$

where the last equality follows because $h_j(t) \in \mathrm{col}(H_j(t))$. Second, because both $H_j^+(t)$ and $h_j(t)$ are continuous at t_0, $H_j^+(t_0)h_j(t_o) - H_j^+(t)h_j(t) \to 0$ as $t \to 0$. Hence, as $t \to t_0$

$$
\begin{aligned}
&(I - H_j^+(t)H_j(t))(q(t_0) + H_j^+(t_0)h_j(t_o) - H_j^+(t)h_j(t)) \\
&\qquad \to \quad (I - H_j^+(t_0)H_j(t_0))q(t_0) = q(t_0),
\end{aligned}
$$

where the last equality follows because $q(t_0) \in \mathrm{Null}(H_j(t_0))$. We now have that

$$
\begin{aligned}
v_j(t) &= H_j^+(t)h_j(t) + (I - H_j^+(t)H_j(t))(q(t_0) + H_j^+(t_0)h_j(t_o) - H_j^+(t)h_j(t)) \\
&\to H_j^+(t_0)h_j(t_0) + q(t_0) \\
&\succ 0,
\end{aligned}
$$

which completes the proof. ■

Lemma 2 connects the local stability of the optimal partition with the continuity of $H_j^+(t)$, and Lemma 3 shows that the Moore-Penrose pseudo inverse is continuous so long as rank is preserved. This result, together with Lemma 1, allow us to use the steady-state behavior of the rank of $H_j(t)$ to show that the optimal partition stabilizes. A proof of the following result is found in [5].

Lemma 3 *The matrix function $H_j^+(t)$ is continuous at t_0 if, and only if, $\mathrm{rank}(H_j(t_0)) = \mathrm{rank}(H_j(t))$, for t sufficiently close to t_0.*

We are now ready to establish that the optimal partition of $LP(t)$ and $LD(t)$ stabilizes for sufficiently large t.

Theorem 1 *Assume that $(A(t), b(t), c(t))$ satisfies Assumptions 1 and 2. Then, there exists a T, such that for all $t \geq T$, $(B(t)|N(t)) = (B^{\phi(T)}|N^{\phi(T)})$.*

Proof: We first note that $H_j(t)v_j = h_j(t)$ has a solution if, and only if, $\mathrm{rank}(H_j(t)) = \mathrm{rank}([H_j(t)|h_j(t)])$. From Assumption 2 and Lemma 1 we have that there is a T_1 such that for all $t \geq T_1$ and all $j \in \{1, 2, \ldots, 2^n\}$,

$$
\begin{aligned}
\mathrm{rank}(H_j(T_1)) &= \mathrm{rank}(H_j(t)) \text{ and} \\
\mathrm{rank}([H_j(T_1)|h_j(T_1)]) &= \mathrm{rank}([H_j(t)|h_j(t)]).
\end{aligned}
$$

Assumption 1 implies that there exists $T_2 > T_1$ such that for $t \geq T_2$, there exists a sufficiently positive solution to $H_{\phi(t)}(t)v_{\phi(t)} = h_{\phi(t)}(t)$. Let $\overline{T} > T_2 > T_1$.

Then, $\text{rank}(H_{\phi(t)}(t))$ is constant and $h_{\phi(t)}(t) \in \text{col}(H_{\phi(t)}(t))$, for $t \geq \overline{T}$. From Lemma 2 we have that there is an open neighborhood, $\mathcal{N}_1$, about $\overline{T}$ such that $T_2 \notin \mathcal{N}_1$ and $(B(t)|N(t)) = (B(\overline{T})|N(\overline{T}))$, for $t \in \mathcal{N}_1$. Let

$$\mathcal{N}_2 = \{\overline{T} + \hat{\delta} : (B(t+\delta)|N(t+\delta)) = (B(\overline{T})|N(\overline{T})), \delta \in [0, \hat{\delta}]\}.$$

Again, from Lemma 2 we have that for any $t \in \mathcal{N}_1 \cup \mathcal{N}_2$, there is an open neighborhood about t over which the optimal partition is stable, which means that $\mathcal{N}_1 \cup \mathcal{N}_2$ is open. Now, let

$$\hat{t} = \inf\{t > \overline{T} : (B(t)|N(t)) \neq (B(\overline{T})|N(\overline{T}))\}.$$

Suppose for the sake of attaining a contradiction that $\hat{t} < \infty$. Since $\mathcal{N}_1 \cup \mathcal{N}_2$ is open, we have that $(B(\overline{T})|N(\overline{T})) \neq (B(\hat{t})|N(\hat{t}))$. From Lemma 2 we know that there exists an open neighborhood, $\mathcal{N}_3$, about $\hat{t}$ such that $(B(t)|N(t)) = (B(\hat{t})|N(\hat{t}))$ for $t \in \mathcal{N}_3$. However, $\mathcal{N}_2 \cap \mathcal{N}_3 \neq \emptyset$, and for any $t \in \mathcal{N}_1 \cap \mathcal{N}_2$ we have the contradiction that

$$(B(\overline{T})|N(\overline{T})) = (B(t)|N(t)) = (B(\hat{t})|N(\hat{t})).$$

Hence, $(B(t)|N(t)) = (B(\overline{T})|N(\overline{T}))$ for all $t \geq \overline{T}$. ■

Theorem 1 shows that the optimal partition stabilizes, and this result allows us to make the following definitions.

Definition 1 *Assuming the data functions adhere to Assumptions 1 and 2, we define the* **asymptotic optimal partition** *to be the unique partition for which there exists T such that $(B(t)|N(t)) = (B(T)|N(T))$, for all $t \geq T$. We denote this partition by $(\bar{B}|\bar{N})$, and we set $\overline{T}$ to be a sufficiently large time so that $(B(\overline{T})|N(\overline{T})) = (\bar{B}|\bar{N})$.*

Definition 2 *Under Assumptions 1 and 2, and for $t \geq \overline{T}$, the* **asymptotic center solution**, *$x^*(t) = (x^*_{\bar{B}}(t), x^*_{\bar{N}}(t)) = (x^*_{\bar{B}}(t), 0)$, is defined so that $x^*_{\bar{B}}(t)$ is the unique solution to*

$$\max\left\{\sum_{i\in\bar{B}} \ln(x_i) : A_{\bar{B}}(t)x = b(t), x_{\bar{B}} > 0\right\}.$$

In this section we have established, under mild assumptions, that the optimal partition attains a steady-state as time proceeds to infinity. This means that the collection of variables that are zero in every optimal solution becomes invariant for sufficiently large time. Using this information, we defined the asymptotic optimal partition and subsequently defined the asymptotic center solution. Properties of this unique solution are studied in the next section.

3 Analytic Properties of the Asymptotic Analytic Center

In this section we exploit the fact that the optimal partition stabilizes to attain analytic properties of the asymptotic center solution. For a fixed $t \geq \overline{T}$, the analytic properties of the central path and the center solution are well studied. For example, the elements of the central path are analytic functions of μ, b, and c, a fact first recognized by Sonnevend [25]. Differential properties of the central path with respect to μ are important for algorithm design and are found in [1, 10, 11, 13, 26, 30]. Analytic properties of the center solution with respect to b and c are studied in [13] and [14]. However, all of these results assume that the coefficient matrix is fixed, and the only papers that consider the more difficult situation of perturbing A are [6] and [9]. Since each of A, b, and c depend on t in the asymptotic linear program, the results of this section are significantly different than those in the literature. Because the results of this section are asymptotic, we assume for linguistic simplicity that $t \geq \overline{T}$ throughout this section.

The main result of this section states that the asymptotic center solution inherits the analytic properties of $A(t)$ and $b(t)$. So, since both $A(t)$ and $b(t)$ are continuous, $x^*(t)$ is continuous, and if $A(t)$ and $b(t)$ are differentiable, $x^*(t)$ is differentiable. The proofs establishing the continuity and differentiability of $x^*(t)$ are handled separately. The reason for the separate arguments is that the vehicle of proof for differentiability is the implicit function theorem, which is not applicable unless the data functions are themselves differentiable. The continuity of $x^*(t)$ is proven through an adaptation of an argument in [6]. To explain this approach, we introduce some notation and generalize the definition of the analytic center. Let $\{U(t), u(t)\}$ be matrix functions in $\mathbb{R}^{m \times n} \times \mathbb{R}^m$, and for each t, suppose that $P(U(t), u(t))$ defined by $\{x : U(t)x \leq u(t)\}$ is bounded. For $x \in P(U(t), u(t))$, define $s = u(t) - U(t)x$ and let $I = \{i : s_i > 0 \text{ for some } x \in P(U(t), u(t))\}$. The analytic center of $P(U(t), u(t))$ is $x^c(U(t), u(t))$ and is the unique solution to

$$\max \left\{ \sum_{i \in I} \ln(s_i) : x \in P(U(t), u(t)) \right\}.$$

The following small example illustrates the difficulty of dealing with a non-constant coefficient matrix. In particular, it shows that the analytic center need not be continuous even if $U(t)$ and $u(t)$ are smooth.

Example 2 *Consider*

$$\{(U(t),u(t))\} \quad = \quad \left\{ \begin{bmatrix} 2-\frac{1}{1+(t-100)^2} & 1 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

For $t \neq 100$ we have that $I = \{4\}$, but for $t = 100$, $I = \{3,4\}$. It is easy to check that $x^c(U(t),u(t)) = (0,1)$, for all $t \neq 100$ (in fact this is the only element in $P(U(t),u(t))$), but that $x^c(U(100),u(100)) = (1/2,1/2)$.

From this example we see that the analytic center is not necessarily continuous with respect to changes in the matrix coefficients. An important observation is that the first two constraints are implied equalities for $t = 100$, but that the first three constraints are implied for $t \neq 100$. Moreover, notice that

$$\operatorname{rank}\left(\begin{bmatrix} 2-\frac{1}{1+(t-100)^2} & 1 \\ -1 & -1 \end{bmatrix} \right)$$

is 2 for $t \neq 100$ and 1 for $t = 100$. What the authors of [6] were able to show is that the analytic center is continuous with respect to matrix perturbations at t_0, so long as the rank of the matrix formed by the implied equalities at t_0 is constant over some sufficiently small neighborhood of t_0. To state this precisely, we partition the rows of $U(t)$ and $u(t)$ at $t = t_0$ as indicated,

$$U(t) = \begin{bmatrix} A^{t_0}(t) \\ B^{t_0}(t) \end{bmatrix} \text{ and } u(t) = \begin{pmatrix} a^{t_0}(t) \\ b^{t_0}(t) \end{pmatrix},$$

where $A^{t_0}(t_0)x = a^{t_0}(t_0)$ for all $x \in P(U(t_0),u(t_0))$ and $B^{t_0}(t_0)x < b^{t_0}(t_0)$ for some $x \in P(U(t_0),u(t_0))$ —i.e. I indexes the rows of the submatrix B. For example, consider $\{U(t),u(t)\}$ from the previous example, and let $t_0 = 100$. Then, the first two inequalities form the collection of implied equalities at t_0, which means that

$$A^{t_0}(t) = \begin{bmatrix} 2-\frac{1}{1+(t-100)^2} & 1 \\ -1 & -1 \end{bmatrix},\ B^{t_0}(t) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$a^{t_0}(t) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \text{ and } b^{t_0}(t) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

However, for $t_1 \neq 100$ we have

$$A^{t_1}(t) = \begin{bmatrix} 2 - \frac{1}{1+(t-100)^2} & 1 \\ -1 & -1 \\ -1 & 0 \end{bmatrix}, \; B^{t_1}(t) = \begin{bmatrix} 0 & -1 \end{bmatrix},$$

$$a^{t_1}(t) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \text{ and } b^{t_1}(t) = \begin{pmatrix} 0 \end{pmatrix}.$$

So, the superscript indicates the time at which the inequalities are partitioned into those that are implied and those that are not implied. Lemma 4 shows that the analytic center of $P(U(t), u(t))$ is continuous at t_0, provided that the rank of the coefficient matrix for the implied equalities is invariant over some neighborhood about t_0. A proof is found in [6].

Lemma 4 *Let $\{U(t), u(t)\}$ be continuous at t_0. Then, the analytic center $x^c(U(t), u(t))$ is continuous at t_0, provided that* $\operatorname{rank}(A^{t_0}(t))$ *is constant for all t sufficiently close to t_0.*

From Lemma 4 we see that the continuity of the analytic center depends on a rank condition dealing with the implied equalities. Since $x \in \mathcal{P}^*(t)$ for $t \geq \overline{T}$ implies that $x_{\bar{N}}(t) = 0$, and there exists $x \in \mathcal{P}^*(t)$ such that $x_{\bar{B}}(t) > 0$, we have that $\bar{N}$ indicates the entire set of implied equalities that define the optimal face. Moreover, we have that the asymptotic center solution is the analytic center of the optimal face. As the next theorem shows, the rank of these implied equalities is constant for all $t \geq \overline{T}$, and hence the asymptotic analytic center solution is continuous for large t.

Lemma 5 *Let $(A(t), b(t), c(t))$ satisfy Assumptions 1 and 2. Then, $x^*(t)$ is continuous for sufficiently large t.*

Proof: Let $t_0 \geq \overline{T}$, and set

$$U(t) = \begin{bmatrix} A_{\bar{B}}(t) & | & A_{\bar{N}}(t) \\ -A_{\bar{B}}(t) & | & -A_{\bar{N}}(t) \\ 0 & | & I \\ 0 & | & -I \\ --- & | & --- \\ -I & | & 0 \end{bmatrix} \text{ and } u(t) = \begin{pmatrix} b(t) \\ -b(t) \\ 0 \\ 0 \\ -- \\ 0 \end{pmatrix},$$

where the row partitioning indicates the submatrices $A^{t_0}(t)$ and $B^{t_0}(t)$. Since $U(t)x \leq u(t)$ is the same as $A_{\bar{B}}x_{\bar{B}} = b$, $x_{\bar{N}} = 0$, and $x_{\bar{B}} \geq 0$, we have that

$P(U(t),u(t)) = \mathcal{P}^*(t)$. So, $x^c(U(t),u(t)) = x^*(t)$, and from Lemma 4 the continuity of $x^*(t)$ follows because

$$\text{rank}\left(\left[\begin{array}{r|r} A_{\bar{B}}(t) & A_{\bar{N}}(t) \\ -A_{\bar{B}}(t) & -A_{\bar{N}}(t) \\ 0 & I \\ 0 & -I \end{array}\right]\right)$$

is stable for sufficiently large t. ■

The use of $c(t)$ differs from the use of $(A(t), b(t))$ in the proof Lemma 5. This is because $c(t)$ is used only to define $\bar{B}$ and $\bar{N}$, and hence, the dependence that $x^*(t)$ has on $c(t)$ is expressed through the asymptotic optimal partition. This is in contrast to the use of $A(t)$ and $b(t)$, which are not only used to define $\bar{B}$ and $\bar{N}$, but are also used to define the polytope $\mathcal{P}^*(t)$. This observation fore-shadows the fact that $x^*(t)$ inherits the differential properties of $A(t)$ and $b(t)$, but that $c(t)$ only needs to be continuous for such an inheritance to work.

As previously mentioned, the proof establishing the differentiability of $x^*(t)$ follows from the implicit function theorem. However, the nonsingular gradient required by the implicit function theorem is not immediately available. The problem is that $A_{\bar{B}}(t)$ need not have full row rank. We overcome this difficulty by allowing $\hat{A}_{\bar{B}}(t)$ be a submatrix of $A_{\bar{B}}(t)$ such that $\hat{A}_{\bar{B}}(t)$ has full row rank and $\text{null}(A_{\bar{B}}(t)) = \text{null}(\hat{A}_{\bar{B}}(t))$. Then, $\mathcal{P}^*(t) = \{x : \hat{A}_{\bar{B}}(t)x_{\bar{B}} = \hat{b}(t), x_{\bar{N}} = 0, x_{\bar{B}} \geq 0\}$, where $\hat{b}(t)$ is the subvector of $b(t)$ corresponding to $\hat{A}_{\bar{B}}(t)$. We now have that $x^*_{\bar{B}}(t)$ is the unique solution to

$$\max\left\{\sum_{i\in\bar{B}} \ln(x_i) : \hat{A}_{\bar{B}}(t)x_{\bar{B}} = \hat{b}(t), x_{\bar{B}} > 0, x_{\bar{N}} = 0\right\},$$

and because $\hat{A}_{\bar{B}}$ has full row rank, $x^*_{\bar{B}}$ is part of the unique solution to

$$\begin{aligned} \hat{A}_{\bar{B}}(t)x_{\bar{B}} &= \hat{b}(t) \\ \hat{A}^T_{\bar{B}}(t)\rho + \gamma &= 0 \\ X_{\bar{B}}\gamma &= e. \end{aligned}$$

What this means is that $x^*(t)$ is not affected by removing redundant equality constraints, and hence, we can remove redundant equality constraints with impunity. Lemma 6 uses this fact to establish that the asymptotic center solution is as smooth as $A_{\bar{B}}(t)$ and $b(t)$.

Lemma 6 *Assume that $(A(t), b(t), c(t))$ satisfies Assumptions 1 and 2. Additionally, for $t \geq \overline{T}$ assume that both $A_{\bar{B}}(t)$ and $b(t)$ are in $\mathcal{C}^n$, for some $n \geq 1$. Then, $x^*(t)$ is in $\mathcal{C}^n$ for all $t \geq \overline{T}$.*

Proof: Since $x_{\bar{N}}(t) = 0$ for all $t \geq \overline{T}$, the proof trivially holds for these components. Let $t_0 \geq \overline{T}$ and let $\hat{A}_{\bar{B}}(t_0)$ be a full row rank submatrix of $A_{\bar{B}}(t_0)$ with the property that $\text{null}(\hat{A}_{\bar{B}}(t_0)) = \text{null}(A_{\bar{B}}(t_0))$. We note that because the determinants of all square submatrices are either fixed or have no roots, the collection of rows used to form $\hat{A}$ is independent of $t \geq \overline{T}$. Let k be the rank of $\hat{A}_{\bar{B}}(t_0)$, and define $\Psi : \mathbb{R}^{2|B|+k+1} \to \mathbb{R}^{2|B|+k}$ by

$$\Psi(x_{\bar{B}}, \rho, \gamma, t) = \begin{pmatrix} \hat{A}_{\bar{B}}(t)x_{\bar{B}} - \hat{b}(t) \\ \hat{A}^T_{\bar{B}}(t)\rho + \gamma \\ X_{\bar{B}}\gamma - e \end{pmatrix},$$

where $\hat{b}(t_0)$ is the subvector of $b(t_0)$ that corresponds with the submatrix $\hat{A}_{\bar{B}}(t_0)$. We point out that the solution to $\Psi(x_{\bar{B}}, \rho, \gamma, t_0) = 0$, $x_B > 0$, and $\gamma > 0$ is unique, which follows because this solution satisfies (8) with $A_{\bar{B}}(t_0)$ and $b(t_0)$ replaced with $\hat{A}_{\bar{B}}(t_0)$ and $\hat{b}(t_0)$. Hence, the $x_{\bar{B}}$ part of this solution is $x^*_{\bar{B}}(t_0)$. We let ρ_0 and γ_0 be the unique solution to $\Psi(x^*_{\bar{B}}(t_0), \rho, \gamma, t_0) = 0$. The gradient of Ψ with respect to $x_{\bar{B}}$, ρ, and γ is

$$\nabla_{(x_{\bar{B}}, \rho, \gamma)}\Psi(x_{\bar{B}}, \rho, \gamma, t) = \begin{bmatrix} \hat{A}_{\bar{B}}(t) & 0 & 0 \\ 0 & \hat{A}^T_{\bar{B}}(t) & I \\ \Gamma & 0 & X_{\bar{B}} \end{bmatrix}.$$

The full row rank of $\hat{A}_{\bar{B}}(t_0)$ implies that $\nabla_{x_{\bar{B}}, \rho, \gamma}\Psi(x^*_{\bar{B}}(t_0), \rho_0, \gamma_0, t_0)$ is invertible (see Theorem II.41 in [23]). The desired analytic property of $x^*_{\bar{B}}(t_0)$ follows from the implicit function theorem. ■

As previously alluded to, the proof of Lemma 5 required that only $c(t)$ be continuous for sufficiently large t. Similarly, there are no differential properties imposed on $A_{\bar{N}}(t)$. Theorem 2 follows directly from Lemmas 5 and 6 and shows that the asymptotic center solution inherits the analytic properties of $A_{\bar{B}}(t)$ and $b(t)$.

Theorem 2 *Assume that $(A(t), b(t), c(t))$ satisfies Assumptions 1 and 2. Then, for sufficiently large t we have that $x^*(t) \in \mathcal{C}^n$, provided that $A_{\bar{B}}(t)$ and $b(t)$ are in $\mathcal{C}^n$, $n \geq 0$.*

4 Economic Applications

In this section we show how to use the asymptotic optimal partition to extend a classic result in economics known as the Nonsubstitution Theorem (a result first proved by the Nobel Laureate Paul Samuelson [24]). This result states that there is a collection of processes in an economy that are optimal, in the sense that the amount of required labor is as small as possible, independent of the demands. The importance of the Nonsubstitution Theorem is highlighted in the following quote [20], "The theorem was received with some astonishment by the authors working in the neoclassical tradition since it flatly contradicted the importance attached to consumer preferences for the determination of relative prices." Indeed, this result has been studied by other Nobel Laureates (Mirrlees [22]) and continues to be investigated [19].

This section is divided into two subsections. Subsection 4.1 begins by developing a simple model of an economy and continues by showing how linear programming techniques are used to select production procedures and calculate prices. After stating the The Nonsubstitution Theorem, we allow the data describing the economy to become dynamic —i.e. dependent on the single parameter of time. Subsection 4.1 concludes with a dynamic version of the Nonsubstitution Theorem, which has a surprising corollary. Subsection 4.2 removes one of the economic assumptions required by the Nonsubstitution Theorem, and develops a similar result under a new set of assumptions.

4.1 A Dynamic Version of the Nonsubstitution Theorem

We consider an economic model where there are constant returns to scale and a single, primary, non-producible, homogeneous labor source. Suppose we want to manufacture n commodities, indexed by j, from m processes, indexed by i. We assume that there is at least one process capable of producing each commodity, which means that $m \geq n$. A process is described by the triple (a^i, b^i, l^i), where

- a^i is a commodity input row-vector for process i (a^i_j is the amount of commodity j required by process i),
- b^i is a commodity output row-vector for process i (b^i_j is the amount of commodity j yielded by process i), and
- l^i is the amount of labor required by process i (we assume that every process requires some labor, and hence, that l^i is positive).

The goal is to achieve a profit rate of r by deciding 1) prices for the commodities, 2) a price for the labor, and 3) a processing technique. We make the following assumption throughout this section.

Assumption 3 *There is no joint production, meaning that a process can only produce a single commodity.*

From Assumption 3 we have that each b^i contains a single positive component. Without loss of generality, we assume that the output of each process is one unit, which means that b^i_j is 1 if process i yields commodity j, and 0 otherwise. Let

$$A = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{bmatrix}, \quad B = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{bmatrix}, \text{ and } \quad l = \begin{bmatrix} l^1 \\ l^2 \\ \vdots \\ l^m \end{bmatrix}.$$

So, A is an $m \times n$ input matrix, B is an $m \times n$ output matrix, and l is an m vector of labor requirements. We order the processes so that B^T has the following form,

$$B^T = \left[\begin{array}{c|c|c|c} 1\,1\,\ldots 1 & & & \\ \hline & 1\,1\,\ldots 1 & & \\ \hline & & \ddots & \\ \hline & & & 1\,1\,\ldots 1 \end{array}\right].$$

The decision variables for the economy are

- x_i — the amount of process i to use (or how long process i runs),
- p_j — the price of commodity j, and
- w — the labor cost.

The process and price vectors are $p^T = (p_1, p_2, \ldots, p_n)^T$ and $x^T = (x_1, x_2, \ldots, x_m)^T$. Two important quantities are Bp and $(1+r)Ap + wl$; the former is a price vector for the commodities we produce, and the latter is a price vector for the amount we wish to recover (Ap prices the commodities used as inputs, wl is the labor costs for the processes, and the multiple $(1+r)$ represents the amount of profit we wish to recover). We say that commodity i has *extra costs* if $(Bp)_i < ((1+r)Ap + wl)_i$, and that it pays *extra profits* if $(Bp)_i > ((1+r)Ap + wl)_i$. Suppose that process i_0 pays an extra profit. Then, for $x_{i_0} > 0$, $x_{i_0}(Bp)_{i_0} > x_{i_0}((1+r)Ap + wl)_{i_0}$, and as $x_{i_0} \to \infty$, the gap between these two quantities grows towards infinity. Since $x_{i_0}(Bp)_{i_0}$ represents the revenue generated by selling the commodity produced by process i_0, and $x_{i_0}((1+r)Ap+wl)_{i_0}$ is greater than our cost to run process i_0 (the actual cost is $x_{i_0}(Ap+wl)_{i_0}$), we see that we can achieve infinite profits by running process i_0. Because this is unrealistic, we assume there are no processes that pay extra profits. That is we assume

$$Bp \leq (1+r)Ap + wl.$$

The triple (x, p, w), where $x \geq 0$, $p \geq 0$, and $w > 0$, is called a *long-period solution* if

$$x^T[B - (1+r)A]p = wx^T l \quad \text{and} \quad x^T B > 0.$$

The first equality guarantees that the processes that are run have no extra costs —i.e. they achieve the sought after profit. The second inequality guarantees that at least one process is used for each commodity. Let d be a positive n-vector, with d_j being the demand for commodity j. Prices and demand are related through the normalization constraint $d^T p = 1$. So, the economy is represented by

$$\begin{aligned}
[B-(1+r)A]p &\leq wl, && (9)\\
x^T[B-(1+r)A]p &= wx^T l, && (10)\\
x^T B &> 0, && (11)\\
d^T p &= 1, && (12)\\
x, p &\geq 0, \text{ and} && (13)\\
w &> 0. && (14)
\end{aligned}$$

Our first objective is to show that long period solutions may be generated by solving a linear program. The following lemma provides conditions for a matrix to be *monotonic*, meaning that the matrix is invertible and that its inverse is non-negative.

Lemma 7 (See Theorem A.3.1 in [20]) *If there exists a non-negative x and a scalar λ such that $x^T[\lambda I - A]$ is positive, then λ is positive and $[\lambda I - A]$ is monotonic.*

A *technique*, denoted by σ, is a collection of processes capable of producing all n commodities such that no two processes produce the same commodity. In what follows, we alter the set subscript notation so that A_σ is the collection of rows, not columns, of A indexed by σ. The initial ordering of procedures means that for any technique σ, $B_\sigma = I$. The proof of Theorem 3 can be found in [20], but we include the proof because we extend it in the following section.

Theorem 3 (See Lemma 5.2 in [20]) *System (9) - (14) is feasible if, and only if, the following primal and dual pair of linear programs is well-posed (meaning that both problems have an optimal solution),*

$$\begin{aligned}
LP_{econ} \quad & \min\{x^T l : x^T[B-(1+r)A] \geq d^T,\ x \geq 0\} \quad \text{and}\\
LD_{econ} \quad & \max\{d^T y : [B-(1+r)A]y \leq l,\ y \geq 0\}.
\end{aligned}$$

Moreover, if x^* and y^* are optimal for LP_{econ} and LD_{econ}, then $x = x^*$, $p = (1/d^T y^*)y^*$, and $w = 1/d^T y^*$ are long-period solutions to system (9) - (14).

Proof: Consider the following equations,

$$
\begin{aligned}
[B-(1+r)A]u &\leq l, && (15)\\
x^T[B-(1+r)A]u &= x^T l, && (16)\\
x^T[B-(1+r)A] &\geq d^T, && (17)\\
x^T[B-(1+r)A]u &= d^T u, \quad \text{and} && (18)\\
d^T u &> 0 && (19)\\
x, u &\geq 0. && (20)
\end{aligned}
$$

Let x^* and y^* be optimal solutions to (LP_{econ}) and (LD_{econ}). Since optimal solutions are complementary, we have that

$$d^T y^* = (x^*)^T[B-(1+r)A]y^* = (x^*)^T l.$$

So, x^* and y^* satisfy equations (15), (16), (17), (18), and (20). Since d and l are positive, every feasible solution to LP_{econ} yields a positive objective value. Hence, $(x^*)^T l = d^T y^*$ is positive, and y^* satisfies equation (19).

If x^* and u^* are solutions to system (15) - (20), equations (15), (17), and (20) show that x^* is feasible to (LP_{econ}) and u^* is feasible to (LD_{econ}). Moreover, from (16) and (18) we have that $(x^*)^T l = d^T u^*$, and the Strong Duality Theorem of linear programming implies that x^* is optimal to (LP_{econ}) and u^* is optimal to (LD_{econ}). So, the primal and dual pair of (LP_{econ}) and (LD_{econ}) being well-posed is equivalent to the consistency of system (15) - (20).

We complete the proof by showing that system (9) - (14) admits a solution if, and only if, system (15) - (20) admits a solution. Let x^* and u^* satisfy system (15) - (20). Setting $\hat{x} = x^*$, $\hat{p} = (1/du^*)u^*$, and $\hat{w} = 1/du^*$, we see that $\hat{x}$, $\hat{p}$, and $\hat{w}$ satisfy equations (9), (10), (12), (13), and (14). Also, $\hat{x}^T B = (x^*)^T B \geq (1+r)(x^*)^T A + d > 0$, and equation (11) is satisfied. So, the consistency of system (15) - (20) implies the consistency of system (9) - (14).

Let (x^*, p^*, w^*) be a solution to system (9) - (14). From equation (11) we know that each commodity is being produced, which means there is a technique σ such that $x^*_\sigma > 0$. From Lemma 7 we know that $[I-(1+r)A_\sigma]$ is monotonic. Set $\hat{x}^T_\sigma = d^T[I-(1+r)A_\sigma]^{-1}$, and embed $\hat{x}_\sigma$ into $\hat{x}$ such that $\hat{x}$ is nonnegative and $\hat{x}^T[I-(1+r)A] = d^T$. Setting $\hat{u} = (1/w^*)p^*$, we see that $\hat{x}$ and $\hat{u}$ are solutions to system (15) - (20). ■

There are numerous economic models similar to system (9) - (14), each arising from a slightly different set of assumptions. A complete discussion of these models is beyond the scope of this work, with our objective being the demonstration of how a model of the economy can be transformed into the realm of linear programming. The economic variations are ultimately equivalent to system (15) - (20), the difference being the interpretation of the data. (see [20] for an explanation). As such, the pair of linear programs LP_{econ} and LD_{econ} is essential to the analysis of these economies. The primal linear program is easy to interpret as minimizing the amount of labor so that demand is satisfied, and the dual problem calculates the rates at which the optimal amount of labor changes with respect to changes in the demand —i.e. if $(x^*(d))^T l$ is the minimum amount of labor for demand d, $\partial(x^*(d))^T l/\partial d_i = y_i^*$ (This follows only because there is a unique solution to LD_{econ}. This is not an obvious fact, and we direct interested readers to Theorem 5.2 in [20]).

Let σ be a technique. As discussed in the proof of Theorem 3, the matrix $[I-(1+r)A_\sigma]$ is monotonic, which means that $(x_\sigma)^T = d^T[I-(1+r)A_\sigma]^{-1}$ is non-negative. Consequently, $(x_\sigma, 0)$ is a basic feasible solution. Moreover, there are no basic feasible solutions other than those induced by techniques. To see this, let ν be a collection of processes that is not a technique. For ν to induce a basic feasible solution, the matrix $[B_\nu-(1+r)A_\nu]$ must be invertible, and hence square. Since ν is not a technique, this means that there is a commodity not produced by any of the processes in ν. Subsequently, there is no nonnegative solutions to $x_\nu^T[B_\nu-(1+r)A_\nu] = d^T$, and ν does not induce a basic feasible solution.

From the Fundamental Theorem of Linear Programming we know that some basic feasible solution is optimal. A technique σ is optimal if $(x_\sigma, 0)$ is an optimal basic solution, and we say that a process is optimal if it is used in some optimal technique. An important result first proved by Samuelson [24] is that there is a technique that is optimal independent of demand.

Theorem 4 (Nonsubstitution Theorem) *Under Assumption 3 there is a technique σ^* that is optimal for every possible demand vector d.*

A technique that is optimal independent of demand is called *demand-independent*, and an optimal processes is demand-independent if it may be used regardless of the demand. We point out that the Nonsubstitution Theorem does not say that σ^* is unique. For example, suppose that there are two identical processes with low labor requirements. A demand-independent optimal technique can only contain one of these processes, but since the two processes are identical, there must be an alternative demand-independent optimal technique that contains the other process. So, calculating a demand-independent optimal technique does not guarantee that all demand-independent optimal processes are found.

This is where the idea of the optimal partition comes to the forefront, and the result developed below captures the concept of partitioning the processes into those that are optimal and those that are not. The difference is that we allow the labor requirements, the profit, the input and output coefficients, and the demand to be dynamic, meaning that they depend on time. To accommodate this, we let $A(t) \geq 0$ be the matrix of material inputs for the processes at time t, $l(t) > 0$ be the labor requirements for the processes at time t, $d(t) > 0$

be the demand at time t, and $r(t) \geq 0$ be the profit at time t. We let $M(t)$ be the partitioned matrix $[B-(1+r(t))A(t) \,|\, -I]$ and $m(t)$ be the partitioned row-vector $(d^T(t) \,|\, 0)$, where the number of zeros augmented to $d^T(t)$ corresponds with the size of the identity augmented to $B - (1 + r(t))A(t)$. We use $M(t)$ and $m(t)$ to get a "standard form" linear program, meaning that the primal is stated with equality constraints, and this form is realized by including surplus variables in LP_{econ} (these variables correspond with the augmented identity). We investigate the dynamic linear programs,

$$LP_{\text{econ}}(t) \quad \min\{x^t l(t) : x^T M(t) = m(t), x \geq 0\} \quad \text{and}$$
$$LD_{\text{econ}}(t) \quad \max\{m(t)y : M(t)y + s = l(t), s \geq 0\}.$$

Notice that because every process may be run simultaneously to strictly satisfy demand, we have that there is a positive x such that $x^T M(t) = m(t)$. Hence, the strict interior of the feasible region of $LP_{\text{econ}}(t)$ is non-empty. Also, the fact that $l(t)$ is positive means that $(y, s) = (0, l(t))$ is in the strict interior of the feasible region of $LD_{\text{econ}}(t)$. So, the strict interior of feasible region of $LD_{\text{econ}}(t)$ is non-empty, and Assumption 1 is satisfied.

The following dynamic extension of the Nonsubstitution Theorem follows directly from Theorem 1.

Theorem 5 *Under Assumptions 2 and 3, the collection of optimal processes stabilizes.*

Comparing Theorem 4 to Theorem 5, we see that Theorem 5 only requires the addition of Assumption 2, which immediately follows if $A(t)$, $d(t)$, $l(t)$, and $r(t)$ are rational. While Theorem 5 is similar to the Nonsubstitution Theorem, it is different. First, our result is stronger in the sense that it allows changes not just in the demand, but also in the input matrix, the labor requirements, and the profit. However, the result we have is that the the collection of optimal processes becomes time-independent, not demand-independent. So, while we allow all the data to vary with respect to time, we do not get a result that is truly independent of all demands.

That Theorem 5 permits a dynamic profit is significant. This is because "the assumption of a given rate of profit radically transforms the substance of [neoclassical] theory" [20]. In fact, modern economists now understand that it is the assumption of a fixed profit that is the underlying support of the Nonsubstitution Theorem (this is because the concepts of "endowment" and "scarcity" are not allowed, see [20]). However, Theorem 5 does not assume a static profit, and hence, leads to the new economic question: Is it possible to economically explain that a dynamic profit can still lead to a stable set of optimal processes? The following Corollary shows that the collection of optimal processes is stable for all sufficiently small profits.

Corollary 1 *Suppose that for a fixed A, l, and d, the economy represented by system (9) - (14) is consistent for every profit $r \in [0, \bar{r}]$. Then, under Assumptions 2 and 3, there exists an $\hat{r} \in [0, \bar{r}]$, such that the collection of optimal processes is stable for all $r \in (0, \hat{r})$.*

Proof: Setting $r(t) = 1/t$, we see that the proof follows immediately from Theorem 5. ■

4.2 Allowing Joint Production

In this section we allow a process to produce multiple commodities. However, we do not remove Assumption 3, but rather replace it with the following Assumption.

Assumption 4 *We allow processes to produce multiple commodities, but only if there is a process for each commodity that produces* **only** *that commodity. Moreover, if process i produces commodities $j_1, j_2, \ldots j_k$, and processes $i_1, i_2, \ldots i_k$ each produce uniquely one of the commodities $j_1, j_2, \ldots j_k$, the commodity inputs for process i are the sums of the commodity inputs for processes $i_1, i_2, \ldots i_k$ —i.e. $a^i = a^{i_1} + a^{i_2} + \ldots a^{i_k}$.*

Assumption 4 allows processes that produce multiply commodities to be added to the economy, but it does not allow single production processes to be removed. The condition on the commodity inputs states that we are able to replace several processes with one process, but that the single process does not alter the input requirements to produce the commodities. As such, we are not allowed to introduce processes that more efficiently use their input commodities. However, we are allowed to introduce multiple output processes that make more efficient use of labor.

We use Assumption 4 to guarantee that Theorem 3 remains valid. The only place where Assumption 3 is used in the proof of Theorem 3 is in the last paragraph, where we show that the consistency of system (9) - (14) implies the consistency of system (15) - (20). Allowing processes to produce multiple commodities means that a technique σ need not have the quality that B_σ is the identity. Hence, we can not use Lemma 7 in the final paragraph of the proof of Theorem 3 to calculate x_σ —i.e. $B_\sigma - (1+r)A_\sigma$ is not necessarily monotonic.

Suppose that process i_0 produces commodities $j_1, j_2, \ldots, j_k$, and suppose that process i_0 is running in technique σ_0. From Assumption 4 we know that there are processes $i_1, i_2, \ldots, i_k$ such that each of these processes produces exactly one of the commodities $j_1, j_2, \ldots, j_k$. We also have from Assumption 4 that we can assign values to $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ such that $x_{i_0} = \sum_{\alpha=1}^{k} x_{i_\alpha}$ and

$$x_{i_0}[B_{\{i_0\}} - (1+r)A_{\{i_0\}}]p = x^T_{\{i_1,i_2,\ldots,i_k\}}[B_{\{i_1,i_2,\ldots,i_k\}} - (1+r)A_{\{i_1,i_2,\ldots,i_k\}}]p.$$

In other words, we can distribute the work load to the processes that only produce a single commodity. Consequently, if σ is a technique such that B_σ is not the identity, we may redistribute the work load to single commodity processes to form a technique σ', where $B_{\sigma'}$ is the identity. This means that Assumption 3 can be replaced by Assumption 4 in Theorem 3 to obtain the following result.

Theorem 6 *Under Assumption 4, system (9) - (14) is feasible if, and only if, the following primal and dual pair of linear programs is well-posed,*

$$LP_{econ} \quad \min\{x^T l : x^T[B-(1+r)A] \geq d^T, x \geq 0\} \quad \textit{and}$$
$$LD_{econ} \quad \max\{d^T y : [B-(1+r)A]y \leq l, y \geq 0\}.$$

Moreover, if x^* *and* y^* *are optimal for* (LP_{econ}) *and* (LD_{econ}), *then* $x = x^*$, $p = (1/dy^*)y^*$, *and* $w = 1/dy^*$ *are long period solutions to system (9) - (14).*

This leads to the following theorem and corollary, which are the first Nonsubstitution type results for a dynamic economy that allows joint production.

Theorem 7 *The collection of optimal processes stabilizes under Assumptions 2 and 4.*

Corollary 2 *Suppose that for a fixed A, l, and d, the economy represented by system (9) - (14) is consistent for every profit* $r \in [0, \bar{r}]$. *Then, under Assumptions 2 and 4, there is an* $\hat{r} \in [0, \bar{r}]$ *such that the collection of optimal processes is stable for all* $r \in (0, \hat{r})$.

5 Conclusions and Directions for Further Research

We have shown under mild conditions that the optimal partition for linear programming stabilizes under parameterization. This result allowed us to define an asymptotic analytic center solution, which we have shown inherits the analytic properties of $A(t)$ and $b(t)$. Furthermore, the existence of the asymptotic optimal partition implies significant extensions of the Nonsubstitution Theorem.

There are many avenues for future research.

- Whether or not there is a demand-independent optimal partition remains an open question.
- The authors of [6] have shown that there is an analytic center that is defined independent of the representation of the polytope. This center is called the *prime analytic center*, and it would be nice to know under what conditions one could define an asymptotic prime analytic center.

- Analytic centers can be defined for regions more complex than polytopes, as in the area of Semidefinite Programming. The difficulty lies in the fact that the optimal partition contains three sets, rather than two. How, and if, these results extend to these broader problems statements appears to be a challenging, yet potentially fruitful pursuit.
- The use of semimonotonic operators, meaning that $A^+ \geq 0$, might allow Theorem 6 to be stated under an assumption that is more general than Assumption 4. Such adjustments would lead to further economic extensions of Theorem 7.
- If the labor source is not homogeneous, the linear program LP_{econ} becomes a multiple objective linear program. An optimal partition for multiple objective linear programming is introduced in [12]. If this partition were shown to stabilize, one could allow non-homogeneous labor sources in the economic results of the last section.

Acknowledgments

The authors would like to thank Harvey Greenberg for originally suggesting the idea of an asymptotic optimal partition.

References

[1] I. Adler and R. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50:29–51, 1991.

[2] E. Altman, K. Avrachenkov, and J. Filar. Asymptotic linear programming and policy improvement for singularly perturbed markov decision processes. *Mathematical Methods of Operations Research*, 59(1):97–109, 1999.

[3] L. Bernard. A generalized inverse method for asymptotic linear programming. *Mathematical Programming*, 43(1):71–86, 1989.

[4] L. Bernard. An efficient basis update for asymptotic linear programming. *Linear Algebra and its Applications*, 184:83–102, 1993.

[5] S. Campbell and Jr. C. Meyer. *Generalized Inverses of Linear Transformations*. Fearon Pitman Publishers Inc., Belmont, CA, 1979.

[6] R. Caron, H. Greenberg, and A. Holder. Analytic centers and repelling inequalities. Technical Report CCM 142, Center for Computational Mathematics, University of Colorado at Denver, 1999. To appear in European Journal of Operations Research.

[7] A. Goldman and A. Tucker. Theory of linear programming. In H. Kuhn and A. Tucker, editors, *Linear Inequalities and Related Systems*, volume 38, pages 53–97. Princeton University Press, Princeton, New Jersey, 1956.

[8] H. Greenberg. *Mathematical Programming Glossary*. World Wide Web, `http://www-math.cudenver.edu/~hgreenbe/glossary/glossary.html`, 1996-2001.

[9] H. Greenberg. Matrix sensitivity analysis from an interior solution of a linear program. *INFORMS Journal on Computing*, 11(3):316–327, 1999.

[10] O. Güler. Limiting behavior of weighted central paths in linear programming. *Mathematical Programming*, 65:347–363, 1994.

[11] M. Halická. Analytical properties of the central path at boundary point in linear programming. *Mathematical Programming*, 84(2):229–245, 1999.

[12] A. Holder. Partitioning multiple objective solutions with applications in radiotherapy design. Technical Report 54, Trinity University Mathematics, 2001.

[13] A. Holder and R. Caron. Uniform bounds on the limiting and marginal derivatives of the analytic center solution over a set of normalized weights. *Operations Research Letters*, 29:49–54, 2000.

[14] A. Holder, J. Sturm, and S. Zhang. Marginal and parametric analysis of the central optimal solution. Technical Report No. 48, Trinity University Mathematics, 1999. To appear in Information Systems and Operational Research.

[15] R. Jeroslow. Asymptotic linear programming. *Operations Research*, 21:1128–1141, 1972.

[16] R. Jeroslow. Linear programs dependent on a single parameter. *Discrete Mathematics*, 6:119–140, 1973.

[17] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[18] L. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akademiia Nauk SSSR*, 244:1093–1096, 1979.

[19] K. Kuga. The non-substitution theorem: Multiple primary factors and the cost function approach. Technical Report Discussion Paper No. 529, The Institute of Social and Economic Research, Osaka Univeristy, Osaka, Japan, 2001.

[20] H. Kurz and N. Salvadori. *Theory of Production: A long-period analysis*. Cambridge University Press, New York, NY, 1995.

[21] L. McLinden. An analogue of Moreau's approximation theorem, with applications to the nonlinear complementarity problem. *Pacific Journal of Mathematics*, 88(1):101–161, 1980.

[22] J. Mirrlees. The dynamic nonsubstitution theorem. *Review of Economic Studies*, 36(105):67–76, 1969.

[23] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach.* John Wiley & Sons, New York, NY, 1997.

[24] P. Samuelson. Abstract of a theorem concerning substitutability in open leontief models. In T. Koopmans, editor, *Activity Analysis of Production and Allocation*, page Chapter 7. Wiley, New York, NY, 1951.

[25] G. Sonnevend. An "analytic centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In A. Prekopa, J. Szelezsan, and B. Strazicky, editors, *Lecture Notes in Control and Information Sciences*, volume 84, pages 866–875. Springer-Verlag, Heidelberg, Germany, 1986.

[26] C. Witzgall, P. Boggs, and P. Domich. On the convergence behavior of trajectories for linear programming. *Contemporary Mathematics*, 114:161–187, 1990.

[27] S. Wright. *Primal-Dual Interior-Point Methods.* SIAM, Philadelphia, PA, 1997.

[28] Y. Ye. *Interior Point Algorithms Theory and Analysis.* John Wiley & Sons, Inc., New York, NY, 1997.

[29] H. Ying. A canonical form for pencils of matrices with applications to asymptotic linear programs. *Linear Algebra and its Applications*, 234:97–123, 1996.

[30] G. Zhao and J. Zhu. Analytic properties of the central trajectory in interior point methods. In D. Du and J. Sun, editors, *Advances in Optimization and Approximation*, pages 362–375. Kluwer Academic Publishers, The Netherlands, 1994.