A THEORETICAL AND COMPUTATIONAL COMPARISON OF "EQUIVALENT" MIXED-INTEGER FORMULATIONS

by

R. R. Meyer

Computer Sciences Technical Report #311

January 1978

A THEORETICAL AND COMPUTATIONAL COMPARISON OF "EQUIVALENT" MIXED-INTEGER FORMULATIONS

by

R. R. Meyer†

Computer Sciences Technical Report #311

January 1978

†This research was sponsored by the National Science Foundation under contract MCS74-20584 A02.

ABSTRACT

This paper provides a theoretical and computational comparison of alternative mixed integer programming formulations for optimization problems involving certain types of economy-of-scale functions. Such functions arise in a broad range of applications from such diverse areas as vendor selection and communications network design. A "non-standard" problem formulation is shown to be superior in several respects to the traditional formulation of problems in this class.

1. Formulations: Equivalent and Optimal

This paper provides a theoretical and computational comparison of alternative mixed integer programming formulations for optimization problems involving certain types of economy-of-scale functions. Such functions arise in a broad range of applications from such diverse areas as vendor selection and communications network design. A "non-standard" problem formulation is shown to be superior in several respects to the traditional formulation of problems in this class.

This first section describes a rigorous approach to formulating certain optimization problems through the use of "minimization models" [4,5,6]. The minimization model concept is then used as the basis for defining a family of "equivalent" formulations as well as means of defining an "optimal" formulation. Sections 2 and 3 establish the optimality of a very compact formulation for functions belonging to a class of economy-of-scale functions. Computational results for a communications network problem are then given to illustrate the superiority of this formulation as compared to a "standard" formulation of the problem.

The economy-of-scale property that we will consider is encountered in a broad variety of cost functions for goods ranging from doughnuts to telecommunications links. Roughly speaking, a function is said to have an economy-of-scale property if the cost (per unit) of a commodity decreases if certain "large" quantities of the commodity are purchased. A simple example of such a cost functions, but one which serves to illustrate some of the properties that we wish to consider is a "cheaper-by-the-dozen" function defined as follows: let y_1 denote the number of single units of a commodity with the cost per single unit being a positive constant c_1; let y_2 denote the (non-negative, integer) number of dozens (groups of 12) purchased, the price per dozen being a positive constant $c_2 < 12c_1$ (so that it is cheaper to purchase a dozen than it is to purchase 12 single units) and let $k(x)$ (see Figure 1 for a typical $k(x)$) denote the "cheaper-by-the-dozen" function that represents the minimum cost of purchasing at least x units (for simplicity in this example, x and y_1 will be assumed to be continuous variables). It is easily seen that $k(x)$ can be compactly represented as:

$$
\begin{aligned}
k(x) \;=\; &\min_{y_1,y_2} && c_1y_1 + c_2y_2 \\
(1.1)\qquad &\text{s.t.} && y_1 + 12\cdot y_2 \geq x \\
& && y_1,y_2 \geq 0,\quad y_2 \text{ integer.}
\end{aligned}
$$

That is, substituting any constant $\bar{x}$ for x in the RHS of (1.1) yields an optimization problem (in the variables y_1 and y_2) whose optimal value is precisely $k(\bar{x})$. Of course, the piecewise-linear function $k(x)$ can be represented in many other ways, but, as will be seen, the representation (1.1) is not only compact but also is useful in formulating optimization problems involving $k(x)$.

The RHS of (1.1) is an example of a mixed-integer minimization model (MIMM), a concept that was described in [4,5,6]. To define this concept, suppose f is a function from $\mathbb{R}^1$ into $(-\infty,+\infty]$, and that the following equation holds for all x belonging to a subset S or $\mathbb{R}^l$:

$$
\text{(1.2)} \qquad
\begin{aligned}
f(x) = \min_{y} \quad & cy \\
\text{s.t.} \quad & Ay = b - \hat{A}x, \\
& y \geq 0 \text{ and } y_i \text{ integer for } i \in I,
\end{aligned}
$$

where I is a subset of $\{1,\ldots,n\}$, b is an element of $\mathbb{R}^m$, and c, A, and $\hat{A}$ are of dimensions $1 \times n$, $m \times n$, and $m \times l$ respectively. (We will assume that the optimization problem on the RHS on (1.2) has an optimal solution if its feasible set is non-empty, and that the "optimal value" is defined to be $+\infty$ if the feasible set is empty.) The expression on the RHS of (1.2) is said to be a MIMM for f on S (for our purposes, it is convenient to assume that S is convex, although for the general theoretical development of MIMM's given in [4], this is not necessary). As noted in [4], the utility of MIMM's arises in part from the fact that for any set $\tilde{S} \subseteq S$ the following two problems are equivalent:

$$
\text{(1.3)} \qquad
\begin{aligned}
\min_{x,z} \quad & f(x) + \hat{f}(x,z) \\
\text{s.t.} \quad & x \in \tilde{S},\ (x,z) \in T,
\end{aligned}
$$

and

$$
\text{(1.4)} \qquad
\begin{aligned}
\min_{x,y,z} \quad & cy + \hat{f}(x,z) \\
\text{s.t.} \quad & x \in \tilde{S},\ (x,z) \in T, \\
& Ay = b - \hat{A}x, \\
& y \geq 0 \text{ and } y_i \text{ integer for } i \in I.
\end{aligned}
$$

The problems (1.3) and (1.4) are equivalent in the sense that (1.3) has a feasible solution if and only if (1.4) has a feasible solution, and (x^*,z^*) is an optimal solution of (1.3) if and only if there exists a y^* such that (x^*,y^*,z^*) is an optimal solution of (1.4). From a computational point of view the transformation from (1.3) to (1.4) may allow the replacement of a piecewise-linear objective function term $f(x)$ by a linear objective function term cy. Thus, if an optimization problem has only linear constraints and objective function terms for which MIMM's exist, then this conversion procedure may be carried out term-by-term until the original problem has been transformed into a mixed integer program (MIP). Note, however, that although this MIP will be equivalent to the original problem, equivalence may be destroyed if the integrality constraints on the newly added variables are deleted, a relaxation which is usually the first step of an algorithm for the solution of an MIP. In particular, the relaxation of the integrality constraints of a MIMM will yield a parametrically defined family of problems (a linear programming minimization model (LPMM)) whose optimal value must be (see [4]) a convex function on all of $\mathbb{R}^1$. Thus, this relaxation will mean that a nonconvex objective function term of the original formulation is replaced by a convex approximation. In algebraic terms, defining

$$f^*(x) \equiv \min_y cy$$

$$\text{s.t. } Ay = b - \hat{A}x, \quad y \geq 0, \tag{1.5}$$

it follows that f^* is <u>convex</u> on $\mathbb{R}^1$, so that if f (as defined in (1.2)) is <u>nonconvex</u> on S (in the sense that there exist points x_1, x_2, $\bar{x} \in S$ and a $\lambda \in (0,1)$ such that $\bar{x} = \lambda x_1 + (1-\lambda)x_2$ and $f(\bar{x}) > \lambda f(x_1) + (1-\lambda)x_2)$, then f and f^* <u>cannot</u> coincide over all of S (in particular, they would not agree at $\bar{x}$). The difference $f(x) - f^*(x)$ (which is always non-negative because of the relaxation of the constraints) will be termed the <u>relaxation error</u> of the LPPMM at x.

In the case of the MIMM (1.1), for example, the optimal value function (for $x \geq 0$) for the LPMM obtained by relaxing the integrality constraints of (1.1) is easily seen to be the <u>linear</u> function $k^*(x) \equiv c_2 x/12$. The relaxation error in this particular case is thus the difference between the values $k(x)$ and $k^*(x)$ (see Figure 1). Note that this difference is <u>positive</u> unless x is an integer multiple of 12. (For $x < 0$, $k(x) = k^*(x) = 0$, but we are concerned here only with non-negative values of x.)

In comparing <u>alternative</u> MIMM formulations, a comparison of the behavior of the relaxation errors establishes the relative accuracy of the approximations used in the first step of the solution of the respective MIP's. Thus, if f^{**} is the optimal value function of the continuous relaxation of a <u>different</u> MIMM for f, and $f^*(x) \geq f^{**}(x)$ for all $x \in S$ (which we write as $f^* \overset{S}{\geq} f^{**}$), then the MIMM (1.2) may be considered to be at least as good (with respect to the relaxation error criterion as the MIMM from which f^{**} was derived. Moreover, if it can be established that the inequality $f^* \overset{S}{\geq} f^{**}$, holds for <u>all</u>

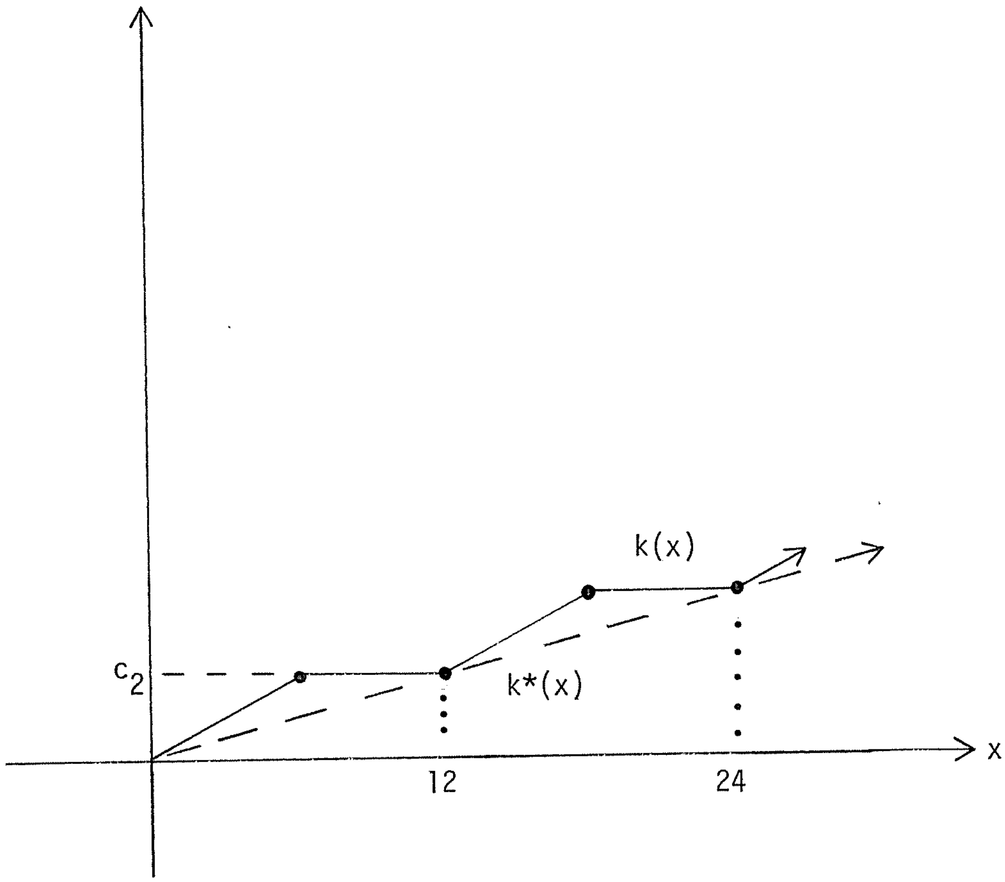

Figure 1: The functions k(x) and k*(x)

convex functions f^{**} satisfying $f \overset{S}{\geq} f^{**}$, then the MIMM giving rise to f^* will be <u>optimal</u> from the standpoint of error in a relaxation solution strategy, and will therefore be said to be <u>relaxation-optimal</u> on S. (As will be seen, a function may have more than one relaxation-optimal MIMM, so additional MIMM criteria also will be considered.) In order to more easily describe results of this type, it is convenient to introduce some additional terminology. If h is a function mapping a convex set T into $[-\infty,+\infty]$, the <u>convex envelope</u> of h on T (which may be thought of geometrically as the <u>largest</u> convex function <u>below</u> h on T), denoted by $c^*(h,x,T)$, is the function satisfying the relations:

(1.6) $\quad c^*(h,x,T) \leq h(x)$ for all $x \in T$,

(1.7) $\quad c^*(h,x,T)$ is convex on T,

(1.8) $\quad$ if $g(x) \leq h(x)$ for all $x \in T$ and g is convex on T, then $g(x) \leq c^*(h,x,T)$ for all $x \in T$.

(In places where reference to the variable is not needed, we will write $c^*(h,T)$ in place of $c^*(h,x,T)$.) Existence and uniqueness of $c^*(h,T)$ easily follow from the fact that the point-wise supremum of a family of convex functions is convex. Defining on T the set of functions

$$C(h,T) \equiv \{g \mid g \text{ is convex on } T,\ g \leq h\},$$

$c^*(h,T)$ is simply the supremum of $C(h,T)$. It might be noted that the domain T plays a very significant role in determining the convex envelope. That is, the value of the convex envelope at a particular point may be different for different choices of T. This aspect of the convex envelope will be taken up in Section 2.

The optimal value function of a LPMM, in addition to being convex, is also piecewise-linear (PL), and it is also convenient to introduce some terminology for piecewise-linear functions of a single variable, which are our principal concern in this paper.

We will say that a real-valued function h defined on a closed interval $[\alpha_0,\alpha_p] \subset \mathbb{R}^1$ is a <u>piecewise-linear function on $[\alpha_0,\alpha_p]$ with breakpoints</u> $\alpha_0 < \alpha_1 \ldots < \alpha_p$ if h is <u>affine</u> on each subinterval $[\alpha_i, \alpha_{i+1}]$ and $[h(\alpha_{i+1})-h(\alpha_i)]/(\alpha_{i+1}-\alpha_i) \neq [h(\alpha_i)-h(\alpha_{i-1})]/(\alpha_i-\alpha_{i-1})$ for $i = 1,\ldots,p-1$ (that is, the slope to the left of α_i differs from the slope to the right of α_i).

The basic result that will be used to establish that certain formulations yield convex envelopes is the sufficiency part of the following theorem:

<u>Theorem 1.1</u>: Let g be a lower semi-continuous (l.s.c.) function mapping $[\alpha_0,\alpha_p]$ into $(-\infty,+\infty]$ with $g(\alpha_i) < +\infty$ for $i = 0,\ldots,p$.

Let g^* be a convex piecewise-linear function on $[\alpha_0,\alpha_p]$ with breakpoints $\alpha_0 < \alpha_1 < \ldots \alpha_p$, and let $g^*(x) \leq g(x)$ for $x \in [\alpha_0,\alpha_p]$. A necessary and sufficient condition for g^* to be the convex envelope of g on $[\alpha_0,\alpha_p]$ is that $g^*(\alpha_i) = g(\alpha_i)$ for $i = 0,\ldots,p$.

<u>Proof</u>: To establish <u>sufficiency</u>, suppose that $\tilde{g} \in C(g,[\alpha_0,\alpha_p])$. Then for any $\bar{x} \in [\alpha_0,\alpha_p]$ there exists at least one pair α_i, α_{i+1} of breakpoints such that $\bar{x} \in [\alpha_i,\alpha_{i+1}]$. Choosing $\lambda \in [0,1]$ such that $\bar{x} = \lambda\alpha_i + (1-\lambda)\alpha_{i+1}$, we have (using the convexity of $\tilde{g}$):
$\tilde{g}(\bar{x}) \leq \lambda\tilde{g}(\alpha_i) + (1-\lambda)\ \tilde{g}(\alpha_{i+1}) \leq$
$\lambda g(\alpha_i) + (1-\lambda)\ g(\alpha_{i+1}) = \lambda g^*(\alpha_i) + (1-\lambda)\ g^*(\alpha_{i+1}) = g^*(\bar{x})$. Thus, $\tilde{g}(x) \leq g^*(x)$ for any $x \in [\alpha_0,\alpha_p]$ establishing that $g^* = c^*(g,[\alpha_0,\alpha_p])$.

To show necessity, suppose that $g(\alpha_0) - g^*(\alpha_0) \equiv \varepsilon_0 > 0$. Since g is l.s.c. and g^* is upper semi-continuous, there exists a $\delta_0 \in (0,\alpha_0)$ such that $\alpha_0 \leq x \leq \alpha_0 + \delta_0$ implies $g(x) \geq g(\alpha_0) - \varepsilon_0/2$ and $g^*(x) \leq g^*(\alpha_0) + \varepsilon_0/2 = g(\alpha_0) - \varepsilon_0/2$.

Now consider the PL function $\tilde{g}$ (see Figure 2) with breakpoints at $\alpha_0, \alpha_0 + \delta_0, \alpha_1, \ldots, \alpha_p$ and function values $\tilde{g}(\alpha_0) = g(\alpha_0) - \varepsilon_0/2$, $g(\alpha_0+\delta_0) = g^*(\alpha_0+\delta_0)$, $\tilde{g}(\alpha_i) = g^*(\alpha_i)$ $(i = 1,\ldots,p)$. Note that $\tilde{g}(\alpha_0) > g^*(\alpha_0)$, but that $\tilde{g}(x) = g^*(x)$ for $x \in [\alpha_0+\delta_0,\alpha_m]$ and that $\tilde{g}$ is a convex function on $[\alpha_0,\alpha_m]$. Finally, the relations
$\tilde{g}(\alpha_0) = g(\alpha_0) - \varepsilon_0/2$ and $\tilde{g}(\alpha_0+\delta_0) = g^*(\alpha_0+\delta_0) \leq g^*(\alpha_0) + \varepsilon_0/2$
$= g(\alpha_0) - \varepsilon_0/2$ imply $\tilde{g}(x) \leq g(\alpha_0) - \varepsilon_0/2$ for $x \in [\alpha_0,\alpha_0+\delta]$, so that $\tilde{g}(x) \leq g(x)$ for $x \in [\alpha_0,\alpha_p]$. Thus, $\tilde{g}(x) \in C(g(x), [\alpha_0,\alpha_p])$ and $\tilde{g}(\alpha_0) > g^*(\alpha_0)$, contradicting the hypothesis that $g^*(x) = c^*(g,x,\ [\alpha_0,\alpha_p])$. A contradiction may be similarly obtained if $g(\alpha_p) > g^*(\alpha_p)$. For an interior breakpoint α_i, the construction of a suitable $\tilde{g}$ is similar (see Figure 3), except that the breakpoints of $\tilde{g}$ (where it coincides with g^*) are taken to be

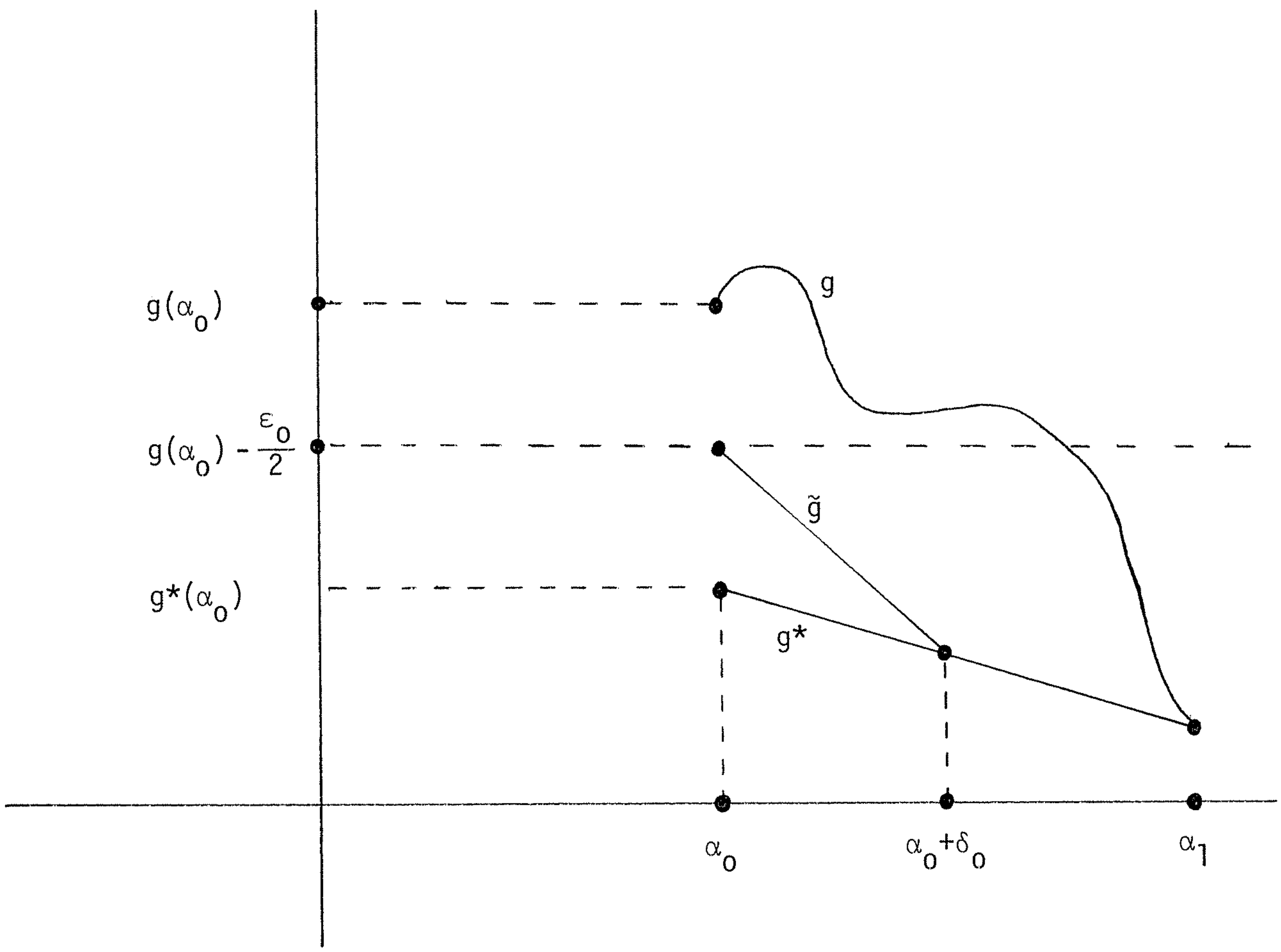

Figure 2. The case: $g^*(\alpha_o) < g(\alpha_o)$.

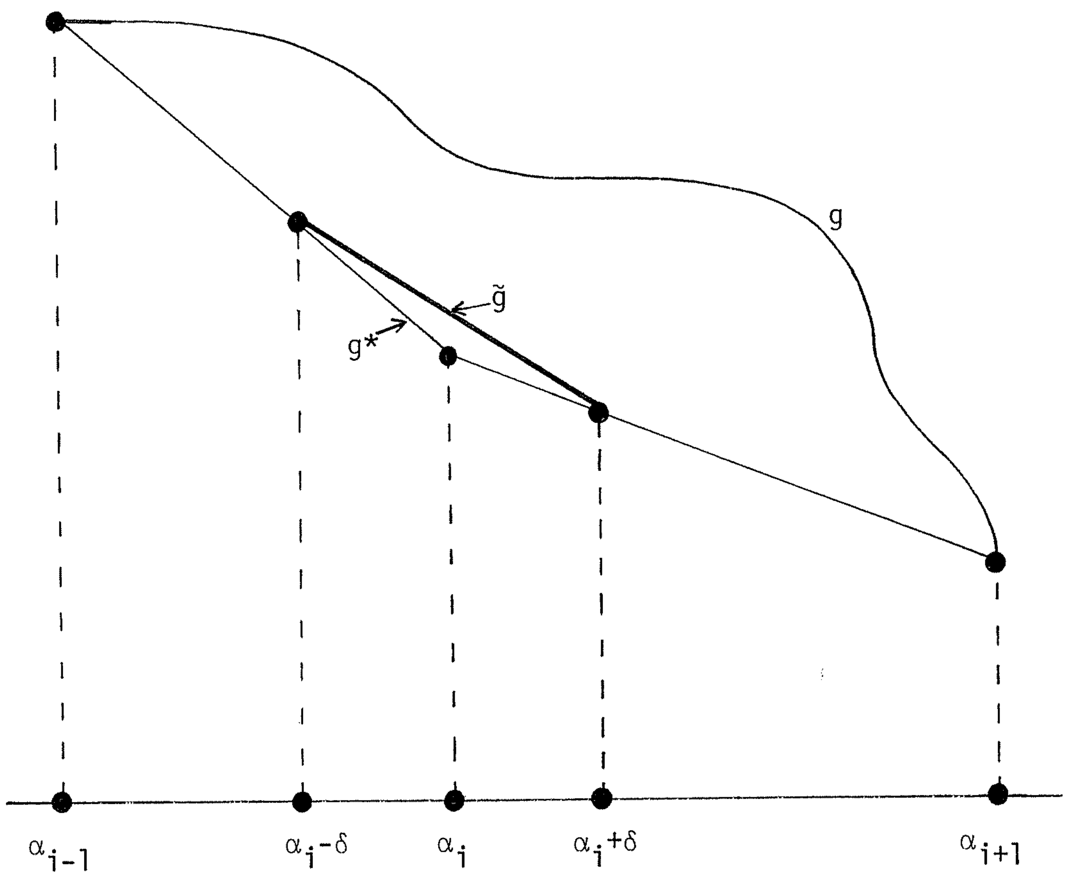

Figure 3. The case: $g^*(\alpha_i) < g(\alpha_i)$, where $0 < i < p$.

$\alpha_0,\ldots,\alpha_{i-1}$, $\alpha_i-\delta_i$, $\alpha_i + \delta_i,\ldots,\alpha_p$, where $0 < \delta_i < \min\{\alpha_i-\alpha_{i-1}, \alpha_{i+1}-\alpha_i\}$ is chosen so that, defining $\varepsilon_i \equiv g(\alpha_i) - g^*(\alpha_i) > 0$, we have, for $x \in [\alpha_i-\delta_i, \alpha_i+\delta_i]$, the inequalities $g(x) \geq g(\alpha_i) - \varepsilon_i/2$ $g^*(x) \leq g^*(\alpha_i) + \varepsilon_i/2$. Because of the change in slope at breakpoints, it may be verified that $g^*(\alpha_i) < \tilde{g}(\alpha_i)$, and thus a contradiction may be obtained. ∎

Note that for sufficiency, lower semi-continuity of g is <u>not</u> required. In this paper we are primarily concerned with the sufficiency part of this theorem, but it should be noted that in [4] the lower semi-continuity of optimal value functions of MIMM's was established under rationality assumptions on the coefficients of the MIMM.

It might also be noted that the argument used in the proof can be used to show that g does <u>not</u> have a PL convex envelope if $g(\alpha_0) = +\infty$ or $g(\alpha_p) = +\infty$, since this would mean that $g^*(\alpha_0) < g(\alpha_0)$ or $g^*(\alpha_p) < g(\alpha_p)$ for any PL function g^*. On the other hand, a PL convex envelope may exist if there are <u>interior</u> points $\bar{x}$ of $[\alpha_0,\alpha_p]$ with the property that $g(\bar{x}) = +\infty$. This allows the domain of g to have "gaps" on which g may be thought of as being $+\infty$. Such gaps often occur in optimal value functions of MIMM's.

From Figure 1, one might conjecture that k^* is the convex envelope of k on R^1_+. This is indeed true, and in Section 2 we will use the approach of Theorem 1.1 to establish a more general result from which this follows as a special case.

2. The Unbounded Case

In this section we will consider MIMM's for a broad class of economy-of-scale functions that includes the economy-of-scale function $k(x)$ of the previous section. Specifically, we will develop relaxation-optimal MIMM's for the class of functions whose elements may be represented as optimal value functions of the following type:

$$
\begin{aligned}
(2.1) \qquad f_1(x) \equiv \ & \min_y \ cy \\
& \text{s.t. } ay \geq x \\
& \qquad y \geq 0, \quad y_i \text{ integer for } i \in I,
\end{aligned}
$$

where $c = (c_1,\dots,c_n) \geq 0$, $a = (a_1,\dots,a_n) > 0$, and I is a subset of $\{1,\dots,n\}$.

(The case in which there are $c_i = 0$ is not of economic interest, but is included for mathematical completeness. The sign restrictions on c and a do serve to guarantee the existence of an optimal solution for all x, but, as shown in Appendix A, could be replaced by this hypothesis. In the next section, where <u>bounds</u> on the y_i are assumed, it will be seen that these sign restrictions have greater significance.) Note that the class of functions representable in the form (2.1) includes fixed-charge functions and economy-of-scale functions allowing several <u>different</u> volume discounts (as opposed to only one in the case of $k(x)$). (The computational results in Section 5 deal with an example in which $n = 3$.) For notational convenience we will assume that the variables have been ordered so that

(2.2) $$c_1/a_1 \equiv r_1 \leq c_2/a_2 \equiv r_2 \leq \cdots \leq c_n/a_n \equiv r_n \ .$$

From a cost viewpoint, this means that, on a per unit basis, the most "economical" purchase quantity is a_1, the next most economical is a_2, etc., and the right-hand side x represents the minimum amount to be purchased.

Consider the continuous relaxation of the MIMM in (2.1), which has the optimal value function defined by

(2.3) $$f_1^*(x) \equiv \min_y \ cy$$
$$\text{s.t. } ay \geq x \ , \quad y \geq 0 \ .$$

The following lemma states that f_1^* is linear on R_+^1, and provides the basis for a proof of the relaxation-optimality of the MIMM on the RHS of (2.1).

Lemma 2.1: For $x \in \mathbb{R}_+^1$, $f_1^*(x) = r_1 \cdot x$.

Proof: Note that, for any $x \geq 0$, the dual of (2.3) may be written as

(2.4) $$\max_v \ vx$$
$$\text{s.t. } va \leq c \ , \quad v \geq 0 \ .$$

By setting $y_1^* = x/a_1$ and $y_2^* = y_3^* = \ldots = y_n^* = 0$ and $v^* = r_1$, we obtain primal and dual feasible solutions with common objective function value $r_1 x$. This is thus the optimal value, $f_1^*(x)$. ∎

Having obtained a closed form representation of $f_1^*(x)$, the relationship between f_1 and f_1^* is easily established.

Theorem 2.2: The following relations hold between f_1 and f_1^*:

(2.5) $$f_1(x) = f_1^*(x) \quad \text{for} \quad x = k \cdot a_1 \quad (k=0,1,\ldots)$$

(2.6) $$f_1^* = c^*(f_1, \mathbb{R}_+^1).$$

Proof: Since $f_1^*(x) \leq f_1(x)$ for $x \in \mathbb{R}_+^1$, (2.5) may be established by showing that, for $x = k \cdot a_1$ $(k = 0,1,2,\ldots)$, (2.1) has a feasible solution with objective function value $f_1^*(ka_1) = r_1 \cdot ka_1 = kc_n$. Such a feasible solution is obtained by setting $y_1 = k$ and $y_2 = y_3 = \ldots = y_n = 0$.

To prove (2.6), it suffices to show that for any $x' \in \mathbb{R}_+^1$, there exist $x_1', x_2' \in \mathbb{R}_+^1$ such that, for some $\lambda' \in [0,1]$, we have $\lambda' x_1' + (1-\lambda') x_2' = x'$ and $f_1^*(x') = \lambda' f_1(x_1') + (1-\lambda') f_1(x_2')$, since any $f \in c^*(f_1, \mathbb{R}_+^1)$ must satisfy $f(x') \leq \lambda' f_1(x_1') + (1-\lambda') f_1(x_2')$. These quantities are obtained by taking $x_1' = 0$, $x_2' = k \cdot a_1$, where k is an integer chosen such that $ka_1 \geq x'$, and λ' such that $(1-\lambda')ka_1 = x'$. Then $\lambda' f_1(x_1') + (1-\lambda') f_1(x_2') = 0 + (1-\lambda') r_1 ka_1 = r_1 x' = f_1^*(x')$. ∎

It is of some mathematical interest to note that the constraint $ay \geq x$ in (2.3) is satisfied as an equality by an optimal solution of (2.3). The observation may be used to establish that f_1^* is also the convex envelope on $\mathbb{R}_+^1$ of the optimal value function in the corresponding <u>equality-constrained</u> case:

$$
\text{(2.8)} \qquad \begin{aligned} f_1^e(x) \equiv \min_y \ & cy \\ \text{s.t. } & ay = x \\ & y \geq 0 , \quad y_i \text{ integer for } i \in I . \end{aligned}
$$

This result follows since $f_1^e(x) = f_1^*(x)$ for $x = k \cdot a_1$ $(k=0,1,\ldots)$. Since f_1^* may be written in the form (2.3) with the constraint $ay \geq x$ replaced by $ay = x$, it follows by the analog of Theorem 2.2 that the modified MIMM is relaxation-optimal in the equality-constrained case (2.8) as well.

On the other hand, it is not always possible to establish relaxation-optimality if a positive <u>constant</u> is added to the RHS of the constraint with RHS x in (2.1) (negative constants pose no difficulty, as we will show in Section 3.) An example illustrating the difficulties that may arise in this case is given in Appendix B. However, it is possible to extend the results of this section to the case in which non-negative <u>bounds</u> are imposed on the variables. This case is taken up in Section 3.

Finally, in the case that the a_i are all <u>rational</u>, Theorem 1.1 is a special case of a result of Blair and Jeroslow [3], who considered a system of constraints and showed that the convex envelope of the optimal value function of the MIMM (for $x \in \mathbb{R}^n$)

$$(2.7) \qquad \begin{aligned} &\min_y \ cy \\ &\text{s.t. } Ay \geq x,\ y \geq 0,\ \ y_i \ \text{ integer for } \ i \in I\ , \end{aligned}$$

coincides with the optimal value function of the continuous relaxation of the MIMM. The thrust of the next section can thus be viewed as an extension of this result to certain cases in which <u>non-zero constants</u> are allowed in the constraints of (2.7). (In general the Blair-Jeroslow result does <u>not</u> extend to the non-homogeneous case, as may be ascertained from the examples in Section B.)

3. Bounds on y

For most integer programming codes, it is necessary to have bounds on the integer variables. If the range of the y variables in (2.3) is restricted by the imposition of bounds, then the corresponding optimal value function on $\mathbb{R}^1_+$ is piecewise-linear (where it is finite), but the relaxation-optimality property of Section 2 may nonetheless be extended to this case. We first consider the case of upper bounds, and then the case of upper and lower bounds. As in Section 2 we assume that $c \geq 0$ and $a > 0$. (By making some obvious extensions, the constraint $a > 0$ may be removed, but as may be seen from an example in Appendix B, sign restrictions on c are needed in the bounded case to guarantee relaxation-optimality.)

Specifically, instead of the MIMM in (2.1) we first consider

$$
\begin{aligned}
f_2(x) = \min_y \quad & cy \\
\text{s.t.} \quad & ay \geq x \\
& 0 \leq y \leq u \\
& y_i \ \text{integer}, \quad i \in I
\end{aligned}
\tag{3.1}
$$

where $c \geq 0$, $a > 0$, the ordering assumption (2.2) is assumed to be satisfied, and the u_i are non-negative constants with u_i integer for $i \in I$. To prove relaxation-optimality we will show that the convex envelope of f_2 on $D \equiv [0,au]$, denoted by $c^*(f_2,D)$, is given by the optimal value function of the continuous relaxation:

$$
(3.2) \qquad \begin{aligned} f_2^*(x) \equiv \min_y \quad & cy \\ \text{s.t.} \quad & ay \geq x \\ & 0 \leq y \leq u \quad . \end{aligned}
$$

(We are not concerned with $x > au$ since $f_2(x) = f_2^*(x) = +\infty$ for such x.)

For notational convenience in stating a closed form expression for $f_2^*(x)$, we make the following definitions:

$$
b_j \equiv \sum_{i=1}^{j} a_i u_i , \quad d_j \equiv \sum_{i=1}^{j} c_i u_i \quad (j=0,\ldots,n),
$$

where it is understood that $b_0 = 0$ and $d_0 = 0$.

The following is the analog of Lemma 2.1:

<u>Lemma 3.1</u>: $f_2^*(x) = r_j(x-b_j) + d_j$ for $b_j \leq x \leq b_{j+1}$

$(j=0,\ldots,n-1)$

<u>Proof</u>: The proof is analogous to that of Lemma 2.1. For any x, the dual of (2.8) is given by

$$
\begin{aligned} \max_{v,w} \quad & vx - wu \\ \text{s.t.} \quad & va - w \leq c, \quad v \geq 0, \; w \geq 0 \quad . \end{aligned}
$$

In addition, for any $x \in D$, the optimal solutions of the primal and dual problems are as follows: if $b_j \leq x \leq b_{j+1}$, set $y_i^* = u_i$ for $i \leq j$, set $y_i^* = 0$ for $i > j+1$, and choose y_{j+1}^* such that $ay^* = x$; set $v^* = r_j$, $w_i^* = r_j a_i - c_i$ for $i < j$, and $w_i^* = 0$ for $i \geq j$. ■

Note from Lemma 3.1 that the breakpoints of f_2^* are contained in the set $\{b_0,\ldots,b_n\}$. By applying Theorem 1.1, we can obtain the following analog of Theorem 2.2:

Theorem 3.2: The following relationships hold between f_2 and f_2^*:

(3.3) $$f_2(\bar{x}) = f_2^*(\bar{x}) \quad \text{if} \quad \bar{x} = b_j \quad (j=0,\ldots,n),$$

(3.4) $$f_2^* = c^*(f_2,D) .$$

Proof: The relation (3.3) follows from considering the feasible solution and $y_i^* = u_i$ for $i \leq j$ and $y_i^* = 0$ for $i > j$. The relation (3.4) then follows directly from (3.3) and Theorem 1.1. ■

In a branch-and-bound algorithm in which the y_i are used as the branching variables, the formulation (3.1) has the additional very nice property of yielding a relaxation-optimal formulation at each node in the tree, since relaxation-optimality is not affected by the imposition of additional integer upper and lower bounds on the y_i in (3.1). This is because introduction of non-negative lower bounds is equivalent to the addition of a negative constant to the RHS of the constraint $ay \geq x$. Since a constraint of the form $ay > x - \gamma$, where $\gamma \geq 0$ implies an optimal value of 0 for $x \in [0,\gamma]$ in both the corresponding MIMM and its relaxation, it is easily shown that a translation of variables leads to the following result (see Appendix C for details):

Corollary 3.3: For $x \geq 0$, let

$$
\begin{aligned}
f_3(x) \equiv \min_y \quad & cy \\
\text{s.t.} \quad & ay \geq x \\
& \ell \leq y \leq u' \\
& y_i \text{ integer}, \quad i \in I,
\end{aligned}
\tag{3.3}
$$

where $\ell \geq 0$ and ℓ_i and u'_i are integer for $i \in I$; then the MIMM in (3.3) is relaxation-optimal on any interval $[\alpha, au']$, where $\alpha \in [0, a\ell]$.

In the next two sections we will compare these results to a "standard" approach to formulation that yields relaxation-optimal MIMM's for quite general piecewise-linear functions.

4. An Alternate Approach

A standard and quite general approach to modelling continuous piecewise-linear non-convex functions is to employ the so-called "λ formulation" of separable programming with the additional restrictions that at most two λ_i are allowed to be positive and that these must be "consecutive". We will see that, while this approach also yields relaxation-optimal models, it can, in contrast to the approach of Section 3, lead to computational difficulties in the absence of special provisions for handling the variables.

Assume that $\hat{f}$ is a piecewise-linear function on $[\alpha_0,\alpha_p]$ with breakpoints $\alpha_0 < \alpha_1 <\ldots< \alpha_p$. (It is possible to deal with l.s.c. "piecewise-linear" functions by a slightly different formulation technique (see [4]), but, aside from the need for more complex notation, the results are essentially the same.) Consider the following MIMM for $\hat{f}$:

$$
\begin{aligned}
\hat{f}(x) \quad = \quad \min_{\lambda_i,\delta_i} \quad & \sum_{i=0}^{p} \hat{f}(\alpha_i)\lambda_i \\
\text{s.t.} \quad & \sum_{i=0}^{p} \alpha_i\lambda_i = x \\
& \sum_{i=0}^{p} \lambda_i = 1, \quad \lambda_i \geq 0 \ (i=1,\ldots,p) \\
& \lambda_0 \leq \delta_1 \\
& \lambda_1 \leq \delta_1 + \delta_2 \\
& \quad \vdots \\
& \lambda_{p-1} \leq \delta_{p-2} + \delta_{p-1} + \delta_p \\
& \lambda_p \leq \delta_p \\
& \sum_{i=1}^{p} \delta_i = 1, \quad \delta_i \geq 0 \ \text{ and integer } (i=1,\ldots,p)
\end{aligned}
$$

and let $\hat{f}^*$ denote the optimal value function corresponding to the continuous relaxation of the RHS of (4.1). Note that $\hat{f}^* \in C(f,[\alpha_o,\alpha_p])$.

Theorem 4.1: The MIMM on the RHS of (4.1) is relaxation-optimal on $[\alpha_o,\alpha_p]$.

Proof: Let $\bar{x} \in [\alpha_o,\alpha_p]$ and let $\bar{\lambda}_i$ be chosen so that $\hat{f}^*(\bar{x})$ is obtained by setting $\lambda_i = \bar{\lambda}_i$ in the corresponding LPMM, so that $\hat{f}^*(\bar{x}) = \sum f(\alpha_i)\cdot\bar{\lambda}_i \geq \sum c^*(\hat{f},\alpha_i,[\alpha_o,\alpha_p])\cdot\bar{\lambda}_i \geq c^*(\hat{f},\bar{x},[\alpha_o,\alpha_p])$. Since $\hat{f}^* \in C(\hat{f},[\alpha_o,\alpha_p])$, this implies that $\hat{f}^*(\bar{x}) = c^*(\hat{f},\bar{x},[\alpha_o,\alpha_p])$ and the conclusion follows. ∎

While Theorem 4.1 implies that the standard MIMM will also be relaxation-optimal for a continuous economy-of-scale function in the class considered in Section 3, the MIMM (4.1) has several computational disadvantages. One obvious disadvantage is its sheer size, since the number of constraints and variables in (4.1) is determined by the number of breakpoints of $\hat{f}$, whereas this is not the case for the formulations of Sections 2 and 3. A more subtle disadvantage is the failure of the integer variables δ_i of (4.1) to directly reflect physical quantities. In particular, the δ_i all have cost coefficients of 0 and, moreover, a 0 "branch" on a δ_i has no effect on the allowable range of x values unless it has the largest or smallest index of any δ_i not yet fixed. While these disadvantages may be alleviated via the use of "Special Ordered Set" (SOS) strategies for branching (see [1]), such strategies are often not available in MIP codes (see [3]). In particular, SOS strategies are not fully implemented on the Univac FMPS-MIP code in use at the Madison Academic Computing Center, and in the next section we compare results obtained with FMPS and the formulation approaches of Sections 3 and 4. (It should be noted that the use of an SOS strategy has the advantage of imposing disjoint upper and lower bounds on the range of the variable x in (4.1) when SOS branching is performed. Branching on the y_i in (3.1) imposes upper bounds on x, but does not directly impose lower bounds. Lower bounds on the range of x may be directly imposed by adding to (3.1) constraints of the form

$$x \geq ay - \breve{a}z,$$

plus additional constraints of the form $z_i \leq y_i$. By selecting the coefficients a to reflect maximum "surpluses" so that for any $\bar{x} \in [0,au]$, a $\bar{y}$ yielding an optimal solution to (3.1) for $x = \bar{x}$ will satisfy $\bar{x} \geq a\bar{y} - \tilde{a}\bar{z}$ for some feasible $\bar{z}$, relaxation optimality will be preserved. This follows easily from the fact that, by assumption, the optimal value function of the MIMM remains $f_2(x)$, while the optimal value of the continuous relaxation, which cannot increase beyond $c^*(f_2,[0,au])$ (in spite of the added constraint) must also remain the same. Some theoretical and computational aspects of such lower bound constraints as well as some other modelling refinements to deal with upper bounds on x are currently under investigation.)

5. A Computational Comparison

In this section we consider a comparison of solution times for different formulations of the following communications network problems: determine the minimum cost networks (see Table 1) that meet specified demands (see Table 2) between six distinct pairs of cities (A,B), (A,C), (A,D), (B,C), (B,C), and (C,D), where the communications traffic between the elements of a city-pair may be routed via any acyclic path between the cities (there are 5 such routes between each city-pair).

Arc	Single Channel	12 Channels	60 Channels
A-B	789.75	7028.77	17690.40
B-C	878.25	7992.07	21341.47
C-D	1407.70	13232.38	42512.54
D-A	654.90	5697.63	13098.00
D-B	1045.60	9619.52	28022.08
C-A	1236.57	11500.10	35860.53

Table 1. Costs

City-Pair	Demand Set I	Demand Set II
A-B	2	4
B-C	10	10
C-D	46	64
D-A	5	5
D-B	2	10
C-A	4	14

Table 2. Two Sets of Communications Demands

Algebraically, this problem has the form:

$$\min_{x,z} \sum_{i=1}^{6} h_i(x_i)$$

$$\text{s.t.} \sum_{j=1}^{5} z_{j,k} = d_k \quad (k=1,\ldots,6)$$

$$\sum_{(j,k)\in A_i} z_{j,k} = x_i \quad (i=1,\ldots,6)$$

$$x_i,\ z_{j,k} \geq 0\ ,$$

where $z_{j,k}$ represents the number of channels on the j^{th} path between the k^{th} city-pair, d_k is the total number of channels needed by the k^{th} city-pair, A_i is the set of pairs (j,k) such that the corresponding path uses arc i, x_i is the total number of channels on arc i, and $h_i(x_i)$ is the minimum cost of leasing at least x_i channels on arc i. (Note that the h_i are economy-of-scale functions of the type considered in Sections 2 and 3 with n = 3. For computational convenience the variables associated with single channels on arcs were assumed continuous. Because of the fixed demands, bounds could be imposed on all variables. General integer variables were decomposed into 0-1 variables, since the FMPS-MIP code requires this.)

The computational results of Table 3 illustrate the dramatic difference in solution behavior and times between the formulation approaches of Sections 3 and 4. The MIP code used was the Univac FMPS-MIP code (level 7R1) and the problems were run on the Madison Academic Computing Center Univac 1110. For demand set I, the Section 3 formulation requires only about 1/4 the computer time of the Section 4 formulation. For demand set II, the solution time for the Section 3 formulation is 15 seconds, whereas the FMPS system was unable to solve the Section 4 formulation. Similar behavior was observed in runs using a locally-developed MIP code, IPMIXD, which successfully solved both I-S and II-S, but failed to solve either I-L or II-L because of storage overflows.

Problem	Rows	Columns	0-1 Variables	Solution Time (Sec.)
I-S*	12	54	18	4
I-L**	76	122	40	15
II-S	12	60	24	15
II-L	116	202	80	***

* I denotes demand set I; S denotes "short" formulation
** L denotes "long" standard formulation
*** FMPS system forced termination of run with message "numerical errors"

Table 3. Problem Sizes and Solution Times

A number of other versions of the problems were run in which some of the cost function terms were modelled via the Section 3 approach and the remainder via the Section 4 approach. In all cases the results were worse then those obtained via the Section 3 approach.

6. Conclusion

For piecewise-linear functions belonging to a broad class of economy-of-scale functions, a compact mixed-integer-programming formulation has been described. This formulation was then shown to behave at least as well as any other mixed-integer formulation of the function in terms of the approximation error resulting from the relaxation of integrality constraints. Moreover, a computational comparison (using a communications network problem as a test problem) showed the superiority of the compact formulation over a standard mixed-integer formulation of the same problem.

Acknowledgement

Jay M. Fleisher of the Madison Academic Computing Center assisted in the development of the test problems and obtained the computational results of Section 5.

References

1. E. M. L. Beale and J. H. H. Forest, "Global Optimization Using Special Ordered Sets", Math Prog., 10 (1976), pp. 52-69.

2. C. E. Blair and R. G. Jeroslow, "The Value Function of a Mixed Integer Program: II", Management Science Research Report 377, GSIA, Carnegie-Mellon University, December 1976.

3. A Land and S. Powell, "Computer Codes for Problems of Integer Programming", to appear in the proceedings of the conference, Discrete Optimization, 1977, held at the University of British Columbia, Vancouver, August, 1977.

4. R. R. Meyer, "Integer and Mixed-Integer Programming Models: General Properties", J. Opt. Theory and Appl., 16 (1975), pp. 191-206.

5. R. R. Meyer and M. V. Thakkar, "Rational Mixed-Integer Minimization Models", Mathematics Research Center Rpt. #1552, University of Wisconsin-Madison, 1976.

6. R. R. Meyer, "Mixed Integer Minimization Models for Piecewise-Linear Functions of a Single Variable", Discrete Math, 16 (1976), pp. 163-171.

Appendix A

To justify the statement in Section 2 that the restrictions $a > 0$ and $c \geq 0$ can be replaced by assuming that (2.1) has an optimal solution for $x \geq 0$, we consider the remaining cases: (1) i such that $c_i \geq 0$ and $a_i \leq 0$ (2) i such that $c_i < 0$ and $a_i > 0$, and (3) i such that $c_i < 0$ and $a_i < 0$.

Case 1: For those i such that $c_i \geq 0$ and $a_i \leq 0$, one may obtain an equivalent problem by deleting the corresponding variables y_i from the problem, since, for any $x \geq 0$, an optimal solution may be obtained in which such $y_i = 0$.

Case 2: If there are i such that $c_i < 0$ and $a_i > 0$, then clearly the objective function of (2.1) must be unbounded from below, so this case is ruled out by the existence of an optimal solution.

Case 3: If, for some i, $c_i < 0$ and $a_i < 0$, then either all $a \leq 0$, in which case (2.1) is infeasible for $x > 0$, or there exists at least one j such that $a_j > 0$. In the latter case let $r^+ \equiv \min \{ {}^{c_k}/a_k | a_k > 0\}$ and $r^- \equiv \max \{ {}^{c_k}/a_k | c_k < 0, a_k < 0\}$. If $r^- \leq r^+$, then, assuming that the variables are ordered so that $a_1 > 0$ and ${}^{c_1}/a_1 = r^+$, it may be seen from obvious extensions of the proofs of Lemma 2.1 and Theorem 2.2 that the desired result holds. On the other hand, if $r^- > r^+$, then the objective function of (2.1) is unbounded from below for all x. This follows by letting $r^+ = {}^{c_1}/a_1$ and $r^- = {}^{c_p}/a_p$, noting that ${}^{c_1}/{-c_p} < {}^{a_1}/{-a_p}$, and choosing a rational $\theta > 0$ such that ${}^{c_1}/{-c_p} < \theta < {}^{a_1}/{-a_p}$, from which it follows that $a_1 + a_p\theta > 0$ and $c_1 + c_p\theta < 0$. Now choose an integer $M > 0$ such that $M\theta$ is integer and note that the relations $a_1 \cdot M + a_p \cdot M\theta > 0$ and $c_1 M + c_p \cdot M\theta < 0$ imply unboundedness.

Appendix B

Here we consider several examples to illustrate the difficulties that can arise when one attempts to extend the results of Sections 2 and 3 by either (1) inserting a positive constant on the RHS of the constraint involving x, or (2) relaxing sign restrictions in the bounded case, or (3) allowing more than constraint involving x in the bounded case.

The following illustrates the difficulties that may arise when a positive constant appears in the RHS of a MIMM (see Figure 4):

$$\begin{aligned} k_1(x) \equiv \min_y \ & y_1 + 10y_2 \\ \text{s.t. } & y_1 + 12y_2 \geq x + 10 \\ & y_1, y_2 \geq 0, \ y_2 \text{ integer.} \end{aligned}$$

In this case, the convex envelope of $k_1(x)$ on R^1_+ is easily seen to have a value of 10 on [0,2], so that it does not coincide at $x = 0$ with the optimal value function of the continuous relaxation of the MIMM as given by:

$$\begin{aligned} k_1^*(x) \equiv \min_y \ & y_1 + 10y_2 \\ \text{s.t. } & y_1 + 12y_2 \geq x + 10 \\ & y_1, y_2 \geq 0 , \end{aligned}$$

since $k_1^*(0) = 10 \cdot \frac{10}{12} < 10$.

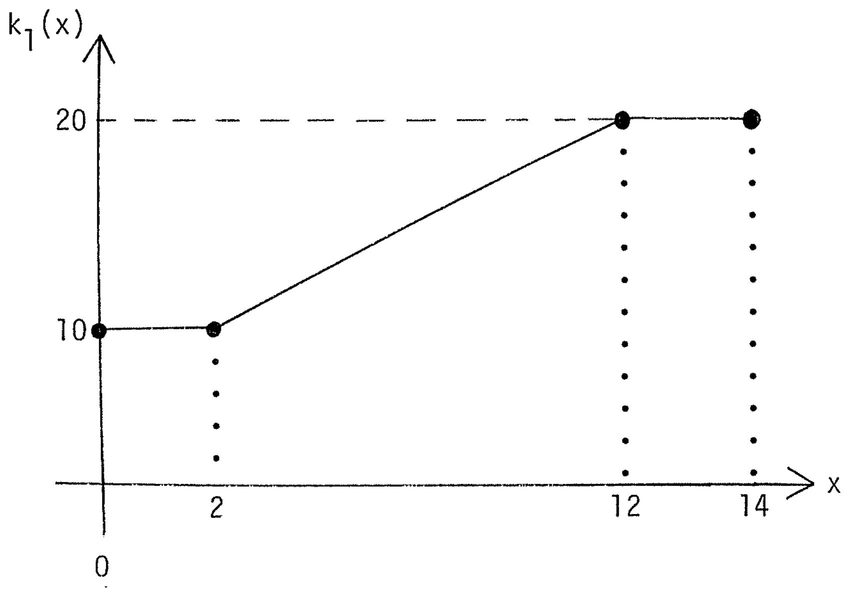

Figure 4. $k_1(x)$ on $[0,14]$

Note also that the addition of <u>bounds</u> does not help, since defining

$$
\begin{aligned}
k_2(x) \equiv \min_{y}\ & y_1 + 10y_2 \\
\text{s.t. } & y_1 + 12y_2 \geq x + 10 \\
& 0 \leq y_1 \leq 10 \\
& 0 \leq y_2 \leq 1 \\
& y_2 \text{ integer}
\end{aligned}
$$

yields $k_1(x) = k_2(x)$ for $x \in [0,12]$, and $k_1(x)$ coincides with its convex envelope on $[0,12]$, whereas the optimal value function of the continuous relaxation is again strictly less than $k_1(x)$ at $x = 0$.

Now consider the following example in which a RHS constant is not present in the constraint involving x, but there are <u>negative</u> coefficients:

$$
\begin{aligned}
k_3(x) \equiv \min_{y_1, y_2}\ & -y_1' + 10y_2 \\
\text{s.t. } & -y_1' + 12y_2 \geq x \\
& 0 \leq y_1' \leq 10 \\
& 0 \leq y_2 \leq 1 \\
& y_2 \text{ integer.}
\end{aligned}
$$

Making the change of variables $y_1' = 10 - y_1$ we have

$$k_3(x) = -10 + \min_{y_1, y_2} \; y_1 + 10y_2$$
$$\text{s.t. } y_1 + 12y_2 \geq x + 10$$
$$0 \leq y_1 \leq 10$$
$$0 \leq y_2 \leq 1$$
$$y_2 \text{ integer,}$$

so that $k_3(x) = -10 + k_2(x)$. It is easily seen that while k_3 coincides with its convex envelope on $[0,12]$, it differs from the optimal value function of the corresponding continuous relaxation at $x = 0$.

In our last example, we consider the case of two constraints with positive coefficients and RHS x:

$$k_4(x) \equiv \min \; y_1 + y_2$$
$$\text{s.t. } 2y_1 + 4y_2 \geq x$$
$$4y_1 + 3y_2 \geq x$$
$$0 \leq y_1, y_2 \leq 1$$
$$y_1, y_2 \text{ integer .}$$

In this case the optimal value function is finite for $x \leq 6$, and is easily seen to have the values:

$$k_4(x) = \begin{cases} 0 & \text{for} \quad x = 0 \\ 1 & \text{for} \quad 0 < x \leq 3 \\ 2 & \text{for} \quad 3 < x \leq 6 \,. \end{cases}$$

Thus the convex envelope of k_4 on $[0,6]$ is simply $x/3$. On the other hand, for $x = 5$ the continuous relaxation of the above MIMM for k_4 is easily seen to have optimal value $^3/2$ for $x = 5$ (choose $y_1 = \frac{1}{2}, y_2 = 1$), and therefore it does <u>not</u> coincide with the convex envelope, which has value $^5/3$ at $x = 5$.

Appendix C

We wish to establish relaxation-optimality in the case of both upper and lower bounds as considered in Corollary 3.3. Define

$$
\text{(C.1)}\qquad
\begin{aligned}
f_3(x) \equiv \min_y \;\; & cy \\
\text{s.t.}\;\; & ay \geq x\,, \\
& \ell \leq y \leq u' \\
& y_i \;\; \text{integer},\; i \in I
\end{aligned}
$$

and

$$
\text{(C.2)}\qquad
\begin{aligned}
f_3^*(x) \equiv \min_y \;\; & cy \\
\text{s.t.}\;\; & ay \geq x \\
& \ell \leq y \leq u'\,,
\end{aligned}
$$

where $\ell \geq 0$ and ℓ_i and u_i' are integer for $i \in I$. By making the substitutions $y = z + \ell$, $x = t + a\ell$, and $\tilde{u} = u'-\ell$, we have

$$
\begin{aligned}
f_3(x) &= c\ell + \min_z \;\; cz \\
&\qquad \text{s.t. } az \geq t,\;\; 0 \leq z \leq \tilde{u},\;\; z_i \text{ integer},\; i \in I \\
&= c\ell + \tilde{f}_2(t) = c\ell + \tilde{f}_2(x-a\ell),
\end{aligned}
$$

where

$$\tilde{f}_2(t) \equiv \min_z \; cz$$

$$\text{s.t. } az \geq t, \quad 0 \leq z \leq \tilde{u}, \; z_i \text{ integer}, \; i \in I$$

Similarly, $f_3^*(x) = c\ell + \tilde{f}_2^*(x-a\ell)$ where

$$\tilde{f}_2^*(t) \equiv \min_z \; cz$$

$$\text{s.t. } az \geq t, \; 0 \leq z \leq \tilde{u} \, .$$

Note that, for $0 \leq x \leq a\ell$, $\tilde{f}_2(x-a\ell) = \tilde{f}_2^*(x-a\ell) = 0$, so that for $0 \leq x \leq a\ell$, $f_3(x) = f_3^*(x) = c\ell$, and for $x > a\ell$, the breakpoints of f_3 are obtained by translations of the breakpoints of $\tilde{f}_2$ by $a\ell$. Relaxation-optimality is then easily established by considering the analog of Theorem 2.2.