LINEAR PROGRAMMING:
ELEMENTARY GEOGRAPHICAL APPLICATIONS OF THE TRANSPORTATION PROBLEM

Alan Hay

ISSN 0306 – 6142
ISBN 0 90224665 8

CATMOG

(Concepts and Techniques in Modern Geography)

CATMOG has been created to fill a teaching need in the field of quantitative methods in undergraduate geography courses. These texts are admirable guides for the teachers, yet cheap enough for student purchase as the basis of class-work. Each book is written by an author currently working with the technique or concept he describes.

Other titles in preparation

This series, Concepts and Techniques in Modern Geography is produced by the Study Group in Quantitative Methods, of the Institute of British Geographers.
For details of membership of the Study Group, write to the Institute of British Geographers, 1 Kensington Gore, London, S.W.7.
The series is published by Geo Abstracts, University of East Anglia, Norwich, NR4 7TJ, to whom all other enquiries should be addressed.

LINEAR PROGRAMMING : ELEMENTARY GEOGRAPHICAL APPLICATIONS OF THE TRANSPORTATION PROBLEM

by

Alan Hay
(University of Sheffield)

Acknowledgement

I am very grateful to Ron Johnston and Doug Watts for comments on the first draft and to a CATMOG referee for comments on the penultimate draft.

Permission to reproduce Tables 7-12 was kindly given by Professor J.M. Henderson and the Harvard University Press.

I INTRODUCTION

(i) The conceptual framework

In this booklet the transportation problem of linear programming is introduced. An understanding of the objectives and procedures can be gained using the most elementary algebra and arithmetic. The validity of the procedures can be demonstrated mathematically but this booklet does not attempt such a demonstration.

The central concept of linear programming is that of optimisation. This is a fairly common concept in every day life as most individuals and organisations will, from time to time, ask "what is the best method (the optimal method) by which we can achieve our ends?". But in every day life the answer will often be "it all depends upon what you mean by 'best'". The answer to this query will vary : for some individuals the 'best' method means the method which involves achieving a stated target with least effort, least time or least cost. For example a gardener might specify the amount of vegetables he wishes to grow and ask how can I arrange my garden to achieve this with minimum effort. Such people optimise by attempting to minimise a specified quantity. On the other hand some people may start from a different angle : "We have this limited garden plot, how can we choose a mix of vegetables to give us a maximum yield?". This maximum must again be defined : in terms of food (maximum calories?), in terms of value (maximum cash value of the produce?) or in terms of net value (cash value after costs have been subtracted). But once such a definition has been made the gardener is clearly a maximiser. Of course in some circumstances the maximisation of one quantity is in effect the minimisation of another, a point to which discussion will return in section III below.

Clearly, in every day life, most people who ask optimising questions are implicitly assuming certain rules which they will not break. For example a car driver who wants to know the quickest route between two places may (or may not) be tacitly assuming that he will obey speed limits; a less law abiding individual may on the other hand be constrained by the top speed of the car which he intends to drive on this trip. Similarly the man who seeks advice on maximising the food yield of his vegetable garden may be implicitly saying But I cannot give it more than 10 hours work a week" : he is placing a constraint on the type of solution which he will accept. This therefore introduces the concept of optimisation subject to constraints as a common every day concept. The task of linear programming is to define the quantity to be maximised (or minimised) exactly, to state the relevant constraints and then to identify a procedure by which the optimal solution can be identified.

Although some optimising problems are capable of direct mathematical solution many of them are not so soluble. It would of course be possible, although extremely tedious to identify all feasible solutions and then to calculate which is the optimum. By feasible is implied a solution which meets all the constraints specified. The number of feasible solutions is often so large that such a procedure is unsatisfactory. Linear programming provides a fairly rapid trial and error method of search characterised by the following features:

a) the method starts from a feasible solution and progresses towards an optimal solution,

b) the method arrives at an optimal solution, and

c) it is clear when an optimal solution has been reached.

A search procedure of this kind, which involves the same steps repeated many times is known as iteration or iterative solution of the problem. In some cases iteration is the only method of solution : in other cases a direct solution would be possible but iteration proves to be more convenient. Many iterative procedures are ideally suited to computer operations.

(ii) Linear programming defined

The phrase 'linear programming' covers a group of techniques for achieving optimal solutions to allocation problems. One of the results of linear programming studies has been to demonstrate the mathematical similarity of what are at first sight extremely dissimilar problems. This booklet focuses upon one aspect of linear programming only - the so called transportation problem - because its solution has proved to have wide relevance in studies of optimal spatial allocation. It should be noted however that linear programming has also been used to study other types of allocation problem including the optimal allocation of land and labour in agricultural systems. The most widely used method of solution for these problems is the so-called simplex method of linear programming. Several of the references introduce these wider aspects of linear programming (Abler et al. 1971, Vajda 1960).

In addition there is the field known as recursive linear programming. Readers interested in that field and its application to geographic problems like von Thunen's land use model should refer to the work of Day and Tinney (1969).

II THE TRANSPORTATION PROBLEM

(i) The basic transportation problem

Consider a situation in which three points of origin (A_1, A_2, A_3) have supplies available to meet needs at three destinations (N_1, N_2, N_3). The amounts available at each of the origins are specified (a_1, a_2, a_3) as also are the amounts needed at each of the destinations (n_1, n_2, n_3). Furthermore the costs of moving goods between origins and destinations can be set up as a table of m_{ij}'s where the subscripts indicate the cell giving the cost of moving from the i'th origin to the j'th destination : so, for example m_{32} is the cost of moving goods from origin A_3 to destination N_2. In a real life problem these quantities must be written in specific units : the supplies available and the needs might be in tonnes, or even in thousands of tonnes; the movement costs would then be in cost units per tonne, for example £/tonne. Suppose now an individual operator is seeking to supply these needs from thee suppliers at a minimum cost.

Table 1. The transportation problem

		Destinations :			Amounts available : (a_i)
		N_1	N_2	N_3	
Origins :	A_1	m_{11}	m_{12}	m_{13}	a_1
	A_2	m_{21}	m_{22}	m_{23}	a_2
	A_3	m_{31}	m_{32}	m_{33}	a_3
Amounts needed (n_j)		n_1	n_2	n_3	

The linear programming problem is to find the optimal (least cost) method of supplying N_1, N_2 and N_3 from A_1, A_2 and A_3. We are seeking a pattern of flows (f_{ij}'s) where f_{ij} is the flow from the i'th origin to the j'th destination:

Table 2. The unknown flows in the transportation problem

		Destinations :			
		N_1	N_2	N_3	
Origins :	A_1	f_{11}	f_{12}	f_{13}	a_1
	A_2	f_{21}	f_{22}	f_{23}	a_2
	A_3	f_{31}	f_{32}	f_{33}	a_3
		n_1	n_2	n_3	

Clearly for this Table 2 to be 'sensible' the elements (f_{ij}) in each row should add up to the row total, the column values should add to the column totals and the sum of the row totals should equal the sum of the column totals. Finally it is clear that any one f_{ij} might be positive or zero, but could not in any sensible manner be interpreted if it were negative. These restrictions can be written as constraints on the linear programming problem which can now be specified.

We are seeking a set of flows from the i'th origin to the j'th destination such that:

$$\left.\begin{aligned} f_{11} + f_{12} + f_{13} &= a_1 \\ f_{21} + f_{22} + f_{23} &= a_2 \\ f_{31} + f_{32} + f_{33} &= a_3 \end{aligned}\right\} \text{ generally } \sum_j f_{ij} = a_i \text{ for all } i \qquad 1.$$

Strictly the form should be $\sum_{j=1}^{3} f_{ij}$ but $\sum_j$ is used to simplify the text with the meaning "sum over all the j columns".

and

$$\left.\begin{aligned} f_{11} + f_{21} + f_{31} &= n_1 \\ f_{12} + f_{22} + f_{32} &= n_2 \\ f_{13} + f_{23} + f_{33} &= n_3 \end{aligned}\right\} \sum_i f_{ij} = n_j \text{ for all } j \qquad 2.$$

and

$$a_1 + a_2 + a_3 = n_1 + n_2 + n_3$$

or more concisely $\sum_i a_i = \sum_j n_j$ 3.

furthermore the flows cannot be negative:

$$f_{ij} \geq 0 \qquad 4.$$

So we now have four constraints.

The total transport cost is given by

$$\left.\begin{aligned} &f_{11}m_{11} + f_{12}m_{12} + f_{13}m_{13} \\ &+f_{21}m_{21} + f_{22}m_{22} + f_{23}m_{23} \\ &+f_{31}m_{31} + f_{32}m_{32} + f_{33}m_{33} \end{aligned}\right\} = \sum_i \sum_j f_{ij}m_{ij}$$

It is this grand sum $\sum_i \sum_j f_{ij}m_{ij}$ which is to be minimised, subject to the constraints already noted. Clearly there are many possible allocations of the flows between the origins and destinations and to consider all of them in the search for the minimum would be extremely tedious.

(iii) A simple example

Consider now an example of the transportation problem where the values of the a_i, the n_i and the m_{ij} are specified:

Table 3. A simple transportation problem

	N_1	N_2	N_3	
A_1	$m_{11} = 7$	$m_{12} = 11$	$m_{13} = 4$	$a_1 = 30$
A_2	$m_{21} = 8$	$m_{22} = 3$	$m_{23} = 2$	$a_2 = 50$
A_3	$m_{31} = 5$	$m_{32} = 5$	$m_{33} = 5$	$a_3 = 20$
	$n_1 = 60$	$n_2 = 25$	$n_3 = 15$	

A feasible solution to this problem can be identified as a first step. To do this there are a number of procedures here we introduce the northwest corner rule. Take the cell_{11} (the "northwest" corner of the table) : compare the values of a_1 and n_1 and take the lesser value (in this case $n_1 = 60$ and $a_1 = 30$ so the lesser value is 30). This lesser value is now entered as f_{11}:

	N_1	N_2	N_3	
A_1	30	..	..	30
A_2	..	..	..	50
A_3	..	..	..	20
	60	25	15	

Inspection makes it clear that all the supplies, available at A, are now exhausted so f_{12} and f_{13} are entered as zeros. Attention is now shifted to the unsatisfied needs of n_1 (n_1 minus the 30 already supplied); this unsatisfied need is for 30 units : the lesser of this and a_2 is therefore entered in cell_{21}, (as 30 is less than a_2 which equals 50) $f_{21} = 30$. N_1 is now fully supplied and the cell f_{31} can be entered as zero.

	N_1	N_2	N_3	
A_1	30	-	-	30
A_2	30	-	-	50
A_3	-	-	-	20
	60	25	15	

Having made up the table thus far attention is paid to the allocation of the 20 units still available at A_2 : these can be allocated to N_2 (using the 'lesser' criterion again). Readers should continue the procedure until all the cells have been entered in this manner with either positive values or zeros. The result should be:

	N_1	N_2	N_3	
A_1	30	-	-	30
A_2	30	20	-	50
A_3	-	5	15	20
	60	25	15	

Finally it is worthwhile to note that this solution is indeed feasible by checking that it meets the four constraints:

$$f_{ij} \geq 0 \qquad 1.$$

$$\sum_j f_{ij} = a_i \qquad 2.$$

$$\sum_i f_{ij} = n_j \qquad 3.$$

$$\sum a_i = \sum n_j \qquad 4.$$

Readers should do this for themselves.

The costs of the feasible solution can now be calculated using the m_{ij} (cost table):

	N_1	N_2	N_3
A_1	7	11	4
A_2	8	3	2
A_3	5	5	5

For this feasible solution $\sum f_{ij}\, m_{ij}$ is given in Table 4. The next stage is to ask whether this is an optimal solution. This is done by introducing the device of 'shadow prices'. These can be conceived initially as being the relative price of the commodity at the various origin and destination locations, relative that is to an arbitrarily chosen 'bench mark' or 'datum level'. For example we might set the datum level as being the value of the commodity at A_1, arbitrarily set at 0. There is in this case a flow from

Table 4. Total costs in a feasible solution

Origin - destination cell	Tonnes f_{ij}	Cost/ tonne m_{ij}	Total cost $f_{ij}m_{ij}$
f_{11}	30	7	210
f_{12}	0	-	-
f_{13}	0	-	-
f_{21}	30	8	240
f_{22}	20	3	60
f_{23}	0	-	-
f_{31}	0	-	-
f_{32}	5	5	25
f_{33}	15	5	75
		$\Sigma f_{ij}m_{ij}$ =	610

A_1 to N_1 : the movement cost is known to be 7 units. The costs of the goods delivered to N_1 will thus be the cost at A_1 plus the cost of movement : in this case 0 + 7. If the cost of goods at N_1 is 7 it tells us that the cost at A_2 must have been less than at A_1 : ie. to make the value at N_1 equal to 7, after 8 transport cost units the value at A_2 must be -1. This reasoning is represented algebraically as follows. Each origin has a shadow price u_i and each destination has a shadow price v_j such that $v_j - u_i = m_{ij}$; for all occupied (non zero cells) in the first solution. For example the cell f_{21} is occupied : $u_2 = -1$, $v_1 = 7$, $m_{ij} = 8$ so : 7 - (-1) = 8. This rule can be used to produce the following pattern - which the reader should check by calculation.

				u_i
	30	-	-	0
	30	20	-	-1
	-	5	15	-3
v_j	7	2	2	

The absolute values of these shadow prices depend upon the choice of a bench mark (which origin and which price) but the relative values are the same regardless of that choice (some readers may choose to check this also for themselves). These shadow prices are then used to calculate what are called shadow costs for the unoccupied cells.

$$c_{ij} = v_j - u_i$$

In this case:

-	2	2	0
-	-	3	-1
10	-	-	-3
7	2	2	

These shadow costs for unoccupied cells are also independent of the choice of bench mark. The shadow costs can be interpreted as the costs of leaving these cells unoccupied. If the cost of leaving a cell unoccupied (c_{ij}) is less than the cost of using it (ie. m_{ij}) there is no need to change the initial solution, but if the reverse is true ($c_{ij}>m_{ij}$) then a new solution must be sought which occupies the cell in question. In this case the $c_{31}=10$ whereas $m_{31}=5$ ie. $c_{31}>m_{31}$ and some way of making f_{31} positive is required.

This is done by searching the third row and first column of the initial solution for a quantity which could be moved to the cell f_{31} and which could be compensated by adding and subtracting from other cells in order to maintain the row and column totals. To achieve this it is necessary to make a sequence of horizontal and vertical transfers within the table starting and ending at the cell in question, alternately adding and subtracting a chosen amount to or from occupied cells in such a way that only one occupied cell is emptied, the chosen unoccupied cell is filled, and such that all the other cells which were initially occupied remain occupied and all other cells which were initially unoccupied remain so. In this case a sequence can be constructed by taking five units from f_{32} (reducing it to zero) adding five units to f_{31}(occupying it for the first time) subtracting five from f_{21}(reducing it from 30 to 25 but it remains positive) and adding five to f_{22} (raising it to 25). The result is a new feasible solution:

30	-	-	30	0
25	25	-	50	-1
5	-	15	20	2
60	25	15		
7	2	7		

with shadow costs compared to real costs:

SHADOW COSTS

-	2	7
-	-	8
-	0	-

REAL COSTS

7	11	4
8	3	2
5	5	5

Here attention is focused upon f_{23} and its excess of shadow costs (8 units) over real costs (2 units) by 6 units. To occupy this cell 15 units are subtracted from f_{33} (making it 0), 15 are added to f_{23} (as required, making it positive) 15 are subtracted from f_{21} (making it now 10) and 15 are added to f_{31} making it 20. That is:

30	-	-	30	0
10	25	15	50	-1
20	-	-	20	2
60	25	15		
7	2	1		

for which shadow costs are

-	2	1
-	-	-
-	0	1

as in no case do these shadow costs exceed real costs an optimal solution has been found. A check reveals that its total costs will be 495 units (compared with 610 units in the first solution).

III THE SCOPE OF THE TRANSPORTATION PROBLEM

(i) Extensions of the transportation problem

The transportation problem is one of the easier linear programming problems to solve and despite its name it has a wide range of possible applications. The first point to note is that the cost table m_{ij} may be set up in any appropriate units. In this case m_{ij} was a money cost and total money cost was minimised. If however m_{ij} is in distance units (km) then the $\sum_{ij}\sum f_{ij}m_{ij}$ would be the total number of tonne-kilometres, which would have to be performed. If the problem concerned optimal use of a vehicle fleet m_{ij} might be the number of vehicle hours taken to move between i and j.

Less evident is the fact that the solution of the transport problem can also be adapted to location problems. Consider a situation where three plants are supplying three markets. The costs of supplying a market from a plant can be written so that m_{ij} is the joint cost of production and transport. Furthermore many location problems may be concerned with determining the amounts produced at locations (subject to an upper constraint which is the maximum capacity available at each location, but where total capacity of all locations exceeds demand). This too can be treated as a transportation problem in which an additional 'destination' is created - a dump which can 'absorb' production at zero cost of production and transport. This dump is occasionally referred to as a 'dummy region' or a 'slack variable'. For example if A_1, A_2 and A_3 have capacity of 40,000, 35,000 and 25,000 tonnes respectively but N_1, N_2 and N_3 only need 90,000 tonnes (30,000, 50,000 and 10,000 respectively) the balance of 10,000 tonnes can be entered in a table as a dump region.

	N_1	N_2	N_3	Dump	
A_1	9	7	3	0	40
A_2	5	5	5	0	35
A_3	2	8	7	0	25
	30	50	10	10	

The costs of production and transport can similarly be entered as before, but the final column (clump) is entered with zeros. An optimal solution to this problem is given by the sequence of tables in Table 5. The end result is given below in Table 6 and the whole problem is illustrated in Figure 1. Revealing that though A_2 and A_3 produce at full capacity A_1 produces 30,000 tonnes only; the remaining 10,000 "produced for the dump" are in reality neither produced nor transported.

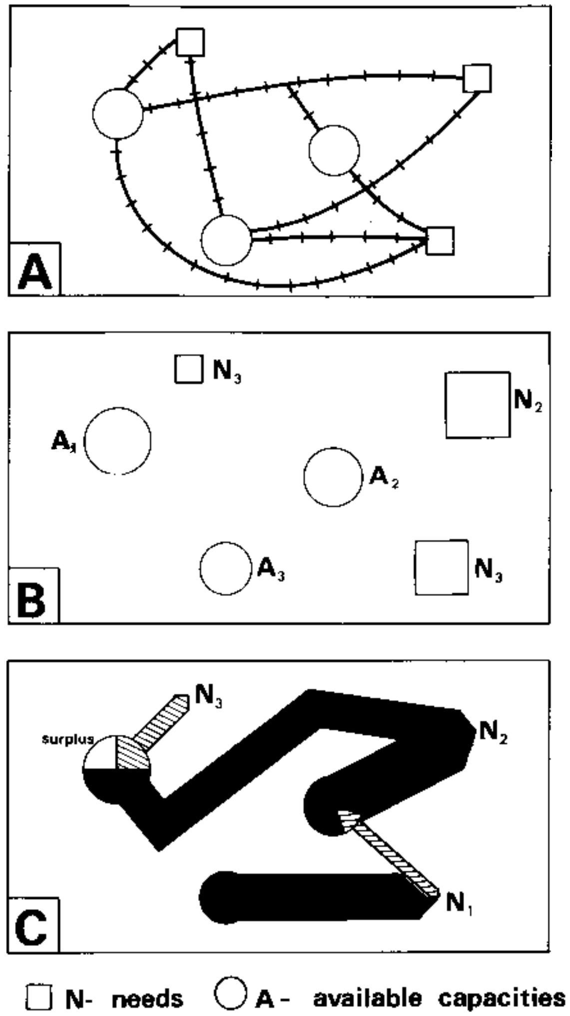

Figure 1 A simple problem : see Table 6.

Note the surplus at A_1 is for the 'dump'.

Table 5. The location problem : transport-and-production costs with a dump region

1st

30	10	-		40	-	-	6	-1	0
-	35	-		35	11	-	8	-3	2
-	5	10	10	25	8	-	-	-	-1
30	50	10	10		9	7	6	-1	

2nd

-	40	-	-	40	7	-	6	-1	0
30	5	-	-	35	-	-	8	-3	2
-	5	10	10	25	8	-	-	-	-1
30	50	10	10		7	7	6	-1	

3rd

-	40	-	-	40	7	-	12	5	0
25	10	-	-	35	-	-	14	3	2
5	-	10	10	25	-	12	-	-	5
30	50	10	10		7	7	12	5	

4th

-	40	-	-	40	7	-	7	5	0
15	10	10	-	35	-	-	-	3	2
15	-	-	10	25	-	2	2	-	5
30	50	10	10		7	7	7	5	

5th

-	30	-	10	40	7	-	7	-	0
5	20	10	-	35	-	-	-	-2	2
25	-	-	-	25	-	2	2	-5	5
30	50	10	10		7	7	7	0	

6th

-	20	10	10	40	7	-	-	-	0
5	30	-	-	35	-	-	1	-2	2
25	-	-	-	25	-	2	-2	-5	5
30	50	10	10		7	7	3	0	

Table 6. An optimal location of production

	Flows				Production	Surplus Capacity
A_1	-	20	10	10	30	10
A_2	5	30	-	-	35	-
A_3	25	-	-	-	25	-

It is only one step further to see that problems with no transport element at all may be set up as 'transportation problems'. Consider three farms A_1, A_2 and A_3 each having a fixed amount of land 80, 100 and 60 units respectively. But three crops are needed using land in the amounts 40, 130 and 70 units respectively. How can the farms share production to produce these needed quantities at least cost (labour units, total inputs etc.) given unit costs for each farm. Here the farms can be treated as origins, the crops as destinations

		Crops			
		N_1	N_2	N_3	
Farms	A_1				80
	A_2				100
	A_3				60
		40	130	70	

and the costs as m_{ij} cells:

12	8	7
10	10	9
7	12	12

The f_{ij} in this case will be areas under production: $\sum_i \sum_j f_{ij} m_{ij}$ will be total costs; the optimal solution will show the least cost solution to the land allocation problem.

				u_i
	0	80	0	0
	0	30	70	-2
	40	20	0	-4
v_j	3	8	7	

In this solution farm A_3 specialises in crop N_1 (but with subsidiary production of N_2), farm A_2 specialises in crop N_3 (but with subsidiary production of N_3 also) while farm A_1 specialises in the production of crop N_2 only.

(ii) Additional features of the transportation problem

Experiments with a number of small transportation problems will reveal some additional features and occasional snags.

Firstly it may be noted that the method of setting up a feasible solution by starting with f_{11} (called the northwest corner rule) may lead to a clearly inefficient solution and hence to a lengthy series of iterations. To avoid this an alternative procedure can be adopted: first select the lowest value of m_{ij} in the table and allocate the lesser of a_i and n_j. In the case already studied this would identify the cell m_{23}(=2) and enter 15 units in f_{23} (15 being n_3, less than a_2). The remaining 35 units in at a_2 can then be allocated to the next lowest m_{ij} in the second row (f_{22}=25) and in this case 10 units remain for cell$_{21}$ (f_{21} = 10).

-	-	-
10	25	15
-	-	-

The next smallest value is m_{13}(=4) but the column O_3 already has the needed 15 units (n_3 = 15). The next smallest in that row is m_{11} (=7) and 30 units can be allocated there exhausting a_1. The sole remaining cell is m_{31}, in which f_{31} is set equal to 20.

30	-	-
10	25	15
20	-	-

In this case a quick check shows that the method has in fact identified the optimal solution directly, but this will not always occur.

Secondly it may be found that the procedure for identifying a feasible solution yields a pattern of f_{ij}'s in which less than the-number-of-rows-plus-the-number-of-columns-minus-one cells are occupied. This feature is known as <u>degeneracy.</u>

Consider the problem below:

	N_1	N_2	N_3	
A_1				20
A_2				40
	20	15	25	

	N_1	N_2	N_3
A_1	2	1	2
A_2	3	4	1

In this case the northwest corner rule yields a pattern:

20	-	-
-	15	25

where only 3 cells are filled (but rows + columns - 1 = 4). If we try to apply the rule for calculating shadow prices: we allocate a price of 0 to A_1 it giving a price of 2 at N_1 but there is no way of then calculating A_2, N_2 and N_3 shadow prices. The method of calculating shadow prices has failed. To circumvent this problem a simple device is adopted : each column total is increased by x and <u>one</u> row total is increased by nx (where x is an algebraic symbol and n is the number of columns). The revised problem can then be set:

	N_1	N_2	N_3	
A_1				20 + 3x
A_2				40
	(20 + x)	(15 + x)	(25 + x)	

	N_1	N_2	N_3
A_1	2	1	2
A_2	3	4	1

for which the northwest corner rule yields a feasible solution:

(20 + x)	2x	-	0
-	(15 - x)	(25 + x)	-3
2	1	-2	

it has shadow costs:

-	-	-2
5	-	-

The shadow cost for cell$_{21}$ is greater than m_{21}. The cell is filled by moving (15 - x) units (f_{21} = 15 - x) decreasing f_{11} by (15 - x) to (5 + 2x) increasing f_{12} (to 15 + x) and reducing f_{22} to zero. F_{23} remains (25 + x).

(5 + 2x)	(15 + x)	-	(20 + 3x)	0
(15 - x)	-	(25 + x)	40	-1
(20 + x)	(15 + x)	(25 + x)		
2	1	0		

with shadow costs:

-	-	0
-	2	0

all of which are less than m_{ij}. The solution is optimal. If x is then set equal to zero the solution becomes

5	15	-
15	-	25

which is optimal.

The third feature which is often noted in simple cases is that shadow costs of this optimal solution are for some cells equal to the m_{ij}'s. These equalities indicate that there exists another solution which equally minimises the total movement costs. In this case there is nothing to choose between the solutions in terms of the presented minimisation problem, though there may then be room to use other criteria to choose between them. Finally it is worthwhile to note that if the quantities supplied and demanded are integers (whole numbers) the f_{ij}'s in the solution will also be integers. The formal proof of this is beyond the scope of this booklet (Trustrum, 1971, p. 36). This explains why it was possible to seek the optimal solution by adding and subtracting an integer value from the cells identified in the feasible solution.

(iii) <u>Shadow prices</u>

Many first time users of the linear programming method having approached the problem with a clear statement of the constraints and of the quantity to be maximised, have solved the problem and been satisfied with the 'answer' so obtained. In geographic studies this is unfortunate because two of the 'by-products' of the linear programming solution have important implications and help in a full understanding of the solution; these are the <u>shadow prices</u> and <u>the dual.</u> (Scott (1971) gives a good account of these two concepts)

In solving the transportation problem use was made of shadow prices as a computational convenience. It was noted that they all relate to a "bench mark" or datum level price, but that the same relative differences will occur whatever bench mark is chosen. Recalling the first problem the shadow prices associated with the solution were:

	N_1	N_2	N_3	Shadow prices u_i =
A_1	30	-	-	0
A_2	10	25	15	-1
A_3	20	-	-	2
Shadow prices v_j =	7	2	1	

while in the dump region case they were:

	N_1	N_2	N_3	D	u_i
A_1	-	20	10	10	0
A_2	5	30	-	-	2
A_3	25	-	-	-	5
v_j	7	7	3	0	

In both cases there are clear differences between the shadow prices at the originating locations. These differences may be interpreted as the relative value of a unit of product due to its locational advantage. In the first case A_3 has locational advantage of 2 value units over A_1 and of 3 units over A_2 : in the second case too A_3 has the locational advantage over both the other suppliers. It is only one step from this interpretation to a rent interpretation : in the first case (all other things being equal) it will be worthwhile for a producer located at A_2 to pay rent equivalent to as much as 1 unit for the privilege of moving to A_1 or as much as 3 units to be allowed to move to A_3. Such a rent interpretation is analogous to the land rent concepts of Ricardo and von Thunen and the application of linear programming to the rent of crops has been explored by Stevens, (1961), Henderson (1959) and Maxfield (1969).

(iv) <u>The Dual</u>

In the introduction it was noted that many maximisation problems have an 'opposite' which is a minimisation problem. The established terminology in linear programming refers to a primal problem and a dual problem. Just as the primal problem in the transportation problem is to identify a set of f_{ij}'s such that $\sum_i\sum_j f_{ij}m_{ij}$ is <u>minimised</u> subject to constraints so the dual problem is to identify a set of shadow prices (u_i and v_j) which <u>maximises</u> $\sum_j n_jv_j - \sum_i a_iu_i$ subject to the constraints:

$$\left.\begin{array}{l} v_1 - u_1 = m_{11} \\ v_1 - u_2 = m_{21} \\ v_1 - u_3 = m_{31} \\ v_2 - u_1 = m_{12} \\ \vdots \\ v_3 - u_3 = m_{33} \end{array}\right\} \text{ generally } v_j - u_i = m_{ij}$$

In some cases the dual has a useful computational role in that it allows the solution of problems which in their primal form are awkward or impossible to solve. The expression $\Sigma_j n_jv_j - \Sigma_i a_iu_i$ has a direct economic interpretation as the value added in transportation: it is this which is maximised, subject to the constraint that the difference in price between any supplying region (A_i) and any consuming region (N_j) is never greater than the transport costs between them. The constraint ignores any addition to the price made by the producer as a result of pricing policies: but if he raises that price obviously and appreciably an alternative supplier can compete with him.

This discussion of the shadow prices has introduced also the important question of competition. The first applications related to an operator wishing to set up an optimal <u>system,</u> but the argument has now developed to include competing individuals within a market system. These two aspects of linear programming lie at the heart of the applications which will be discussed in the next section.

IV <u>APPLICATIONS</u>

At first sight optimising methods appear peculiarly inappropriate to studies of aggregate patterns in a real world which appears to be lacking in spatial organisation of this kind. There are however two ways in which the linear programming approach has proved to be of value in geography.

(i) <u>Testing the efficiency of real-world systems</u>

The first application is essentially normative : the geographer uses the tools of linear programming to expose the inefficiency of a system as it now exists. Such an exposure may either underline the inefficiency of a system in achieving its 'own' goals or it may show how grossly inefficient the system is at achieving goals which are thought to be desirable by the research worker.

In the first case it is often difficult to identify fairly all the goals and constraints of which the real life decision makers are aware. Indeed real life decision makers may have multiple goals and constraints which are mutually incompatible. In such a case linear programming will help to identify these inherent contradictions in the system. But if the system under study is a centralised co-ordinated system with declared goals then linear programming will be highly appropriate : as in the study of centrally planned economies (see Barr 1970) or the study of monopolies (perhaps nationalised) in capitalist societies (eg. the National Coal Board or the British Steel Corporation). One word of warning is appropriate here - the decision makers may have themselves used linear programming and an independent academic repetition of the same analysis is in danger of being trivial. Chisholm and O'Sullivan for example show that the movements of oil products in the United Kingdom conform closely to the linear programming model, scarcely surprising when it is known that the few major oil companies not only use linear programming but frequently act in concert on such problems.

Nevertheless a number of useful studies have been able to underline the inefficiencies of the system as it is presently operated (Osayimwese 1974 But even such inefficiencies may need closer investigation. For example if the behaviour of the operators is considered they will frequently be responding to day-to-day or week-to-week patterns of demand. Most academic studies will use the aggregate data for much longer periods in which the short term fluctuations have been concealed. For this reason optimal short term behaviour may appear as sub-optimal in the longer term. For example Osayimwese investigated in Nigeria the flow of export crops (specifically groundnuts) from production areas to export ports. In so doing he identified certain inefficiencies in the existing pattern the optimal solution indicates that Apapa and Port Harcourt were underutilised by 17 and 15 per cent respectively, 58 per cent of the total shipments should have been by road, whereas less than 25 per cent actually arrived in Apapa by road ...". (Osayimwese, 1974, p.62). But the responsible authority, the Marketing Board, was monitoring week to week conditions of stocks, available railway rolling stock and available shipping. These short term considerations necessarily included the effects of strikes, breakdowns, and a multitude of similar local effects, which probably contributed much to the 'inefficiency' identified by Osayimwese.

(ii) Equilibrium interpretations of linear programming

Consider again the pattern of flows arrived at by linear programming in Section IIi. At first sight this appears to be possible only if an overall monopolistic decision maker was subjecting the best interests of origin A_1 (to give its supplier to N_3) to those of the system. It was however recognised at an early stage in the history of linear programming that such a solution will also arise under perfect competition. Suppose each of the suppliers were 'bidding' for the right to supply given markets. The supplier bidding most for the privilege of supplying N_2 and N_3 is A_2 not only because it is 'well located' for such a role but because of the penalty it would pay (in terms of high costs) if it were forced to supply N_1;A_3 in contrast has little to lose by being forced to supply N_1 and would not therefore bid much for the privilege of supplying N_2 or N_3. This interpretation of the linear programming flows as resulting from pure competition is, as has already been suggested, analogous to the ideas of land quality rent and location rent resulting from the competition between crops in studies of agricultural location.

A second interpretation of the linear programming solution is often made in terms of spatial equilibrium. Recall the final solution:

7*	11	4	0
8*	3*	2*	-1
5*	5	5	2
7	2	1	

Consideration of this solution reveals that for the occupied cells the prices in the destination regions differ from those in the origin regions by exactly the transport costs between them. This is consistent with the equilibrium interpretation of interregional trade summarised by Morrill and Garrison:"Trade will take place between two regions if the difference between regional prices is greater than the amount of transportation costs between them. The amount traded will be that required to shift prices through supply and demand relationships until prices are separated by interregional transportation costs" (Morrill and Garrison, 1960, p. 118).

Finally, note can be taken of the proof by Samuelson that the linear programming solution can be used to analyse spatial price equilibrium if the quantity to be maximised is stated as the sum of consumers' plus producers' surplus minus transport costs (Samuelson, 1952). Such a solution determines the amounts produced, the amounts consumed, prices in supply and consumer regions and the flow pattern. It is therefore a general solution of the location problem although it is so restricted by its assumptions and data requirements that it has not been widely applied.

The final point to be made in this context is the limited nature of any appraisal made on efficiency criteria alone. Even if Samuelson's model fits the data and is correct it maximises consumer and producer surplus jointly. It says nothing about the distribution of that surplus between producers and consumers. Does efficiency mean that the peasant receives a higher price or just that an affluent townsman can drink his coffee more cheaply? The critical point is thus one of distributions between individuals, between classes or between regions for which it is difficult to define a quantity to be maximised or minimised.

(iii) Comparing results with reality

Whatever the chosen application of linear programming, it is clear that it will be desirable to compare its solution with the patterns in an existing system. Such a comparison may focus either upon aggregates or upon individual values. The simplest aggregate comparison is to compare the quantity which has been maximised (or minimised) with that attained in reality. Such a comparison may then express the efficiency of the system : for example a transportation problem might calculate the percentage of excess costs:

$$\frac{\Sigma f_{ij}\, m_{ij} \text{ (actual)}}{\Sigma f_{ij}\, m_{ij} \text{ (solution)}} \times 100$$

It must however be used cautiously as it takes no note of the inherent complexity (or simplicity) of the problem under study. Some problems are such that the optimal solution is unique and no other feasible solution remotely approaches the optimum in terms of efficiency.

Another point to be borne in mind when considering a linear programming solution as the 'explanation' of a geographic pattern is that there may be many feasible solutions which are only slightly sub-optimal. Consider the problem which was solved earlier as

	N_1	N_2	N_3
A_1	30	-	-
A_2	10	25	15
A_3	20	-	-

with total $\Sigma f_{ij} m_{ij}$ = 495. This could be amended very slightly by re-allocating only one unit:

29	1	-
11	24	15
20	-	-

$\Sigma f_{ij} m_{ij} = 504$

or

30	-	-
11	25	14
19	-	1

$\Sigma f_{ij} m_{ij} = 501$

etc.

Similarly two units might be re-allocated to yield yet other slightly sub-optimal solutions. More difficult to visualise is the fact that a solution which is only slightly sub-optimal (or indeed one of several equally optimal solutions) may have a set of cell values, and by implication a geographic pattern which is radically different. It is for this reason that the measure of total efficiency is scarcely a good measure of the fit (or explanation achieved) of a linear programming application to a geographic pattern observed in reality.

Comparison of individual values in the optimal solution (e for expected) with actual values (a) is therefore more often favoured. Such a comparison may be arithmetic : some authors favour the distribution of the differences $(f_{ij_e} - f_{ij_a})$. These will have a mean of zero and it is the pattern of values (plus and minus) around this mean which distinguishes a 'good' or 'bad' fit. Alternatively the average differences can be calculated using the expression:

$$\frac{\Sigma \left| f_{ij_a} - f_{ij_e} \right|}{n}$$

where n is the total number of pairs. In this case the mean deviation is calculated. Finally some authors prefer the expression:

$$\frac{\Sigma (f_{ij_a} - f_{ij_e})^2}{n}$$

in which the sum of the squares of the deviations is divided to give a mean square deviation.

It is only a small step from these techniques to drawing the results as a scatter graph in which the axes are observed and expected values. A perfect fit will occur if all the points lie on the line $f_{ij_e} = f_{ij_a}$,

which will of course exactly bisect the right angle. Unfortunately some authors, (e.g. Cox, 1965), have been seduced by this similarity into using the correlation coefficient r_{xy} to measure the closeness of observed and expected values. Although used in basic geographical references this method is unacceptable. Logically if the values are identical the regression line would have the form:

$$y = ax + b$$

with a = 1 and b = 0 ie. y = x.
It is however possible to have a high correlation coefficient where $a \neq 1$ and $b \neq 0$. Furthermore, the high interdependence among the expected values and among the observed values and the frequency of zero in the expected values makes the use of correlation and associated inferential statistics inappropriate. This approach to testing should therefore be ignored. Chisholm and O'Sullivan (1973, pp.76-78) use the technique but are clearly aware of its shortcomings.

(iv) The United States coal mining industry : an example

An interesting example of the application of linear programming to a production and transport cost minimisation was prepared by J.M. Henderson (1958). The study concerned the supply of coal from eleven mining regions in the United States to supply fixed demands in the same eleven regions and in three additional consumption regions. The study covered 1947, 1949 and 1951 but this account will focus on the 1947 data only. Unit extraction costs were known for both shaft and open cast mining as also were productive capacities. The costs of transport were also known, giving a transportation problem with 22 origins (eleven regions and two mining systems in each case) and fourteen destinations. Before the problem could be expressed in soluble form it was necessary to adjust for the varying quality of coal : affecting both the value of the delivered product and the effective costs of transport. This adjustment was achieved by expressing all quantities in 10^{10} British Thermal Units. Demand and production capacity were expressed in these units (Table 7 and Figures 2 and 3) : while costs of production and transport were expressed in US dollars per 10^{10} BTU.

Table 7. Demand and capacity. (10^{10} B.t.u.)

	Demand	Capacities (shaft)	(Opencast)
1	221,400	225,598	142,982
2	66,276	381,733	97,282
3	139,708	270,831	40,506
4	107,089	49,918	7,613
5	165,421	64,661	52,341
6	283,305	165,348	74,997
7	62,773	18,828	27,682
8	15,763	690	3,365
9	28,548	36,690	7,382
10	18,16.	20,424	927
11	10,522	1,388	609
12	60,186		
13	123,221		
14	47,005		

Source : Henderson op. cit. Tables 12 and 13

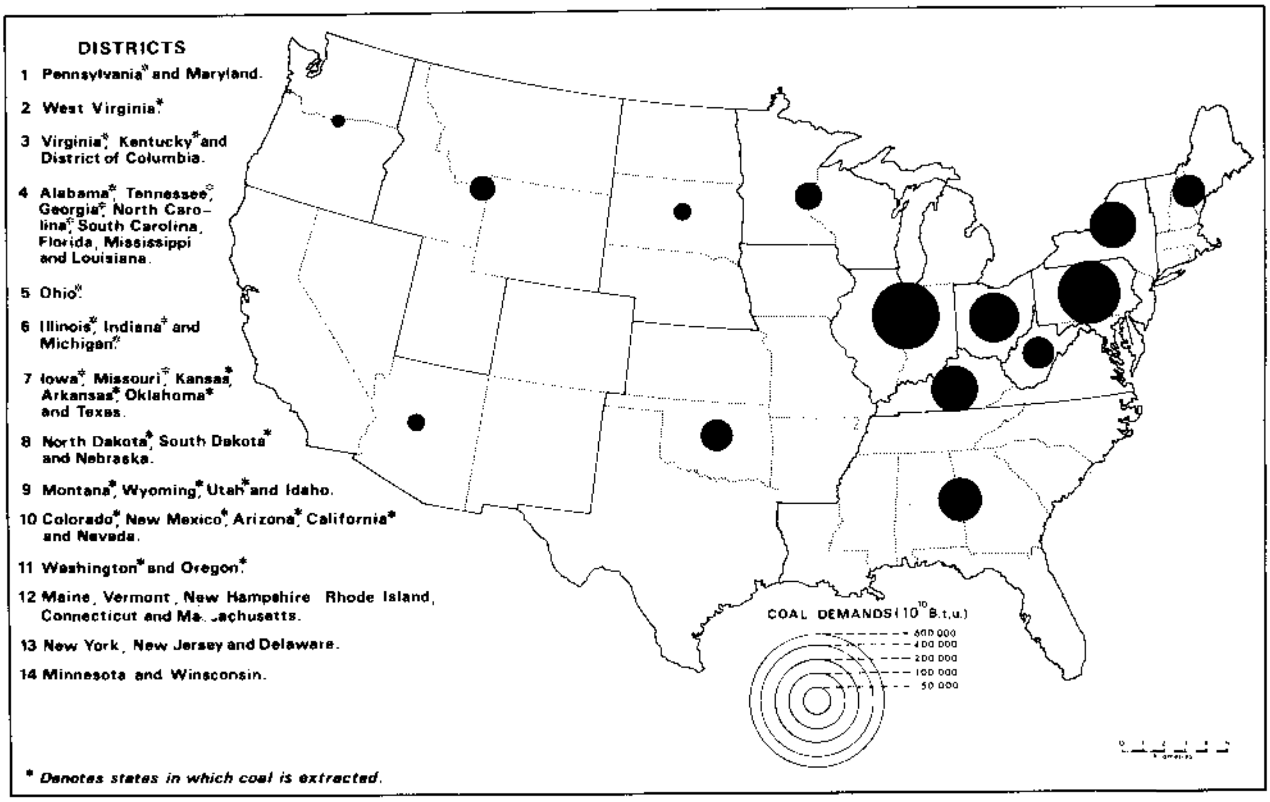

Figure 2 Demand for coal in the U.S.A. (Henderson's 1947 data)

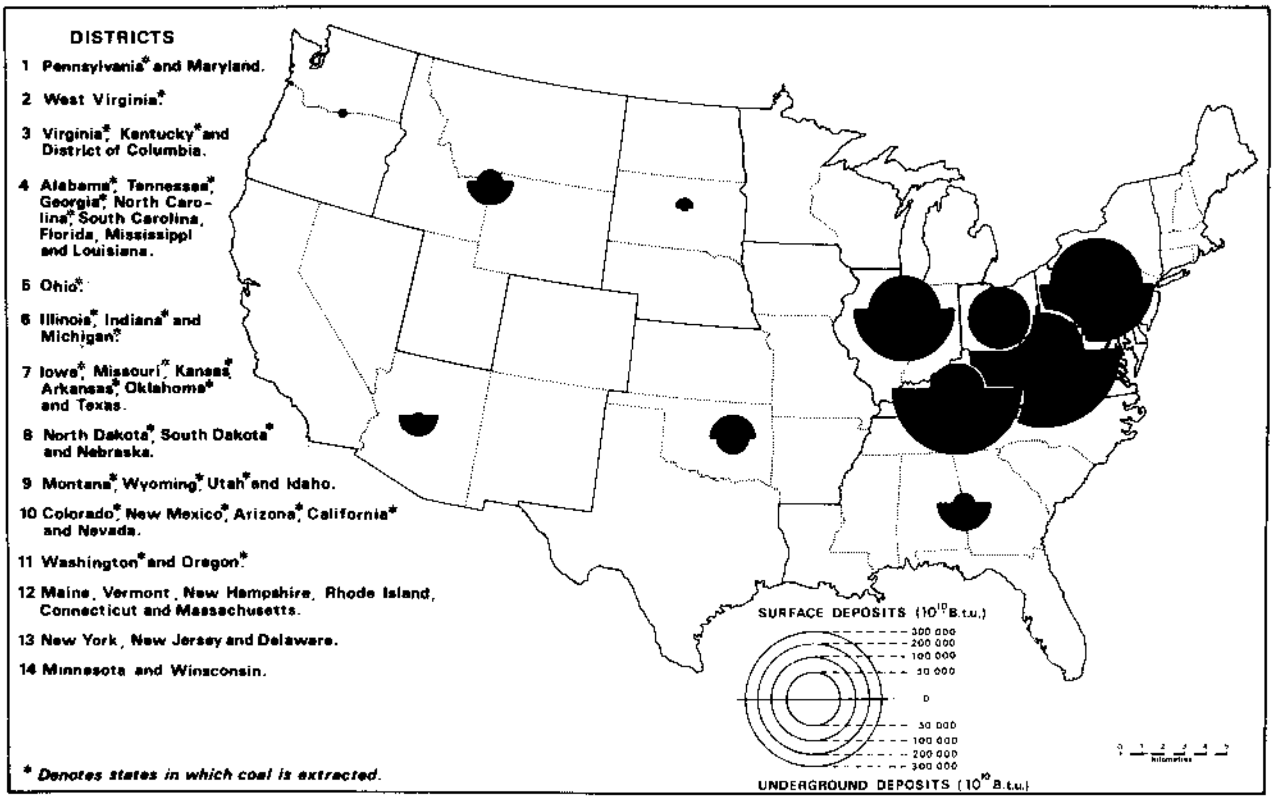

Figure 3 Production capacity for coal in the U.S.A. (Henderson's 1947 data)

Table 8. The optimal solution identified by Henderson.

	From district -	1	2	3	4	5	6	7
SHAFT	1 U	78,418				48,419		
	2 U							35,091
	3 U			99,202	99,476		42,960	
	4 U							
	5 U					64,661		
	6 U						165,348	
	7 U							
	8 U							
	9 U							
	10 U							
	11 U							
OPENCAST	1 S	142,982						
	2 S		66,276					
	3 S			40,506				
	4 S				7,613			
	5 S					52,341		
	6 S						74,997	
	7 S							27,682
	8 S							
	9 S							
	10 S							
	11 S							
	Demands	221,400	66,276	139,708	107,089	165,421	283,305	62,773

Source : Henderson op. cit. Table 19

Solution of the transportation problem yielded the optimal pattern of flows shown in Table 8 and Figure 4. It is clear that this example reflects many of the features already noted. Most of the flows are within regions : only regions 1 (Pennsylvania and Maryland), 2 (West Virginia) and 3 (Virginia and Kentucky) appear as major exporters. Unused capacity appears most marked in West Virginia, Pennsylvania and Maryland and the Alabama regions : the unused capacity in the optimal solution is the higher cost shaft mining capacity, whereas all the open cast capacity would be fully utilised. Finally it can be noted that the unutilised capacity tends to reflect the joint effects of high costs and competition.

It was not possible to compare the flows in the optimal solution with actual flows (data were not available) but the production from each region was compared for both the model and reality. Table 9 shows the difference between actual production and that demanded by the optimal solution. From this table it is clear that there was a general tendency for shaft mining to be maintained even where it was inefficient. Furthermore this 'excess' production was concentrated in regions 1 (Pennsylvania and Maryland) and 2 (West Virginia).

Table 8 (contd.)

From district	8	9	10	11	12	13	14	Unused capacity	Capacities
1 U							47,005	51,756	225,598
2 U					29,180	123,221		194,241	381,733
3 U	11,708							17,485	270,831
4 U								49,918	49,918
5 U									64,661
6 U									165,348
7 U								12,828	12,828
8 U	690								690
9 U		21,166		8,525				6,999	36,690
10 U			17,181					3,243	20,424
11 U				1,388					1,388
1 S									142,982
2 S					31,006				97,282
3 S									40,506
4 S									7,613
5 S									52,341
6 S									74,997
7 S									27,682
8 S	3,365								3,365
9 S		7,382							7,382
10 S			927						927
11 S				609					609
	15,763	28,548	18,108	10,522	60,186	123,221	47,005	336,470	1,685,795

Table 9. Contrasts between optimal and actual production patterns

Difference (actual - optimal)	Underground	Opencast	Total
1	+58,432	-41,101	+17,331
2	+94,805	-35,919	+58,886
3	+ 6,486	-10,980	- 4,494
4	-11,288	- 1,053	-12,341
5	-13,626	- 8,455	-22,081
6	-27,998	- 2,428	-30,426
7	+ 7,906	- 3,496	+ 4,410
8	- 29	- 147	- 176
9	+ 673	- 370	+ 303
10	- 2,555	- 329	- 2,884
11	+ 472	- 170	+ 402

Source : Henderson op. cit. Table 28

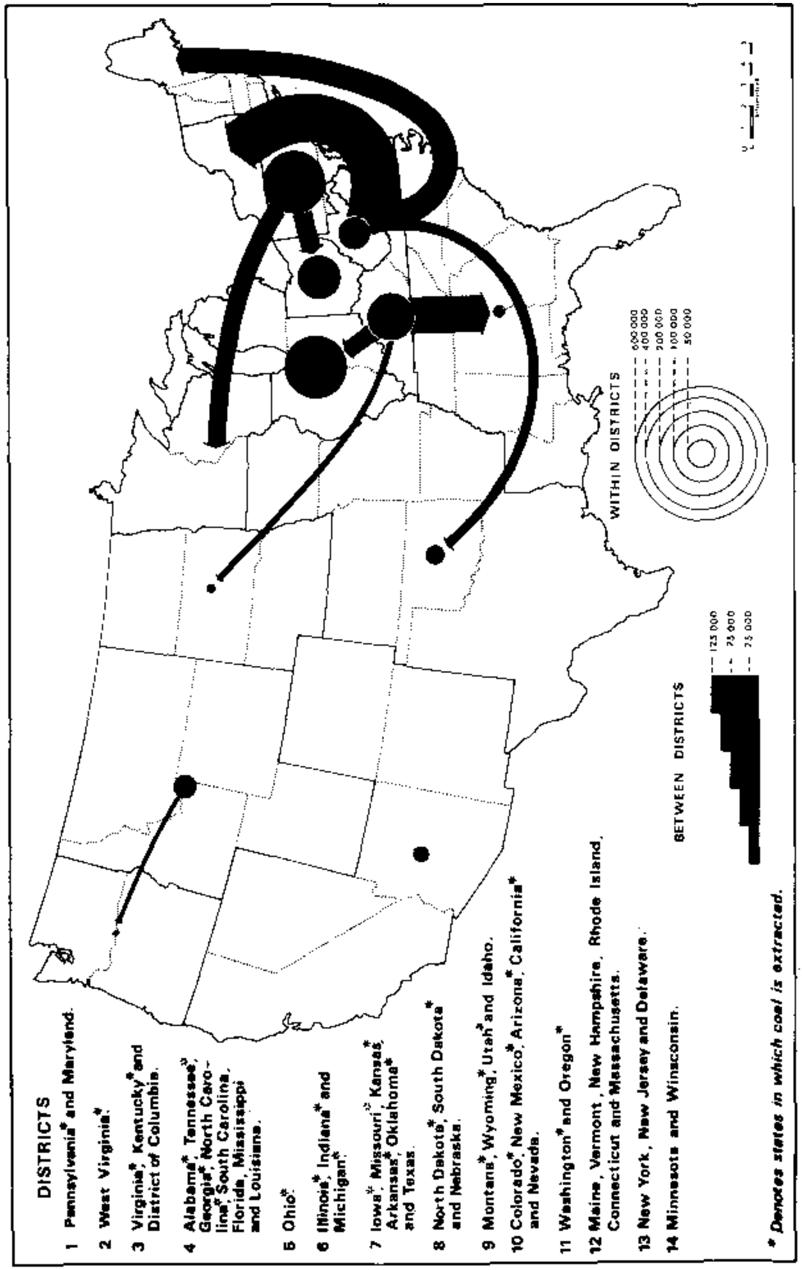

Figure 4 The production and flow of coal : an optimal solution. (Henderson's 1947 data)

A final comparison made by Henderson was for total costs of the production and transport system. From Table 10 it can be seen that actual production and movement costs were almost 9% above the costs which would have been incurred by an optimal solution.

Table 10. Total costs of production and transport

		Actual	Optimal
Costs : ($ m)			
	extraction	1,725	1,634
	transport	1,125	991
	total	2,850	2,625

Source : Henderson op. cit. Table 29

The dual of this problem was interpreted by Henderson in terms of 'competitive unit royalties' and 'competitive delivered prices'. The royalties (shadow prices in producing regions minus production costs) for open cast deposits show the pattern in Table 9 : it is clear that once again the joint effects of production costs and relative location affect the dual. On the one hand the producers in Oregon and Washington would have high royalties (because there are no competitors near at hand) despite one of the highest production costs : on the other hand the low cost producers of Virginia and ,Kentucky (region 3) had only moderate royalties. The second column in the table gives the shadow prices in consuming regions showing marked differences between these prices in the highly competitive market of 2 and 3 (Virginia, West Virginia etc.) : prices are correspondingly high in 11 (Washington and Oregon), 12 (the New England states) and 8 (the Dakotas).

Table 11. Interpretation of the dual ($ per 10^{10} BTU)

District	Royalties (open cast only)	Delivered (shadow) prices
1	675	2,004
2	436	1,456
3	698	1,895
4	528	2,357
5	798	2,170
6	1,054	2,551
7	1,197	2,887
8	908	2,933
9	475	2,200
10	483	2,073
11	1,454	3,506
12		2,935
13		2,469
14		2,453

Source : Henderson op. cit. Tables 25-26

V BIBLIOGRAPHY

A) Elementary introductions

Abler, R., Adams, J.S., Gould, P.R., (1971), *Spatial Organization,* Prentice Hall, Englewood Cliffs, Chapter 12.

Higgins, J.C., (1970), *Linear Programming,* Editype, Hove.

Scott, A.J., (1971), *An introduction to spatial allocation analysis,* Association of American Geographers, Commission on College Geography, Research Paper No. 9, Washington.

Taaffe, E.J. and Gauthier, H.L, (1973), *Geography of Transportation,* Prentice Hall, Englewood Cliffs, Chapter 6.

Trustrum, K., (1971), *Linear Programming,* Routledge and Kegan Paul, London.

Vajda, S., (1960), *An introduction to linear programming and the theory of games,* Methuen, London.

B) Basic references

Dantzig, G., (1963), *Linear programming and extensions,* Princeton, University Press.

Dorfman, R., Samuelson, P.A. and Solow, R.M., (1958), *Linear programming and Economic Analysis,* McGraw Hill, New York.

C) Case studies (n.b. not all these studies use the same method of solution as has been adopted in II(i) above)

Barr, B., (1970), *The Soviet wood processing industry,* University Press, Toronto.

Casetti, E., (1966), Optimal locations of steel mills serving the Quebec and Southern Ontario Steel Market, *Canadian Geographer,* 10, 27-38.

Chisholm, M. and O'Sullivan, P., (1973), *Freight flows and spatial aspects of the British Economy,* University Press, Cambridge.

Cox, K.R., (1965), The application of linear programming to geographic problems, *Tijdschrift voor Economisch en Sociale Geografie,* 56, 228-235.

Ghosh, A., (1965), *Efficiency in location and inter-regional flows : the Indian Cement Industry during the Five Year Plans 1950 - 1959,* North Holland Publishing Company, Amsterdam.

Gould, P.R. and Leinbach, T.R., (1966), An approach to the geographical assignment of hospital services, *Tijdschrift voor Economisch en Sociale Geografie,* 57, 203-206.

Henderson, J.M., (1958), *The efficiency of the coal industry, an application of linear programming,* Harvard University Press, Cambridge, Mass.