

MULTIMODAL IMPLICATIONS OF A UNIMODAL INVESTMENT:
THE BENEFITS MEASUREMENT

by

Sergio Rodolfo Jara Díaz

Ingeniero Civil, Universidad de Chile
(1974)

Magister en Planificación Urbana y Regional,
Universidad Católica de Chile
(1977)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

MASTER OF SCIENCE IN
CIVIL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1980

© Massachusetts Institute of Technology 1980

Signature of Author


Department of Civil Engineering


Certified by


Fred Moavenzadeh


Terry Friesz

Thesis Supervisors

Accepted by


C. Allin Cornell
Chairman, Departmental Committee on Graduate Students
Department of Civil Engineering

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

APR 25 1980 ¹

LIBRARIES

MULTIMODAL IMPLICATIONS OF A UNIMODAL INVESTMENT:

THE BENEFITS MEASUREMENT

by

SERGIO RODOLFO JARA DIAZ

Submitted to the Department of Civil Engineering
on January 3, 1979 in partial fulfillment of the
requirements for the Degree of Master of Science in
Civil Engineering .

ABSTRACT

The traditional measure of consumers' surplus variation is no longer valid when dealing with interrelated demands. This is particularly important in transportation analysis because in general willingness to pay for transport on one particular mode depends on the price on the other available modes. In an economic sense, these modes are substitutes and changes brought about by investment in any one of them will affect the whole system. The literature on this subject has concentrated on the analysis of two competing highways, which constitutes conceptually the same problem provided that perceived costs rather than fares or monetary costs are used. The shifts in modal demands caused by investment in one mode call for a different treatment of the measure of benefits; this has generated at least three different approaches: Williams (1976) and Agnello (1977) use the generalized version of consumers' surplus, the line integral formulation, due to Hotelling (1938) in order to provide an answer to this problem, resulting in different propositions; Mohring (1976) proposed a different approach including explicitly supply considerations. Moreover, the assumed direction of shift in the modal demands differs among these authors and no clear explanation for this phenomenon is offered.

This thesis presents a new approach to the measurement of benefits due to a modal investment, based primarily on a full explanation of the nature of the modal demands. For purposes of presentation, analysis and comparison, the case of two competing modes providing transportation between two points is developed. It is first shown that the production and consumption structure associated with each node generates an aggregate and stable transportation demand function. Modal demands, functions which relate flow and willingness to pay on each mode, are obtained from this framework as analytical expressions involving the aggregate demand and modal supply functions. Assuming that a modal investment improves quality and capacity, it is shown that the amount of users willing to travel at each perceived cost level increases on the improved mode, and decreases on the competing one. The explanation of the nature of the

modal demands allows for a clear measure of consumers' surplus variation when they shift, which provides a better basis for the use of the so-called rule-of-a-half as an approximate measure than the existing one. It is also shown that the conditions for uniqueness of the line integral expressing the generalized consumers' surplus variation are not fulfilled; as this latter has a unique value, only one integration path is correct.

Thesis Supervisor: Dr. Fred Moavenzadeh
Title: Professor of Civil Engineering

Thesis Supervisor: Terry L. Friesz
Visiting Assistant Professor
of Civil Engineering

ACKNOWLEDGEMENTS

I wish to thank my thesis supervisors, Professors Fred Moavenzadeh and Terry Friesz for their guidance and support throughout the course of this research.

I am extremely grateful to Professor Ann Friedlaender for her advice, orientation and encouragement.

Further thanks to Dr. Enrique Fernandez, with whom I had rewarding hours of technical discussion while working for the Egypt Project, and to Francisco Tourreilles and Bernard Francois, my roommates in Room 1-041, always ready to provide help and support. Enya Gracechild typed the final version of this thesis and Pat Vargas provided useful help; thank to them for their kindness.

This thesis really belongs to my parents, who always put the happiness of their children as first priority, and to my wife Momy and my son Pedro, who contributed with their love and support during all this period; the time I did not share with them is the real opportunity cost of this work.

TABLE OF CONTENTS

	<u>Page</u>
TITLE	1
ABSTRACT	2
ACKNOWLEDGEMENTS	4
TABLE OF CONTENTS	5
LIST OF FIGURES	7
CHAPTER 1. INTRODUCTION	8
CHAPTER 2. ECONOMIC SURPLUS AS BENEFIT MEASURE	12
2.1 The Concept of Consumers' Surplus	12
2.2 Producer's Surplus	16
2.3 An Application: Unimodal Transportation	18
CHAPTER 3. THE MULTI-MODAL CASE	23
3.1 Interrelated Demands	23
3.2 Existing Approaches	27
3.3 Some Comments	33
APPENDIX TO CHAPTER 3. Generalization of Consumer's Surplus	37
CHAPTER 4. THE DEMAND FOR TRANSPORTATION	39
4.1 The Aggregate Transportation Demand	39
4.2 Benefits on Production and Consumption	43
4.3 Modal and Aggregate Supply Curves	46
4.4 Derivation of the Modal Demands	49
4.5 Some Comments	52
APPENDIX TO CHAPTER 4. Derivation of the Slope of the Modal Demand Curves	54

TABLE OF CONTENTS (CONTINUED)

	<u>Page</u>
CHAPTER 5. BENEFITS OF A MODAL INVESTMENT	55
5.1 Supply Effects and Changes in Equilibrium	55
5.2 The Shift in Modal Demands	57
5.3 Measurement of Benefits: Average Cost Supply Curves Case	58
5.4 Comparison with Other Approaches	61
5.5 Extension of the Framework to Marginal Cost Supply Functions	64
APPENDIX TO CHAPTER 5. Direction of Shift in Mode 1's Demand Curve	69
CHAPTER 6. SYNTHESIS, EXTENSIONS AND CONCLUSIONS	72
6.1 Synthesis	72
6.2 On Extensions of the Approach	74
6.3 Equilibrium and Evaluation	77
6.4 The Income Effect	81
6.5 Final Comments and Conclusions	84
BIBLIQGRPAHY	88

LIST OF FIGURES

	<u>Page</u>
1. Compensating and Equivalent Variations	14
2. Marshallian and Compensated Demands	14
3. Economic Surplus	19
4. Total Surplus Variation After a Modal Investment	21
5. Shifts in Modal Demand Curves Due to Investment in One Mode	26
6. Consumers' Surplus Variation in the Inerrelated Modes Case	28
7. Mohring's Measure of Benefits After Improvement in Mode 1	32
8. Non-uniqueness of the Line Integral (8)	35
9. Derivation of the Demand for Transportation	41
10. Transportation Consumers' Surplus Equivalencies	45
11. A Simple Network	47
12. Modal and Aggregate Transportation Supply Curves	47
13. Derivation of the Modal Demands	51
14. Changes in the Modal and Total Supplies and New Equilibrium Levels After an Investment in Mode 1	56
15. Total and Modal Supply and Demand Curves Before and After an Investment is Made in Mode 1	59
16. Consumers' Surplus Variation Due to Improvement in Mode 1	62
17. Total Surplus Variation Under Marginal Pricing Policy	67
AP.5 Direction of Shift of Mode 1's Demand Curve	70
18. Three Markets Case	76
19. The Modal Demands Step by Step	82

CHAPTER 1. INTRODUCTION

Study and research on transport project evaluation techniques are acquiring more and more importance in this last decade, both in qualitative and quantitative terms. The justification for the increasing level of intellectual effort devoted to this subject relies on a number of interrelated facts:

- i) the substantial relative incidence of transport investment with respect to total public investment in developing countries (20 to 40 per cent), both to satisfy increasing levels of demand and to "promote development";
- ii) the important amount involved in loans by international agencies to Third World countries in the transportation sector, which provokes the necessity of demanding accurate feasibility studies;
- iii) the existence of a sometimes intuitive notion of the relation between transportation and economic development; in some author's opinion, this notion has caused overinvestment in the past;
- iv) the spatial dimension of growth, which has been explicitly taken into account by nearly all nations since the sixties;
- v) the availability of techniques for the forecast of demands and the estimation of equilibrium in transportation networks which provides a more reliable physical base for economic evaluation;

- vi) the possibility of an explicit transportation planning scheme in developing regions, which allows for a mixture of direct actions developed by the state and a set of indicative stimulus to guide the market forces toward an optimum.

All these facts have caused increasing interest in the problems of best allocation of resources in the transport sector.

From an economic or social perspective, two effects are central in the evaluation of an investment project: the change in utility level of the different groups affected and the change in real resources consumption level. The latter effect has been traditionally approached, in terms of measurement, by modifying the monetary value (price) given by the market to each unit of resource consumed through certain techniques which leads to the so-called shadow prices or "real" values; this constitutes a whole branch in economics. The evaluation of the former effect requires the conversion of changes in people's utility into monetary terms; the economic concept associated to this conversion is the so-called consumers' surplus.

There is more than one approach to the problem of turning utility variation into monetary units, as will be shown in Chapter 2. The traditional concept, i.e. the Marshallian consumers' surplus, is based on the concept of willingness to pay associated to the common market demand; its measure is unambiguous when dealing with one good in ceteris paribus conditions, but becomes less clear when related goods are taken into account, because generally a series of shifts in the demand curves will arise, difficulting the aggregate measure of willingness to pay.

In the transportation case it happens to be unclear, to say the least,

how to actually measure the benefits caused by changes in the network due to investment in one link, over the whole transportation network and the entire economy. This difficulty arises from the fact that demands for modes which are complements or substitutes are indeed interrelated and all links affected should be taken into account; partial analysis is not likely to be adequate. The correct measurement of benefits due to a modal investment in the presence of interrelated modal demands has been approached in different ways in the literature; different authors have given different answers to this problem. As we have pointed out in general, the main difficulties arise from the fact that the simple measure of consumers' surplus variation in the one mode case becomes unclear when willingness to pay depends on the conditions prevailing on other modes, because a modal demand shifts when changes occur in the supply of some mode involved. The main purpose of this work is to contribute to the analysis of this problem by looking at it from a different perspective that takes advantage from the fact that transportation demand is a derived one, in order to understand the joint behavior of modal demands and to propose a new method, or to support some existing one, for measuring economic surplus variation under different supply conditions, i.e., marginal and average costs schedules.

We have intended this work to be, in some sense, self-contained, devoting Chapter 2 to the presentation of the economic surpluses as measurements of benefits, focusing particularly on the problem of assigning monetary equivalencies to changes in the level of consumption and, implicitly, in utility. The unimodal case is presented there as an application.

The multimodal case, which introduces demand interrelation, is presented

in Chapter 3, where we also develop the three most interesting approaches to the subject, showing the discrepancies and similarities among them.

The theoretical derivation of an aggregate demand for transportation from the economic environment is performed in the fourth chapter, using this derivation to show also that the consumers' surplus associated to this transportation demand summarizes both producers and consumers' surpluses associated to the generating markets. Modal demands are obtained from this aggregate figure and the modal and aggregate supply curves, both graphically and analytically assuming only two modes. The explanation of the nature of this modal demand is the basis for the analysis of the shifts caused by an investment in one of the modes, which is developed in Chapter 5. The measurement of benefits (economic surplus) is obtained separately in terms of the aggregate, initial and final demands, providing a sounder theoretical basis for the so-called rule-of-a-half in its generalized version, than the previously existing for the case of average costs supply curves. The marginal costs case is also developed, showing that partial analysis is approximately correct.

Chapter 6 summarizes the theoretical approach developed from Chapters 2 to 5, adding some ideas in order to extend the approach toward a greater number of nodes (potential origins or destinations) or modes. A special section is devoted to present the relation among measurement of benefits and equilibrium, focusing particularly on the conditions for existence. Income effect implications are also stated in this chapter, which was intended not to "end" a work but to motivate further elaboration, although the main conclusions are also included.

CHAPTER 2. ECONOMIC SURPLUS AS BENEFIT MEASURE

Assume a price change takes place in the market for a good and no other market is affected. In general there will be a change in the utility level of the consumers and in the cost of production related to this good. In this chapter the concept of consumer and producers' surplus as a way to assign a monetary value to the change in welfare implied by a change in one market's equilibrium is presented. Both concepts are then applied to the unimodal transportation case, assuming an investment takes place causing a change in the short run supply function. All this gives a basis for the treatment of the multimodal case presented in Chapter 3 and analyzed in Chapter 5.

2.1 The Concept of Consumer's Surplus

As long as the quantity individually demanded for a good usually increases when the price of this good decreases, it makes sense to try to measure the consumer's gain in such a case since a greater amount consumed leads to a higher level of utility. This basic idea has led economists to create the concept of consumer's surplus; however, at least three approaches have been proposed in order to capture in monetary terms consumer's utility variation: the so-called Marshallian consumer's surplus, and Hicksian compensating and equivalent variations.

The most widely used concept of consumer's gain (or loss) due to changes in one good's price is the Marshallian consumer's surplus, which relies on the concept of a market demand as the function that relates different levels of consumption to the amount an individual is willing to

pay for it. Formally, define U as utility level; x_i amount of good i consumed; P_i price of good i (willingness to pay) and Y income; then an individual's behavior is assumed to be:

$$\begin{aligned} & \text{Max } U(x_1, x_2, \dots, x_n) \\ & \text{subject to } \sum_{i=1}^n p_i x_i \leq Y \end{aligned}$$

The solution to this problem will give $x_i = X_i(p_1, p_2, \dots, p_n, Y)$ which is the Marshallian demand for good i . Marshallian consumer's surplus at a price level p_i^0 is defined as:

$$\text{MCS} = \int_0^{x_i} p_i dx - p_i^0 x_i \quad (1)$$

and is trying to capture the excess payment an individual would have been prepared to make with respect to the actual expenditure on this good.

Given this definition, a simple graphical representation is given by the area under the demand curve, above the price level p_i^0 . Thus, the variation in consumer's surplus caused by a ceteris paribus change in price of x_i from p_i^0 to p_i^1 is given by

$$\Delta \text{MCS} = - \int_{p_i^0}^{p_i^1} x_i dp_i \quad (2)$$

Hicksian compensating and equivalent variations were conceived primarily as income equivalencies to changes in utility due to changes in the price of one commodity. Following Figure 1, an individual whose income is Y_0 buys x_1^0 at a price p_1^0 reaching a utility level U_0 (point A). If p_1 falls to p_1^1 , the individual can reach a utility level U_1 buying x_1^1 (point C).

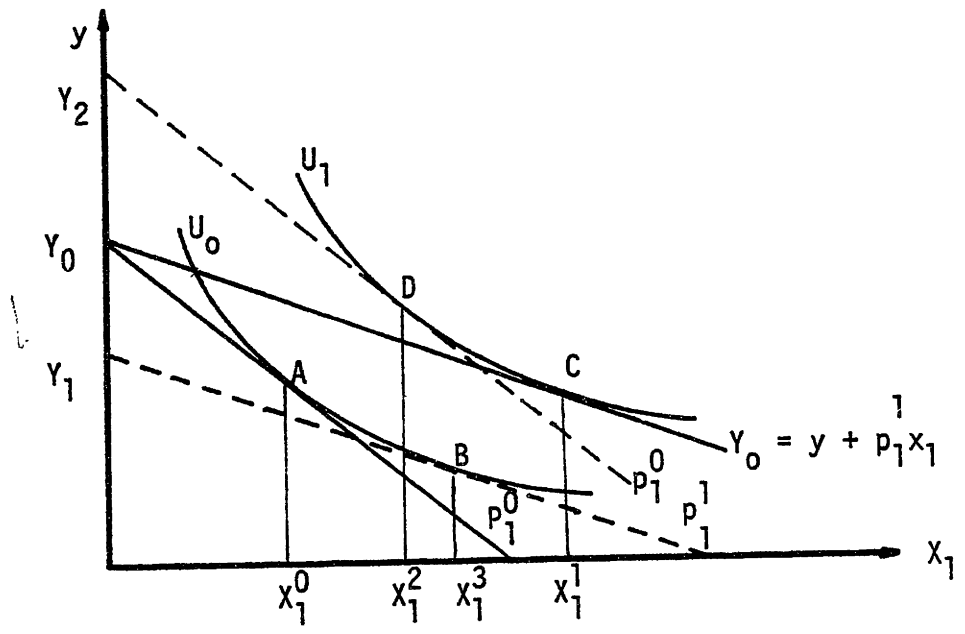


FIGURE 1: Compensating and Equivalent Variations
 $(y = \sum_{i \neq 1} p_i x_i)$

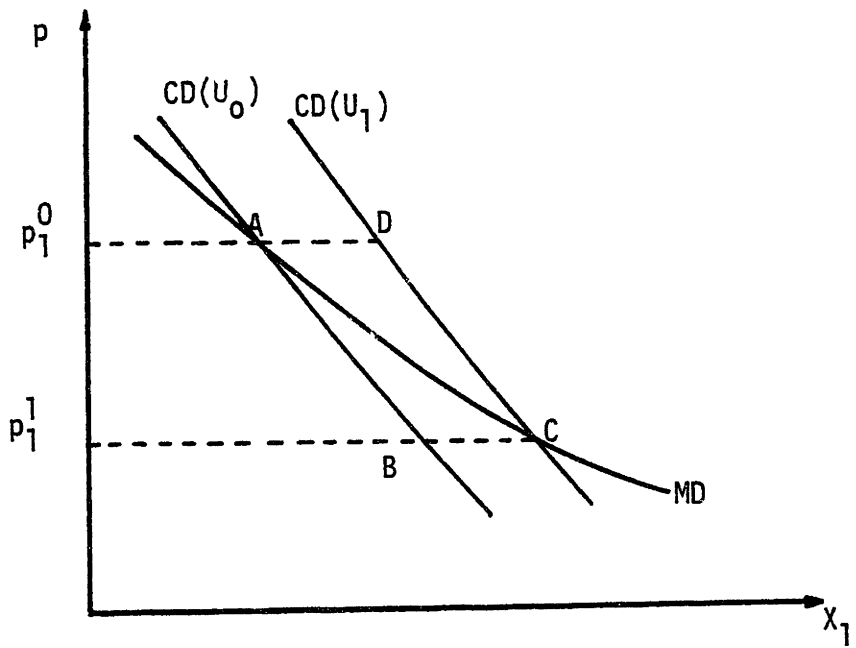


FIGURE 2: Marshallian(MD) and Compensated (CD) Demands

The level U_1 can be reached also by keeping price at p_1^0 and increasing income to Y_2 ; the amount $Y_2 - Y_0$ is called equivalent variation (EV). On the other hand, at a price p_1^1 a change in utility level from U_0 to U_1 is equivalent to a change in income from Y_1 to Y_0 ; the amount $Y_0 - Y_1$ is called compensating variation (CV).

There is general consensus in accepting both compensating and equivalent variations as the best theoretical monetary measures of a change in utility level due to a change in some good's price. However, in general EV and CV will be different for the same change in price, unless the income effect is zero.¹ Both variations can be also represented in terms of areas under demand curves, but in reference to what is known as a compensated demand, which relates quantity consumed (demanded) with different levels of price but allowing for income to vary in such a way that a given utility level is maintained. This way, each indifference curve generates one compensated demand curve. Thus, pairs (x_1^0, p_1^0) and (x_1^3, p_1^1) belong to the compensated demand associated to U_0 ; similarly (x_1^1, p_1^1) and (x_1^2, p_1^0) belong to the compensated demand related to U_1 .

Figure 2 shows the corresponding compensated demands at U_0 and U_1 levels, the former having a common point with the Marshallian demand at A and the latter at C. It can be shown that CV is equal to the area $P_1^0 A B P_1^1$ and that EV can be measured as the area $P_1^0 D C P_1^1$. Therefore, although both the CV and EV have a very good theoretical base to represent in monetary terms the gain in utility due to a decrease in price, the measures differ and there is no reason to prefer one over the other. In this sense, the

¹That means both points A and D correspond to x_1^0 and both points B and C to x_1^1 in Figure 1.

variation in the Marshallian consumer's surplus can be used as good surrogate or compromise approximation to a single measure of the benefits, since $CV \leq MCS \leq EV$ in this case; it should be noticed that the opposite inequalities hold for a price increment. Equalities hold when the income effect is zero. Note that this whole reasoning assumed constant income (Y_0) and, therefore, a stable Marshallian demand.

Assume now that every individual change in utility due to a ceteris paribus change in one good's price is equally important and that utilities are independent among individuals; then the aggregate variation in the Marshallian consumer's surplus, which can be called consumers' surplus, can be obtained from the market demand curve (which is the horizontal summation of individual demands) using the corresponding measure given in (2). Similarly, the aggregate compensating variation can be derived from the aggregate compensated demand which is in turn obtained by horizontally summing the compensated demands derived for the initial utility level of each individual. The aggregate equivalent variation can be obtained in a similar fashion, regarding the final levels of satisfaction. It should be noticed that in the aggregate case, a zero income effect will cause the market and compensated demands to coincide, but this is not a necessary condition for coincidence to happen.

2.2 Producer's Surplus

What is generally understood as producer's surplus is the area above the supply curve below the prevailing price. From an individual competitive producer's point of view, this area represents the net benefits in the short run, since the supply curve coincides with the marginal cost curve and therefore:

$$PS = p_i^0 x_i - \int_0^{x_i} MC dx \quad (3)$$

where PS is producer's surplus, p_i^0 the selling price, x_i the quantity produced, and MC the marginal costs. In the long run the competitive firm will sell at a price equal to the minimum average cost and therefore the PS vanishes.

If firms in a certain market are under competition but have different cost structures, then even in the long run a producer's surplus appears, since the so-called infra-marginal firms² make a profit; in this case, the measurement of the PS is the same as in (3) and the aggregate producers' surplus can be obtained in a similar form from the market supply curve.

It is worthwhile to make some brief comments about the notion of producer's surplus. First, the PS accruing to a perfectly competitive firm in the short run is equivalent to the quasirent obtained due to the existence of fixed factors of production. Second, in those cases when, for some reason, the market supply curve is the average cost curve, then the PS vanishes.³ Third, in an imperfect market some firms may receive a surplus in the long run due to some degree of market power. Because of these and several other reasons, some authors have preference to call PS an economic rent, which seems perfectly reasonable.

When both consumers' and producers' surpluses are taken into account,

²Those firms which have a minimum long-run average cost less than the prevailing (equilibrium) market price.

³This may be the case when externalities appear, e.g., when the increment in production by one firm increases the cost of the others; then the individual marginal cost equals the average cost at a market level.

the addition of (1) and (3) gives

$$\text{MCS} + \text{PS} = \int_0^{x_i} p_i dx - \int_0^{x_i} \text{MC} dx \quad (4)$$

which represents the consumers' valuation of the quantity being bought minus the total cost of producing it; (4) is graphically given by the area between demand and marginal cost curves from 0 to x_i^0 , as shown in Figure 3. Note that the total surplus can be obtained by either adding MCS and PS, or subtracting from the total willingness to pay (area under the demand curve) the total cost (area under the supply curve).

2.3 An Application: Unimodal Transportation

Transportation supply is characterized by an unusual fact, that is, one of the factors of production, namely the travel time, is provided by the "consumers" (passenger or goods). In the short run transportation facilities can be considered as fixed, so the variable costs can be simply stated as the sum of operating costs and time-related cost. A great amount of literature has shown the existence of what is called transport congestion, a short run phenomenon, which means that above a certain level each new user will increase the cost for each other user, that is, average cost (AC) is an increasing function of flow.⁴ This particular case of externalities determines then that the "market" supply curve, the one that gives the "price" actually charged at each level of "production", is the average cost curve. In this case, each additional user is not paying the additional

⁴This is particularly well stated in Beckman, McGuire and Winsten (1956).

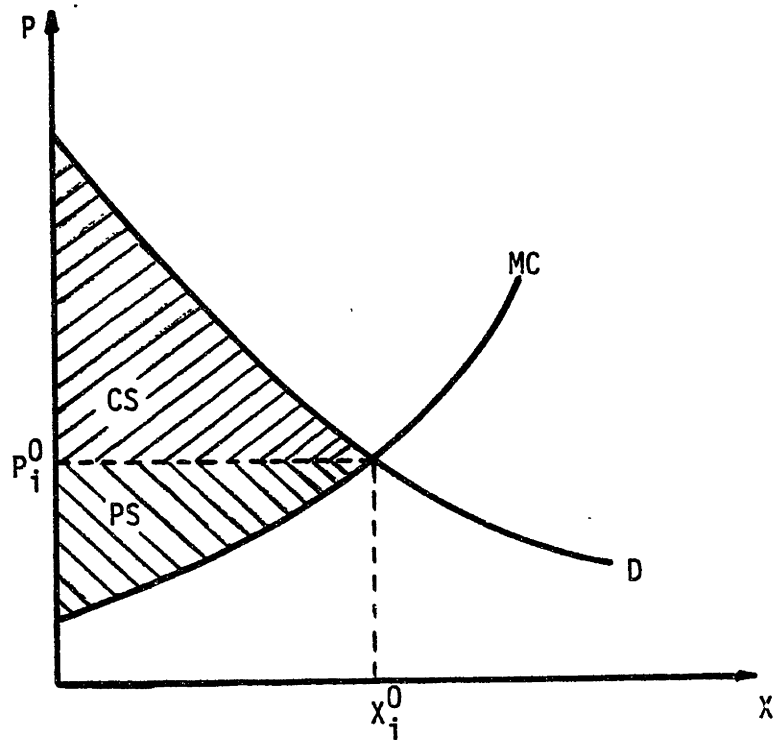


FIGURE 3: Economic Surplus

cost he causes to the other users, namely the marginal cost (MC), but less than that. From a resource allocation point of view, it would be desirable to overcome the situation; one way to do it is to charge a toll such that at the equilibrium flow the marginal cost prevails. As a toll equal to the difference between marginal and average cost can be adjusted at every level of users, the "supply" curve becomes the marginal cost curve. Of course, if congestion does not occur, unit transport cost is constant and both cost curves coincide.

Assume transportation demand between two points is known and only one mode with given facilities' characteristics is available to satisfy it. Assume also that the average cost function including time cost and operating cost is known. If no tolls are charged, then the producers' surplus as was defined in the last section is zero at every equilibrium point. In this case, the benefit variation (total surplus variation) after an investment that changes the short run average cost curve, will be given by the variation in the consumers' surplus. If "ideal tolls" were charged, the definition of the producers' surplus would lead to an amount exactly equal to the total tolls collected.⁵ Total surplus variation after the modal investment will be given by consumers' surplus variation plus variation in toll revenues. Figure 4 shows both cases. The change in economic surplus in the second case (b) can be explained through the increment in producers' surplus, which is equal to the sum of areas 1 and 2, and the increment in producers' surplus, which varies from 1 plus 3

⁵Define T as the ideal toll, C total cost. Then $T = MC - AC$; if q is the equilibrium flow, toll revenues = $Tq = (MC - AC)q = MCq - C = PS$. This is equivalent to the quasirent due to a fixed factor, in this case the facilities' characteristics.

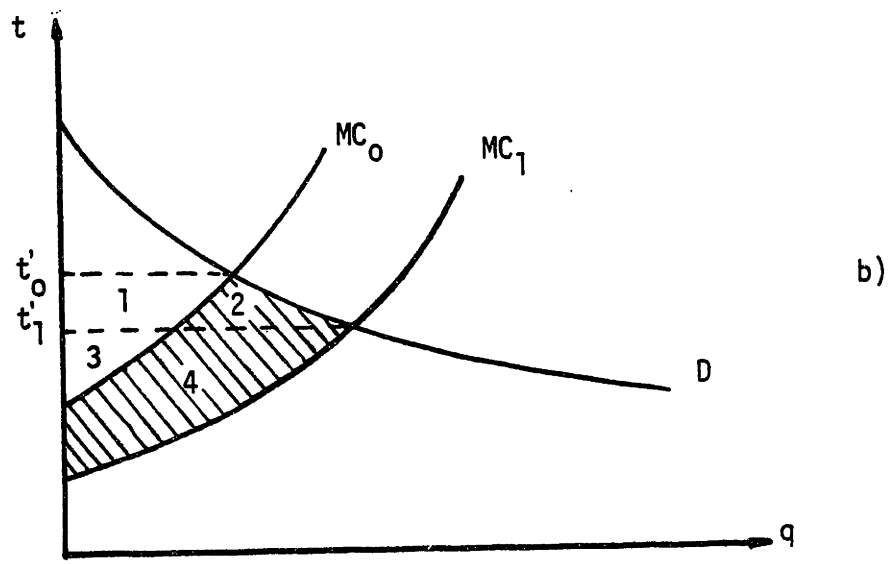
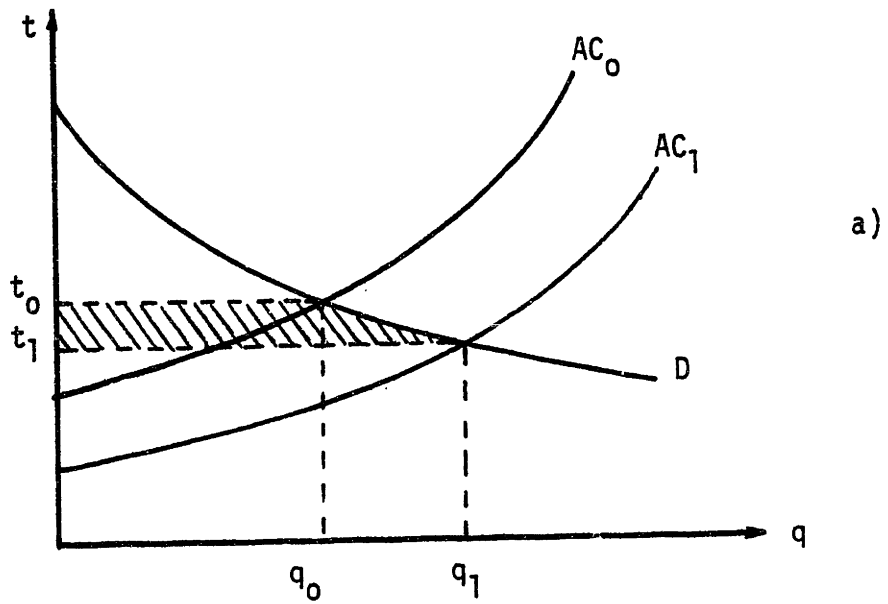


FIGURE 4: TOTAL SURPLUS VARIATION AFTER A MODAL INVESTMENT
 a) NO TOLLS CHARGED; b) IDEAL TOLLS CHARGED

to 3 plus 4 and, therefore, equals the difference between areas 4 and 1. The addition of both variations is equal to the sum of the shaded areas 4 and 2.

CHAPTER 3. THE MULTIMODAL CASE

The main characteristic of the multimodal case, namely, that situation in which more than one mode is available, is that each modal demand is a function not only of the modal price but also of the price prevailing in all alternative modes. For this case, the surplus measurement is not as simple as in the unimodal case. In this chapter the problem of two modal interrelated demands is presented and a summary of the existing approaches to the measurement of benefits after a modal investment is given, showing that some differences and incompatibilities exist among them.

3.1 Interrelated Demands

The measurement of the variation in consumers' surplus considering all related markets constitutes a problem of aggregation over commodities. It is intuitively clear that a prohibitive increment in one good's price, say tea, would hurt people more if there were no substitutes for it than if say coffee could be bought as an alternative. In general, "if the commodity in question has close substitutes or complements on the demand or supply side, a partial analysis is not likely to be adequate. For example, if a commodity A has a close substitute, B, on the demand side, then a change in price of A will shift the demand curve for B, if the market price of B changes, then producers of B will be affected. This ought to be taken into account. In addition, the demand curve for A, constructed on the assumption that the price of B is given, will also shift. If this leads to a further change in the price of A, the demand curve for B will shift once again and so on until equilibrium is restored" [Currie et. al., 1971].

Thus, transportation demands for modes which are complements or substitutes are indeed interrelated. The existing literature on this subject has focussed on the competing modes case, or strictly speaking, analyzes the case of only one mode with two alternative paths. In fact, the problem is conceptually the same in both cases if the price is treated in an appropriate way. The basic idea of the existing approaches is that the ordinary area of concept of consumers' surplus is no longer valid in the case of interdependent transportation demands, because the willingness to pay for one mode depends on the prices of related modes; in the literature modal demands are stated as a function of the perceived costs t_i on each alternative mode. This perceived cost, also called inclusive or generalized cost, is assumed to take into account all kinds of factors in addition to operating costs and value of time in such a way that comparison among modes is possible in equivalent units. Let us call Q_i the flow on each mode and assume only two modes for simplicity. Then:

$$Q_1 = D_1(t_1, t_2) \quad \text{and} \quad Q_2 = D_2(t_1, t_2) \quad (5)$$

are the travel demand functions. Assume also two supply functions that relate perceived cost to volume on each mode:

$$Q_1 = S_1(t_1) \quad \text{and} \quad Q_2 = S_2(t_2) \quad (6)$$

Then it becomes clear that the demand curves, seen as $Q_i = F_i(t_i) = D_i(t_i, t_j)$, will shift when investment in one mode takes place. Let us denote by superscripts 1 and 0 the values of t_i and Q_i with and without the project respectively. Traditionally, consumers' surplus variation

due to an investment in mode i would be stated as

$$\Delta CS_i = \int_{t_i^0}^{t_i^1} -F_i(t_i) dt_i \quad (7)$$

but in this case this measure loses sense, since a change in the initial equilibrium state in mode 1 from t_1^0 to $t_1^0 - \Delta t_1$ due to an improvement, i.e. a change of s_1 to s_1^1 with $s_1^1(t_1) > s_1(t_1) \forall t_1$, is going to be reflected in a change in $D_2(t_1^0, t_2)$ to a new demand curve $D_2(t_1^0 - \Delta t_1, t_2)$, causing a new equilibrium $t_2^0 - \Delta t_2$ which will affect Q_1 and so forth till equilibrium is reached. The initial and final states are shown in Figure 5, assuming an investment takes place in mode 1.⁶ It should be noticed that in the case of a perfectly elastic supply function in mode 2, the perceived cost in the improved mode will change (t_1) causing a shift in mode 2 demand curve, but since t_2 will remain constant, the demand for mode 1 will not shift. In other words, the shifting sequence arises only in the case of congestion in both modes.

The question then is what is the correct measurement of the variation in the economic surplus, in particular in the consumers' surplus since producers' surplus measure presents no theoretical problems when supply functions depend only on each mode's perceived cost.

⁶We will show in Chapter 5 that the final shifts in the two modes case are those indicated in Figure 5, this is, the demand for the improved mode shifts upward and the demand for the competing mode shifts downward; this contradicts Williams' "perverse demand curve effect" (1976) and Mohring (1976), and is in accordance with Agnello's graphical representation (1977).

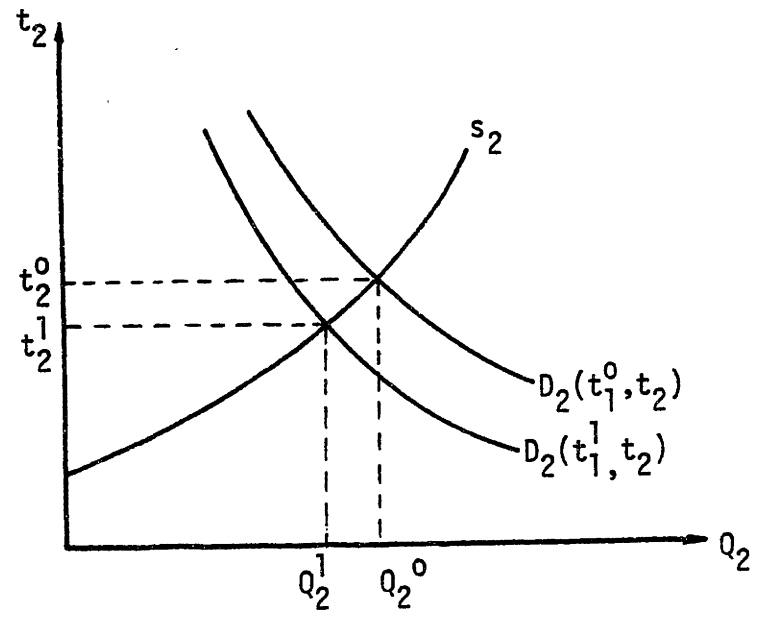
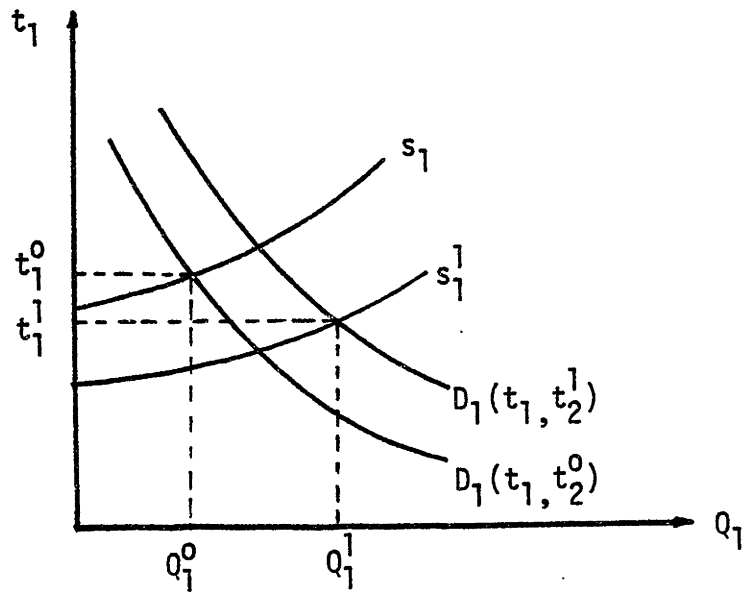


FIGURE 5. Shifts in Modal Demand Curves due to Investment in One Mode

3.2 Existing Approaches

Williams (1976) uses Hotelling's (1938) extension of the Marshallian measure to a general equilibrium context, in which consumers' surplus variation is given by a line integral which in the case of two modes has the form:

$$\Delta CS = \int_{(t_1^0, t_2^0)}^{(t_1^1, t_2^1)} \sum_{i=1}^2 D_i(t_1, t_2) dt_i \quad (8)$$

The value of this line integral depends on the path of integration, which corresponds to the consumers' continuous adjustment to changes in t_1 and t_2 from (t_1^0, t_2^0) to (t_1^1, t_2^1) . The rule of a half,

$$\Delta CS \approx 1/2 \sum_{i=1}^2 (Q_i^0 + Q_i^1)(t_i^0 - t_i^1) \quad (9)$$

commonly used as a measure of perceived user benefit's variation, is recognized by Williams as the result of evaluating the generalized surplus (8) along a linear path obtained through the Taylor expansion of each demand function neglecting terms in the integrand of second and higher order, that is, neglecting curvature effects along the path. Williams' central proposition is the use of higher order terms in the Taylor series expansion when changes in the Q_i are nonmarginal, that is, when linearization may not be a good approximation. In any case, linear or not, the rule-of-a-half gives a measure of the consumers' surplus variation equal to the sum of the shaded areas of Figure 6a. It is important to point out that the previous justification of this rule is somewhat intuitive, in

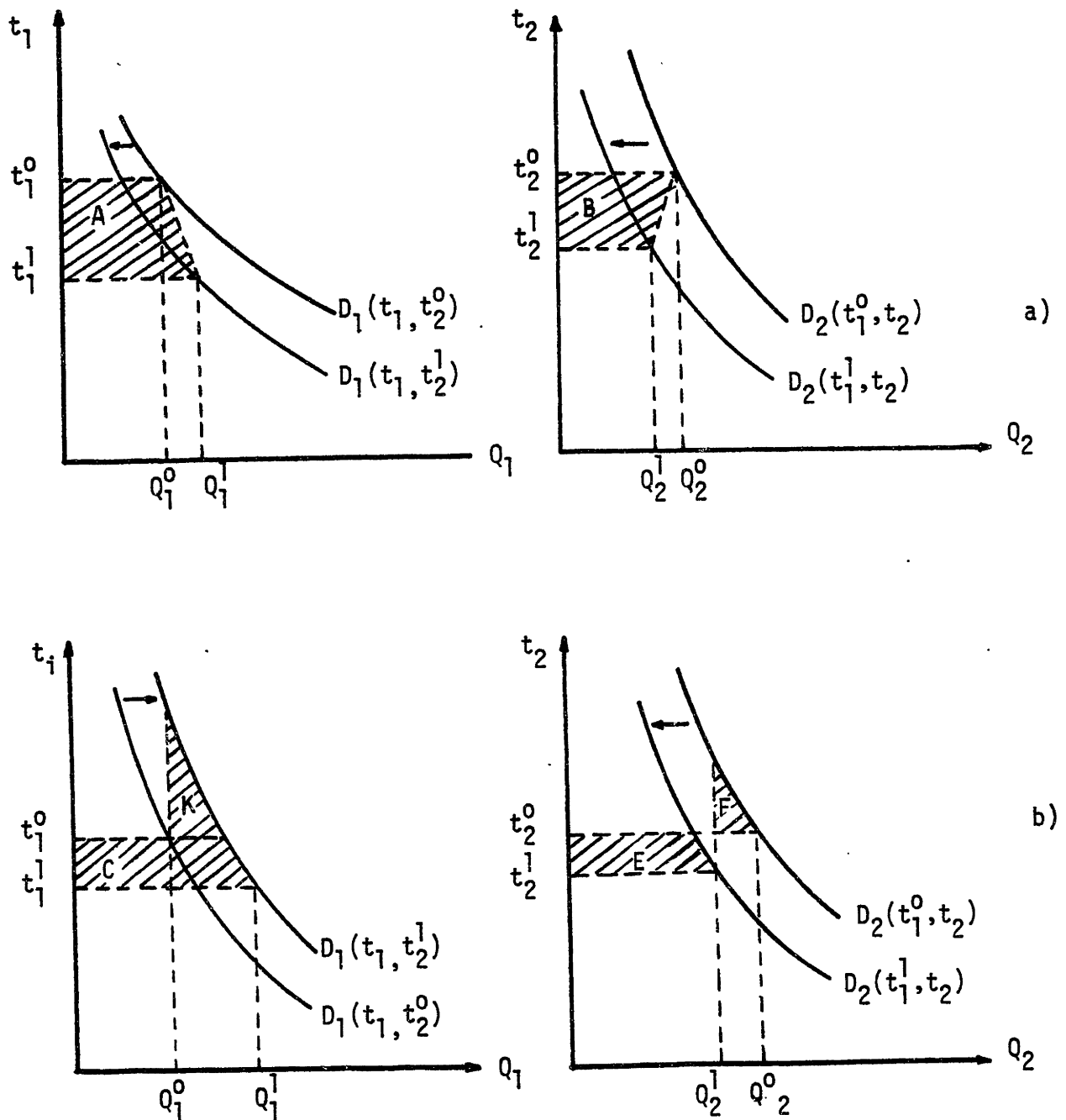


FIGURE 6. CONSUMERS' SURPLUS VARIATION IN THE INTERRELATED MODES CASE

(a) Rule-of-a-Half, William's version (1976) = A+B

(b) Agnello's approach (1977): C+E-F+K

the sense that it is supposed to take into account the full benefit of the reduction in perceived cost by the remaining or "old" users of each mode, that is $(t_i^0 - t_i^1) Q_i^0$, plus the average of the maximum and minimum "possible benefits" for the users who shift, this is, $1/2 (t_i^0 - t_i^1)(Q_i^1 - Q_i^0)$. See, for instance, Neuberger (1971) for a justification of the use of this rule in the case of non-fixed total number of trips.

Agnello (1977) developed another approach starting with the original formulation of Hotelling (1938), who stated that in the case of n related commodities "the natural generalization of the integral representing total benefit, of which consumers' surplus is a part", is a line integral whose limits of integration are the initial and final quantity vectors. In our notation, Agnello's formulation of the "joint valuation" is

$$B = \int_{(Q_1^0, Q_2^0)}^{(Q_1^1, Q_2^1)} \sum_{i=1}^2 D_i^{-1}(Q_1, Q_2) dq_i \quad (10)$$

Again the value depends on the path of integration except when assuming identical income elasticities of demand for the highways whose demands are interrelated.⁷ This condition is assumed and the line integral is reduced to the sum of ordinary integrals by taking the path of integration along quantity axes.

⁷In general, a line integral $\int_{x^0}^{x^1} \sum_{i=1}^n f_i(x_1, x_2, \dots, x_n) dx_i$ has a unique value, i.e., is independent of the path of integration, under the necessary and sufficient integrability conditions $\frac{\partial f_j}{\partial x_i} = \frac{\partial f_i}{\partial x_j} \quad \forall i, j = 1, \dots, n.$

Then, Agnello reduces (10) to:

$$B = \int_{Q_2^0}^{Q_2^1} D_2^{-1}(Q_2, Q_1^0) dQ_2 + \int_{Q_1^0}^{Q_1^1} D_1^{-1}(Q_1, Q_2^1) dQ_1 \quad (11)$$

Then the following adjustments are made to get the net user benefits:

(i) adding the costs no longer incurred by users diverted from mode 2 to mode 1, this is, $(Q_2^0 - Q_2^1) t_2^1$;

(ii) adding consumers' surplus increment "of those still using highway (mode) 2", measured over the final demand curve for mode 2;⁸

(iii) adding the cost reduction of mode 1's old users, this is, $(t_1^0 - t_1^1) Q_1^0$;

(iv) subtracting the cost payments of additional users of mode 1, this is $(Q_1^1 - Q_1^0)t_1^1$.

After all these adjustments, the variation in consumers' surplus after a modal investment would be given by the shaded areas of Figure 6b, which summarizes Agnello's propositions.

The approach developed by Mohring (1976) explicitly includes supply curves for each mode (in fact he considers competitive roads) analyzing the effects of a modal investment in two cases:

(i) no tolls are charged and, therefore, the supply curves represent average cost (see 2.3);

⁸It is not clear why this method of addition is used.

(ii) ideal tolls are charged and, therefore, the supply curves represent marginal costs.⁹

In Mohring's first case above, which is similar to Williams' and Agnello's, there is no variation in the producers' surplus and the benefits attributable to the project are only those represented by the consumers' surplus variation. Mohring states that, as the perceived costs are inter-related, it is possible to express the perceived cost in mode 2, t_2 , as a function of the perceived cost in mode 1, t_1 . Therefore, $D_1(t_1, t_2)$ can be expressed as $D_1[t_1, t_2(t_1)]$ which represents a curve in the (Q_1, t_1) space passing through the initial and final equilibrium points in mode 1. The corresponding $D_2[t_1, (t_2), t_2]$ coincides with the average cost curve (supply) of mode 2 (AC_2) joining the two equilibrium points. Consumers' surplus variation is measured with respect to these two modified demands. In the marginal cost curves case, a similar construction is used to measure consumers' surplus variation, and the total tolls collected (which are equivalent to the producers' surplus measure as shown in 2.3) are added; toll collections on mode 2 reduces exactly in the amount the surplus with respect to $D_2[t_1(t_2), t_2]$ increases. Therefore, Mohring concludes, "the benefit equals the sum of the change in consumers' surplus and toll revenues on the improved link," ignoring traffic effects on other links. Benefits are shown in Figure 7 for the average (a) and marginal (b) cost cases. Note that Mohring's measure of benefits in the first case gives a justification for the rule-of-a-half as an approximate measure, i.e. $A + B$ (Figure 6a) $\approx A' + B'$.

⁹In fact, a third case is analyzed by Mohring which we omit from this review: the inefficient tolls case; i.e., tolls are charged such that the "supply price" actually paid is greater than the average costs but less than the marginal costs.

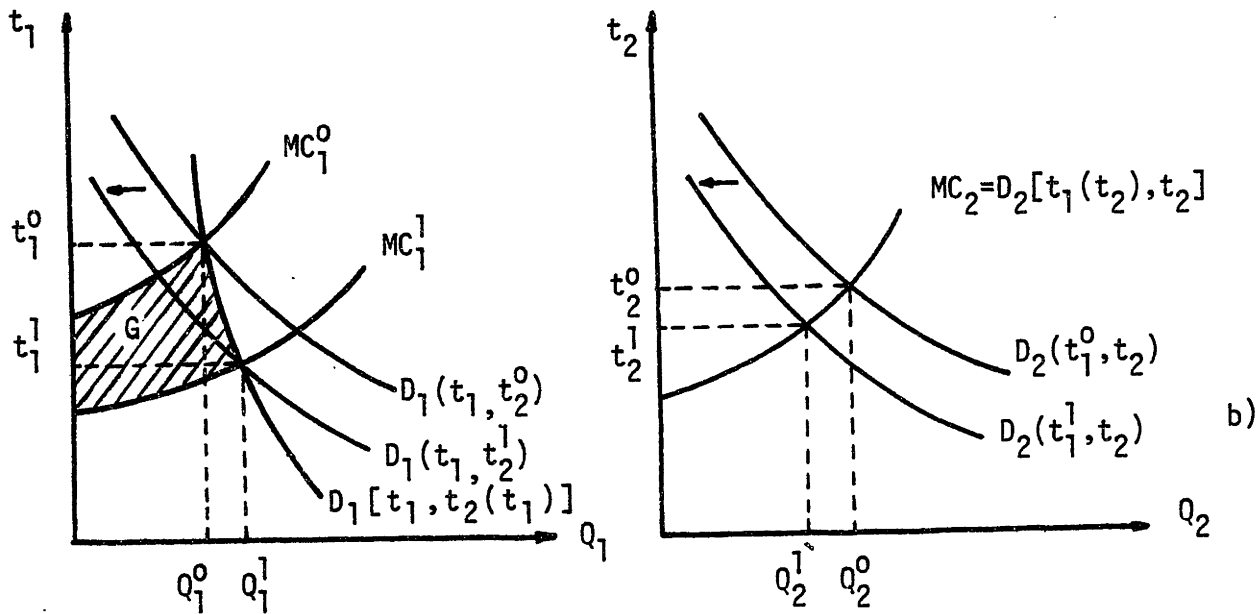
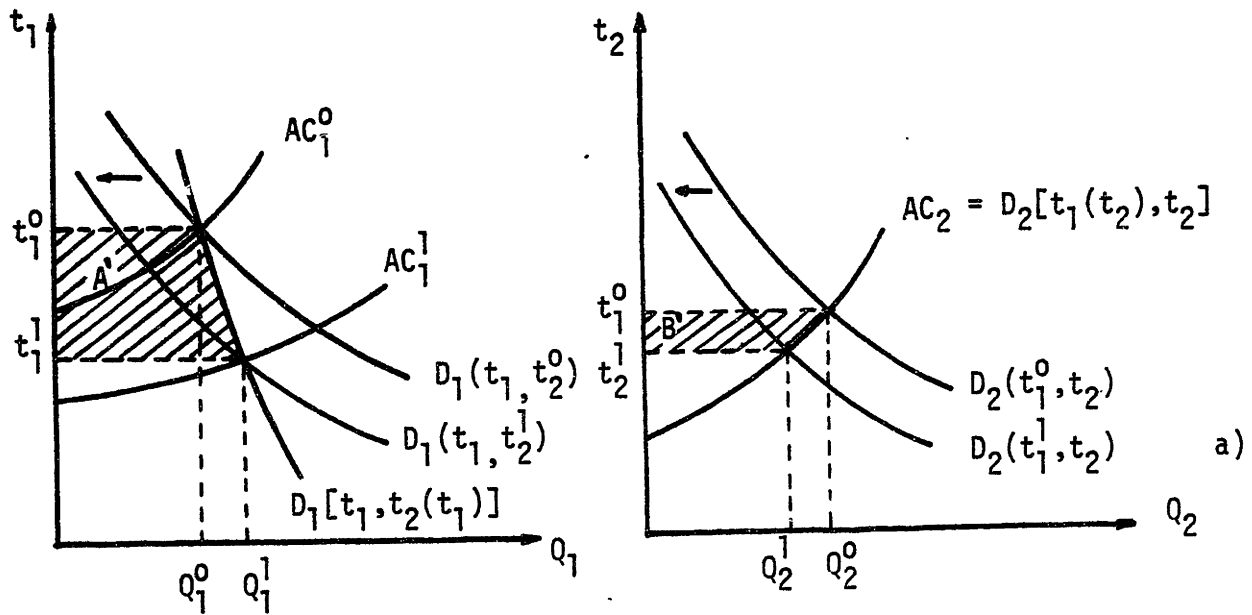


FIGURE 7. MOHRING'S MEASURE OF BENEFITS AFTER IMPROVEMENT IN MODE 1

a) Average Cost Supply Curves ($A' + B'$)

b) Marginal Cost Supply Curves. (G)

3.3 Some Comments

Now that we have introduced the problem of how to measure benefits due to modal improvements in the interrelated modal demands case, and presented the three most interesting approaches to solve it, it is possible to make some comments in order to point out the differences, contradictions and, if any, complementary or common aspects of these approaches with the intention of clarifying the main ideas and to facilitate the work toward a unified (and hopefully correct) approach.

A first, and at this point obvious, fact is that the existing approaches differ at least in two ways:

- (i) the assumed direction of the shifts in the modal demands differ among authors and, moreover, is not adequately justified by any of them;
- (ii) the proposed measure of benefits after a modal improvement is also different.

These two divergences can be easily seen from the comparisons among figures 6 and 7, in which we have kept the original propositions as presented by the authors, although using a common notation in all figures. There are also some points in common among the three approaches, points that are implicitly stated in the theoretical treatment, although having a clear but intuitive justification. These are:

- a) a reduction in the equilibrium perceived cost in both modes, and
- b) an increment in Q_1 and a reduction in Q_2 after an improvement in mode 1.

It is surprising to note that the direction of shifts of the modal demands is just assumed without further justification. This fact happens to be important in the following way; if the condition for the line integral (10) to have a unique value is assumed, that is,

$$\frac{\partial D_1^{-1}}{\partial Q_2} = \frac{\partial D_2^{-1}}{\partial Q_1} \quad , \quad (12)$$

then if D_i is continuously differentiable it also holds

$$\frac{\partial D_1}{\partial t_2} = \frac{\partial D_2}{\partial t_1} \quad (13)$$

It means that, in general, the line integral (8) will also have a unique value when (12) holds and, therefore, any path of integration should lead to the same value of ΔCS . However, under the assumption of modal demands shifting in different directions ("a la Agnello") this line integral has not a unique value as easily verified following Figure 8, in which two seriatim paths of integration have been chosen; the first one hold t_1 constant at the initial level t_1^0 while varying t_2 from t_2^0 to t_2^1 , then t_1 varies from t_1^0 to t_1^1 holding t_2 constant at the final level t_2^1 . The second path is constructed in a similar fashion but varying t_1 first. Therefore, if the direction of shift of the modal demands "a la Agnello" happens to be correct, the condition of integrability does not hold.

It is important to recognize also that the kind of function used by Mohring in order to express t_i as $t_i(t_j)$ allowing for the construction of a "real" demand curve, is equivalent to the choice of a single path of integration in (8), because it expresses the continuous adjustment in

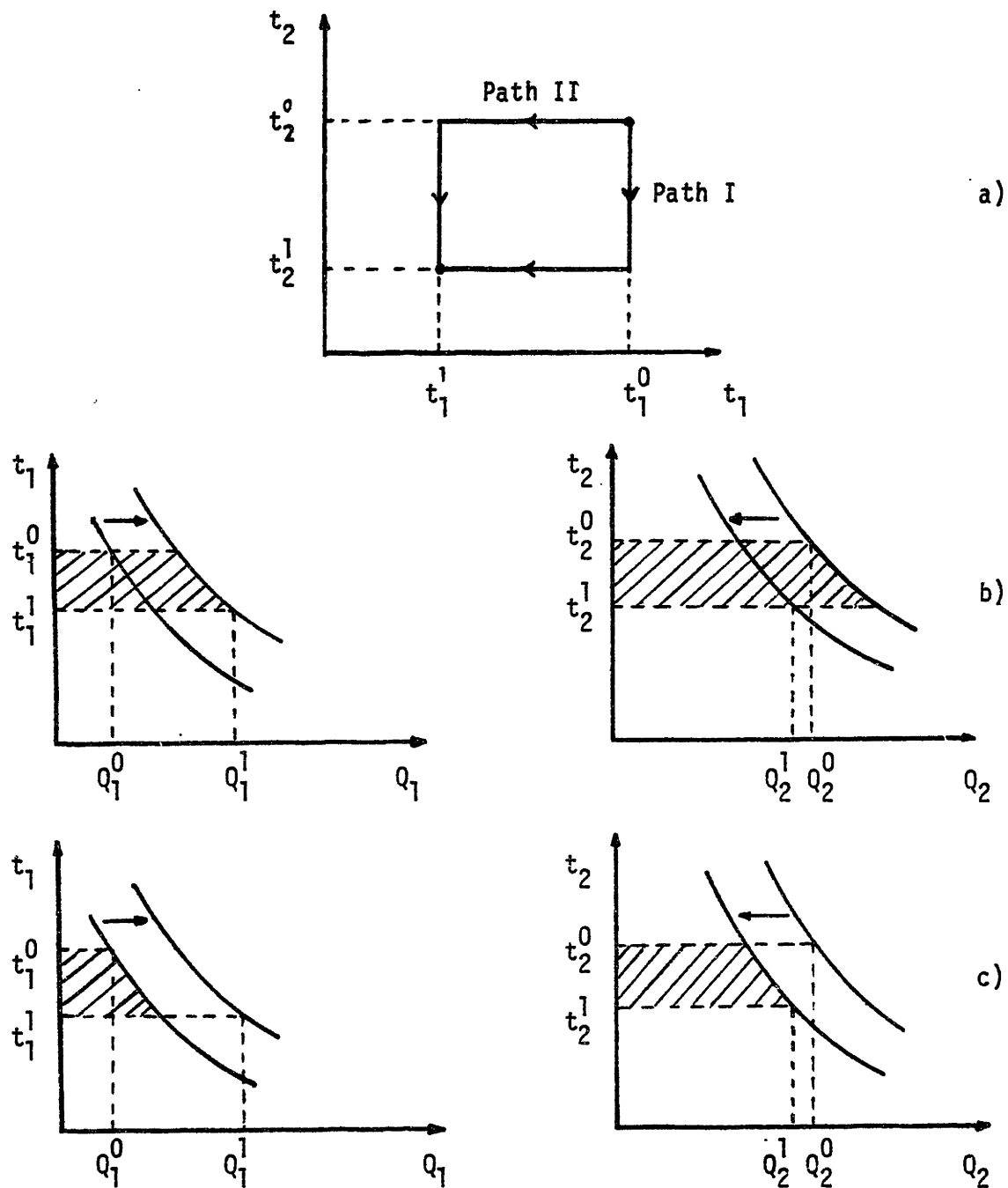


FIGURE 8. Non-uniqueness of the Line Integral (8)

- a) paths of integration
- b) value of the line integral following path I.
- c) value of the line integral following path II.

perceived costs from (t_1^0, t_2^0) to (t_1^1, t_2^1) as a single curve in the (t_1, t_2) space.

Another necessary comment is related to the last one in the sense that a continuous adjustment in terms of relative perceived cost between the two modes, defines a relation that involves not only people's reaction but also the "answer" of the transportation system in terms of variation in the perceived cost due to supply effect, especially when congestion takes place. In other words, it seems that certain behavioral assumptions (a properly demand phenomenon) plus supply characteristics have to be taken into account at the same time in order to get the modal demands. Following this idea, it does not seem to make much sense to "hold t_i constant", in order to get the demand for mode j as $D_j(t_i, t_j)$ because it would be equivalent to obtain a modal split assuming an infinitely elastic supply function for mode i .

Our final comment refers to the nature of transportation demand. As said, it is by definition a derived demand, but no approach takes into account this fact in order to understand the simultaneous behavior of the modal demands although some implicit statements suggest the existence of a "common demand schedule", as can be deduced from different authors when they talk about "travellers who shift" and "new travellers". We hope to clarify all these aspects in the following chapters.

APPENDIX TO CHAPTER 3. Generalization of Consumers' Surplus

In Chapter 2 we have pointed out that total surplus can be obtained in two different ways:

- i) adding consumers' and producers' surplus;
- ii) subtracting total costs from the "total valuation of consumption."

This total valuation or total benefit is given by the area under the commodity's demand curve. Of course, consumers' surplus is part of this area, the difference being actual payments made by the consumers. When consumption changes from X_i^0 to X_i^1 , total valuation change in

$$\int_{X_i^0}^{X_i^1} P_i(x_i) dx_i \quad (a)$$

assuming ceteris paribus conditions. When trying to extend this concept to n goods, Hotelling (1938) stated that "the natural generalization of the integral representing total benefit is the line integral"

$$\int_{\bar{X}^0}^{\bar{X}^1} (P_1 dX_1 + P_2 dX_2 + \dots + P_n dX_n) \quad , (b)$$

where P_i represents willingness to pay or the inverse of the demand function in terms of all goods consumed, that is,

$$P_i = f_i(X_1, X_2, \dots, X_n) \quad . \quad (c)$$

It should be noticed that income is implicitly assumed constant.

If the variation in actual payments made by consumers is subtracted from (b), the resulting expression would represent a generalization of consumers' surplus variation. It is worthwhile to point out that Hotelling did not present such an expression, but rather the variation in total surplus by replacing P_i in (b) by

$$P_i \rightarrow f_i - g_i \quad , \quad (d)$$

where g_i is the marginal costs function for production of X_i .

CHAPTER 4. THE DEMAND FOR TRANSPORTATION

It is universally accepted that transportation demand is a derived one; but, unfortunately, this basic and simple fact is frequently forgotten when trying to obtain such a demand or to understand theoretically its interrelation with the economic environment. This happens to be a crucial point in the development of a conceptual framework in order to better understand the interrelated modal demands problem. In this chapter we derive a stable transportation demand curve from the long-run economic characteristics prevailing at two points (markets) in the space, showing later that the benefits caused by transportation on production and consumption are summarized by the transport market. Equilibrium in this market and in the overall economy is obtained by constructing modal and aggregate supply curves for the two-modes case. In the final section we show that modal demands can be obtained from this framework.

4.1 The Aggregate Transportation Demand

Let us consider one good X produced and consumed at two spatially separated markets M_1 and M_2 under competitive conditions. Because of technological differences and amount of available resources, the aggregate supply curve for X will be generally different at each market; the same reasoning holds for the demand for X , because of different individual income and tastes. This leads to different equilibrium points at M_1 and M_2 if considered in isolation. Let us define

- P = price of X ; q = quantity of X
 P_i = equilibrium price at M_i isolated
 ψ_i = demand for X as a function of P at M_i
 Ω_i = long run supply of X as a function of P at M_i
 ES_i = excess supply of X at M_i
 ED_i = excess demand for X at M_i
 \bar{Q} = quantity of X transported between the two markets
 P_i^* = the price of X at M_i when \bar{Q} is transported between the two markets
 t = unit transportation cost.

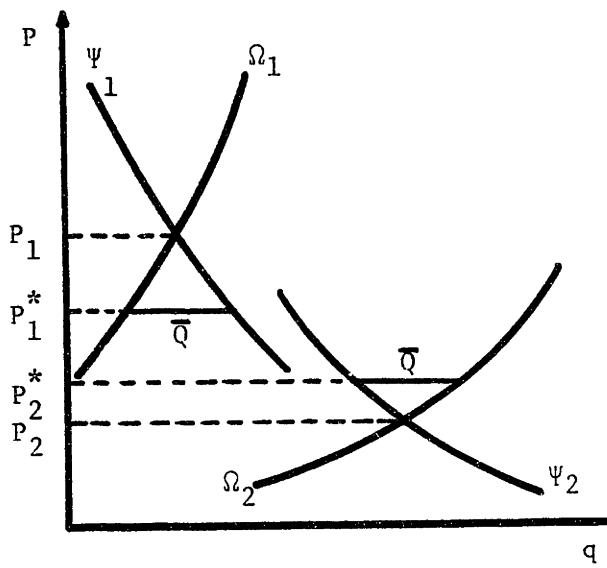
Following Figure 9.a, since $P_2 < P_1$, some people will be interested in the interchange between markets; in particular, consumers at M_1 would like to buy X at a lower price and producers at M_2 would like to sell X at a higher price. If X is a common transportable good, people will be willing to pay some amount to move units of X from M_2 to M_1 ; if X is not transportable, consumers at M_1 will be willing to pay some amount to go to M_2 .¹⁰ Let us assume the former case for simplicity. At every level of \bar{Q} , three types of equilibrium conditions must hold:

- (i) consumption at M_1 should be equal to production at M_1 plus the amount of X carried from M_2 to M_1 , that is

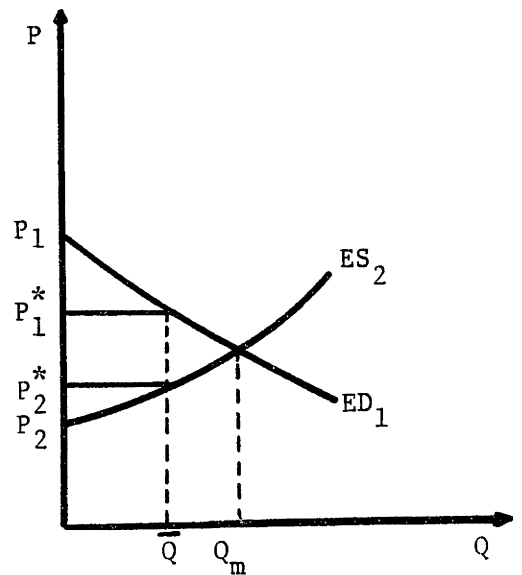
$$\psi_1 = \Omega_1 + \bar{Q} ; \quad (14)$$

- (ii) production at M_2 should be equal to consumption at M_2 plus the amount of X carried from M_2 to M_1 , that is

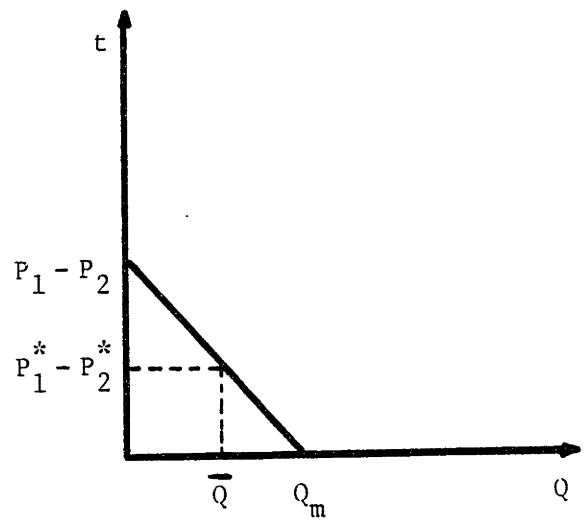
¹⁰This would be the case of rest or touristic places.



(a)



(b)



(c)

FIGURE 9. Derivation of the Demand for Transportation

$$\Omega_2 = \psi_2 + \bar{Q}$$

(iii) the existence of \bar{Q} changes the price level at each market, and the new price at M_1 should be equal to the new price at M_2 plus the unit transportation price, that is:

$$P_1^* = P_2^* + t \quad (16)$$

because otherwise a price differential would appear again.

Equations (14) to (16) define a relation between \bar{Q} and t , that is, a demand for transportation which can be better understood by restating the equations in the following way:

$$\bar{Q} = \psi_1 - \Omega_1 = ED_1 \quad (17)$$

$$\bar{Q} = \Omega_2 - \psi_2 = ES_2 \quad (18)$$

$$t = P_1^* - P_2^* = \Omega_1^{-1} - \Omega_2^{-1} = \psi_1^{-1} - \psi_2^{-1} \quad (19)$$

Relations (17) and (18) are also shown in Figure 3.b, where excess demand (ED_1) and excess supply (ES_2) curves have been built for M_1 and M_2 respectively. This way, each level of transportation \bar{Q} is associated with certain levels of "dynamic prices" P_1^* and P_2^* in each market and, therefore, with a level of t given by the price differential. Graphically,

the vertical difference between ED_1 and ES_2 gives the level of t associated with the various levels of \bar{Q} , generating the transportation demand curve for the good X between M_2 and M_1 , shown in Figure 9.c (Galvez and Jara, 1975).

It should be noticed that ψ_i and Ω_i are defined as long run curves which, in the supply case, assumes that the X market is competitive but firms have different cost structures so that the aggregate supply is upward sloping as discussed in 2.2. Therefore, the derived transportation demand as a curve is stable with respect to possible investments in the modes joining both the markets.

4.2 Benefits on Production and Consumption

It is evident that a change in the equilibrium conditions at M_1 and M_2 caused by any positive value of \bar{Q} has welfare implications. Let us compare the situation "without" transportation, i.e., markets in isolation, with that in which equilibrium occurs at prices P_1^* and P_2^* . It is important to note that:

$$\int_{P_1^* - P_2^*}^{P_1 - P_2} \bar{Q}(t) dt = \int_{P_1^*}^{P_1} (ED_1) dP + \int_{P_2}^{P_2^*} (ES_2) dP$$

$$= \int_{P_1^*}^{P_1} \psi_1 dP - \int_{P_1^*}^{P_1} \Omega_1 dP + \int_{P_2}^{P_2^*} \Omega_2 dP - \int_{P_2}^{P_2^*} \psi_2 dP; \quad (20)$$

that is, the consumers' surplus measured using transportation demand summarizes both consumers' and producers' surpluses variation in both M_1 and M_2 when going from the state (P_1, P_2) to (P_1^*, P_2^*) . This is graphically given by the equivalence of the shaded areas in Figure 10.a. Note that in M_1 the consumers' surplus increases while producers' surplus decreases but in a smaller absolute amount, so that the net variation in total surplus is positive and equivalent to area . Similarly, producers' surplus increment is greater than consumers' surplus decrease exactly in . Then in each market total surplus variation is positive but there is always a group affected negatively.

This surplus equivalency implies that it is not necessary to measure the effects of transportation over production and consumption of other goods when transportation demand is available. Adding these effects to users' surplus variation would lead to double counting when evaluating a transportation project because of the derived nature of the demand involved. In spite of this, there are some cases which call for the direct measure of benefits through the analysis of the transported goods' market, as in the so-called "penetration roads," where the pre-existing flows are zero and the generated flow and the associated willingness to pay can be better understood in terms of production and consumption variations in the integrated zones. The general case, however, presents pre-project flows different from zero, as assumed in Figure 10.b. The property shown by (20) can be easily extended to general flow variations; intuitively, the reduction in equilibrium perceived cost from t to $t - \Delta t$ causes an increment in flow , associated to a lower price in M_1 and a higher price in M_2 . The sum of the absolute

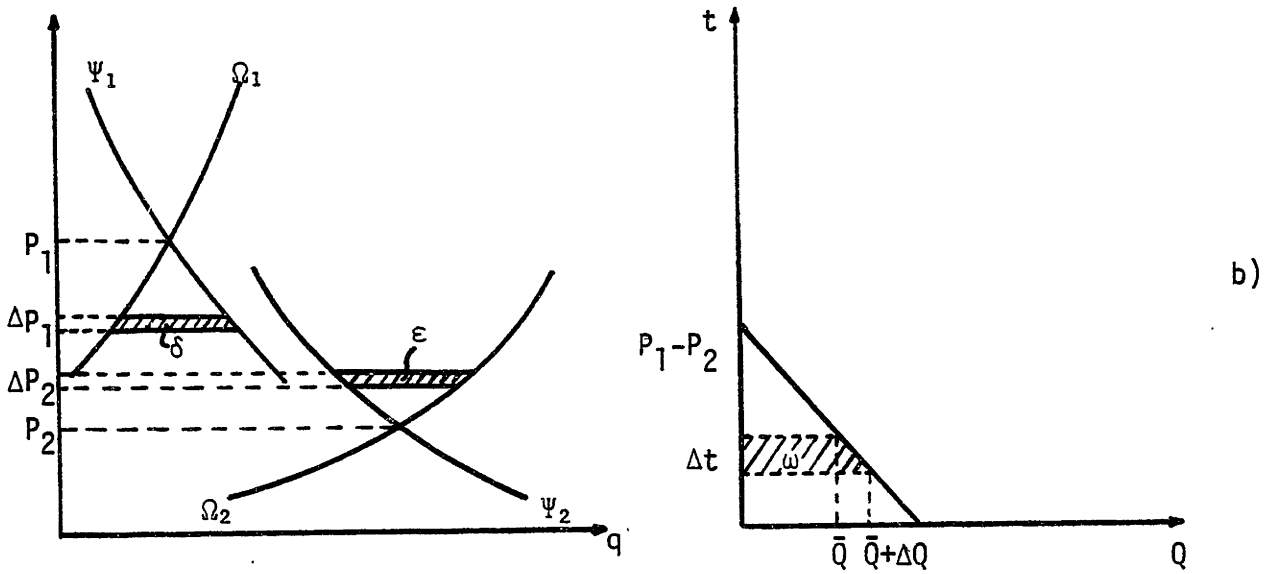
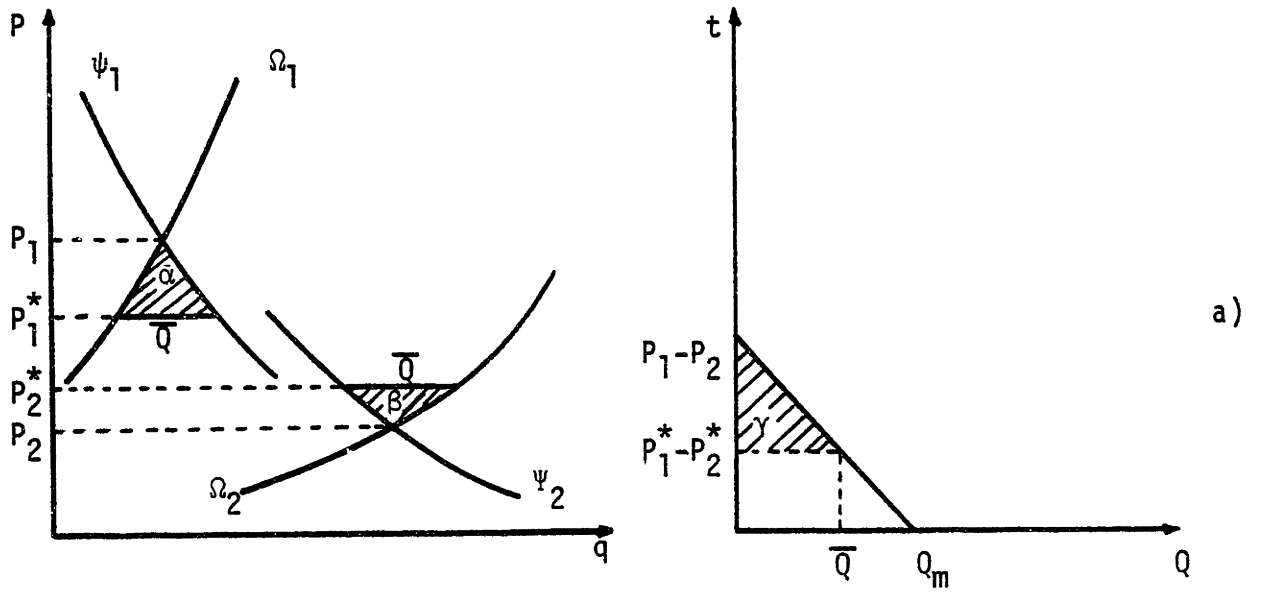


FIGURE 10. Transportation Consumers' Surplus Equivalencies

a) $\alpha + \beta = \gamma$

b) $\delta + \epsilon = \omega$

variation in prices is equal to Δt and δ plus ϵ equals ω ; this is, the already stated equivalency among surpluses variation also holds in this case.

4.3 Modal and Aggregate Supply Curves

Let us assume for analytical purposes that it is possible to obtain an aggregate demand by summation over different single commodity O-D demands. From the derivation of this aggregate function it can be clearly seen that it is stable with respect to investments in the modes joining both markets because the existence of such a transportation demand is independent of the actual transportation supply. The existence or not of an actual quantity transported will depend, as usual, on the intersection of the derived O-D demand and the aggregate supply function that represents the network involved.

In order to analyze the two competing modes case, a simple network of two links representing the two modes between two nodes A and B will suffice. Figure 11 shows this elemental network. The modal supplies in terms of the cost perceived by the users at each flow level, are link-oriented; we will assume the most general case, that is, congestion takes place in both modes as explained in 2.3.

Let us define

$$Q_i = S_i(t_i) \quad (21)$$

as the supply function associated to mode i .¹¹ Both are graphically shown in Figure 12.a. Later it will be necessary to make a distinction between "Marginal Cost" and "Average Cost" supply curves in order to

¹¹The most usual explicit forms of the supply functions are stated as

$$t_i = S_i^{-1}(Q_i).$$

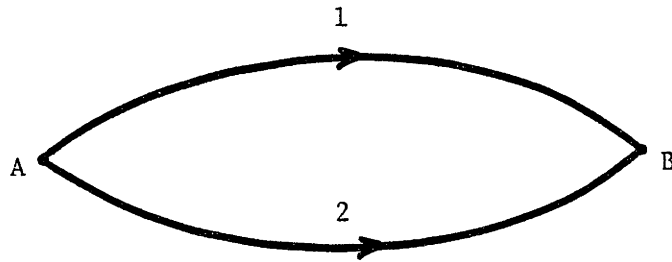


FIGURE 11.A Simple Network

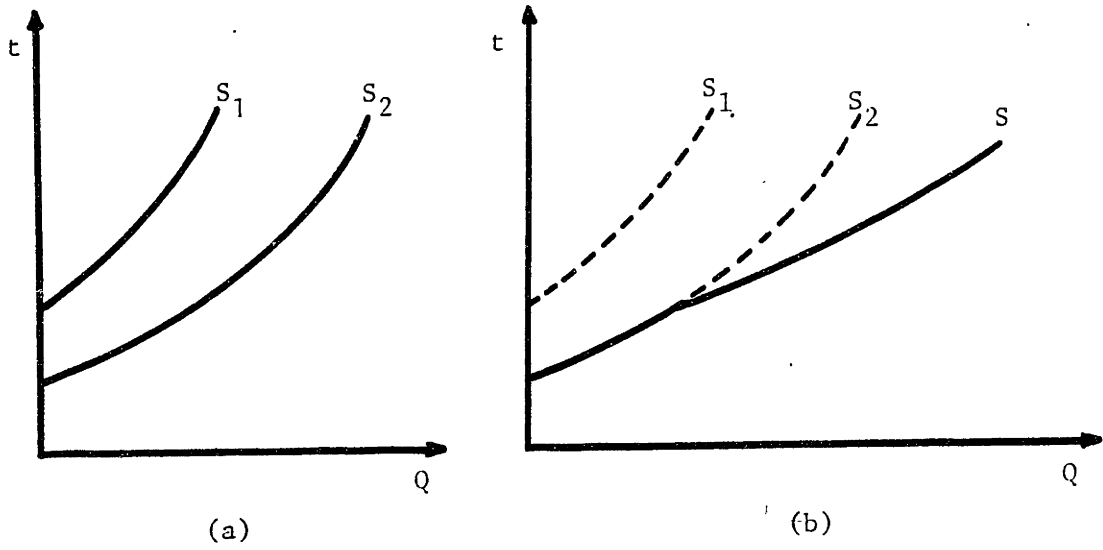


FIGURE 12. Modal and Aggregate Transportation Supply Curves

consider total benefits derived from a unimodal investment. By now, let us think about the perceived cost as being affected by the volume of users on each mode.

In order to obtain an aggregate supply which represents the network between A and B, we must make assumptions on the behavior of the users of the system. For this purpose we are going to assume that users are identical, divisible and that each unit chooses the less expensive mode in terms of perceived cost; that is, a user will shift to the competing mode if the actual perceived cost is higher. This behavioral assumption, known as Wardrop's first principle or user optimization rule (Potts and Oliver, 1972), implies that for a given number of users, the total flow going from A to B splits in such a way that the perceived cost is the same in both modes; otherwise, some units using the more expensive mode would shift until it no longer represents an advantage. In terms of the transportation market, these assumptions are equivalent to state that the aggregate supply which represents the network between A and B is obtained by horizontal summation of the modal supplies specified by (21); that is

$$Q = S_1(t) + S_2(t) = S(t) \quad (22)$$

Figure 12.b shows this aggregate function.

It is worthwhile to point out that it is always possible to construct an aggregate transportation supply representing the network between A and B, even if it were a more complex one. Through a reasoning similar to the one followed to get the aggregate supply for the competing

modes case, we can get an equivalent supply for two complementary modes, e.g., two successive links on a path; in this case it is constructed by adding both functions vertically; that is, the "price" charged at a given flow level is the sum of the prices charged on each link at the same level of flow. By adding complementary and competitive link-supplies as indicated, an aggregate supply can be obtained for any network between A and B, given that each elementary supply is stated in terms of perceived cost and assuming that all users perceive this cost in the same manner.

4.4 Derivation of the Modal Demands

Let us associate nodes A and B to spatially separated markets that generate a demand for transportation from A to B as derived in 4.1, defined as

$$Q = D(t) \quad (23)$$

The equilibrium total flow between A and B will be given by the solution (t_ℓ, Q_ℓ) to the system of equations determined by (22) and (23). The flow on each mode will be given by (21) with $t_i = t_\ell$. The question to be solved now is whether modal demands such as the ones presented in Chapter 3 can be derived from this framework. The next step shows that it is not only possible but consistent with partial (modal) equilibria.

By definition, a point on a modal demand has to associate the amount per unit flow people are willing to pay to send a quantity Q_i by mode i , with this quantity. Assume a certain flow Q_0 from A to B; the value t people are willing to pay to send Q_0 is given by the aggregate

demand function as

$$t = D^{-1}(Q_0) \quad (24)$$

On the other hand, an amount Q_0 will be actually charged

$$t' = S^{-1}(Q_0) \quad (25)$$

to go from A to B; thus, the modal flows will be given by

$$Q_i = S_i[S^{-1}(Q_0)] \quad (26)$$

This way, people are willing to pay $D^{-1}(Q_0)$ to send $S_i[S^{-1}(Q_0)]$ by mode i .
Therefore

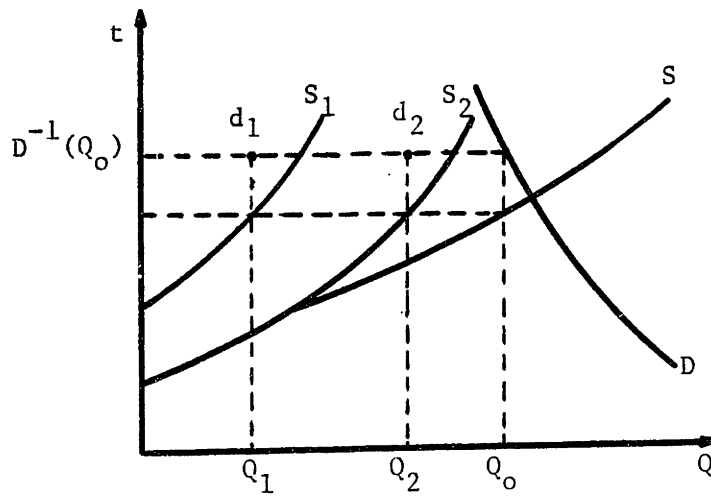
$$\{D^{-1}(Q_0), S_1[S^{-1}(Q_0)]\} \in D_1$$

and

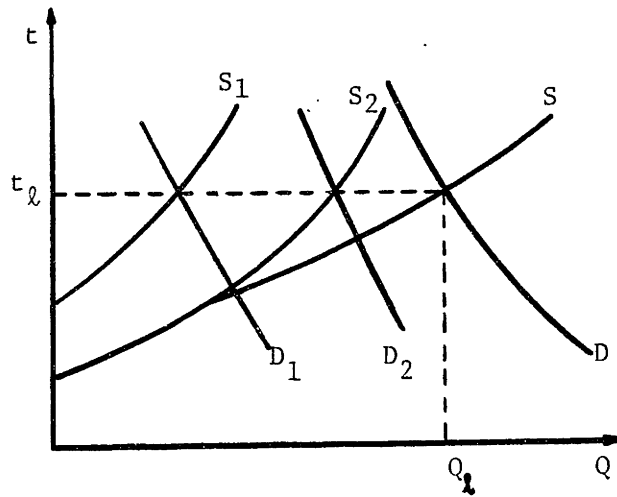
$$\{D^{-1}(Q_0), S_2[S^{-1}(Q_0)]\} \in D_2 \quad (27)$$

where D_i is the modal demand. Varying Q_0 the whole modal demands can be obtained. Figure 13.a shows the derivation of $d_1 \in D_1$ and $d_2 \in D_2$; Figure 13.b shows the whole set of demand and supply curves. Analytically, each modal demand is given by

$$t = D^{-1}\{S[S_i^{-1}(Q_i)]\} = D_i^{-1}(Q_i) \quad , \quad (28)$$



(a)



(b)

FIGURE 13. Derivation of the Modal Demands

which gives the amount people are willing to pay per unit to send Q_i units from A to B by mode i . Note that by construction it is true that

$$D(t) = D_1(t) + D_2(t) \quad . \quad (29)$$

Moreover, as $S^{-1}(Q_\ell) = t_\ell$ by definition, then applying (26)

$$D_i(t_\ell) = S_i[S^{-1}(Q_\ell)] = S_i(t_\ell) \quad , \quad (30)$$

that is, modal demand equals modal supply when total demand and supply are also equal. As an important remark, note also that each D_i depends on S_1 , S_2 and D ; therefore, a change in one modal supply curve will change both modal demands.

4.5. Some Comments

The framework developed in this chapter is sufficient to clarify at least one question posed in Chapter 3, the one related to the problem of expressing each modal demand as $D_j(t_i, t_j)$ holding t_i constant. Following Figure 13.b, a totally different modal demand schedule would be obtained if, for instance, S_2 were changed to a different supply curve S_2' keeping only one point in common with S_2 , the point (t_ℓ, Q_2^0) . This way the same equilibrium perceived cost and modal split would be obtained but the modal demand curves would be different.

Another reason for deleting the idea of a modal demand given a fixed perceived cost in the competing mode, is related with the very concept of what a perceived cost is, a concept which is used by all the authors

analyzed in Chapter 3. This concept is that of a generalized cost that includes every factor which influences the user's decision and translates it to monetary terms; this way it allows for comparison among modes directly. If we kept constant this generalized cost in one mode, say mode 1, and we try to think about the demand for mode 2 as the amount of users willing to travel by this latter at different levels of t_2 , we will conclude that the modal demand would have two segments: it would be zero at every level of $t_2 > t_1$, and it would coincide with the aggregate demand D at every level of $t_2 < t_1$, assuming a rational behavior of the users (which is equivalent to the assumption of Wardrop's first principle to hold).

Therefore, it seems to be more appropriate to state a modal demand as $D_i(t_i, t_j)$ but not with t_i fixed but having a relation with t_j reflecting a certain behavior of the users, as Wardrop's $t_i = t_j$ reflects. This function should not be mistaken for Mohring's $t_i(t_j)$ which reflects the continuous adjustment of perceived costs after a change in one mode's supply function.

APPENDIX TO CHAPTER 4. Derivation of the Slope of the Modal
Demand Curves

From (28) and (29) we can state

$$t = D^{-1} \{S_1[S_1^{-1}(Q_1)] + S_2[S_1^{-1}(Q_1)]\} \quad (a)$$

$$t = D^{-1} \{Q_1 + S_2[S_1^{-1}(Q_1)]\} \quad (b)$$

Taking derivative with respect to Q_1 we get

$$\frac{dt}{dQ_1} = \frac{dD^{-1}}{d\{Q_1 + S_2[S_1^{-1}(Q_1)]\}} \cdot \left\{ 1 + \frac{d S_2}{d[S_1^{-1}(Q_1)]} \frac{d[S_1^{-1}(Q_1)]}{dQ_1} \right\}. \quad (c)$$

The first term of the right hand side is always negative, because it represents the slope of the aggregate transportation demand at some point. The expression in parenthesis has a positive sign provided the supply curves are upward sloping at all points. Therefore the slope of the modal demands is also negative.

CHAPTER 5. BENEFITS OF A MODAL INVESTMENT

Using the basic framework established through the preceding chapter we are able to analyze now the impact of a unimodal investment on the whole system, putting it in terms of measures involving modal supplies and demands. In this chapter we present the effects of an investment on the modal supply, on the aggregate supply and on the equilibrium state; then we derive the new modal demands showing that they shift in opposite directions. The variation in consumers' surplus as a measure of benefits when the supply curves represent average cost is stated in terms of the aggregate and modal functions. Finally, the marginal cost supply curves case is presented.

5.1 Supply Effects and Changes in Equilibrium

Let us assume an investment is made in mode l ; in general such an action will be reflected in a reduction in the perceived cost actually "charged" at every level of users of this mode, with respect to what was "charged" before. This can be adequately represented by a shift downward of mode l 's supply curve. By virtue of the additive nature of the aggregate supply curve, this latter will also shift downward, causing a change in the aggregate level of flow using both modes. Figure 14 presents all these changes graphically, where S_l^1 , S' , t_l^1 and Q_l^1 denote the new mode l 's supply curve, aggregate supply, equilibrium perceived cost and total flow level respectively.

A modal investment may have two effects on the associated supply function: quality and/or capacity effects (Fernández, 1979). The quality

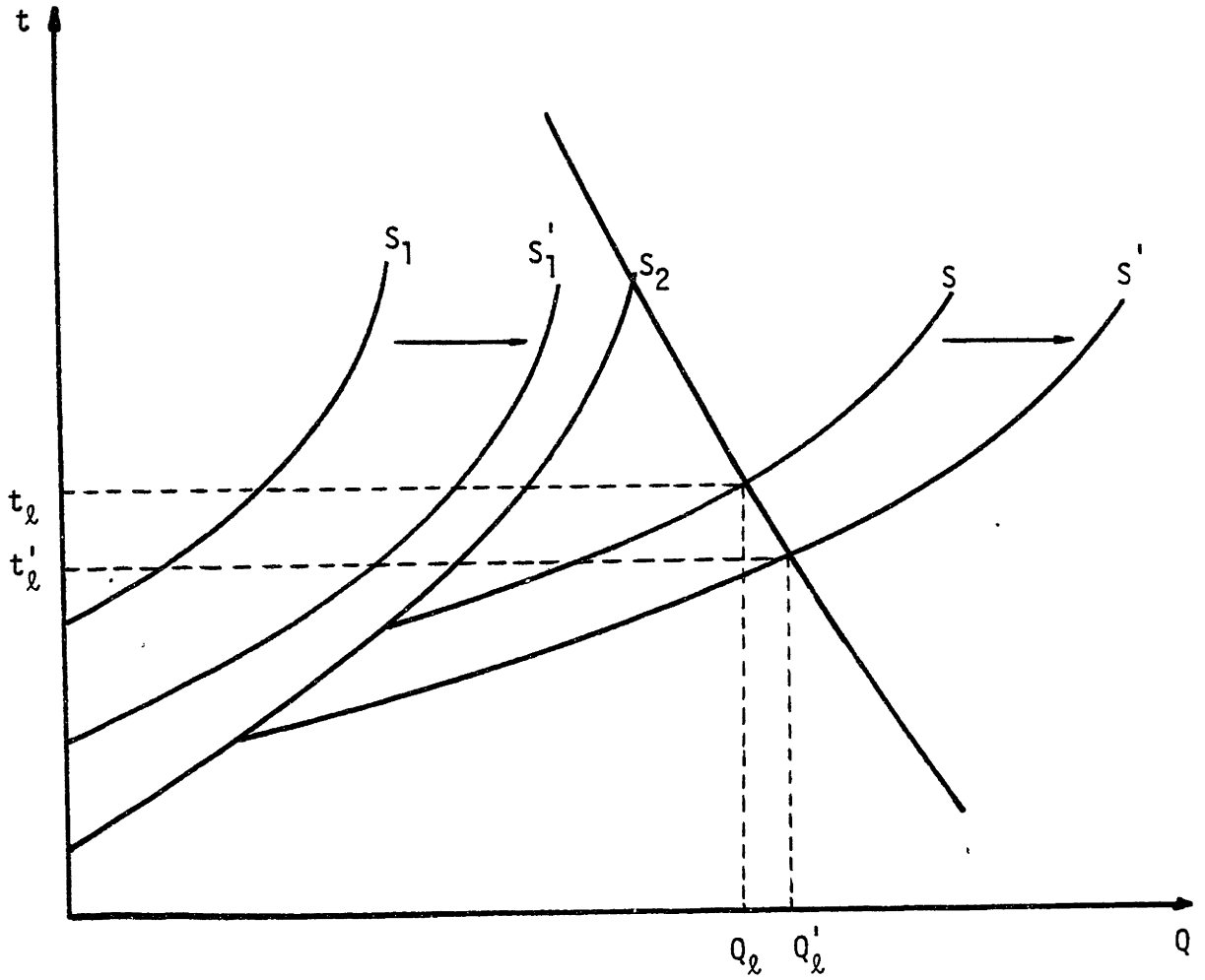


FIGURE 14. Changes in Modal and Total Supplies and New Equilibrium Levels After an Investment in Mode 1.

effect is reflected by a reduction in the perceived cost at very low flow level, i.e., when no congestion can be assumed;¹² the capacity effect is reflected by an increment in the level of flow at which congestion becomes important, i.e., when $\frac{\partial t}{\partial Q} \gg 0$. We have assumed here that both effects take place after investing in mode 1.

5.2 The Shift in Modal Demands

It was shown in Chapter 4 that transportation demand as a function relating willingness to pay and flow level between two points, depends only on the economic environment and, therefore, that this demand function was stable with respect to changes in the transportation network. We have also seen how the modal demands can be derived from the information provided by this aggregate curve and the modal associated supplies. Therefore, the modal demands after investing in mode 1 can be obtained in the same manner as in 4.3. Let us call D_1' and D_2' these new curves. As we have assumed that $S_1'(t) > S_1(t)$ and $S'(t) \geq S(t)$ (which is equivalent to say that $S^{-1}(Q_0) \geq S'^{-1}(Q_0)$), and also that every supply curve is increasing, that is, $\frac{\partial t}{\partial Q} > 0$, it can be shown that (see Appendix)

$$s_1[S^{-1}(Q_0)] < s_1'[S'^{-1}(Q_0)] \quad \forall Q_0 \quad . \quad (31)$$

This means that at the same perceived cost level, given by $D^{-1}(Q_0)$, the amount of users willing to travel by mode 1 is greater after than before the investment was made. Therefore

¹²This is what is usually called "free flow" level in highways.

$$D_1'(t) > D_1(t) \quad \forall t \quad . \quad (32)$$

As in (29), it must hold that

$$D_1'(t) + D_2'(t) = D(t) \quad . \quad (33)$$

Therefore, under our assumptions,

$$D_2'(t) < D_2(t) \quad \forall t \quad . \quad (34)$$

Figure 15 shows all the curves and relevant points before and after the investment was made. The demand for the improved mode has increased at every t level, and the demand for the competing mode has shifted inward. In other words, the amount of users willing to travel by the improved mode at every perceived cost level increases, while users willing to travel by the competing mode diminishes. These directions of shift in the modal demands are intuitively more appealing than those implicitly assumed by Mohring and Williams.

5.3 Measurement of Benefits: Average Cost Supply Curves Case.

Under the assumption of the supply curves being unaltered, this is, no tolls are charged over what users perceive as a cost, this function represents average costs from an aggregate perspective due to the externalities involved in the presence of an additional user in the system.¹³

¹³For a very good explanation of this phenomenon in general, see Bishop's manuscript, especially the example of fishermen in a lake.

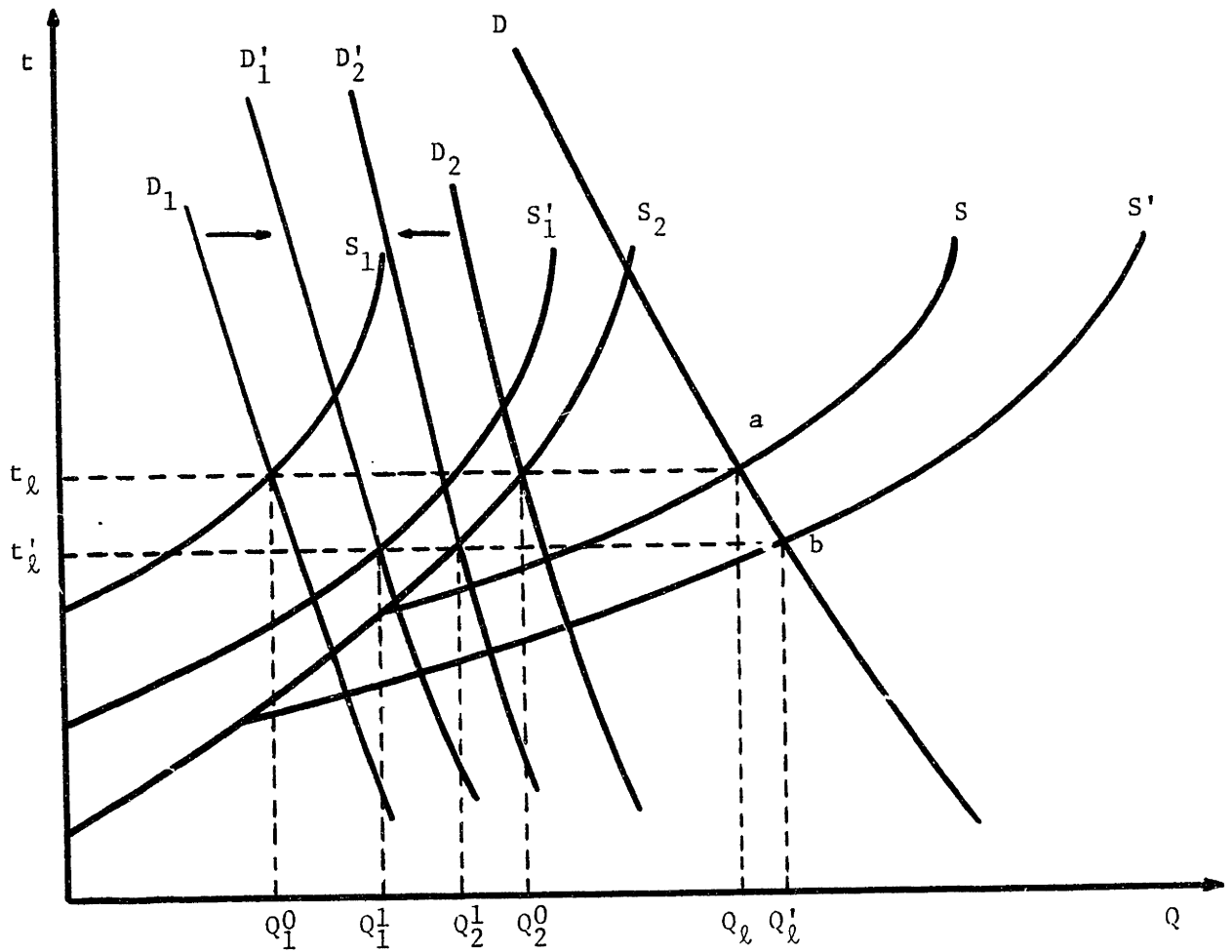


FIGURE 15. Total and Modal Supply and Demand Curves Before and After an Investment Is Made in Mode 1.

In this case the producers' surplus measure vanishes, as already explained in 2.2, and the measurement of benefits is reduced to the calculation of consumers' surplus variation.

Actual consumers' surplus variation is given by

$$ACS = \int_{t'_\ell}^{t_\ell} D(t)dt \quad (35)$$

which is equivalent to area $t_\ell a b t'_\ell$ in Figure 15 and, due to the additive property (see (29) and (33)), it can be expressed also in terms of modal demands as

$$ACS = \int_{t'_\ell}^{t_\ell} D_1(t)dt + \int_{t'_\ell}^{t_\ell} D_2(t)dt = \int_{t'_\ell}^{t_\ell} D'_1(t)dt + \int_{t'_\ell}^{t_\ell} D'_2(t)dt \quad (36)$$

Therefore, the real consumers' surplus variation can be obtained from the modal demands as a simple addition of the common measure using alternatively the "initial" modal demands (D_1 and D_2) or the "final" ones (D'_1 and D'_2). As shown in Chapter 4, this variation summarizes the effects of transportation investment on production and consumption at A and B. The measure of ACS in terms of the original or final modal demands, as obtained in (36), is shown in Figure 16. Note that (36) implies that areas J and H are equal in this Figure.

5.4 Comparison With Other Approaches

Most of the literature on this subject has been focused on the correct measure of the consumers' surplus variation, which corresponds to what we have done in the last section. This is why it is worthwhile to compare at this stage the different approaches, postponing for a while the analysis of the marginal cost supply curves case.

We showed in section 3.3 that the uniqueness of the line integral representing consumers' surplus variation was incompatible with modal demands shifting in opposite directions. We could have stated then that Agnello's approach was probably erroneous because both conditions were assumed simultaneously. But, in fact, the direction of shift was used only for graphical presentation. At this point we can say that the modal demands actually shift this way and that the condition for uniqueness of the line integral does not hold. In fact, the existence of a common demand schedule for the competing modes allows us to state that the Marshallian consumers' surplus has a unique value (given by t_ℓ a b t'_ℓ in Figure 15) and no further assumption are needed. Another point that seems to have been missed in Agnello's approach is that two phenomena after the investment in mode 1 take place: diverted users from mode 2 to mode 1 and increment in the total number of users in $Q'_\ell - Q_\ell$ units (see Figure 15); this can be clearly seen after the general Q-D demand has been obtained, but is not obvious when working directly with modal demands. Stating explicitly that

$$Q_1^0 + Q_2^0 = Q_\ell \quad \text{and} \quad Q_1^1 + Q_2^1 = Q'_\ell \quad (37)$$

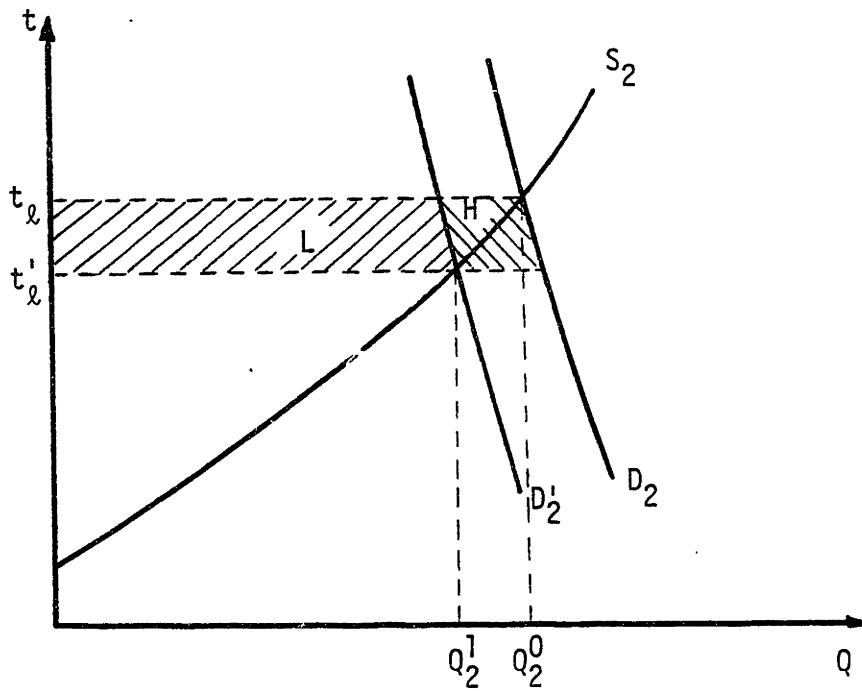
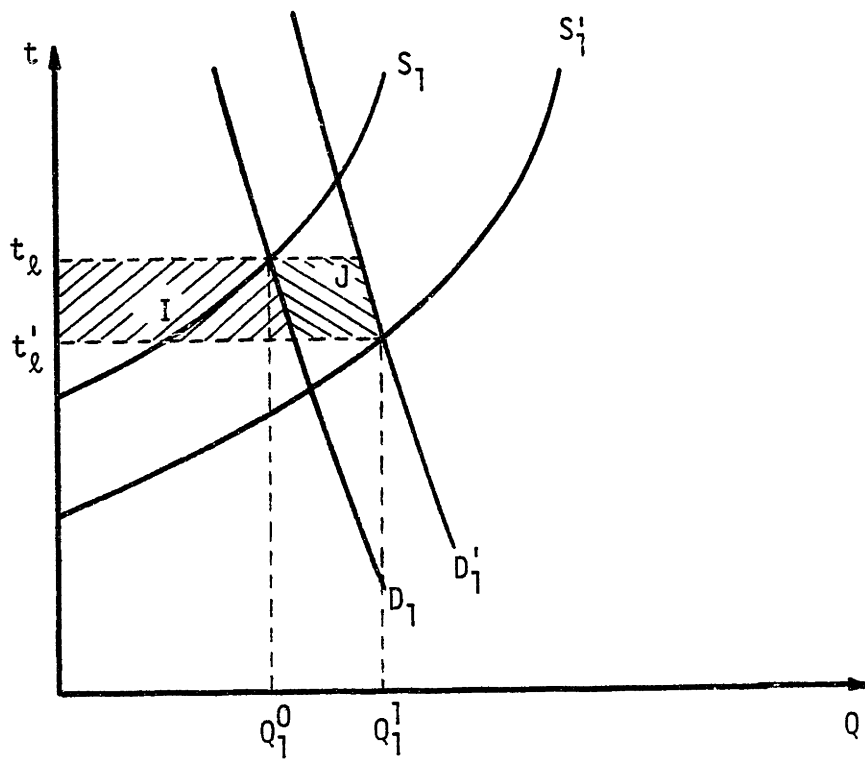


FIGURE 16. Consumers' Surplus Variation Due to Improvement in Mode 1;
 (ACS = I + L + H = I + J + L)

it can be easily shown after some manipulation that

$$Q_1^1 = Q_1^0 + (Q_\ell^1 - Q_\ell) + (Q_2^0 - Q_2^1) \quad (38)$$

That is, the final flow in mode 1 is given by the flow before the investment plus the so-called induced flow in the system $(Q_\ell^1 - Q_\ell)$ plus the diverted flow from mode 2. This induced flow is not considered when setting out costs in Agnello's paper; he refers to "old users" and "additional users" without differentiating among these latter (where we can recognize a diverted and an induced component) leading to an overestimation of the consumers' surplus variation.

A similar analysis can be made regarding the rule-of-the-half. This measure has a very weak interpretation, as stated in Williams' paper, but our approach gives the necessary economic foundation in order to use it as a good practical approximation; we have shown J and H are equal (Figure 16); we have also shown that

$$ACS = \int_{t_\ell^1}^{t_\ell} D(t)dt = I + L + H; \quad (39)$$

but H can be expressed as:

$$H = \frac{1}{2}J + \frac{1}{2}H \quad (40)$$

which leads to

$$ACS = I + \frac{1}{2}J + L + \frac{1}{2}H \quad (41)$$

The area $I + \frac{1}{2}J$ can be evaluated approximately as

$$I + \frac{1}{2}J \approx Q_1^0 (t_\ell - t'_\ell) + \frac{1}{2}(Q_1^1 - Q_1^0)(t_\ell - t'_\ell) \quad (42)$$

and

$$L + \frac{1}{2}H \approx Q_2^1 (t_\ell - t'_\ell) + \frac{1}{2}(Q_2^0 - Q_2^1)(t_\ell - t'_\ell) \quad (43)$$

Adding (42) and (43) and after some manipulation we get

$$ACS \approx \frac{1}{2}(t_\ell - t'_\ell) (Q_1^0 + Q_1^1) + \frac{1}{2}(t_\ell - t'_\ell) (Q_2^0 + Q_2^1) \quad (44)$$

which is the rule-of-a-half. By further manipulation of (44) we can express ACS as

$$ACS \approx Q_\ell (t_\ell - t'_\ell) + \frac{1}{2}(Q'_\ell - Q_\ell)(t_\ell - t'_\ell) \quad (45)$$

Therefore, the approximation is based ultimately on the linearity of the relevant portion of the aggregate transportation demand. There is no possible interpretation relating some "average of the upper and lower bounds on the benefit for those who do shift" as suggested in Neuberger (1971), because in fact the half measure is related to the induced demand and not to the shifting demand as was shown here.

5.5 Extension of the Framework to Marginal Cost Supply Functions.

Let us now turn to the case where the modal "supply" curves are repre-

mented by the marginal costs (MC) associated with the operation of each mode.¹⁴ The whole analysis developed to construct the modal demands is still valid, but now the equality between modal marginal costs implied by the horizontal summation of S_1 and S_2 leads to an economic optimum (efficiency) in a Paretian sense.¹⁵ Although this is a general result from welfare economics, the intuitive idea behind this is that for every level Q the minimum total transportation cost is given by Q_1 and Q_2 such that $\text{Marginal Cost}_1 = S_1^{-1}(Q_1) = S_1^{-1}(Q_2) = \text{Marginal Cost}_2$; any other pair Q_1', Q_2' such that $Q_1' + Q_2' = Q$ would give a higher total cost. This idea is known as Wardrop's second principle or systems' optimization rule (Potts and Oliver, 1972).

For evaluation purposes, we must consider now variations in both consumers' and producers' surpluses in order to get total benefits due to an improvement in mode 1. Consumers' surplus variation is given by (39) and producers' surplus variation is

$$\text{APS} = t'_l Q'_l - \int_0^{Q_l} s'^{-1}(Q) dQ - \{t_l Q_l - \int_0^{Q_l} s^{-1}(Q) dQ\} \quad (46)$$

¹⁴This would be the case if a toll were charged to the users such that average perceived cost plus this toll equalled marginal cost.

¹⁵This is not true when prices in markets related to close complements or substitutes are not equal to the marginal costs; second best theory suggests other procedures for optimal pricing in such cases. See Turvey (1971).

or, in terms of t , by

$$APS = \int_0^{t'_\ell} S'(t)dt - \int_0^{t_\ell} S(t)dt \quad (47)$$

Expressing the second integral of (47) in terms of $D(t)$ we get

$$APS = \int_0^{t'_\ell} S'(t)dt - \left\{ \int_0^{t'_\ell} S(t)dt + \int_{t'_\ell}^{t_\ell} [D(t) - (D(t) - S(t))]dt \right\} \quad (48)$$

and rearranging terms,

$$APS = \int_0^{t'_\ell} [S'(t) - S(t)]dt - ACS + \int_{t'_\ell}^{t_\ell} [D(t) - S(t)]dt \quad (49)$$

Therefore, total surplus variation turns out to be

$$\Delta TS = APS + ACS = \int_0^{t'_\ell} [S'(t) - S(t)]dt + \int_{t'_\ell}^{t_\ell} [D(t) - S(t)]dt \quad (50)$$

which is equivalent to areas R and W respectively in Figure 17. By construction of the modal demands and the aggregate supply, ΔTS can be expressed in modal terms as:

$$\Delta TS = \int_0^{t'_\ell} [S'_1 - S_1]dt + \int_{t'_\ell}^{t_\ell} [D_1 - S_1]dt + \int_{t'_\ell}^{t_\ell} [D_2 - S_2]dt \quad (51)$$

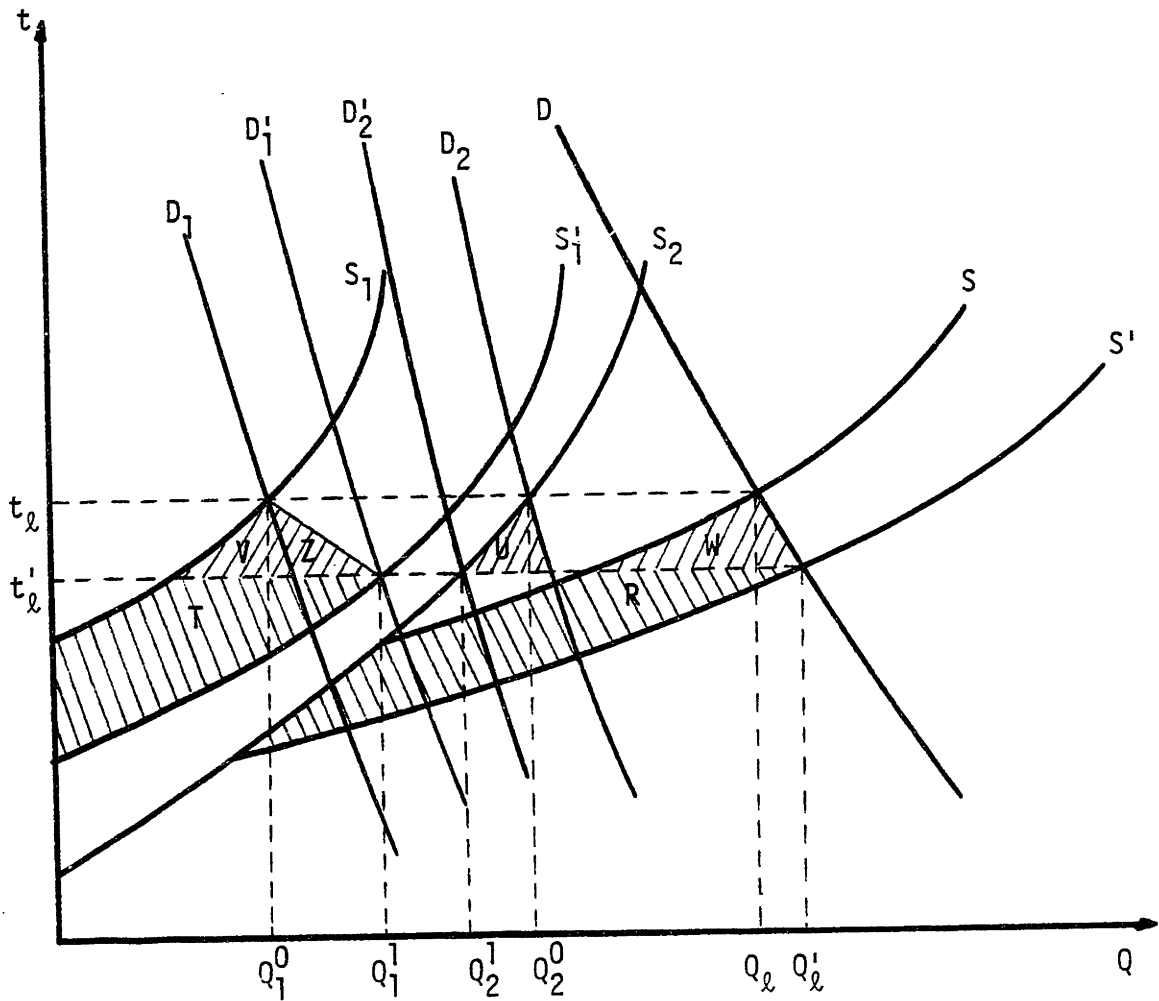


FIGURE 17. Total Surplus Variation Under Marginal Pricing Policy

$$\Delta TS = R + W \approx T + V + Z$$

which is equivalent to areas T, V and U respectively in Figure 17. Assuming linearity in D_1 , D_2 , and S_2 , U is equal to Z. Therefore, an approximate measurement of the exact ΔTS given by (51) is

$$\Delta TS \approx T + V + Z \quad (52)$$

which has the advantage of dealing with only one mode.

The approach developed in the preceding paragraph gives a measure for the variation in total surplus very similar to the one due to Mohring presented in Chapter 3. Remember that Mohring (1976) develops a "real modal demand curve" that goes through $D_1 \cap S_1$ and $D_1' \cap S_1'$, by expressing the perceived cost in mode 2 as a function of the perceived cost in mode 1. Therefore $D_1[t_1, t_2] = D_1[t_1, t_2(t_1)] = D_1^*(t_1)$. If area Z were constructed using this concept instead of a linear approximation, (52) would give an exact measure of ΔTS . In fact, Mohring's function $t_2(t_1)$ implies that unimodal analysis is not enough, although apparently it is. What is true however, is that the approximate measure of ΔTS given by (52) requires calculation related only to mode 1.

APPENDIX TO CHAPTER 5. Direction of Shift in Mode 1's
Demand Curve

For any flow Q_0 , the willingness to pay is $D^{-1}(Q_0)$ and the associated mode 1's flow (see (27)) is given by

$$Q_1 = S_1(S^{-1}(Q_0)) \quad . \quad (a)$$

The figure in the following page may help in visualizing the next steps. Assume a small investment in mode 1 such that $S_1' = S_1 + \Delta S_1$, where ΔS_1 , should be looked at as a function. Let's define

$$P = S^{-1}(Q_0) \quad . \quad (b)$$

For any given level of Q_0 we are interested in knowing the effect of ΔS_1 on Q_1 . (b) implies

$$S_1(P) + S_2(P) = Q_0 \quad . \quad (c)$$

also

$$[S_1 + \Delta S_1](P + \Delta P) + S_2(P + \Delta P) = Q_0 \quad . \quad (d)$$

Subtracting (c) from (d) we get

$$\frac{\partial S_1}{\partial p} \Delta P + \frac{\partial S_2}{\partial p} \Delta P + \Delta S_1 = 0 \quad . \quad (e)$$

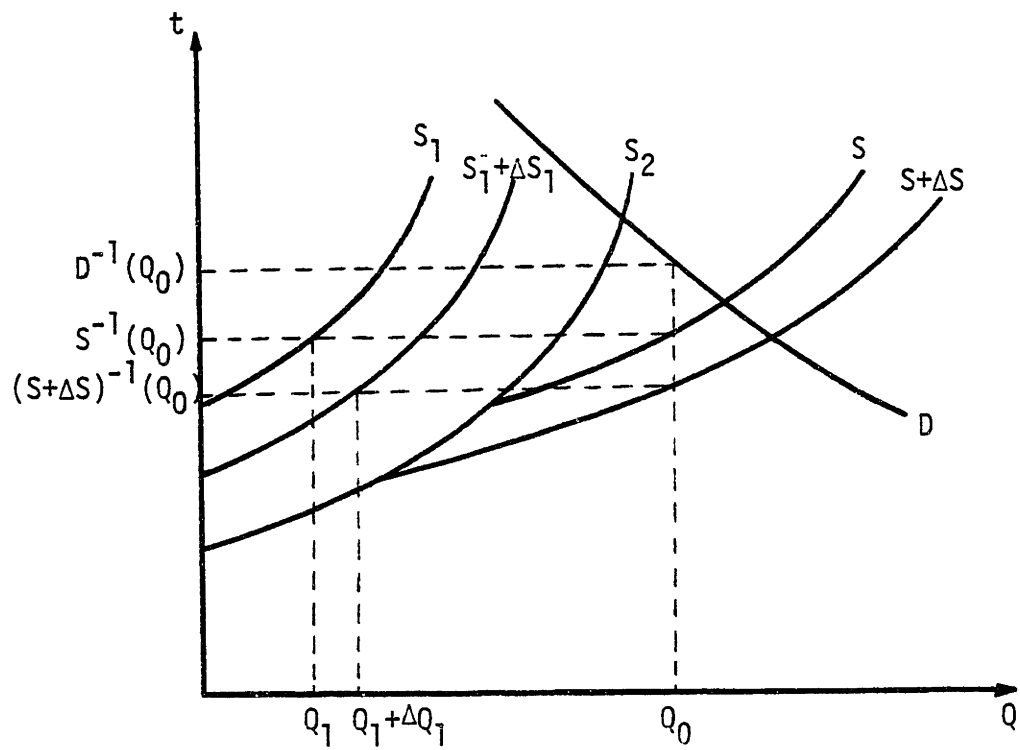


FIGURE Ap. 5. Direction of Shift of Mode 1's Demand Curve

Solving for ΔP

$$\Delta P = - \frac{\Delta S_1}{\partial S / \partial p} \quad . \quad (f)$$

On the other hand,

$$Q_1 = S_1(P) \quad \text{and} \quad (g)$$

$$Q_1 + \Delta Q_1 = [S_1 + \Delta S_1](P + \Delta P) \quad ; \quad (h)$$

subtracting (g) from (h) we get to

$$\Delta Q_1 = \frac{\partial S_1}{\partial p} \Delta P + \Delta S_1 \quad . \quad (i)$$

Replacing the result in (f),

$$\Delta Q_1 = \Delta S_1 \left[1 - \frac{\partial S_1 / \partial p}{\partial S / \partial p} \right] \quad . \quad (j)$$

Noting that $S = S_1 + S_2$, we finally get

$$\Delta Q_1 = \Delta S_1 \frac{\partial S_2 / \partial p}{\partial S / \partial p} \quad . \quad (k)$$

All the elements of the right hand side of (k) have been assumed to be positive, therefore $\Delta Q_1 > 0$, i.e., the willingness to travel by mode 1 increases at every perceived cost level.

CHAPTER 6. SYNTHESIS, EXTENSIONS AND CONCLUSIONS

Chapters 2 to 5 have been devoted to the analysis and solution of the problem of how to approach the benefits measurement conceptually when interrelated modes are considered. The created framework calls for some extensions reaching related fields in transportation. In this chapter, we are going to synthesize what was said throughout this work in order to present later some ideas in relation with generalization of this approach and also to point out certain special aspects; in particular, the relation between equilibrium and evaluation is going to be partially discussed, followed by a brief analysis of the role of the income effect. The last section of the chapter summarizes the main conclusions.

6.1 Synthesis

The traditional measure of consumers' surplus variation is no longer valid when dealing with interrelated demands. This general fact becomes particularly important in transportation analysis from an evaluation perspective when more than one mode is involved, because in general the willingness to pay for moving certain number of units by one mode will depend on the price on the other mode or modes. The literature on this subject has concentrated on the analysis of two competing highways, which constitute conceptually the same problem as the one posed by two competing modes provided that we are working with perceived, generalized or inclusive costs instead of fares or monetary transportation prices alone. In an economic sense, these two modes are substitutes; and, therefore, variations

in the conditions prevailing at equilibrium in one of them will cause changes in the other. It can be said that, in particular, any change in the supply for one mode such that the equilibrium perceived cost is affected will cause a series of shifts of the demands for related modes provided that modal supplies are not infinitely elastic; i.e., when congestion is assumed to occur, till a new equilibrium is reached. These shifts in the modal demands call for a different treatment of the measure of consumers' surplus variation after a modal investment or, in general, after a variation in some modal supply due to certain policy. Williams (1976) and Agnello (1977) use the concept of extended consumers' surplus due to Hotelling (1938) in order to provide an answer to this problem, resulting in different propositions. Mohring (1976) proposed a different approach including explicitly supply considerations. The assumed direction of shift in the modal demands differs among these authors and no clear explanation for this phenomenon is offered.

We have presented a new approach to the measurement of the consumers' surplus variation which requires a full explanation of the nature of the modal demands and their behavior under new supply conditions. For purposes of presentation, analysis and comparison, the case of two competing modes providing transportation between two points was developed. It was shown that the production and consumption structure associated with each node generates an aggregate and stable transportation demand function. The total flow between the two nodes can be obtained from the interaction between this aggregate demand and an aggregate transportation supply, which is constructed by horizontal summation of the individual (modal) supplies

under the assumption of rational users' behavior, also known as Wardrop's first principle or user optimization rule. Modal demands i.e., functions which relate flow and willingness to pay on each mode, can be obtained from this framework as analytical expressions involving the aggregate demand and the modal supply functions. Assuming that a modal investment improves both quality and capacity, we showed that the demand for the improved mode shifts outward; i.e., the amount of users willing to travel at each perceived cost level increases; the demand for the competing mode shifts inward. At this stage, the nature of the aggregate and modal demands has been explained, and consumers' surplus variation can be obtained in a straightforward manner from the aggregate figure and applying some properties of the modal functions, in particular the additive property, the same result can be obtained by simple summation of the consumers' surplus variation regarding the initial or the final modal demands. As further conclusions we have shown: (i) that the conditions for the line integral expressing the generalized consumers' surplus variation to have a unique value are not fulfilled in this competing mode case, although consumers' surplus variation does have a unique value; and (ii) that this framework gives a much better theoretical foundation for the use of the generalized rule-of-a-half as an approximate measure of benefits than the existing one.

6.2 On Extensions of the Approach

In previous chapters we discussed at length the case of two modes, a single generalized good and a single origin-destination pair. This problem provided important insights for understanding the nature of the

problem and for exploring techniques to analyze the multimodal impacts of unimodal investments. It is apparent that this analysis may be extended to consider any number of modes. Extensions to consider multiple goods and multiple origins and destinations require more thought and are not so easily accomplished in the sense of understanding modal demands.

Let us consider the problem of multiple origins and destinations. Our first inclination is to argue for the extension of the method for constructing aggregate transportation demands introduced previously in order to get a set of O-D specific curves. This extension is not readily performed for a fully general transportation network. The three spatial markets case helps to illustrate this point. Following Figure 18, a stable transportation demand function between markets 1 and 2 could be obtained if market 3 were isolated; i.e., t_{31} and t_{23} had huge values. However, if only t_{31} was reduced to appropriate levels, two possible aggregate flows appear: Q_{31} and Q_{21} . Evidently, the demand for transportation between 2 and 1, as a function relating Q_{21} and willingness to pay, depends on the prevailing values of t_{31} and t_{21} and, therefore, on the transportation supplies associated to the network involved.^{16/} The complication then arises from the fact that any event which alters demand for one O-D pair will generally alter demand for other O-D pairs due to the existence of congestion and the fact that any network O-D path will usually have links in common with other O-D paths. In this sense stable aggregate Samuelsonian type demands cannot be readily constructed by simple extensions of the single O-D analysis.

^{16/} See the section Comparative Statistics in Samuelson (1952) for a discussion related to this topic (although the transportation demand is never mentioned).

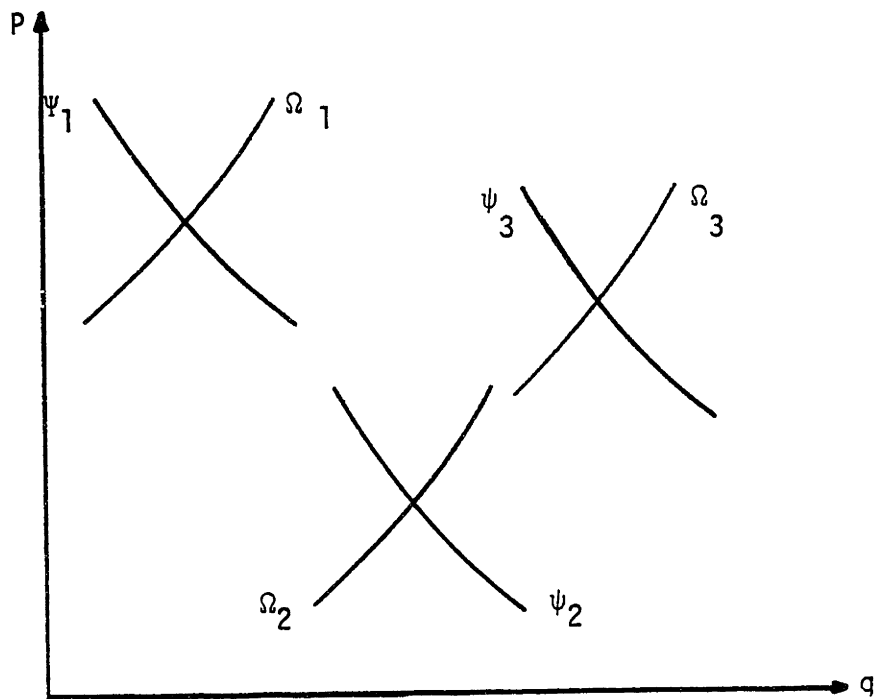


FIGURE 18. Three Markets Case

One possible approach to analyze the multiple markets case can be obtained by looking at the net or excess supply associated to each node (market) which can be negative denoting a positive excess demand. In fact, Samuelson finds the equilibrium of the system associated to infinitely elastic transportation supply function (t_{ij} constant). The congested problem, including the possibility of a network rather than a single link between O-D pairs, is a matter of special presentation. It is interesting to note that the so-called generation model, which results in a set of values expressing the total inflow and outflow associated with each node (zone) can be considered as a particular case of previously defined set of excess supplies and/or demands node-specific. The type of models which begin with this step, then, recognize explicitly the derived and aggregate nature of transportation demand.

6.3 Equilibrium and Evaluation

At this stage it is clear that any extension of the conceptual framework built in the preceding chapters requires network analysis. Moreover, the analysis performed gives a sounder theoretical basis for the so-called rule-of-a-half used in estimating user benefits than had existed previously in the literature. This rule, because it provides a good approximation for changes in surplus due to investments (cost changes) in the network, may be used to describe modal impacts.

Use of the rule-of-a-half to determine modal impacts requires that the effect of investments on equilibrium traffic volumes on all links of the network be ascertained. This effect is obtained by introducing the changes of the appropriate network supply functions due to the investment into the network description and then computing the resultant

traffic equilibrium volumes on all links.

The prevailing perspective with respect to computing traffic equilibria is based on what is sometimes called the "equivalent minimization problem" generally attributed to Beckman et. al. (1956). This approach converts the user equilibrium problem to a mathematical programming problem to which various exact numerical procedures may be applied. It was shown then that a convex minimization problem could be constructed which gives a user equilibrium assignment when solved. They created an objective function which when minimized subject to supply equals demand constraints, non-negativity constraints and certain definitional constraints yields a user equilibrium. This approach assumes as given a set of stable O-D demand functions which, as we have seen, is not a generalization of the one O-D analysis but a particular case of a more general problem. The objective function is equivalent to the maximization of the sum, along all O-D pairs, of the areas between demand and average cost (supply) curves. The equivalence of the minimization problem to the user equilibrium problem may be shown by application of the Kuhn-Tucker necessary conditions for mathematical programming problems. Although one is tempted to give an economic interpretation to the objective function used to define the equivalent minimization problem, Beckman et. al. (1956) strongly cautioned against this; indeed, it seems preferable to consider the objective function as a mathematical construction which leads to the desired result. This point of view applies, of course, to all equivalent minimization formulations for the user equilibrium problem. Note, however, that the maximization of the sum of similar areas regarding marginal cost supply

curves is equivalent to maximize the total economic surplus.¹⁷

The first attempt to generalize the equivalent minimization approach was made by Dafermos (1972). She extended the formulation to the so-called multiclass user case in which physical links of the network may contain multiple classes of users who are not identical insofar as costs are concerned. This concept of multiple user classes is accommodated for within a generalized network perspective where in principle each link's supply function may depend on the full vector of link flows of the network. She showed that the minimization problem proposed was equivalent to the user equilibrium problem and had a solution if and only if, using our notation of the previous chapters:

$$\frac{\partial S_i}{\partial t_j} = \frac{\partial S_j}{\partial t_i} \quad (53)$$

It is at first sight surprising that this condition can be shown to be equivalent to the necessary and sufficient condition for the line integral expressing the generalized consumers' surplus in the two modes case to have a unique value, independently of the integration path. This condition was:

¹⁷Remember this surplus provides a measure of welfare assuming the compensation principle to hold.

$$\frac{\partial D_i}{\partial t_j} = \frac{\partial D_j}{\partial t_i} \quad (54)$$

Let us prove it. We have shown that the modal demands can be expressed as:

$$t = D^{-1} \{S[S_i(Q_i)]\} \quad (55)$$

assuming users' rationality (Wardrop's first principle) to hold. This can be rewritten as:

$$Q_i = S_i\{S^{-1}[D(t)]\} \quad (56)$$

Remember user optimizing $t_1 = t_2 = t$, therefore

$$\frac{d Q_1}{d t_2} = \frac{d Q_1}{d t} = \frac{d S_1}{d\{S^{-1}[D(t)]\}} \frac{d S^{-1}[D(t)]}{d[D(t)]} \frac{d D(t)}{dt} \quad (57)$$

By analogy,

$$\frac{d Q_2}{d t_1} = \frac{d Q_2}{dt} = \frac{d S_2}{d\{S^{-1}[D(t)]\}} \frac{d S^{-1}[D(t)]}{d[D(t)]} \frac{d D(t)}{dt} \quad (58)$$

Remember that $D(t)$ and $S(t)$ are the aggregate demand and supply transportation functions respectively. For a level t (willingness to pay), $D(t)$ gives the total flow. Therefore, by simple inspection (Figure 19 may help for this purpose), $\frac{d Q_1}{dt}$ equals $\frac{d Q_2}{dt}$ at every t level if and only if $\frac{d S_1}{dt}$ equals $\frac{d S_2}{dt}$ for every t level. As $t_i = t_j = t$, condition (53) implies condition (54). Evidently, real situations rarely present symmetry in the crossed effects on the supply curves, which shows that uniqueness of the generalized expression for the variation in consumers' surplus rarely happens. In general the true value corresponds to a particular integration path.

6.4 The Income Effect.

The well known Slutsky equation can help us to understand better the relation among modal demands. Remember that in general the change in quantity of good y demanded when the price of x , P_x , changes, can be expressed as:

$$\frac{\partial y}{\partial P_x} = \frac{\partial y}{\partial P_x} \Big|_{U_0} - x \frac{\partial y}{\partial I} \quad (59)$$

That is, the total effect can be divided in the cross-substitution effect at U_0 level of utility plus the income effect. In the case of interrelated modal demands we have, then,

$$\frac{\partial Q_1}{\partial t_2} = \frac{\partial Q_1}{\partial t_2} \Big|_{U_0} - Q_2 \frac{\partial Q_1}{\partial I} \quad , \text{ and} \quad (60)$$

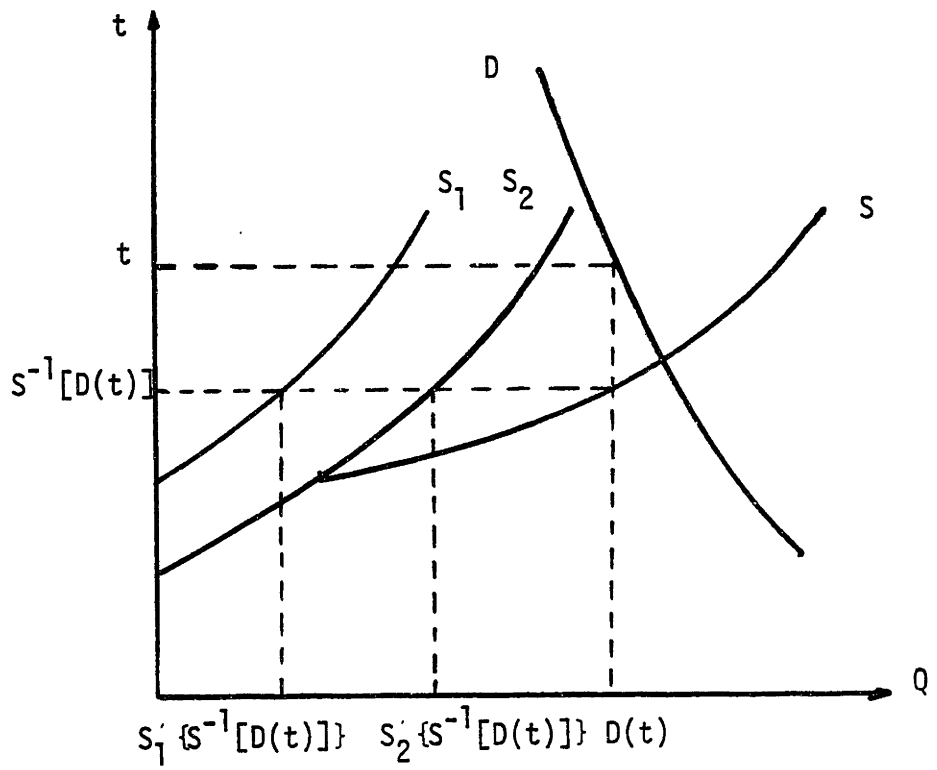


FIGURE 19. The Modal Demands Step by Step

$$\frac{\partial Q_2}{\partial t_1} = \frac{\partial Q_2}{\partial t_1} \Big|_{U_0} - Q_1 \frac{\partial Q_2}{\partial I} \quad (61)$$

Therefore, the condition for the generalized consumers' surplus variation to have a unique value, given by (54), implies the equality of the right hand sides of (60) and (61). Since the matrix of substitution terms is symmetric (see Varian, 1978, p. 99), that is,

$$\frac{\partial Q_1}{\partial t_2} \Big|_{U_0} = \frac{\partial Q_2}{\partial t_1} \Big|_{U_0} \quad (62)$$

then condition (54) implies that the income effects should be equal; i.e.,

$$Q_2 \frac{\partial Q_1}{\partial I} = Q_1 \frac{\partial Q_2}{\partial I} \quad (63)$$

which is equivalent to say that the income elasticity is the same for both modes. On the other hand, the Slutsky equation applied to the aggregate transportation demand states that

$$\frac{\partial Q}{\partial t} = \frac{\partial Q}{\partial t} \Big|_{U_0} - Q \frac{\partial Q}{\partial I} \quad (64)$$

By noting that $Q_1 + Q_2 = Q$, this last equation can be turned into

$$\frac{\partial Q_1}{\partial t} + \frac{\partial Q_2}{\partial t} = \frac{\partial Q_1}{\partial t} \Big|_{U_0} + \frac{\partial Q_2}{\partial t} \Big|_{U_0} - (Q_1 + Q_2) \frac{\partial (Q_1 + Q_2)}{\partial I} \quad (65)$$

Recalling that we have assumed $t_i = t_j = t$, condition (54) and property (62) imply

$$2 \frac{\partial Q_1}{\partial t} = 2 \frac{\partial Q_1}{\partial t} \Big|_{U_0} - Q_1 \frac{\partial Q_2}{\partial I} - Q_2 \frac{\partial Q_1}{\partial I} - (Q_1 \frac{\partial Q_1}{\partial I} + Q_2 \frac{\partial Q_2}{\partial I}) \quad (66)$$

But we have shown that condition (54) also implies equality (63), therefore

$$2 \frac{\partial Q_1}{\partial t} = 2 \frac{\partial Q_1}{\partial t} \Big|_{U_0} - 2 Q_2 \frac{\partial Q_1}{\partial I} - (Q_1 \frac{\partial Q_1}{\partial I} + Q_2 \frac{\partial Q_2}{\partial I}) \quad (67)$$

Equation (60) implies that the expression in parenthesis should be zero.

Therefore,

$$Q_1 \frac{\partial Q_1}{\partial I} = - Q_2 \frac{\partial Q_2}{\partial I} \quad (68)$$

which sign contradicts equality (63) unless

$$\frac{\partial Q_1}{\partial I} = \frac{\partial Q_2}{\partial I} = \frac{\partial Q}{\partial I} = 0 \quad (69)$$

Therefore, for condition (54) to hold, we require the income effect of the aggregate transportation demand to be zero.

6.5 Final Comments and Conclusions.

The variation in utility level due to changes in consumption is a central aspect in economic analysis. Unfortunately, the ordinal nature that the theory assigns to the "system of preferences" makes it difficult to develop analytic work with these concepts. However, surrogate monetary

measures can be justified as a way to capture utility effects which are, ultimately, the direct consumers' perception of what is going on in the different markets of the economy. We have justified the use of the Marshallian consumers' surplus as an operative approximation to overcome the problem of utility variation measurement. One can state that all goods in an economy are interrelated, that it is not possible to define a good without generic or specific complements and substitutes; in spite of this fact, *ceteris paribus* analysis can be found all through economic textbooks, introducing "rational" concepts and ideas which appear perfectly reasonable from isolated perspectives, as one individual, one firm, one market, etc.; but the behavior of the system as a whole cannot be obtained by simple addition of these micro-behaviors of its components. However, one can assume that small changes in one market are not going to cause "disturbances" but in a limited economic neighborhood. Through the present work we have had the flavor of the kind of difficulties which can be postulated to be consubstantial to any analysis with pretensions of being more than partial. We have seen that the generalization of the consumers' surplus measure (which was transparent for an isolated good) in order to incorporate the effects in all related markets, does not capture the essence of the problem, leading to ambiguous measures of utility variation. But we have also seen that the study of the nature of transportation demand led us to an approach which takes advantage of the simplicity of the one-good concept while taking into account the effects on related markets, just by looking at transportation as a "derived service".

The theoretical derivation of any good's demand function appear to be independent of supply characteristics, but we have seen that a modal demand, as a function, depends on the supply function associated with the same and related modes. This result seems to break the intuitive rule that states independence among these functions because demand is based on utility and supply on costs.¹⁸ In fact, this is a "deviation" caused by a ceteris paribus style of analysis; once the smallest step toward a more general approach is attempted, the conclusions are different (see Robinson, 1959).

The problem we have intended to solve is important from an evaluation perspective. The main conclusion relates to the necessity of looking at transportation demand as a derived one. This idea, or this fact, is basic for the understanding of the shifting modal demands and provides an evaluation approach for a specific problem which is considered important in the literature on transportation economics. We recognize the fact that a general case, the one of any intermediate good, could have been posed as a challenge; but our sole intention was to make a contribution to an area of knowledge which has attracted special attention. Extensions of this approach to multimodal effects can be developed in many directions - multiple O-D's, multiple commodities, multiple modes, etc. The economic analysis of this case should take advantage of the mathematical developments associated to the study of the correspondent equilibrium problem, which has been shown to be related.

¹⁸A number of "demand" models developed in the past apparently describe equilibrium more than willingness to pay or to travel.

As a final comment, we would like to recall that the generalization of this framework to multiple spatial markets (origins and destinations) is not possible in terms of stable transportation demands. This naturally calls for the reformulation of existing equilibration approaches which require origin-destination specific demands; neither the "pure network" nor the "pure economics" perspectives seem to be enough to go further in this field. We think that the "joint perspective" is the most rewarding. This conclusion should be taken as a by-product of this work on transport project evaluation.

BIBLIOGRAPHY

1. Agnello, Richard J. "Economic Evaluation of Highway System Benefits," Transportation Research, Vol. 11 (1977).
2. Beckmann, Martin, C. B. McGuire and Christopher Winsten. Studies in the Economics of Transportation. Yale University Press (1956).
3. Bishop, Robert R., Manuscript (unpublished mimeo), Massachusetts Institute of Technology.
4. Currie, John M., J. A. Murphy and A. Schmitz, "The Concept of Economic Surplus and its Use in Economic Analysis," The Economic Journal, No. 324, Vol. 81 (1971).
5. Dafermos, S.C., "The Traffic Assignment Problem for Multiclass-User Transportation Network," Transportation Science, Vol. 7, (1972).
6. Fernández, Enrique, "Optimal Dynamic Investment Policies for Public Facilities: the Transportation Case," Ph.D. Dissertation, Massachusetts Institute of Technology (1979).
7. Gálvez, Tristán, and Sergio Jara-Díaz, "Generación y Efectos del Mercado de Transporte," Pub. ST-inv/01/75, Sección Ingeniería de Transporte, Universidad de Chile, (1975).
8. Hotelling, Harold, "The General Welfare in Relation to Problems of Taxation and Railway and Utility Rates," Econometrica, Vol. 6 (1938).
9. Marshall, Alfred. Principles of Economics, London (1930).
10. Mishan, Edward J. Cost-Benefit Analysis, Preager Publishers (1976).
11. Mohring, Herbert. Transportation Economics, Ballinger Publishing Company (1976).
12. Neuberger, H., "User Benefit in the Evaluation of Transport and Land Use Plans," Journal of Transport Economics and Policy, Vol. V, (1971).
13. Nicholson, Walter. Microeconomic Theory, The Dryden Press, (1978).
14. Potts, Renfrey B. and R. M. Oliver. Flows in Transportation Networks, Academic Press (1972).
15. Samuelson, Paul. Foundations of Economic Analysis, Atheneum, (1965), (originally published 1946).

16. Samuelson, Paul A., "Spatial Price Equilibrium and Linear Programming," American Economic Review, Vol. 42, (1952).
17. Turvey, Ralph. Economic Analysis and Public Enterprises. George Allen and Unwin, Ltd. (1971).
18. Varian, Hal. Microeconomic Analysis, Norton, (1978).
19. Walters, Alan, "The Theory and Measurement of Private and Social Cost of Highway Congestion," Econometrica, Vol. 29 (1961).
20. Williams, H.C.W.L., "Travel Demand Models, Duality Relations and User Benefit Analysis," Journal of Regional Science, Vol. 16, No. 2 (1976).
21. Willig, Robert D., "Consumer's Surplus Without Apology," American Economic Review, Vol. 66, No. 4 (1976).
22. World Bank, "The Economic Analysis of Rural Road Projects," World Bank Staff Working Paper No. 241 (1976).