Economic Approach to Value of Time and Transportation Choice

Reuben Gronau, Hebrew University, Jerusalem

Travel time was recognized to affect the demand for transportation before it was formally incorporated in economic theory. No traveler, producer of transportation services, or transportation-oriented policy maker needs an economist to make this fact known. The major contribution of economics in this context is in formulating the problem and creating a framework that allows one to measure the effect of travel time on the demand for trips. In this framework, the price of the trip includes both the money expenditures and the opportunity cost of time. Thus travel time affects the choice of destination, the choice of mode, and the number of times one travels. In this paper, I shall try to describe the current state of economic theory regarding the value of time, emphasizing some of the difficulties involved in its empirical application.

The first attempts to incorporate the opportunity cost of time in economic theory date back to the early 1960s (7, 8, 9, 10). The most general and far-reaching formulations can be found in Becker's theory of the allocation of time (2), which hypothesizes that the initial source of utility is the activity (or in Becker's terminology the commodity). Each activity involves the combination of goods purchased in the market, the household member's time, and (sometimes) intermediate activities. For example, the activity meal combines the capital services of the dining room and its fittings, the participants' time, and the activities of cooking and serving. The activity visit to another city involves the money expenditures for accommodations and food, the time spent in that city, and the activity trip. The household tries to maximize the utility derived from all the activities engaged in subject to two constraints: the budget constraint, which specifies that total money expenditures cannot exceed income, and the time constraint, which stresses that the time involved in all activities is limited.

Formally, the utility function is as follows:

$$U = U(Z_1, \ldots, Z_n) \tag{1}$$

where Z_i denotes the ith activity; Z_i in turn depends on the household production function,

$$G(Z_1, \ldots, Z_n, X, T) = 0 \tag{2}$$

where

X = vector of market goods and services, and
T = vector of time units, since time is not necessarily a homogeneous input.

We may distinguish hours of the day, days in the week, and months in the year. Daytime may be used extensively for work while sleep is produced at night. Summertime may be a prevalent input in the activity going to the beach, and wintertime figures extensively in the production of skiing. Furthermore, equation 2 allows for joint production. For example, a mother may engage simultaneously in cooking and child care, and an air passenger may travel and watch a movie at the same time.

The maximization of utility is subject to two constraints: the budget constraint

$$PX = W(Z_n) + V \tag{3}$$

and the time constraint

$$T = T_0 \tag{4}$$

where

P = price vector;
$W(Z_n)$ = earnings that are a function of the activity work, Z_n;
V = other sources of income; and
T_0 = a vector of total units of time available (the components of T_0 may differ since there are, for example, more workday than weekend hours and more day than night hours in summer).

The maximization of the utility function with respect to these two constraints yields the optimum combination of activities, the optimum allocation of time and goods, and the value people place on their time.

To analyze this optimum, let us assume for simplicity that there are no intermediate activities, that there is no joint production, and that the production function is continuous throughout the relevant range. Given these simplifying assumptions, equation 2 can be rewritten as

$$Z_i = F_i(X_i, T_i) \tag{5}$$

where

X_i = vectors of goods, and
T_i = time inputs involved in the production of activity i.

The budget and time constraints have to be rewritten

$$\sum_{i=1}^{n} P_i X_i = (Z_n) + V \quad \text{and} \quad \sum_{i=1}^{n} T_i = T_0 \tag{6}$$

Defining the Lagrangian,

$$L = U(Z_1, \ldots, Z_n) + \lambda\left[W(Z_n) + V - \sum P_i X_i\right] + \mu\left(T_0 - \sum T_i\right) \tag{7}$$

and maximizing with respect to Z_i yield the necessary conditions for an optimum

$$u_i = \lambda[P_i x_i + (\mu/\lambda)t_i] \quad (i = 1, \ldots, n-1) \qquad u_n = \mu t_n - \lambda w \tag{8}$$

where

$u_i = \partial U/\partial Z_i$ $(i = 1, \ldots, n)$ = marginal utility of activity i,
$x_i = \partial X_i/\partial Z_i$ = marginal inputs of goods,
$t_i = \partial T_i/\partial Z_i$ = time in the production of Z_i, and
$w = [\partial W(Z_n)/\partial Z_n] - P_n x_n$ = marginal wage rate.

Differentiating the Lagrangian with respect to X_i and T_i yields the optimum combination of inputs in production,

$$\frac{\partial Z_i}{\partial T_i} \bigg/ \frac{\partial Z_i}{\partial X_i} = \frac{x_i}{t_i} = \frac{\mu/\lambda}{P_i} = \frac{K}{P_i} \quad (i = 1, \ldots, n) \tag{9}$$

where $K = \mu/\lambda$. The scalar λ is the marginal utility of income. The vector μ denotes the marginal utility of the various time units. $K = \mu/\lambda$ is therefore a vector denoting the (money) value placed by the household on the different units of time. Thus, by equation 9 the marginal rate of substitution in the production of activity i (i.e., the ratio of the marginal products of T and X) equals the input price ratio.

Rewriting equation 8, we have

$$u_i = \lambda(P_i x_i + K t_i) = \lambda \Pi_i \quad (i = 1, \ldots, n-1)$$

$$K = \mu/\lambda = [w + (u_n/\lambda)]/t_n \tag{10}$$

The first of these equations states that in equilibrium the marginal utility derived from activity i is proportional to its marginal cost of production Π_i. By the second of these equations, the shadow price of time K depends on the marginal wage rate w; the money equivalent of the utility of work, u_n/λ; and the marginal product of the time unit in the production of the activity work $1/t_n$.

In the past, the value of time was usually identified with the wage rate. Equation 10 indicates that even under our set of simplifying assumptions this equality can be regarded only as a crude approximation. The value of time depends on the marginal wage rate, $w = [\partial W(Z_n)/\partial Z_n] - P_n x_n$, and not on the reported average wage. It has been argued that the marginal productivity of labor and the hourly wage rate change as the daily number of working hours varies (Figure 1). Thus, if the daily number of working hours falls short of t_n^* so that the marginal productivity of labor is still increasing (1), the average wage is an underestimate of the marginal wage rate. If on the other hand the number of working hours exceeds t_n^*, the reverse is true. Furthermore, to obtain an estimate of w, one has to deduct the marginal money cost incurred from the marginal wage rate. It is frequently argued that, at least in the case of married women, these costs (e.g., the cost of babysitters and other forms of child care) are substantial.

Even if we were able to obtain accurate data about w, we would have to correct the data for the (unknown) value of the marginal utility of work so that an estimate of the value of time could be obtained. The marginal wage rate w is an overestimate of the value of time K when work involves marginal disutility and is an underestimate when work involves positive utility.

Finally, up to this point, it has been implicitly assumed that all units of time are used for work (i.e., $t_n > 0$). This is clearly an inaccurate assumption: A large fraction of the adult population (mainly housewives) is not part of the labor force, and even those participating in the labor force cannot change their working hours freely (e.g., they cannot substitute night for day working hours). In the extreme case, all working hours are determined institutionally and are not subject to the household's decisions in the short run. Whenever units of time are not used for work or cannot be substituted freely for working time (i.e., when the appropriate elements of the time vector satisfy $\partial Z_n/\partial T_n = 0$), the marginal wage rate becomes irrelevant for determining the value of these time units. The value of these units is determined in this case by their scarcity, i.e., by their supply and demand.

How is the value of time affected by changes in wages and other sources of income? An increase in the marginal wage rate directly affects the value of time of those units that are freely substitutable for working hours. This increase may also affect the second component of the value of time by changing the number of working hours, the marginal utility of income, and the money equivalent of the marginal utility of work, u_n/λ. The resulting change in income increases time scarcity and the value of time units that are not freely substitutable for work. The value of these time units need not increase by the same rate, the change being dependent on the income elasticity of the various activities: The value of those time units that figure extensively in the production of income-elastic activities is more sensitive than others to changes in income. Thus, the effect of a change in wage rates on the value of day hours may differ from the effect on night hours, and weekend hours may be affected differently from workday hours.

Similarly, changes in other sources of income, V, may affect the value of time by changing the number of working hours and the marginal wage rate, the value of the marginal utility of work, and the scarcity of time. Note that this change may have opposite effects on the value of time that can be used for work and time that cannot. Thus, although an increase in other sources of income is expected to increase the value of time, which cannot be converted into working time, it may result in a decline of the value of working hours if the reduction in these hours results in a decline in the marginal wage rate.

In conclusion, even under this set of simplified assumptions it cannot be argued that the value of time equals the average wage. There is no unique value of time. The set of values of the various time units may be positively affected by wages and other sources of income but is not equal to the wage rate.

Changes in the value of time, K, affect both the optimum combination of inputs in the production of each activity and the optimum combination of activities. An increase in the value of time results in a substitution of goods for time and a shift from time-intensive activities (whose relative price rises) to goods-intensive activities.

VALUE OF TIME ADAPTED TO TRANSPORTATION

Adapting the value of time to transportation calls for the removal of some of our simplifying assumptions. The demand for trips is usually a derived demand, the utility derived from the trip itself being only a part (and usually a small part) of the benefits accruing to the traveler; most of the benefits originate in the stay at the destination. Thus, a trip can be regarded as an intermediate activity in the production of a visit. First, to analyze the demand for trips, one must therefore introduce intermediate activities into the model. Second, most modes of travel (and in particular public transportation) do not preclude travelers' engaging in other activities (e.g., conversation, reading, and sometimes working). The assumption of no joint production of activities must therefore be removed. Finally, the assumption that the production function of trips is continuous must be released. In general, travelers cannot affect the traveling time and costs of a given mode. Each mode involves spending money and time in fixed proportions, and the production function is discontinuous. I shall remove these restrictive assumptions one by one.

Let it be assumed (4) that there are only four activities: a visit to some city, Z_V; a trip by mode A to that city, Z_A; a trip to it by mode B, Z_B; and all other activities, including work, $\overline{Z}$. The production functions of Z_A, Z_B, and $\overline{Z}$ are a function of market inputs and time (equation 5) and are continuous, and there is no joint production. The fourth production function, that of the visit Z_V, is somewhat different since it involves the intermediate activities Z_A and Z_B:

$$Z_V = F_V(X_V, T_V, Z_A, Z_B) \tag{11}$$

Maximizing the utility function

$$U = U(Z_V, Z_A, Z_B, \overline{Z}) \tag{12}$$

subject to the time and budget constraints yields the optimum combination of the trips by the two modes:

$$\frac{\partial Z_V/\partial Z_A}{\partial Z_V/\partial Z_B} = \frac{z_B}{z_A} = \frac{\Pi'_A}{\Pi'_B} = \frac{P_A x_A + Kt_A - (u_A/\lambda)}{P_B x_B + Kt_B - (u_B/\lambda)} \tag{13}$$

Since both modes are equally efficient in conveying passengers to their destinations they can be regarded as perfect substitutes in the production of a visit $z_B/z_A = 1$. The choice of mode therefore depends on the relative price, Π'_A/Π'_B. This in turn depends on the money expenditures, the opportunity cost of time, and the money equivalent of the marginal utility derived from the trip (u_i/λ for i = A, B). If the marginal utilities derived from the trip (u_A and u_B) are sufficiently small [i.e., if $u_i < \lambda(P_i x_i + Kt_i)$ for i = A, B], the price line in the relevant range slopes downward. Moreover, if the marginal utilities decline with the number of trips, the price line is concave to the origin (Figure 2).

Travelers who go Z_V° times split their trips: They go Z_A° times by mode A and Z_B° times by mode B. If $\Pi_A = P_A x_A + Kt_A$ is sufficiently different from $\Pi_B = P_B x_B + Kt_B$, or if u_i (for i = A or B) is sufficiently large (in the extreme case u_i may be large enough for the slope of the price line to become positive), the travelers may specialize, taking all their trips by one mode. In this case, mode A is preferred if

$$(1/\lambda)(u_A - u_B) + (P_B x_B - P_A x_A) + K(t_B - t_A) > 0 \tag{14}$$

where $Z_V^\circ = Z_A^\circ$ and $Z_B^\circ = 0$. Note that, in this inequality so often used in predicting modal choice and in estimating the value of time, there is nothing to ensure that the first term, $(1/\lambda)(u_A - u_B)$, is the same for different individuals and is uncorrelated with other terms of the equation. It may vary with income (resulting in differences in λ), with the total number of trips taken Z_V°, and with the length of the trip t_i.

The second assumption to be released is the assumption that there is no joint production of activities (6). Thus, let it be assumed that there are only three activities

$$U = U(Z_1, Z_2, \overline{Z}) \tag{15}$$

where Z_1 is the activity trip, Z_2 is reading, and $\overline{Z}$ are all other activities. Let

$$Z_1 = F_1(X_1, T_1) \qquad Z_2 = F(T_1) \tag{16}$$

where market inputs used in the production of reading are ignored. Maximizing the utility function subject to the constraint inplies that the optimum time spent in traveling is attained when

$$[(u_1/\lambda)/t_1] + [(u_2/\lambda)/t_2] = K \tag{17}$$

i.e., when the value of the marginal product of time in travel and reading equals the value of time. Thus, as long as the marginal unit of time yields utility in addition to the utility of the trip, travelers will be ready to pay less than K for any unit of time saved.

Finally, the assumption of continuity must be removed. The combination of time and money expenditure associated with any given mode is usually (in particular, in the case of public carriers) given to travelers and is not affected by their decisions. Put differently, a trip by a given mode is produced with fixed proportions of time and market goods. Travelers can change the proportions of time and goods only by switching to a different mode. Thus, if one ignores the direct utility derived from the trip and differences in the joint outputs of different modes, then all modes can be regarded as perfect substitutes. The combination of time and money expenditure associated with each mode can be regarded as a point on the isoquant of the activity trip. Let P_i denote the money expenditures (i.e., the units of market inputs are defined so that $x_i = 1$) and let T_i be the time involved in traveling by mode i. Let there be five modes, of which A is the fastest and most expensive and E the slowest and cheapest. In this case, the isoquant consists of five points A, B, C, D, and E (Figure 3). Since it is assumed that $u_i = 0$, the criterion for preferring mode A to mode B (equation 14) becomes

$$(P_B - P_A) + K(T_B - T_A) > 0 \tag{18}$$

i.e., mode A is preferred if the money differential between the two modes is offset by the value of time differential. Alternatively if i is a faster and more expensive mode than j, i is preferred to j if

$$K > (P_i - P_j)/(T_j - T_i) = K^*_{ij} \tag{19}$$

K* is the money differential divided by the time differential. Put differently, it is the amount of money travelers have to forgo to save one unit of time and can therefore be called the price of time. Thus, if their values of time exceed the price of time, travelers prefer the faster mode; otherwise, they choose the slower mode.

The slope of the price line is K, and the optimum combination of time and money expenditure is that where the price line touches the isoquant (mode B, Figure 3).

The optimum point is not a point of tangency, the isoquant being discontinuous at that point. The lack of tangency has in the past created some confusion about what is the value of time, how one should evaluate time savings, and how one can estimate the value of time (3, 5). It is clear from our analysis that there is only one way to evaluate time saving, namely, to use the household's evaluation. The discontinuity in the isoquant results in a difference between the value of time and the price of time. The household's evaluation is determined by its value of time. The price of time can serve only as an upper or lower boundary of that value. Thus, in Figure 3, the value of time is bounded by the two prices K^*_{AB} and K^*_{BC} ($K^*_{BC} < K < K^*_{AB}$). Had the travelers chosen the fastest mode A, the price of time could have served only as a lower boundary, $K > K^*_{AB}$. Had they chosen the slowest mode E, the price of time would have served only as an upper boundary, $K^*_{DE} > K$.

MODAL-CHOICE CHANGE AND DISTANCE OF TRIP

Both the value and the price of time may vary with distance of the trip so that modal choice changes with the distance of the trip. Thus, if it is assumed that the marginal utility of the trip and the marginal product of joint activities (e.g., reading, conversation, work) decrease with the length of the trip, the corresponding increase in the value of time should result in substitution of a faster for a slower mode.

The distance of the trip may also affect the price of time, if the time intensity of the various modes changes with distance. Money expenditures and time elapsed can, in general, be approximated as a linear function of distance M:

$$T_i = \alpha_{0i} + \alpha_{1i}M \qquad P_i = \beta_{0i} + \beta_{1i}M \tag{20}$$

where

- α_{0i} = fixed time component of a trip by mode i, e.g., access and egress time, waiting time at the terminal;
- α_{1i} = marginal time per kilometer, which depends on the speed of the mode;
- β_{0i} = fixed money cost component, e.g., access and egress costs, the fixed component of the fare; and
- β_{1i} = the marginal cost per kilometer (the marginal change in the fare).

By equation (19), the faster mode, i, is preferred if

$$K > \frac{(\beta_{0i} - \beta_{0j}) + (\beta_{1i} - \beta_{1j})M}{(\alpha_{0j} - \alpha_{0i}) + (\alpha_{1j} - \alpha_{1i})M} = K^*_{ij} \tag{21}$$

The fixed time component plays a major role in the choice of mode. Differences in access and egress time may offset any advantage a mode has in terms of marginal speed. Thus, mode i will never be chosen when $T_i > T_j$, i.e., assuming $\alpha_{1i} < \alpha_{1j}$, when

$$M < (\alpha_{0i} - \alpha_{0j})/(\alpha_{1j} - \alpha_{1i}) \tag{22}$$

The price of time K^*_{ij} is inversely related to the distance of the trip when an increase in distance increases the time differential ($T_j - T_i$) at a faster rate than the increase in the money differential ($P_i - P_j$),

$$\frac{\partial K^*_{ij}}{\partial M} < 0 <=> \frac{\partial(T_j - T_i)/\partial M}{T_j - T_i} > \frac{\partial(P_i - P_j)/\partial M}{P_i - P_j} \tag{23}$$

i.e., when

$$(\beta_{0i} - \beta_{0j})/(\beta_{1i} - \beta_{1j}) > (\alpha_{0j} - \alpha_{0i})/(\alpha_{1j} - \alpha_{1i}) \tag{24}$$

Assuming $\alpha_{1j} > \alpha_{1i}$ and $\beta_{1i} > \beta_{1j}$, a sufficient condition for equation (24) to be satisfied is $\beta_{0i} > \beta_{0j}$ and $\alpha_{0i} > \alpha_{0j}$, i.e., the fixed money and time components of the faster mode exceed those of the slower mode. Given the location of terminals (airports, rail, and bus stations) and the amount of waiting time required for things such as baggage handling and security checks, it seems that this sufficient condition is satisfied at least in the case of air travel versus ground travel. In this case the passenger does not use the faster mode unless

$$M > \frac{(\beta_{0i} - \beta_{0j}) + (\alpha_{0i} - \alpha_{0j})K}{(\beta_{1j} - \beta_{1i}) + (\alpha_{1j} - \alpha_{1i})K} = M^*_{ij} \tag{25}$$

The price of the trip consists of the money costs P and the opportunity cost of time KT, such that $\Pi = P + KT$. An increase in the value of time K increases the price of the trip by all modes (a shift from Π to Π' in Figure 4); however, it has a greater effect on the price of the time-intensive mode j, resulting in an inverse relationship between the value of time and the switching distance M*. The switching distance of travelers with a high value of time is smaller than that of low-value-of-time travelers ($M^*_1 < M^*_0$).

The relationship between K and M* specified by equation 25 can be described graphically by a rectangular hyperbola (Figure 5). Travelers with a value of time of K_0 use the faster mode only for distances exceeding M^*_0, and travelers with a value of time of K_1 ($K_1 > K_0$) switch to the faster mode already at a distance of M^*_1. Alternatively, the faster mode is chosen for a trip of M^*_1 kilometers only by travelers whose value of time exceeds K_1, and, for a trip of M^*_0 kilometers, the faster mode is preferred by everyone whose value of time exceeds K_0. Finally, the faster mode will never be used by travelers whose value of time falls short of

$$K < (\beta_{1i} - \beta_{1j})/(\alpha_{1j} - \alpha_{1i}) \tag{26}$$

CONCLUSIONS

I have shown that the economic approach that regards travel time as one of the determinants of the price of the trip provides us with a tool for analyzing the demand for transportation and modal split. However, to use this tool for policy decisions or forecasts of the demand for travel by the various modes one must know the value of time. The average wage can be used as, at best, a very crude approximation of the value of time. The difference between the average and the marginal wage rate, the cost incurred through work, the marginal disutility of work, the marginal utility of travel, the possibility of engaging in other activities while traveling, and institutional barriers to changes in the number of working hours may result in a significant divergence of the value of time from the average wage rate. To evaluate the size of this divergence and to obtain a better estimate of the value of time, one must rely on statistical estimation.

The theory suggests two possible approaches to the estimation problem: derivation of the value of time by observing the person's (or the community's) choice between various modes (or routes) and deduction of this value from the demand for a specific mode (trying to isolate the components of the price of the trip Π). Both methods abound with statistical and conceptual difficulties. Since these problems are described in detail in other studies, it is sufficient to mention only one conceptual problem that is emphasized by this analysis.

Figure 1.

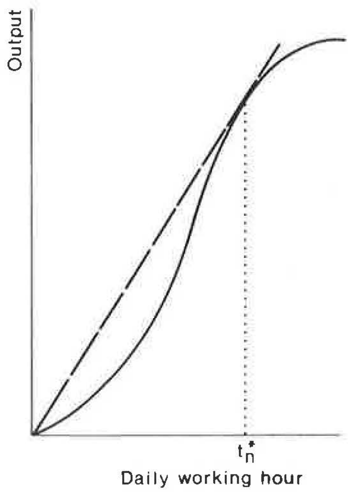


Figure 2.

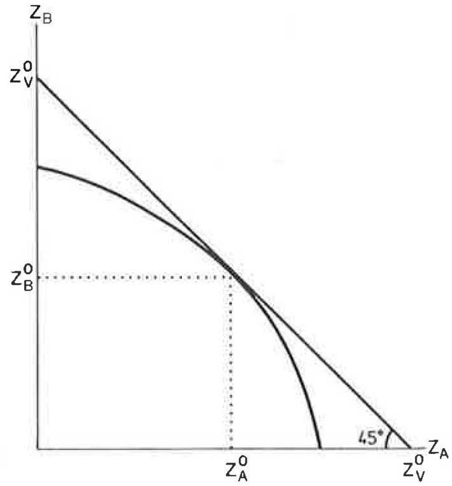


Figure 3.

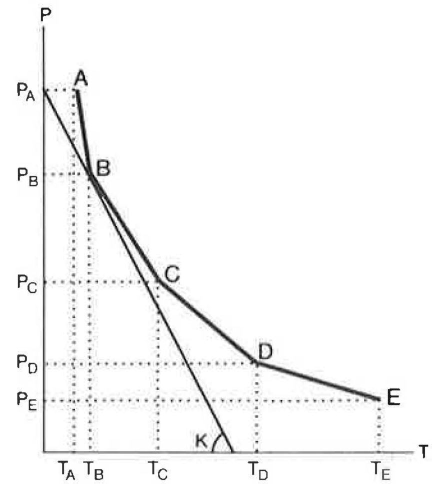


Figure 4.

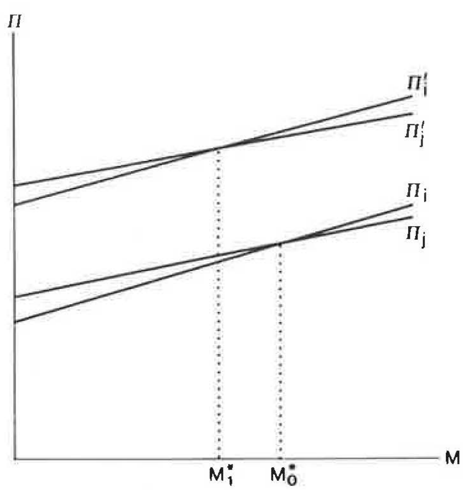
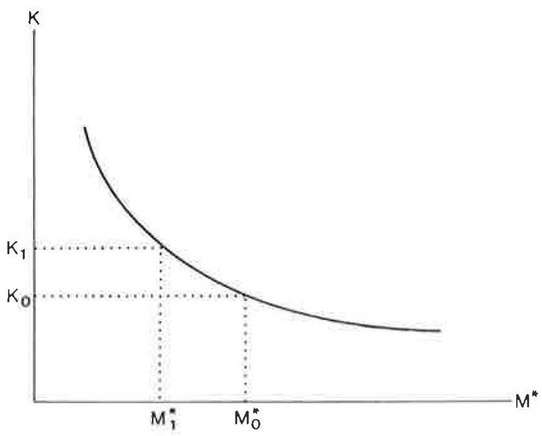


Figure 5.



There is no unique value of time. The value of time varies among individuals according to their income, wage rate, age, education, and family composition. Even for the same individual, the value of time may vary with the purpose and urgency of the trip, the time of day, and the season. Finally, travel decisions are affected by the value of time as well as by the direct utility generated by the trip, which is affected by the convenience, safety, and prestige of the mode of travel and the attributes (e.g., scenery) of the route. Most studies (if not all) fail to separate the direct utility from the value of time. For these reasons, one should not be surprised by the dispersion of the empirical findings. We may have to refine our tools, but this drawback should not diminish the usefulness of this new approach.

REFERENCES

1. Y. Barzel. The Determination of Daily Hours and Wages. Quarterly Journal of Economics, May 1973.
2. G. S. Becker. A Theory of the Allocation of Time. Economic Journal, Sept. 1965.
3. A. C. DeSerpa. A Theory of the Economics of Time. Economic Journal, Dec. 1971.
4. R. Gronau. The Value of Time in Passenger Transportation: The Demand for Air Travel. National Bureau of Economic Research, Columbia Univ. Press, New York, Occasional Paper 109, 1970.
5. R. Gronau. Price and Value of Time. HRB, Special Rept. 149, 1974, pp. 180-183.
6. M. Grossman. The Economics of Joint Production in the Household. Center for Mathematical Studies in Business and Economics, Rept. 7145, Sept. 1971.
7. B. M. Johnson. Travel Time and the Price of Leisure. Western Economic Journal, Spring 1966.
8. J. Mincer. Market Prices, Opportunity Costs, and Income Effects. In Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld (C. F. Christ, ed.), Stanford Univ. Press, Calif., 1963.
9. H. Mohring. Land Values and the Measurement of Highway Benefits. Journal of Political Economy, June 1961.
10. L. N. Moses and H. F. Williamson, Jr. Value of Time, Choice of Mode, and the Subsidy Issue in Urban Transportation. Journal of Political Economy, June 1963.