63

Reprinted from

TRANSPORTATION RESEARCH

Vol. 12, *No.* 4, *pp.* 257–289

ESTIMATING THE FUNCTIONAL FORM OF TRAVEL DEMAND MODELS

MARC J. I. GAUDRY and MICHAEL J. WILLS

Centre de Recherche sur les Transports, Université de Montréal, Montréal, Canada
and
Strategic Planning Group, Transport Canada, Ottawa, Canada

PERGAMON PRESS
NEW YORK · OXFORD · PARIS · FRANKFURT
1978

Transpn Res. Vol. 12, pp. 257–289
© Pergamon Press Ltd., 1978. Printed in Great Britain

0041-1647/78/0801-0257/$02.00/0

ESTIMATING THE FUNCTIONAL FORM OF TRAVEL DEMAND MODELS

MARC J. I. GAUDRY and MICHAEL J. WILLS
Centre de recherche sur les Transports, Université de Montréal, Montréal and Strategic Planning Group, Transport Canada, Ottawa, Canada

(*Received* 22 *November* 1977)

Abstract—This paper shows how many particular models found in the travel demand literature, such as the linear, multiplicative, logit forms and their formulation in terms of first differences, can be derived as special cases of the specification of the fixed and stochastic parts of more general models. Estimates of the appropriate functional form of the two equations of a cross-sectional inter-city travel demand model and of the two equations of a time-series urban travel demand model are presented. It is demonstrated, using current maximum likelihood techniques extended to take multiple-order auto-correlation into account, that estimating the functional form dominates more restricted procedures on grounds both of the statistical significance and of the economic reasonableness of the results. Notably, a standard logit market share equation and log-linear total demand equations appear to be clearly inferior to less restrictive specifications of these equations.

1. INTRODUCTION

1.1 *The significance of functional form*

During the past twenty years, numerous passenger travel demand models have been developed and used to forecast the demand which could be expected in potentially new or modified travel conditions. Typically, the parameters of a model or of a set of models with fixed mathematical form are estimated from historical data and the resulting values are used to forecast the impacts of various actions. Among those parameters, the price and service level parameters, or the elasticities derived from them, are of particular interest because some policy decisions are often based on the responses which they imply.

Yet these demand elasticities are known to be sensitive to the mathematical form assumed for the demand function; reasonable results for certain kinds of problems become associated with specific functional forms: sequential urban trip generation equations are linear, intercity direct demand equations are log-linear, and so on. In the absence of clear theoretical underpinning concerning the form to be used, it sometimes happens that numerous forms will be tried and the "best" results retained. This state of affairs suggests that demand elasticities obtained from arbitrarily selected functional forms are suspect and that the use of an incorrect functional form may constitute a specification error of unknown importance.

Our purpose in this paper is to explore this specification problem in three ways. Firstly, we show that several existing demand models are special cases of a more general fixed and stochastic model specification. Secondly, the exact functional form of two models, each of which consisting of two equations, is estimated from reliable data sets simultaneously with the other parameters of these models. Finally, the dependence of demand elasticities on the functional form is brought out not only in the intercity model estimated from cross-sectional data but also on the urban transit model estimated from time-series data. We conclude that incorrect functional form specification may lead not only to incorrect elasticities but even to erroneous signs of important parameters.

1.2 *Methodology: general form of the transformations*

The class of admissible functions for travel demand equations is intractably wide. One way to restrict it is to derive forms theoretically from more fundamental propositions. Although such methods have proved fruitful in the physical sciences, their application has been noticeably less productive in the socio-economic fields. An attractive alternative is to express the estimation problem in as general a form as possible and to let the data themselves place restrictions on this form.

Although existing models differ, many are in fact variations on either a linear form, in which the variables have independent additive effects, or on a product form, in which complete interaction is presumed to occur. It is useful to realize that these linear and product forms are in fact just specific points on a continuum which allows for varying amounts of independance and interaction among the variables. If no theory specifies *a priori* the optimal degree of that interaction, or if such a theory exists and should be tested, one might wish to estimate the degree of interaction using to that end a class of transformations known as the power family. These monotonic transformations have been studied successively by Anscombe and Tukey (1954), referred to in Tukey (1957), Box and Tidwell (1962), Box and Cox (1964), and by Zarembka (1974) in single equation contexts and applied by Spitzer (1977) in a simultaneous equations framework. We shall limit our investigation to a single-equation context.

Consider the general linear regression problem

$$y = X\beta + e, \tag{1}$$

where y is an $N \times 1$ vector of observations on the dependent variable, X is an $N \times (1+K)$ matrix of observations on a constant and K fixed or exogenous independent variables, β is a $(1+K) \times 1$ vector of parameters and e is an $N \times 1$ vector of independent, normally and identically distributed errors with mean and variance-covariance matrix as follows:

$$\begin{aligned} E(e) &= 0, \\ E(ee') &= \sigma^2 I, \end{aligned} \tag{2}$$

where I is the identity matrix of order $N \times N$. Covariance matrix $\sigma^2 I$ will henceforth be referred to as a scalar matrix for expository purposes. The Box–Tukey transformation of any variable z, where z denotes the dependent variable y or the kth non Boolean independent variable X_k, is defined for the transformation parameter λ, as

$$(z+\mu)^{(\lambda)} = \begin{cases} \dfrac{(z+\mu)^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \quad z+\mu > 0, \\ \ln(z+\mu), & \lambda \to 0, \quad z+\mu > 0. \end{cases} \tag{3}$$

The standard Box–Cox transformation is obtained as a special case by setting $\mu = 0$:

$$z^{(\lambda)} = \begin{cases} \dfrac{z^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \ln z, & \lambda \to 0. \end{cases} \tag{4}$$

The μ is a parameter introduced by Tukey (1957) in the context of the power model. Box and Cox (1964) have showed that the transformation is continuous at $\lambda = 0$ if it is specified as above. Both of these models are given with further details in the Box and Cox paper. Schlesselman (1971) has emphasized the importance of including a constant term in the regression model in order to preserve the invariance of the transformation to units of measurement.

In its most general formulation, the transformation can add two parameter vectors of $1+K$ elements to the original set $\{\beta, \sigma^2\}$: one for the location parameters, $\mu = (\mu_y, \mu_1, \ldots, \mu_k, \ldots, \mu_K)$, and the other for the power parameters, $\lambda = (\lambda_y, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_K)$. The set of parameters under consideration is therefore $\{\mu, \lambda, \beta, \sigma^2\}$. In our time-series section, we will further consider a vector ρ of autoregressive parameters.

1.3 *Nomenclature*

In order to embed the models of interest in this general context, we will adopt a consistent nomenclature. It consists of a four-letter code where the first letter distinguishes the transformation family, whether Box–Tukey (T) or Box–Cox (C) and the second specifies if these transformations are limited (L) to certain variables or extended (E) to all; the third letter indicates whether the transformation parameters λ and/or μ are constrained (C) to particular values, thereby giving special cases, or whether they are unconstrained (U). The final letter summarizes the structure of the covariance matrix which can be either scalar (S), diagonal (D), first (F) or multiple (M) order serial autoregressive, combined diagonal and autoregressive (C) or some more general (G) form.

As an example of this nomenclature, the most common forms of these transformation have been the CLUS and CECS models. These are, respectively, the Box–Cox transformation limited to the dependent variable but unconstrained with respect to the value of its parameter λ_y and using a scalar covariance matrix; and secondly, the Box–Cox transformation applied to all variables with the restriction that the transformation parameters contained in the vector λ are constrained equal and using a scalar covariance matrix as in (2).

1.4 *Scope*

1.4.1 *Selection of models to be estimated.* As our nomenclature implies and as we will note in our classification of existing cross-sectional and time-series models, the joint use of transformation parameters on observed variables or groups of variables and of parameters to specify the distribution of the error terms of these models generates a great variety of interesting candidates among which no more than a few can reasonably be investigated in one paper. When the time comes to choose among and estimate the parameters of models, certain types will perforce be omitted: in the cross-sectional analysis, we will make no investigation of alternative forms of the covariance matrix; the time-series part corrects this shortcoming by undertaking a fairly thorough examination of serially autocorrelated errors, but at the expense of restricting the analysis of the transformations. We will not formally consider situations of heteroskedasticity in which the variance of the disturbance, σ^2, is not constant.

1.4.2 *Criteria for the comparison of estimated models.* 1.4.2.1 *The likelihood function.* Within the context adopted here, the fundamental criterion for the comparison of models is how well they are able to explain the data. One way to achieve this is to maximize the likelihood of the sample. Under the assumptions stated in (2) above for the disturbance vector e, the likelihood of the original observations y for the CECS transformation can be formulated as:

$$\mathscr{L}(y; \beta, \sigma^2, \lambda) = (2\pi\sigma^2)^{-N/2} \times \exp\left[-\frac{(y^{(\lambda)} - X^{(y)}\beta)'(y^{(\lambda)} - X^{(\lambda)}\beta)}{2\sigma^2} \right] . |J(\lambda; y)|, \tag{5}$$

where $|J(\lambda; y)|$ is the absolute value of the Jacobian of the transformation from $y_t^{(\lambda)}$ to the actually observed y_t, or

$$|J(\lambda; y)| = \left| \det\left[\frac{\partial y_t^{(\lambda)}}{\partial y_{t'}} \right] \right| = \prod_{t=1}^{N} y_t^{\lambda - 1}. \tag{6}$$

Because $J(\lambda;y)$ is positive, we will henceforth use it instead of $|J(\lambda;y)|$. As emphasized by Zarembka (1974), the Box–Tukey transformation can be handled similarly to the procedure outlined here.

It is convenient to maximize the log concentrated likelihood function. Equation (5) is concentrated by obtaining an analytical estimate for σ^2, which is $S(\lambda)/N$ for a given λ, where $S(\lambda)$ is the residual sum of squares. The maximized log likelihood, for fixed λ, is then

$$L(y;\lambda) = -\frac{N}{2}[\ln(2\pi)+1] - \frac{N}{2}\ln\frac{S(\lambda)}{N} + (\lambda-1)\sum_{t=1}^{N}\ln y_t \tag{7}$$

the last term of which vanishes if, as Box and Cox suggest, the normalized variate is used:

$$W^{(\lambda)} = \frac{y^{(\lambda)}}{J(\lambda;y)^{1/N}} = \frac{y^{\lambda}-1}{\lambda\left[\left(\prod_{t=1}^{N} y_t\right)^{1/N}\right]^{\lambda-1}} = \frac{y^{\lambda}-1}{\lambda\dot{y}^{\lambda-1}}, \tag{8}$$

where $\dot{y}$ is the geometric mean of the original observations.

1.4.2.2 *Statistical and practical tests.* The natural test of the results as they pertain to the functional form parameters is to compare the values of the maximized likelihood function by means of the likelihood ratio test. If Ω is the parameter space under a maintained hypothesis and ω the restricted space generated by a null hypothesis, the test statistic is well known to be

$$2[L(\Omega) - L(\omega)], \tag{9}$$

where $L(\Omega)$ is the maximized log likelihood under the maintained hypothesis and $L(\omega)$ is the maximized log likelihood under the restricted hypothesis. The test statistic (9) has a χ^2 distribution with q degrees of freedom, one for each restriction appearing in ω but not in Ω.

Because the values of the parameters which govern the functional form are not independent from those of the parameters which specify the distribution of the residuals, they should be examined jointly. However, in the absence of a readily available test for heteroskedasticity, the impact of the functional form on the skewness and kurtosis of the cross-sectional model residuals will simply be illustrated and its impact on the variance of the time-series model residuals will be conjectured upon; the time-series analysis will contain a joint examination of functional form and autoregressive structure parameters.

Another way of comparing models is to study the size and significance of the β parameters in (5). This will lead us to examine their t-statistics and the derived demand elasticities; the variation in the elasticities computed from models of different functional forms is, from an applied point of view, probably the most crucial criterion for distinguishing the properties of one model from those of another. Except in the purely multiplicative model, the elasticities are not immediately apparent but can be derived analytically in a straightforward manner.

2. APPLICATION TO A CROSS-SECTIONAL MODEL OF INTERCITY TRAVEL DEMAND

Although numerous cross-sectional models have been developed in various travel contexts, our perspective is restricted to the class of direct and quasi-direct models.

2.1 *Context and theoretical framework.*

2.1.1 *Notation and definitions.* The cross-sectional models of interest use distinct or composite variables on travel flows, socio-economic variables and transportation mode attributes. We shall use the following definitions:

(a) travel flow variables: T_{ijm} = travel demand between i and j by mode m; $T_{ij} = \sum_m T_{ijm}$ = total travel between i and j; $S_{ijm} = T_{ijm}/T_{ij}$ = share of travel by mode m between i and j.

(b) socio-economic and activity variables: A_{ijr} = observations on the rth socio-economic activity variable for the (i,j)th city pair. Elements of this may include population sizes, income levels and distributions, cultural aspects (particularly linguistic composition), employment level and composition, and functions of these elements.

(c) transportation attribute variables: C_{ijms} =observations on the sth "level of service", generalized cost, distance or other impedance variable for mode m from i to j. Elements of this set include travel time, waiting or transfer time, fare or other money costs, frequency of departure, comfort, convenience, distance and functions of these elements, including perception transformations where applicable. U_{ijm} = a calculated generalized cost or impedance term for mode m from i to j, typically a function of a group of C_{ijms}'s.

2.1.2 *Description and classification of some previous cross-sectional models.* Direct demand models are of two types: purely direct, using a single estimated equation to relate the travel demand by mode directly to all variables considered, and quasi-direct, employing a form of separability between mode split and total travel demand problems. Within this classification, a further distinction may be made between mode-specific, mode-abstract and quasi-mode-abstract types.

2.1.2.1 *Direct models.* (a) Direct and specific models. One of the earliest and best known of this class is the SARC (Kraft, 1963) model, the implicit form of which may be written as

$$T_{ijm} = \beta_{mo}\prod_{r}^{R} A_{ijr}^{\beta_{mr}} \prod_{n}^{M}\prod_{s}^{S} C_{ijns}^{\alpha_{mns}}, \quad (m = 1, 2, \ldots, M). \tag{10}$$

This is clearly a multiplicative model with potentially very many parameters. Generalized impedance is obtained by an exponentially weighted product of modal attributes. With $n = m$, own attribute elasticities are obtained and, for $n \neq m$, cross elasticities. As is well-known for multiplicative models of the Cobb–Douglas type, the elasticities are constant and independent of the level of the variables. Alternative forms, containing linear and exponential terms in addition to multiplicative ones, have been estimated by Domencich, Kraft and Valette (1968) in an urban context.

(b) Direct and abstract models. In 1966, two papers appeared which argued that it is the attributes of commodities which are of interest rather than the commodities themselves (Lancaster; Quandt and Baumol). The ideas of Quandt and Baumol were implemented in an "abstract mode" demand model the coefficients of which were relieved of their modal subscripts, yielding parameters which were estimated by pooling the observations on modes. This device maps observations from mode space into attribute space, thereby producing a single demand function for all modes. Modes differ only by their relative weights in attribute space. Having freed the equation of identification with specific modes, it was held that it could be used for new modes without respecification.

The functional form of the original model of Quandt and Baumol (1966) may be summarized as:

$$T_{ijm} = \beta_0 \prod_r A_{ijr}^{\beta_r} \prod_s C_{ijbs}^{\gamma_s} \prod_s (C_{ijms}/C_{ijbs})^{\alpha_s}, \tag{11}$$

where C_{ijbs} is the "best" mode in the sth attribute from i to j. This dependency on the best mode for mode split and generalized impedance is a weakness which may be partly removed, however, by use of the geometric mean over modes in place of the best mode (Crow, Young and Cooley, 1973).

Although widely used, the validity of the pure abstract mode model can be questioned. The specification error introduced by imposing what amounts to equality constraints on the modal coefficients may be significant. In fact, Young (1969) has shown as much. Nonetheless, the abstract model, even in its pure form, does seem to be able to reveal the structure of the travel market as a whole. Yet it can be debated whether the resulting elasticities, since they pertain to an average of the travel market, are sufficiently representative of any individual mode, including a new mode. Many variants of the abstract mode model have been attempted on a heuristic basis. Some of these may have led to the quasi-direct models which are sometimes quasi-abstract as well and to which we shall presently turn.

(c) Classification. The models described in eqns (10) and (11) are special cases of CECS models in which $\lambda = 0$. The model formulated by Domencich *et al.* was also estimated under the assumption that $\lambda = 1$ and under mixed sets of constraints

$$\lambda = (\lambda_y = 0;\quad \lambda_1, \ldots, \lambda_k = 0;\quad \lambda_{k+1}, \ldots, \lambda_k = 1).$$

and

$$\lambda = (\lambda_y = 1;\quad \lambda_1, \ldots, \lambda_k = 0;\quad \lambda_{k+1}, \ldots, \lambda_K = 1).$$

Young's work was centered on a CECG model with $\lambda = 0$.

2.1.2.2 *Quasi-direct models.* Quasi-direct models are so named because they replace the direct estimation of travel by mode by the successive estimation of the market share function and of the total demand for travel irrespective of mode. Hence

$$T_{ijm} = T_{ij} \cdot S_{ijm}, \tag{12}$$

in which the dot separates total demand and market share parts.

(a) Quasi-direct and specific models. One of the first of these models is McLynn's well known composite analytic model (McLynn *et al.*, 1968, 1969). In order to explain the market share of each mode, a mode split equation is estimated; the parameters so obtained are then fixed and the denominator of the market share model is used as a generalized impedance term in the total demand equation which is thereafter estimated. The form of this model may be written as

$$T_{ijm} = \beta_0 [\prod_r A_{ijr}^{\beta_r}] \Big[\sum_m U_m\Big]^\gamma \cdot \Big[U_m / \sum_m U_m\Big], \tag{13}$$

where $U_m = \alpha_{m0} \prod_s C_{ijms}^{\alpha_{ms}}$ and the right hand expression is written, like (12), such as to distinguish between what determines total demand and what explains modal shares. The advantage of this model over the SARC model is that the number of parameters is reduced while retaining the essence of mode specificity. There still remains, however, a considerable set of parameters to be estimated as separable sub-sets in a two-step procedure.

(b) Quasi-direct and quasi-abstract. A relatively unknown model has been developed by Monsod (1967, 1969): it closely resembles McLynn's model but uses a base mode for the mode split problem and its rationale is predominantly abstract. Again, using the denominator of the market share part as a generalized impedance term leads to

$$T_{ijm} = \beta_0 [\prod_r A_{ijr}^{\beta_r}] \Big[\sum_m U_m\Big]^\gamma \cdot \Big[U_m / \sum U_m\Big], \tag{14}$$

where the dot separates the explanation of total demand from the explanation of market shares, where $U_m = \alpha_{m0} \prod_s C_{ijms}^{\alpha_s}$ and $\alpha_{m0} = 1$ if m is the base mode. This device is due to the fact that mode split is estimated as a ratio of the base mode's attributes: the constant in the mode split equation differs from one only for non-base modes. In the original formulation, Monsod constrained the α_{m0}'s of the mode split part to be equal except for the base mode $m = b$, and then set all α_{m0}'s in the impedance term of the total demand part equal to one.

(c) Classification. As we shall presently see, both eqns (13) and (14), which differ only by the constraint imposed on cost or level of service exponents α_{ms}'s contain market share and total demand parts which respectively constitute a CLCS model with $\lambda_k = 1$ and a CECS model with $\lambda = 0$. We therefore turn to the explicit exploration of the functional form of both market share and total demand equations.

2.1.3 *A family of models and their associated demand elasticities.* We shall explicitly consider families of mode split and total demand models. The market share models may or may not be destined to yield information for use in an eventual generalized impedance term of the total demand equation; the basic structure of the discussion of the total demand equation will be applicable to other contexts. We will drop i, j subscripts to simplify the presentation.

2.1.3.1 *The generalized market share model and special cases.* Let the modal generalized cost be defined as

$$U_m = \exp\left(\alpha_{m0} + \sum_k \alpha_{mk}(C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right),$$

$$k = 1, \ldots, K. \quad (15)$$

where

$$\lambda_1 = (\lambda_{1,y}, \lambda_{1,1}, \ldots, \lambda_{1,k}, \ldots, \lambda_{1,K})$$

and

$$\mu_1 = (\mu_{1,y}, \ldots, \mu_{1,k}, \ldots, \mu_{1,K})$$

are transformation parameter vectors associated with the modal share equation and

$$(C_{mk} + \mu_{1,k})^{(\lambda_{1,k})} = \frac{(C_{mk} + \mu_{1,k})^{\lambda_{1,k}} - 1}{\lambda_{1,k}}$$

according to (3). Then the market share demand equation

$$S_m = U_m \left[\sum_m U_m\right]^{-1} \quad (16)$$

becomes, by substitution of (15) into (16) and by transformation of the dependent variable

$$(S_m + \mu_{1,y})^{(\lambda_{1,y})} = \frac{\exp\,(\alpha_{m0} + \sum_k \alpha_{mk}(C_{mk} + \mu_{1,k})^{(\lambda_{1,k})})}{\sum_m \exp\,(\alpha_{m0} + \sum_k \alpha_{mk}(C_{mk} + \mu_{1,k})^{(\lambda_{1,k})})}. \quad (17)$$

To obtain known forms from the general market share model proposed in (17), one must first set $\mu_{1,y} = 0$ and $\lambda_{1,y} = 1$ in order to use as dependent variable the untransformed market share. The use of that constraint and the imposition of the following equality constraint

$$\alpha_{mk} = \alpha_k, \qquad m = 1, 2, \ldots, M; \quad k = 1, 2, \ldots, K, \quad (18)$$

to (17) results in a quasi-abstract model:

$$S_m = \frac{\exp\left(\alpha_{m0} + \sum_k \alpha_k (C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right)}{\sum_m \exp\left(\alpha_{m0} + \sum_k \alpha_k + \mu_{1,k})^{(\lambda_{1,k})}\right)}. \quad (19)$$

This model may be further restricted to a pure abstract mode format by the additional equality constraint on the mode specific constants in (19)

$$\alpha_{m0} = \alpha_0, \qquad m = 1, 2, \ldots, M. \quad (20)$$

In the absence of constraint (18), the own elasticity of market share derived from (19) is found to be an expression containing not only the parameters but also the market share, S_m, and the value of the cost variable:

$$n_s = \frac{\partial S_m}{\partial C_{mk}}\left(\frac{C_{mk}}{S_m}\right) = \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}(1 - S_m). \quad (21)$$

At this point, it is of interest to highlight special cases contained in the subclass defined by model (19) by setting the remaining transformation parameters of vectors λ_1 and μ_1 to specified values. Setting $\mu_{1,k} = 0$ for all k, the more restricted Box–Cox logit form is obtained:

$$S_m = \frac{\exp\left(\alpha_{m0} + \sum_k \alpha_k C_{mk}^{(\lambda_{1,k})}\right)}{\sum_m \exp\left(\alpha_{m0} + \sum_k \alpha_k C_{mk}^{(\lambda_{1,k})}\right)}. \quad (22)$$

If, in addition, the $\lambda_{1,k}$'s $\to 0$ for all k, the multiplicative or product form is derived:

$$S_m = \frac{\exp(\alpha_{m0})\prod_k C_{mk}^{\alpha_k}}{\sum_m \exp\,(\alpha_{m0})\prod_k C_{mk}^{\alpha_k}}. \quad (23)$$

That basic form is used in the mode split parts of McLynn's and of Monsod's models mentioned above. Finally, if, instead of letting $\lambda_{1,k}$'s $\to 0$, they are set to unity, the standard linear in parameters logit model is obtained:

$$S_m = \frac{\exp\left(\alpha_{m0} + \sum_k \alpha_k C_{mk}\right)}{\sum_m \exp\left(\alpha_{m0} + \sum_k \alpha_k C_{mk}\right)}. \quad (24)$$

The own market share elasticity corresponding to (22) is

$$n_S = \alpha_k C_{mk}^{\lambda_{1,k}}(1 - S_m); \quad (25)$$

for the corresponding multiplicative model (23), it is

$$n_S = \alpha_k(1 - S_m) \quad (26)$$

and the same elasticity for the standard logit model given by (24) is

$$n_S = \alpha_k C_{mk}(1 - S_m). \quad (27)$$

2.1.3.2 *The generalized total demand model and special cases.* The estimation of the total demand for travel between a given city pair involves a set of socioeconomic activity variables A and the general impedance or cost obtained from market share equation estimates. This is given by

$$T = \left[\beta_0 + \sum_k \beta_k (A_k + \mu_{2,k})^{(\lambda_{2,k})} + \beta_{k+1}(U + \mu_{2,k+1})^{(\lambda_{2,k+1})}\right]^{(\lambda_{2,y},\mu_{2,y})^{-1}}, \quad (28)$$

where $\lambda_2 = (\lambda_{2,y}, \lambda_{2,1}, \ldots, \lambda_{2,k+1})$ and $\mu_2 = (\mu_{2,y}, \mu_{2,1}, \ldots, \mu_{2,k+1})$ denote transformation parameter vectors associated with the total trip equation, where

$$U = \sum_m U_m,$$

$$U_m = \exp\left(\alpha_{m0} + \sum_k \alpha_{mk}(C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right)$$

and $(\lambda_{2,y}, \mu_{2,y})^{-1}$ indicates the inverse of the Box–Tukey transformation applied to the left hand side of (28). As in the market share equation, several special cases may be obtained by applying restrictions. Irrespective of the functional form adopted in the market share equation, any of the following are simply derived. Setting $\mu_2 = 0$ leads to the Box–Cox case:

$$T = \left[\beta_0 + \sum_k \beta_k A_k^{(\lambda_{2,k})} + \beta_{k+1} U^{(\lambda_{2,k+1})}\right]^{(\lambda_{2,y})^{-1}}. \quad (29)$$

A special case of eqn (29) was considered by Kau and Sirmans (1976). If, in addition, all elements of λ_2 tend to zero, a product form is obtained as in models (10), (11) and in the total demand part of models (13) and (14). This form includes the gravity model:

$$T = \exp(\beta_0)[\prod_k A_k^{\beta_k}]U^{\beta_{k+1}}. \quad (30)$$

Setting all elements of λ_2 equal to 1 leads to a linear form for the total travel equation:

$$T = \beta_0 + \sum_k \beta_k A_k + \beta_{k+1} U, \quad (31)$$

whereas maintaining $\lambda_{2,y} \to 0$ leads to an exponential form

$$T = \exp\left(\beta_0 + \sum_k \beta_k A_k + \beta_{k+1} U\right). \quad (32)$$

The exponential function used to obtain (30) and (32) is clearly the inverse of the Box–Cox transformation for $\lambda_{2,y} \to 0$.

As in the market share equation, the own elasticity with respect to travel cost or attribute contains not only parameters but varies with market share and the value of the cost variable considered:

$$\begin{aligned}\eta_T &= \frac{\partial T}{\partial C_{mk}}\left(\frac{C_{mk}}{T}\right)\\ &= \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}\beta_{k+1} U_m\\ &\quad \times (U + \mu_{2,k+1})^{\lambda_{2,k+1}-1}(T + \mu_{2,y})^{1-\lambda_{2,y}} T^{-1}, \quad (33)\end{aligned}$$

where

$$T = \left[1 + \lambda_{2,y}\left(\beta_0 + \sum_k \beta_k (A_k + \mu_{2,k})^{(\lambda_{2,k})} + \beta_{k+1}(U + \mu_{2,k+1})^{(\lambda_{2,k+1})}\right)\right]^{\lambda_{2,y}^{-1}} - \mu_{2,y}$$

and U and U_m are as defined for eqn (28).

If the transformation $(\lambda_{1,k}, \mu_{1,k})$ for the market share equation is given, the total travel own elasticities of demand for special cases (29) to (32) can be expressed as follows. For the Box–Cox case, (33) simplifies to

$$\eta_T = \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}\beta_{k+1} U_m U^{\lambda_{2,k+1}-1} T^{-\lambda_{2,y}} \quad (34)$$

where

$$T = \left[1 + \lambda_{2,y}\left(\beta_0 + \sum_k \beta_k A_k^{(\lambda_{2,k})} + \beta_{k+1} U^{(\lambda_{2,k+1})}\right)\right]^{\lambda_{2,y}^{-1}}.$$

The corresponding elasticity for the product form (30) is obtained by, in addition, letting $\lambda_{2,k} \to 0$ for all k, thus reducing (34) to

$$\eta_T = \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}\beta_{k+1} S_m; \quad (35)$$

and, for the linear model, the same elasticity is

$$\eta_T = \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}\beta_{k+1} U_m T^{-1}, \quad (36)$$

where

$$T = (\beta_0 + 1) + \sum_k \beta_k A_k + \beta_{k+1} U,$$

whereas, for the exponential model given by (32), it is

$$\eta_T = \alpha_{mk}(C_{mk} + \mu_{1,k})^{\lambda_{1,k}-1} C_{mk}\beta_{k+1} U_m. \quad (37)$$

2.2 *Model selection and estimation procedure*

2.2.1 *Model considered.* The basic model format selected for investigation is that obtained by combining the quasi-abstract mode split eqn (19) with total travel demand eqn (28), namely the following formulation:

$$T = \left[\beta_0 + \sum_k \beta_k (A_k + \mu_{2,k})^{(\lambda_{2,k})} + \beta_{k+1}(U + \mu_{2,k+1})^{(\lambda_{2,k+1})}\right]^{(\lambda_{2,y},\mu_{2,y})^{-1}} \quad (38)$$

$$\times \frac{\exp\left(\alpha_{m0} + \sum_k \alpha_k (C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right)}{\sum_m \exp\left(\alpha_{m0} + \sum_k \alpha_k (C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right)},$$

where

$$U = \sum_m \exp\left(\alpha_{m0} + \sum_k \alpha_k (C_{mk} + \mu_{1,k})^{(\lambda_{1,k})}\right).$$

This format is a generalization of the particular model ($\lambda_{1,y} = 1, \lambda_{1,k} = 0$, $\mu_1 = 0$; $\lambda_2 = \mu_2 = 0$) found in Wills (1974) and in a Canadian Transportation Commission study (1976).

The explanatory variables in the mode split equation are, in addition to the mode-specific constants, the fare F, the travel time H and travel frequency D of four intercity travel modes: car, air, rail and bus. The sample consists of 368 observations (4 modes × 92 city pairs) on these 7 variables. The variables used in the total demand equation are, in addition to the constant and the general impedance term U, a population product variable $A_1 = P_iP_j$ and a linguistic pairing index $A_2 = L_{ij}$; the latter factor is a measure of linguistic homogeneity defined as $L_{ij} = 100 - |L_i - L_j|$ where L_i is the proportion of the population speaking English in the ith city. The sample consists of 92 observations on these 4 variables for the year 1972 in Canada.

2.2.2 *Rearrangements for estimation and models estimated.* As required by the quasi-abstract format of eqn (38), the mode split part is estimated before and separately from the total demand part.

2.2.2.1 *Market share models.* Taking logarithms of ratios of the market shares by mode and of the modal attributes leads to the following form for the mode split equation

$$\ln(T_m/T_b) = \alpha_{m0} + \sum_k \alpha_k((C_{mk} + \mu_{1,k})^{(\lambda_{1,k})} - (C_{bk} + \mu_{1,k})^{(\lambda_{1,k})} + e, \quad (39)$$

where e is assumed to have the properties defined in (2), b denotes a base mode and α_{b0}, the logarithm of 1, equals zero. This transformation reduces the effective sample by 25% from 368 to 276 observations and yields a form which can be estimated as a regression problem. We shall select for scrutiny the following models which are all of the L type defined in our nomenclature because the dependent variable T_m/T_b in the basic specification (38) is untransformed.

(a) CLCS-1 model:

$$\ln(T_m/T_b) = \alpha_{m0} + \alpha_1(F_m^{(\lambda_{1,1})} - F_b^{\lambda_{1,1})}) + \alpha_2(H_m^{(\lambda_{1,2})} - H_b^{(\lambda_{1,2})}) + \alpha_3(D_m^{(\lambda_{1,3})} - D_b^{(\lambda_{1,3})}) + e, \quad (40)$$

where $\lambda_1 = (\lambda_{1,y} = 0,\ \lambda_{1,1} = \lambda_{1,2} = \lambda_{1,3})$ is vector of constraints.

(b) CLCS-2 model:

As in (40), but allowing $\lambda_{1,1} \neq \lambda_{1,2} \neq \lambda_{1,3}$. (40′)

(c) TLCS-1 model:

$$\ln(T_m/T_b) = \alpha_{m0} + \alpha_1((F_m + \mu_{1,1})^{(\lambda_{1,1})} - (F_b + \mu_{1,1})^{(\lambda_{1,1})}) + \alpha_3((D_m + \mu_{1,3})^{(\lambda_{1,3})} - (D_b + \mu_{1,3})^{(\lambda_{1,3})}) + e, \quad (41)$$

where $\lambda_1 = (\lambda_{1,y} = 0,\ \lambda_{1,1} = \lambda_{1,2} = \lambda_{1,3})$ and $\mu_1 = (\mu_{1,y} = 0,\ \mu_{1,1} = \mu_{1,2} = \mu_{1,3})$ as vectors of constraints.

(d) TLCS-2 model:

As in (41), but allowing $\lambda_{1,1} \neq \lambda_{1,2} \neq \lambda_{1,3}$. (42)

2.2.2.2 *Total demand models.* The total demand equation will be estimated as

$$(T + \mu_{2,y})^{\lambda_{2,y}} = \beta_0 + \sum_k \beta_k(A_k + \mu_{2,k})^{(\lambda_{2,k})} + \beta_{k+1}(U + \mu_{2,k+1})^{(\lambda_{2,k+1})} + e, \quad (43)$$

where e is assumed to have the properties defined in (2) and which is also a linear regression problem for given λ_2 and μ_2 vectors.

We shall estimate the following models which are of the E type because both dependent and independent variables are transformed.

(a) CECS-1 model:

$$T^{(\lambda_{2,y})} = \beta_0 + \beta_1 A_1^{(\lambda_{2,1})} + \beta_2 A_2^{(\lambda_{2,2})} + \beta_3 U^{(\lambda_{2,3})} + e, \quad (44)$$

where

$$U = \sum_m \exp(\alpha_{m0} + \alpha_1(F_m + \mu_{1,1})^{(\lambda_{1,1})} + \alpha_2(H_m + \mu_{1,2})^{(\lambda_{1,2})} + \alpha_3(D_m + \mu_{1,3})^{(\lambda_{1,3})}),$$

with $\{\alpha_{m0}, \alpha_k, \lambda_{1,k}$ and $\mu_{1,k}\}$ set at their optimal values from the market share equation and subject to $\lambda_{2,y} = \lambda_{2,1} = \lambda_{2,2} = \lambda_{2,3}$.

(b) CECS-2 model:

As in (44), but allowing $\lambda_{2,y} \neq (\lambda_{2,1} = \lambda_{2,2} = \lambda_{2,3})$. (45)

(c) CEUS model:

As in (45), but allowing $\lambda_{2,1} \neq \lambda_{2,2} \neq \lambda_{2,3}$. (46)

2.2.3 *Estimation procedure.* In the case of the market share equation, the log concentrated likelihood function for errors distributed as in (2) is, except for a constant term

$$L_1(\lambda_1, \mu_1) = -\frac{N}{2}\ln\left(\sum_t e_t^2/N\right) + \ln J(\lambda_{1,y}), \quad (47)$$

where

$$e_t = \ln(T_m/T_b) - \left[\hat{\alpha}_{m0} + \sum_k \hat{\alpha}_k((C_{mk} + \mu_{1,k})^{(\lambda_{1,k})} - (C_{bk} + \mu_{1,k})^{(\lambda_{1,k})})\right]$$

and

$$J(\lambda_{1,y}) = \prod_{t=1}^{N} y_t^{-1}$$

because $\lambda_{1,y}$ of $y = T_m/T_b$ is equal to 0. For the total travel equation, the appropriate likelihood function is

$$L_2(\lambda_1, \mu_1, \lambda_2, \mu_2) = -\frac{N}{2}\ln\left(\sum_t e_t^2/N\right) + \ln J(\lambda_{2,y}, \mu_{2,y}) \quad (48)$$

where

$$e_T = (T + \mu_{2,y})^{(\lambda_{2,y})} - \left[\hat{\beta}_0 + \sum_k \hat{\beta}_k(A_k + \mu_{2,k})^{(\lambda_{2,k})} + \hat{\beta}_{k+1}(U + \mu_{2,k+1})^{(\lambda_{2,k+1})}\right]$$

and

$$J(\lambda_{2,y}, \mu_{2,y}) = \prod_{t=1}^{N} (y_t + \mu_{2,y})^{\lambda_{2,y}-1}.$$

Likelihood functions (47) and (48) are maximized by nonlinear methods or more simply by scanning.

2.3 *Results for the mode split equation*

The mode split or market share equation is estimated by maximizing the likelihood function given by (47) for the CLCS, TLCS and usual models. Likelihood ratio tests are used to test the significance of transformation vectors λ_1 and μ_1. The impact of these transformations on the distribution of errors is examined in terms of the skewness and kurtosis of the residuals. It is also analyzed in the context of the statistical significance of the variables and of the modal attribute own elasticities of market share. To simplify the notation, subscript 1 will be dropped from $\lambda_1 = (\lambda_1, \lambda_{1,k})$ and $\mu_1 = (\mu_{1,y}, \mu_{1,k})$.

2.3.1 *Estimation and tests of the functional form.* In order to test hypotheses about functional form parameters, the log likelihood test referred to in (9) will be used. Within this framework, two types of tests will be carried out: conditional tests, which assume a value for one of the parameters, and then test the other; unconditional tests, which make no assumption about the untested parameter which is left unrestricted. We will successively consider TLCS-1 and TLCS-2 models.

2.3.1.1 *Model TLCS*-1. The TLCS-1 model is estimated for a grid of λ_k's and μ_k's; the corresponding log likelihood function (L_1) is plotted in Fig. 1. Clearly there is considerable variation due to different values of λ_k, reaching a well-defined maximum for each value of μ_k. At $\mu_k = 0$, $L_{1\,\max} = -0.2$ indicating strong interaction effects among the variables. Whereas the likelihood function is symmetrical at $\mu_k = 0$, increasing values of μ_k lead to increasing asymmetry. That the effectiveness of μ_k is a function of λ_k is clearly shown. At $\lambda_k = 1$, non-zero values of μ_k result in a simple translation of all variables by the same value, a property which leaves the results unaltered. The same result is shown on a portion of the log likelihood surface near the optimum, in Fig. 2.

Three specific hypotheses summarized in Table 1 are tested. Firstly, given μ_k, λ_k is tested for being significantly different from a given value $\tilde{\lambda}_k$. This test is simply

$$C(\lambda_k) = L_1(\hat{\lambda}_k, \mu_k = \tilde{\mu}_k) - L_1(\lambda_k = \tilde{\lambda}_k, \mu_k = \tilde{\mu}_k) > \frac{1}{2}\chi_\alpha^2(1), \tag{49}$$

with 1 degree of freedom and significance level α. A second test reverses the roles of λ_k and μ_k in (49):

$$C(\mu_k) = L_1(\hat{\mu}_k, \lambda_k = \tilde{\lambda}_k) - L_1(\mu_k = \tilde{\mu}_k, \lambda_k = \tilde{\lambda}_k) > \frac{1}{2}\chi_\alpha^2(1). \tag{50}$$

Finally, both λ_k and μ_k are tested simultaneously:

$$U(\lambda_k, \mu_k) = L_1(\hat{\lambda}_k, \hat{\mu}_k) - L_1(\lambda_k = \tilde{\lambda}_k, \mu_k = \tilde{\mu}_k) > \frac{1}{2}\chi_\alpha^2(2); \tag{51}$$

there are two degrees of freedom since two new restrictions are imposed by null hypothesis.

Results are given in Tables 2–4. In Table 2, it is shown that, conditional upon $\mu_k = 0$, λ_k is equal to -0.2 at the optimum and also that the interval between -0.3 and -0.1 contains λ_k with a probability of 95%; that λ_k is zero is rejected at the 0.5% level and λ_k equal to one is unambiguously rejected. Similar conclusions are obtained for $\mu_k = 10$ and 50. Consequently, the model

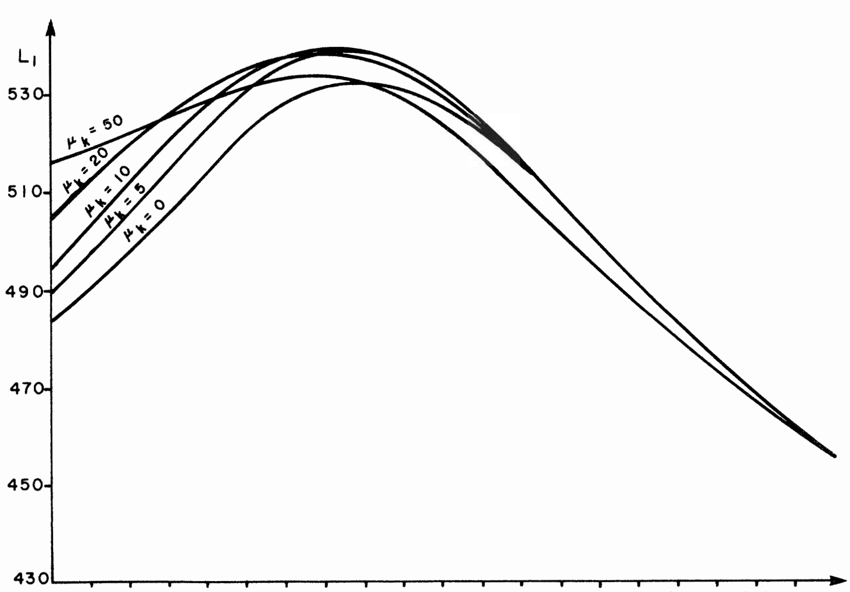

Fig. 1. Mode split equation; model TLCS-1; log likelihood function.

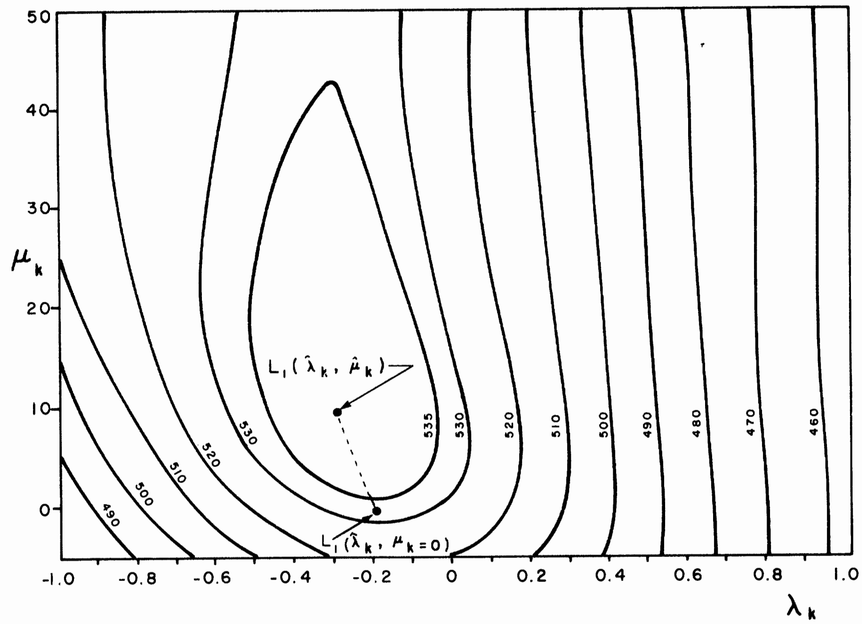

Fig. 2. Mode split equation; model TLCS-1; log likelihood surface.

Table 1. Mode split equation; model TLCS-1; likelihood ratio tests for functional form

Test	Hypothesis	Alternative	Test statistic	Critical Region
$C(\lambda_k)$	H_0: $\lambda_k=\tilde{\lambda}_k$ given $\mu_k=\tilde{\mu}_k$	H_1: $\lambda_k\neq\tilde{\lambda}_k$ given $\mu_k=\tilde{\mu}_k$	$L_1(\hat{\lambda}_k, \mu_k=\tilde{\mu}_k) - L_1(\lambda_k=\tilde{\lambda}_k, \mu_k=\tilde{\mu}_k)$	$C(\lambda_k) > \frac{1}{2}\chi^2_\alpha$ (1)
$C(\mu_k)$	H_0: $\mu_k=\tilde{\mu}_k$ given $\lambda_k=\tilde{\lambda}_k$	H_1: $\mu_k\neq\tilde{\mu}_k$ given $\lambda_k=\tilde{\lambda}_k$	$L_1(\hat{\mu}_k, \lambda_k=\tilde{\lambda}_k) - L_1(\mu_k=\tilde{\mu}_k, \lambda_k=\tilde{\lambda}_k)$	$C(\mu_k) > \frac{1}{2}\chi^2_\alpha$ (1)
$U(\lambda_k,\mu_k)$	H_0: $\lambda_k=\tilde{\lambda}_k$, $\mu_k=\tilde{\mu}_k$	H_1: $\lambda_k\neq\tilde{\lambda}_k$, $\mu_k\neq\tilde{\mu}_k$	$L_1(\hat{\lambda}_k,\hat{\mu}_k) - L_1(\lambda_k=\tilde{\lambda}_k, \mu_k=\tilde{\mu}_k)$	$U(\lambda_k,\mu_k) > \frac{1}{2}\chi^2_\alpha$ (2)

with functional form given by ($\lambda_k = -0.2, \mu_k = 0$) or by ($\lambda_k = -0.3$, $\mu_k = 10$) is clearly superior to either multiplicative form ($\lambda_k = \mu_k = 0$) or the logit form ($\lambda_k = 1, \mu_k = 0$).

Conditional tests for μ_k, given λ_k, are given in Table 3. For $\lambda_k = 0$, the estimated value of 5 for μ_k is significantly different from 0. For $\lambda_k = -0.2$ the null hypothesis of $\mu_k = 0$ is rejected at the 0.1% level. The larger negative that λ_k becomes, the more strongly is the hypothesis of zero μ_k rejected. However, for $\lambda_k = 1$, no distinction can be made between alternative values of μ_k.

The simultaneous unconditional test given in Table 4 provides a more valid test of the functional form. Whereas the conditional test for λ_k accepted $\lambda_k = -0.2$, given $\mu_k = 0$, this is rejected by the unconditional test at the 0.5% level. The only forms which are accepted by this test in Table 4 lie between $\lambda_k = -0.4$ and -0.2, for $\mu_k = 10$. The special case of the multiplicative (log-linear) form with $\lambda_k = \mu_k = 0$ is unequivocally rejected, as is the log model with $\mu_k = 10$. The CLCS-1 model, which corresponds to $\mu_k = 0$. is clearly inferior to the TLCS-1 model.

2.3.1.2 *Model TCLS*-2. Up to this point, the λ_k's have been constrained equal when applied to the independent variables. This constraint is now relaxed in TLCS-2 and CLCS-2. In addition, the λ_k's and μ_k's, previously estimated to the nearest 0.1, are now estimated with 5-figure accuracy and an optimal estimate is calculated with distinct λ_k's and identical μ_k's. The results obtained from these estimation problems are given in Table 5. Column A contains TLCS-2, the most general form estimated, whereas the remaining columns present a sequence of nested special cases. Thus CLCS-2 in column C is a special case of TLCS-2 with the restriction $\mu_k = 0$, and CLCS-1 includes the further restriction that the λ_k's be equal. Finally, in *E*, these λ_k's are set equally to zero. For completeness, the popular standard logit model is given in column F.

The most significant feature of the estimates presented is that the transformation parameters are relatively different once the equality constraint is removed; whereas the fare's λ_1 is similar, the time and frequency λ_k's are quite distinct. However, the likelihood function is presumably not steeply sloping at this point since the

Table 2. Mode split equation; model TLCS-1; conditional test for λ_k

$C(\lambda_k)$	$\mu_k=0$		$\mu_k=10$		$\mu_k=50$	
λ_k	$L_1(\lambda_k,\mu_k)$	Stat.	$L_1(\lambda_k,\mu_k)$	Stat.	$L_1(\lambda_k,\mu_k)$	Stat.
-1.0	484.10	48.86	495.03	44.78	515.95	18.16
-0.9	491.18	41.78	503.36	36.45	518.62	15.49
-0.8	498.86	34.10	511.98	27.83	521.58	12.53
-0.7	506.86	26.10	520.39	19.42	524.83	9.28
-0.6	514.77	18.19	527.92	11.89	528.21	5.90
-0.5	521.97	10.99	533.95	5.86	531.30	2.81
-0.4	527.78	5.18	537.99	1.82	533.48	0.63
-0.3	531.57	1.39	539.81	0	534.11	0
-0.2	532.96	0	539.27	0.54	532.73	1.38
-0.1	531.96	1.00	536.39	3.42	529.25	4.86
0	528.71	4.25	531.38	8.43	523.94	10.17
0.1	523.51	9.45	524.71	15.10	517.32	16.79
0.2	516.82	16.14	516.96	22.85	509.95	24.16
0.3	509.20	23.76	508.67	31.14	502.28	31.83
0.4	501.14	31.82	500.24	39.57	494.66	39.45
0.5	492.97	39.99	491.95	47.86	487.29	46.82
0.6	484.95	48.01	483.98	55.83	480.28	53.83
0.7	477.21	55.75	476.41	63.40	473.69	60.42
0.8	469.84	63.12	469.27	70.54	467.51	66.60
0.9	462.87	70.09	462.58	77.23	461.72	72.39
1.0	456.32	76.64	456.32	83.49	456.32	77.79

Percentage points of $\frac{1}{2}\chi^2$ distribution with 1 degree of freedom

Level	0.1	0.05	0.025	0.01	0.005	0.001	0.0001
Value	1.35	1.92	2.51	3.32	3.94	5.41	7.57

Table 3. Mode split equation; model TLCS-1; conditional test for μ_k

$C(\mu_k)$	$\lambda_k=-1.0$		$\lambda_k=-0.2$		$\lambda_k=0$		$\lambda_k=1.0$	
μ_k	$L_1(\lambda_k,\mu_k)$	Stat.	$L_1(\lambda_k,\mu_k)$	Stat.	$L_1(\lambda_k,\mu_k)$	Stat.	$L_1(\lambda_k,\mu_k)$	Stat.
-1.0	482.41	33.54	530.16	9.29	526.48	5.43	456.32	0
0	484.10	31.85	532.96	6.49	528.71	3.20	456.32	0
1.0	485.38	30.57	535.11	4.34	530.12	1.79	456.32	0
5.0	489.62	26.33	539.03	0.42	531.91	0	456.32	0
8.0	492.81	23.14	539.45	0	531.72	0.19	456.32	0
10.0	495.03	20.92	539.27	0.18	531.38	0.53	456.32	0
12.0	497.32	18.63	538.94	0.51	530.95	0.96	456.32	0
15.0	500.78	15.17	538.33	1.12	530.27	1.64	456.32	0
20.0	506.17	9.78	537.27	2.18	529.14	2.77	456.32	0
50.0	515.95	0	532.73	6.72	523.94	7.97	456.32	0

likelihood ratio test fails to reject TSCS-1 at the 1 per cent level of significance, although it succeeds at the 5 per cent level. In the comparison of the CLCS-2 and CLCS-1 models, the latter is rejected at the 0.1% level. It is also interesting to note that the CLCS-2 model is inferior to the TLCS-1 model. Indeed the significance of $\mu_k \neq 0$ is demonstrated by a comparison of TLCS-2 and CLCS-2 models, the latter being rejected at well over the 0.5% level by the likelihood ratio test.

In summary, the TLCS-2 model is a superior model which dominates the TLCS-1, CLCS-2 and CLCS-1 restricted models. Moreover, all these generalized

Table 4. Mode split equation; model TLCS-1; unconditional test for λ_k and μ_k

$\tilde{\lambda}_k$	$\tilde{\mu}_k$	$L_1(\lambda_k, \mu_k)$	Stat.
-1.0	0	484.10	55.71
-0.5	0	521.97	17.84
-0.4	0	527.78	12.03
-0.3	0	531.57	8.24
-0.2	0	532.96	6.85
-0.1	0	531.96	7.85
0	0	528.71	11.10
0.5	0	492.97	46.84
1.0	0	456.32	83.49
-1.0	10	495.03	44.78
-0.5	10	533.95	5.86
-0.4	10	537.99	1.82
-0.3	10	539.81	0
-0.2	10	539.27	0.54
-0.1	10	536.39	3.42
0	10	531.38	8.43
0.5	10	491.95	47.86
1.0	10	456.32	83.49

Percentage points of $\frac{1}{2}\chi^2$ distribution with 2 degrees of freedom

Level	0.1	0.05	0.025	0.01	0.005	0.001	0.0001
Value	2.30	3.00	3.69	4.61	5.30	6.91	9.21

models dominate the conventional log linear form, which in turn dominates the standard logit model, apparently an inappropriate model, at least for this set of data.

2.3.2 *Effects of the functional form on the distribution of the residuals.* Transformation of the original variables of a model into new variables affects the distribution of the error term. Although the purpose of the transformation is to make this distribution normal, this purpose is not commonly achieved in practice. It is known that the maximum likelihood estimation of the functional form can have the property of regularizing the distribution: in this context, Zarembka (1974) has devised an approximate measure of the way in which the Box and Cox procedure may correct for heteroskedasticity in the error term. Here we shall take note of the fact that a similar phenomenon might be occurring because the estimation of the functional form regularizes the distribution of the residuals in terms of skewness and kurtosis. It is apparent in the graph shown in Fig. 3 that the maximum likelihood estimate of λ_k in model CLCS-1 coincides with minimum values of skewness σ_3 and kurtosis σ_4; for the normal distribution, these statistics are equal to zero.

2.3.3 *Effects of the functional form on the significance of variables.* The choice of functional form may be critical for the significance level of variables in the equation. Indeed, specification error with respect to form may lead to the erroneous inclusion or exclusion of variables and thereby to spurious results. That such an outcome in reality is shown by Fig. 4 which presents results for model CLCS-1 in which the vector of location parameters μ equals zero. The significance of fares is unequivocal even though there is considerable variation in the level of significance. In the case of travel time, however for $\lambda_k = 1$ the t-statistic implies significance whereas for $\lambda_k = -1$ it is insignificant. Thus the improper prior choice of form could lead to travel time being eliminated as an explanatory variable.

Although t-statistics are biased below and above the point of maximum likelihood $\lambda_k = -0.2$, it is of some interest to note that the statistical significance of values of the modal constants is at a minimum around that point. Despite the bias, it appears that the specification error in form might be taken up in these constants intended to capture qualities intrinsic to the modes themselves.

Similarly, allowing some elements of the vector μ, which are equal to zero in Fig. 4, to differ from zero in Fig. 5, damps down the wide fluctuations in the t-statistics. If, again, model TLCS-1 is the true model and model CLCS-1 is misspecified, the t-statistics presented in Fig. 4 are biased and a rigorous comparison of the results is not feasible. Yet the apparent reduction in the statistical significance of the values of the modal

Table 5. Mode split equation; comparison of models

PARAMETERS	MODELS					
	A	B	C	D**	E	F
	TLCS-2	TLCS-1	CLCS-2	CLCS-1	Log linear	Logit
(F) $\lambda_{1,1}$	-0.2399	-0.2660	-0.2626	-0.1930	0.0	1.0
(H) $\lambda_{1,2}$	-1.0982	-0.2660	-0.0513	-0.1930	0.0	1.0
(D) $\lambda_{1,3}$	0.0298	-0.2660	0.5712	-0.1930	0.0	1.0
μ_k	35.757	8.6862	0.0	0.0	0.0	0.0
(air) α_{10}	4.7986	0.7288	21.762	0.9343	1.3910	113.00
	(3.629)	(1.044)	(5.001)	(1.462)	(1.695)	(3.486)
(rail) α_{20}	5.0241	0.5087	21.115	0.0380	1.0136	112.28
	(4.230)	(1.124)	(4.957)	(0.097)	(1.618)	(3.459)
(bus) α_{30}	4.2821	-0.3085	20.354	-0.8106	0.2516	111.76
	(3.761)	(-0.793)	(4.837)	(-2.504)	(0.452)	(3.448)
(car) α_{40}	0.0	0.0	0.0	0.0	0.0	0.0
	(-)	(-)	(-)	(-)	(-)	(-)
(F) α_1	-1.8164	-1.7231	-1.8274	-2.2254	-2.9653	-0.841×10^{-3}
	(-14.13)	(-14.96)	(-17.09)	(-18.70)	(-15.57)	(-11.31)
(H) α_2	-0.0153	-0.3932	-0.8358	-0.3605	-1.3148	-0.0141
	(-5.941)	(-5.105)	(-4.596)	(-4.144)	(-5.576)	(-9.092)
(D) α_3	1.3779	0.4414	13.503	0.1331	0.4221	0.0114
	(5.735)	(5.067)	(5.368)	(5.022)	(4.607)	(3.521)
$L_1(\lambda,\mu)$	543.87	539.95	538.66	532.97	528.71	456.32
$L_1(\tilde{\lambda},\tilde{\mu}) - L_1(\tilde{\lambda},\tilde{\mu})$	0	3.94	5.21	10.90	15.16	87.55
R^2	0.7301	0.7223	0.7197	0.7079	0.6987	0.4909
Skewness	3.155	3.331	3.268	3.466	3.731	6.888
Kurtosis	0.566	0.711	0.624	0.765	0.990	2.667
D.F.	0	2	1	3	4	4

*Numbers in parentheses are t-statistics.
**Results computed with higher accuracy than those of Table 2 for $\mu_k = 0$.

Percentage points of $\frac{1}{2}\chi^2$ distribution with 4 degrees of freedom

Level	0.1	0.05	0.025	0.01	0.005	0.001	0.0001
Value	3.89	4.74	5.57	6.64	7.43	9.23	11.76

constants can suggest that further error specification in form has been eliminated. The vector μ, which is relatively ineffectual with respect to the likelihood function, clarifies in Fig. 5 the significance of mode-specific constants relative to that of fare and time service level variables for all values of λ_k. This impact of μ might then be used to justify its presence on grounds different from those provided by the likelihood ratio tests.

2.3.4 *Effects of the functional form on elasticities.* A very significant comparison between models is that with respect to their elasticities. Models with different elasticities have different properties. Where models are to be used for out-of-sample extrapolation in order to simulate or forecast, the correct estimation of elasticities is of primary importance.

Figures 6 and 7 show market share elasticities for fare and travel time for four modes as a function of λ_k, given vector $\mu = 0$, that is with model CLCS-1. There is a considerable variation in these elasticities which, to some extent matches the pattern found with the t-statistics for fare and time. The generally lower elasticities for car travel are largely due to the existing large market share taken by the car. Bus and rail travel have remarkably similar elasticities and this is all the more remarkable that their interzonal flow data were derived from widely different sources. In terms of the overall pattern, air

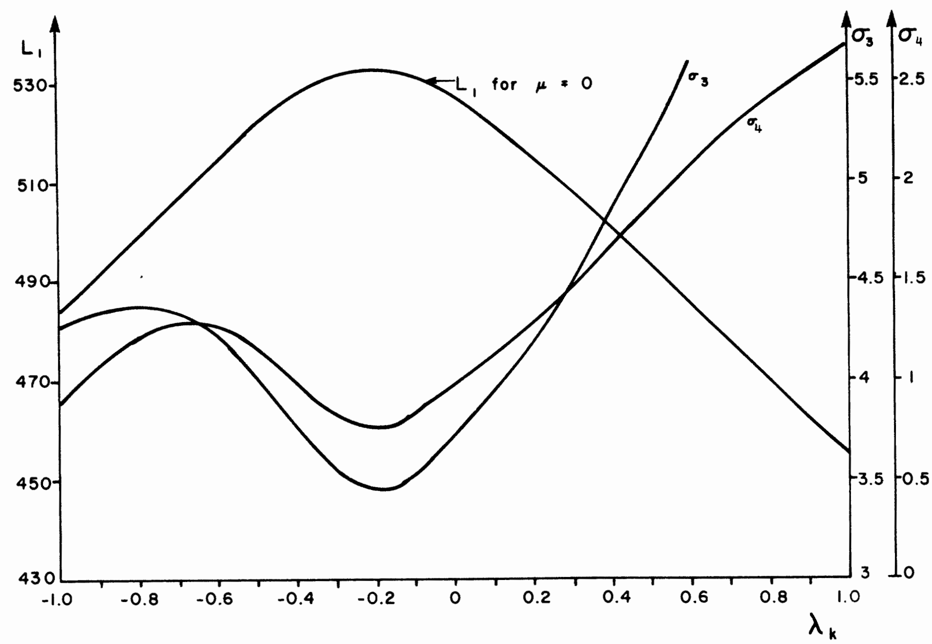

Fig. 3. Mode split equation; model CLCS-1; log likelihood function with associated skewness σ_3 and kurtosis σ_4 of residuals.

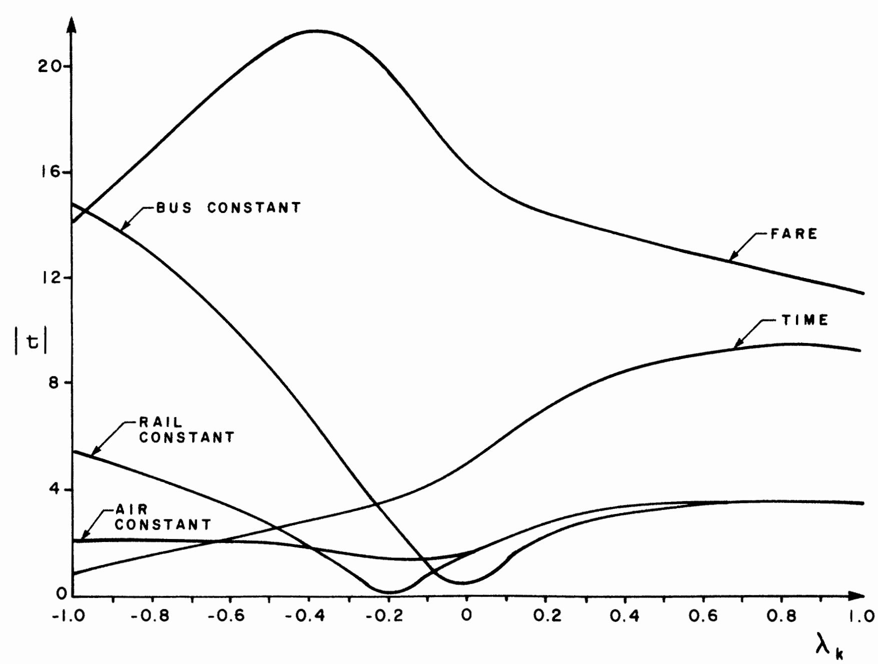

Fig. 4. Mode split equation; model CLCS-1; absolute values of t-statistics.

travel appears to be the odd man out. For fares, air is rising when the other modes are falling whereas, for time, air is falling when the others are rising. Further details on the elasticities computed from various λ_k's and μ_k's are shown in Tables 6–8. Elasticities are computed at the mean value of the observations and, using Baumol's convention, have a sign opposite to the sign of the estimated parameters.

2.4 *Results for the total travel demand equation*

The remainder of this section is devoted to estimation of the class of models given by (43). Again, use is made of the family of monotone power transformations defined by vectors λ_2 and μ_2. A feature of interest stems from the substitution of the denominator of the estimated mode split equation as an argument of the total demand equation according to (38). Thus the likelihood function for the total travel demand equation is a function not only of the parameters of that equation but also of the prior mode split equation. For the purposes of the present analysis, these first stage parameters are not allowed to vary in the total travel function: they are fixed

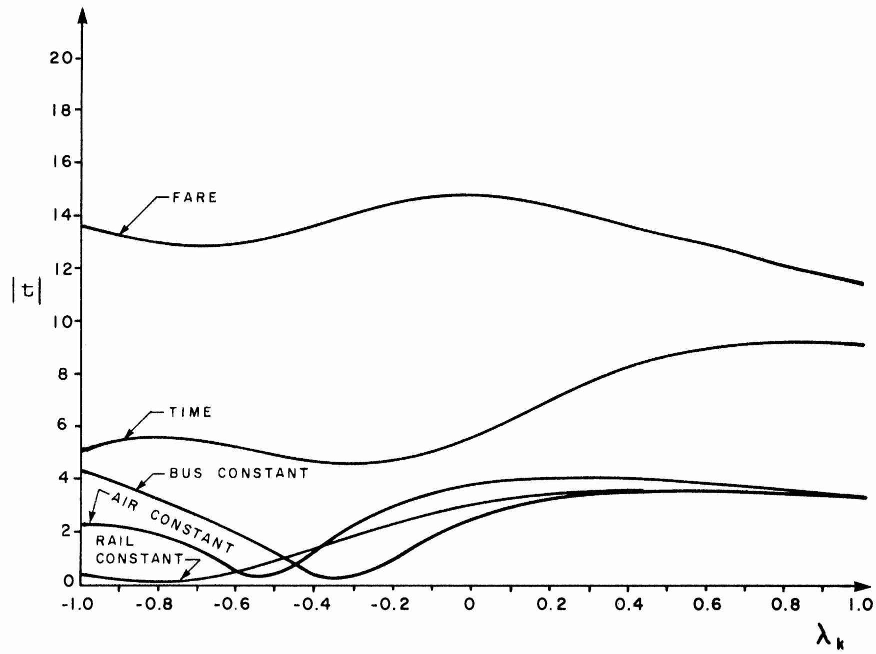

Fig. 5. Mode split equation; models TLCS-1; absolute values of t-statistics for $\mu_{1,k} = 20$.

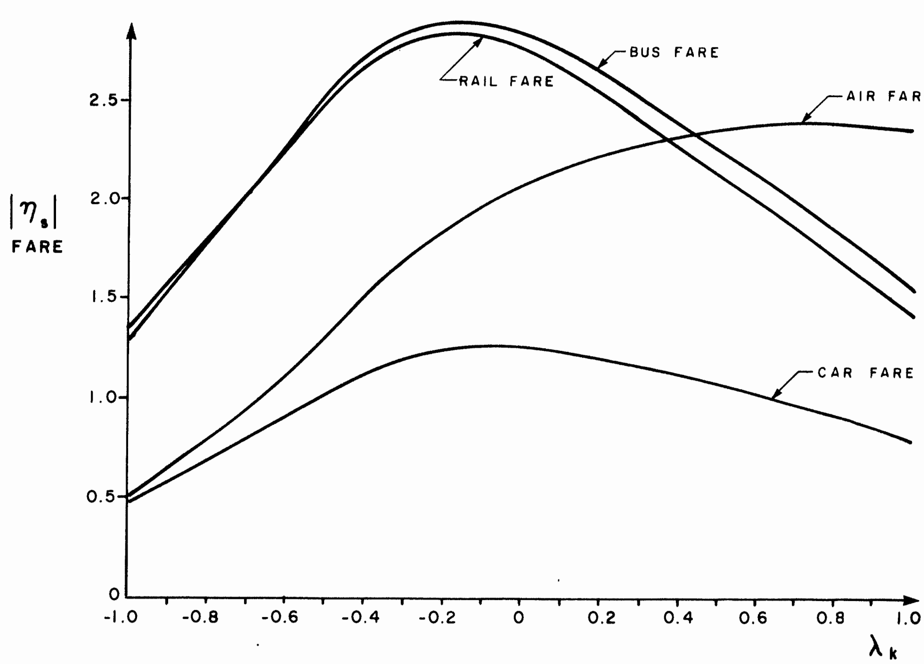

Fig. 6. Mode split equation; model CLCS-1; absolute value of fare elasticities.

at their optimal values from the mode split equation models found in Table 5. This imposed separability results in a considerable simplification of the estimation problem.

2.4.1 *Estimation and tests of the functional form.* The shape of the logarithmic concentrated likelihood function is shown in Fig. 8 for model CECS-2 defined in (45), given optimal values from the TLCS-1 model at the first stage. Clearly, the surface is quite regular with a slight asymmetry reflected in steeper slopes due to high values of $\lambda_{2,y}$ combined with low values of $\lambda_{2,k}$'s. Contours form nearly concentric ellipses rather than circles, the major axis of the ellipses being orientated along the plane where $\lambda_{2,y} = \lambda_{2,k}$. In fact the optimal solution for $\lambda_{2,y} = \lambda_{2,k}$ appears close to the optimal solution for $\lambda_{2,y} \neq \lambda_{2,y}$ leading to the subjective expectation that the additional free parameter is slightly superflous. That this conclusion is premature is shown in Table 9 in which all models use first-stage TLCS-2 values presented in column A of Table 5. A comparison of models B and C

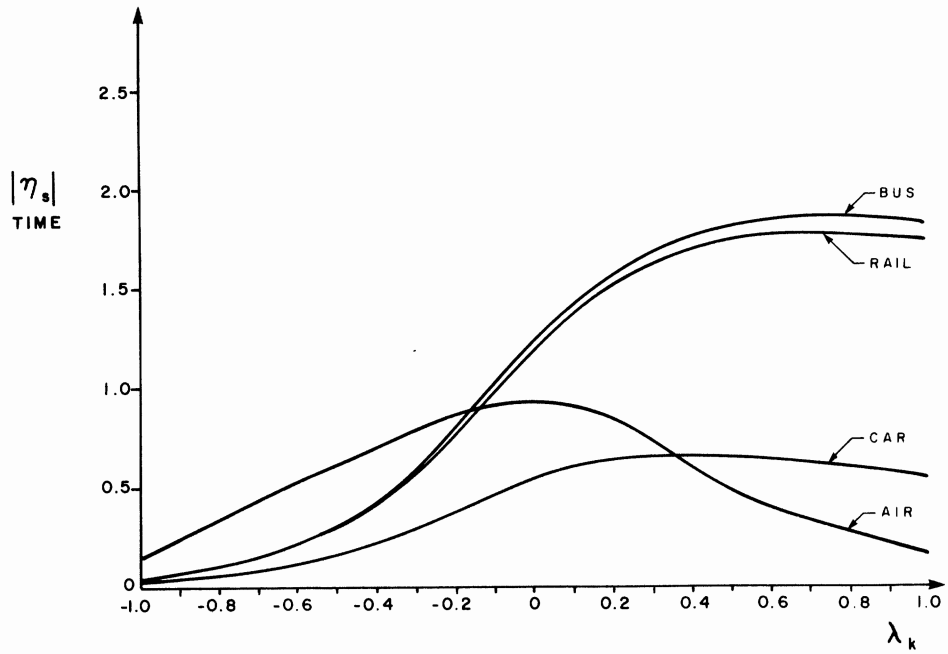

Fig. 7. Mode split equation; model CLCS-1; absolute value of time elasticities.

Table 6. Mode split equation; model CLCS-1; market share elasticities

μ_k=0	CAR			AIR		
λ_k	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
-1.0	0.48	0.02		0.50	0.16	-0.14
-0.5	1.03	0.16		1.31	0.61	-0.24
-0.2	1.25	0.37	N/A	1.86	0.86	-0.28
0	1.26	0.56		2.07	0.92	-0.29
0.2	1.21	0.65		2.22	0.77	-0.31
0.5	1.09	0.66		2.37	0.48	-0.32
1.0	0.82	0.56		2.36	0.17	-0.26

μ_k=0	RAIL			BUS		
λ_k	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
-1.0	1.34	0.03	-0.52	1.32	0.03	-0.24
-0.5	2.48	0.29	-0.53	2.50	0.29	-0.36
-0.2	2.84	0.75	-0.46	2.91	0.77	-0.40
0	2.74	1.21	-0.39	2.83	1.26	-0.40
0.2	2.53	1.52	-0.33	2.64	1.59	-0.40
0.5	2.14	1.74	-0.25	2.27	1.82	-0.38
1.0	1.46	1.75	-0.12	1.59	1.84	-0.28

*By definition, elasticities are computed with the opposite of the sign of the estimated coefficient.

Table 7. Mode split equation; model TLCS-1; market share elasticities

μ_k=10	CAR			AIR		
λ_k	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
-1.0	0.48	0.18		0.49	0.83	-0.32
-0.5	0.88	0.51		1.09	1.08	-0.32
-0.2	1.14	0.53	N/A	1.65	0.76	-0.34
0	1.22	0.56		1.99	0.63	-0.35
0.2	1.21	0.60		2.23	0.53	-0.35
0.5	1.09	0.63		2.39	0.38	-0.33
1.0	0.82	0.56		2.36	0.17	-0.26

μ_k=10	RAIL			BUS		
λ_K	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
-1.0	1.31	0.30	-0.58	1.29	0.30	-0.49
-0.5	2.11	0.94	-0.43	2.12	0.96	-0.46
-0.2	2.55	1.09	-0.37	2.61	1.12	-0.46
0	2.65	1.24	-0.32	2.73	1.28	-0.46
0.2	2.53	1.45	-0.28	2.64	1.50	-0.43
0.5	2.15	1.69	-0.21	2.28	1.77	-0.38
1.0	1.46	1.75	-0.18	1.59	1.84	-0.28

*By definition, elasticities are computed with the opposite of the sign of the estimated coefficient.

in Table 9 shows that this additional free parameter, by which model CECS-2 differs from model CECS-1, has a significant impact on the value of the likelihood function.

Further relaxation of the constraint that all $\lambda_{2,k}$'s be equal specifies model CEUS defined in (46). The results for this model which appear in column A of the same table show that this model should be preferred not only to both models CECS-1 and CECS-2 but also to the standard log linear and linear models.

2.4.2 *Other effects of the functional form.* Having clearly established the importance of functional forms in the analysis of the mode split results and having obtained from tests of the functional form parameters indications that analogous effects on residuals, t-statistics and elasticities could also be obtained from the total demand equation, we will proceed to the time-series model.

3. APPLICATION TO A TIME-SERIES MODEL OF URBAN TRAVEL DEMAND

After the examination of travel demand models estimated from cross-sectional data sets, we now consider the smaller class of models estimated from time-series data. The examples readily found in the literature are direct and mode-specific.

3.1 *Context and theoretical framework*

3.1.1 *Notation and definitions.* We will use the following definitions of travel flow, socio-economic and

Table 8. Mode split equation; model TCLS-1; market share elasticities

$\lambda_k = -.2$	CAR			AIR		
μ_k	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
0	1.25	0.37		1.86	0.86	-0.28
2	1.20	0.47		1.76	0.96	-0.30
5	1.16	0.53	N/A	1.68	0.93	-0.31
10	1.14	0.53		1.65	0.76	-0.34
20	1.16	0.47		1.68	0.51	-0.38
50	1.23	0.36		1.81	0.25	-0.40
90	1.28	0.30		1.93	0.17	-0.39

$\lambda_k = -.2$	RAIL			BUS		
μ_k	FARE	TIME	FREQUENCY	FARE	TIME	FREQUENCY
0	2.84	0.75	-0.46	2.91	0.77	-0.40
2	2.71	0.95	-0.43	2.77	0.98	-0.42
5	2.60	1.08	-0.40	2.66	1.11	-0.44
10	2.55	1.09	-0.37	2.61	1.12	-0.47
20	2.60	0.99	-0.33	2.65	1.02	-0.49
50	2.77	0.81	-0.27	2.82	0.83	-0.49
90	2.90	0.72	-0.23	2.96	0.75	-0.46

*By definition, elasticities are computed with the opposite of the sign of the estimated coefficient.

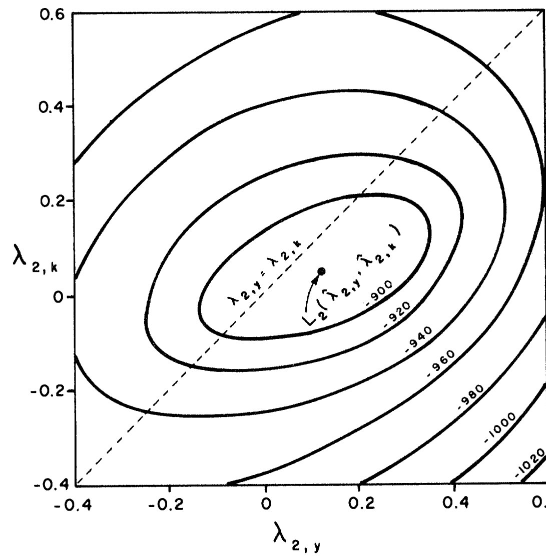

Fig. 8. Total travel equation; model CECS-2 with optimal TLCS-1 values from first stage; likelihood surface.

Table 9. Total travel equation; comparison of models using TLCS-2 first-stage results

PARAMETERS	MODELS				
	A	B	C	D	E
	CEUS	CECS-2	CECS-1	Log linear	Linear
$\lambda_{2,y}$	0.1685	0.1310	0.06	0.0	1.0
$(A_1)\ \lambda_{2,1}$	0.1973	0.0550	0.06	0.0	1.0
$(A_2)\ \lambda_{2,2}$	1.9396	0.0550	0.06	0.0	1.0
$(U)\ \lambda_{2,3}$	0.0355	0.0550	0.06	0.0	1.0
β_0	24.462	10.8875	9.230	6.9768	-21125.
	(29.57)	(7.244)	(12.803)	(15.68)	(-0.4319)
$(A_1)\ \beta_1$	0.8060	1.4124	0.6172	0.4789	13.213
	(27.69)	(23.00)	(21.53)	(19.77)	(7.768)
$(A_2)\ \beta_2$	0.0014	1.6025	0.7358	0.4764	943.81
	(8.360)	(6.266)	(6.061)	(5.638)	(1.610)
$(U)\ \beta_3$	2.5130	1.8088	0.8289	0.3595	11634.
	(36.23)	(31.10)	(29.06)	(27.22)	(3.694)
$L_2(\lambda,\mu)$	-863.53	-877.23	-880.81	-885.05	-1080.8
$L_2(\hat{\lambda},\hat{\mu}) - L_2(\tilde{\lambda},\tilde{\mu})$	0	13.70	17.28	21.52	217.27
R^2	0.9544	0.9360	0.9277	0.9165	0.4516
Skewness	4.954	4.442	2.192	1.066	N/C
Kurtosis	0.065	0.380	0.448	0.011	N/C
D.F.	0	2	3	4	4
Significant figures	5	5	2	2	2

*Numbers in parentheses are t-statistics transportation attribute variables: T_n = travel flow by mode n on a particular origin-destination pair or aggregate of a group of origin-destination pairs. A_r = the rth socio-economic or activity level variable for a particular origin-destination pair or for a group of origin-destination pairs. C_{ms} = the sth "level of service" (fare, waiting time, travel time, etc) variable for mode m for a particular origin-destination pair or for a group of origin-destination pairs.

The part of each of these definitions which applies obviously depends on the level of aggregation considered in each case; i, j subscripts are not necessary because actual studies have typically considered *individual* origin-destination pairs for pairs *aggregated* over a certain area: in that sense time-series models have not to date been explicitly spatial.

3.1.2 *Description and classification of some previous time-series models.* Because time-series models have been direct and mode-specific, a single formulation can summarize them all; their general form has been

$$T_{n_t}^{(\lambda)} = \beta_{n0} + \sum_r \beta_{nr} A_{r_t}^{(\lambda)} + \sum_m \sum_s \beta_{nms} C_{ms_t}^{(\lambda)} + u_{n_t}, \quad m, n = 1, \ldots, M$$

$$u_{n_t} = \sum_{l=1}^{R} \rho_{nl} u_{n_{t-l}} + e_{n_t}, \qquad t = 1, \ldots, N, \tag{52}$$

where the transformation function is defined in (4), the stationary autoregressive process of the error term u_{n_t} is of an order greater than one and constitutes an AR (R) process in the sense of Box and Jenkins (1970) and the error term e_{n_t} has properties stated in (2).

The literature contains several special cases of formulation (52), all estimated for a single mode. At least two authors have used the linear form constraint ($\lambda = 1$) with varying restrictions on the vector of autoregressive parameters $\rho = (\rho_1, \ldots, \rho_l, \ldots, \rho_R)$: Fairhurst and Morris (1975), in an analysis of public transport money receipts in London, used first differences, i.e. $\rho = (\rho_1 = 1,\ \rho_l = 0$ for $l \neq 1)$, which amounts to specifying a CECF model; Gaudry (1973, 1975, 1977) estimated CECM models, also with the constraint $\lambda = 1$.

Other authors have used the logarithmic constraint ($\lambda = 0$). Fisher (1962), in a study of rail travel between Boston and New York, used first differences of the logarithms of the variables, as did aggregate intercity CECF models by the U.S. Civil Aeronautics Board (1966) and Straszheim (1969), both summarized in Meyer and Straszheim (1971); Huang's CECS urban transit model (1973) neglected serial correlation altogether (ρ = 0).

3.1.3 *A family of total demand models, their first difference format and their associated demand elasticities.* All of those models are similar in form to the total demand part of eqn (38). For that reason they belong to the family of models defined in sub-Section 2.1.3.2 above, which need only be modified by reinterpretation of the term U as standing for any of the modal costs instead of a composite cost and by addition of autoregressive errors. In effect, the general family of models with autocorrelated errors is given by

$$(y_t + \mu_y)^{(\lambda_y)} = \beta_0 + \sum_{k=1}^{K} \beta_k (X_{k_t} + \mu_k)^{(\lambda_k)} + u_t, \quad t = 1, \ldots, N \tag{53}$$

with

$$u_t = \sum_{l=1}^{R} \rho_l u_{t-l} + e_t.$$

Special cases are the Box–Cox form (29) or (52), the multiplicative or Cobb–Douglas gravity form (30), the linear (31) and the exponential (32) forms. Further special cases are first difference specifications of these forms obtained by setting $\rho = (\rho_1 = 1,\ \rho_l = 0$ for all $l \neq 1)$.

In general, the elasticity of demand for y with respect to any variable X_k is the analog of (33):

$$\eta = \beta_k (X_k + \mu_k)^{\lambda_y - 1} X_k [(y + \mu_y)^{\lambda_y - 1} y]^{-1}, \tag{54}$$

where

$$y = \left[1 + \lambda_y \left(\beta_0 + \sum_{k=1}^{K} \beta_k (X_k + \mu_k)^{(\lambda_k)}\right)\right]^{\lambda_y^{-1}} - \mu_y.$$

The analogs of expressions (34) and (35), the Box–Cox and multiplicative cases are

$$\eta = \beta_k X_k^{\lambda_k} y^{-\lambda_y} \text{ and } \eta = \beta_k. \tag{55}$$

The analogs of expressions (36) and (37), the linear and exponential cases, are

$$\eta = \beta_k X_k y^{-1} \text{ and } \eta = \beta_k X_k. \tag{56}$$

3.2 *Model selection and estimation procedure*

3.2.1 *Model considered.* Our purpose here is to examine the impact on Gaudry's (1977) model of partially relaxing its current functional form restrictions. These restrictions applied to TEUM formulation (53) are

$$\begin{aligned} \mu &= (\mu_y, \mu_1, \ldots, \mu_k, \ldots, \mu_K) = 0, \\ \lambda &= (\lambda_y, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_K) = 1. \end{aligned} \tag{57}$$

The model consists of adult and schoolchildren demand equations for public transit trips in Montreal; in each equation, the A_r terms refer to income and various activity levels (work, shopping, school, special events) and the C_{ms} terms refer to price, service level and comfort variables of public transport, competing modes and non-transportation goods. Each equation also contains other terms which we will group under OTHER$_q$: the adult and schoolchildren equations contain respectively 28 and 27 variables; the sample consists of 181 consecutive monthly observations for the period December 1956 to December 1971.

3.2.2 *Models estimated.* Each of the market equations of interest contains approximately two dozen explanatory variables. For that reason the estimation of unconstrained Box–Tukey, using for instance the Box and Tidwell (1962) interative least-squares, or of Box–Cox models, is not practical. We shall therefore envisage the following relaxation of the set of constraints (57) for each market equation:

(a) CECM model: We will first maintain $\mu = 0$ but allow λ to differ from 1 whilst restricting all the elements of that vector to be equal:

$$T_t^{(\lambda)} = \beta_0 + \sum_q \beta_q \text{ OTHER}_{qt}^{(\lambda)} + \sum_r \beta_r A_{rt}^{(\lambda)} + \sum_m \sum_s \beta_{ms} C_{ms_t}^{(\lambda)} + u_t \quad (58)$$

and

$$u_t = \sum_{l=1}^{R} \rho_l u_{t-l} + e_t.$$

(b) TLCM model: We allow for the presence of a location parameter on the dependent variable only, thus using $\mu = (\mu_y, \mu_k = 0)$ instead of $\mu = 0$, and permit the vector of identical elements of vector λ to differ from zero:

$$(T_t + \mu_y)^{(\lambda)} = \beta_0 + \sum_q \beta_q \text{ OTHER}_{qt}^{(\lambda)} + \sum_r \beta_r A_{rt}^{(\lambda)} + \sum_m \sum_s \beta_{ms} C_{ms_t}^{(\lambda)} + u_t \quad (59)$$

and

$$u_t = \sum_{l=1}^{R} \rho_l u_{t-l} + e_t.$$

We will give particular attention to the way in which the autoregressive structure is affected by the joint estimation of the functional form parameters.

3.2.3 *Estimation procedure*: 3.2.3.1 *A simplified likelihood function in the presence of multiple autocorrelation.* The likelihood function (5) or its associated log likelihood (7), which we were able to apply in (47) and (48) for the cross-sectional models, is not appropriate for our models (58) and (59): in the latter models, the variance-covariance matrix of the error term u_t is not $\sigma^2 I$ but $\sigma^2\Omega$ where Ω is positive definite. Owing to the computational burden involved in obtaining and using the full maximum likelihood function in the presence of multiple auto-correlation, we shall circumvent such difficulties by deriving a simplified maximum likelihood function which will be derived by neglecting a few of the sample observations. To achieve this, we will write the full likelihood formulation for the first order case in a way which brings out the nature of our intended simplification; we will then simplify the likelihood function for the first order case and extend the procedure to higher order cases. For expository purposes we will consider the following model:

$$y_t^{(\lambda)} = \sum_{k=1}^{K} \beta_k (X_{k_t})^{(\lambda)} + u_t \quad (60)$$

$$u_t = \sum_{l=1}^{R} \rho_l u_{t-l} + e_t, \quad (61)$$

in which it is understood that the constant term, implicitly included among the X_k's, is not transformed and in which the variance of e_t is σ^2.

(A) *The full formulation in the presence of first order of first order autocorrelation.* When the autoregressive process (61) simplifies to

$$u_t = \rho_1 u_{t-1}, \quad |\rho_1| < 1, \quad (61)$$

the error vector e can be expressed, as in Kadiyala (1968), Goldfeld and Quandt (1972) and elsewhere, in the form

$$e = Qu, \quad (63)$$

where the $N \times N$ transformation matrix Q is defined as

$$Q = \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & . & . & . & 0 \\ -\rho_1 & 1 & . & . & . & 0 \\ 0 & . & . & . & . & 0 \\ 0 & 0 & . & . & -\rho_1 & 1 \end{bmatrix}. \quad (64)$$

If $y_t^{(\lambda)}$ and $X_t^{(\lambda)}$ were observed, λ being considered as a constant, the likelihood of the vector e would be

$$\mathscr{L}(e; \beta, \sigma^2, \rho_1) = (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{e'e}{2\sigma^2}\right) \quad (65)$$

and, since $|J|$, the Jacobian of the transformation from e to u, is det $Q = (1-\rho_1^2)^{1/2}$, the likelihood of the vector u would be

$$\mathscr{L}(e; \beta, \sigma^2, \rho_1) = (2\pi\sigma^2)^{-N/2}(1-\rho_1^2)^{1/2} \exp\left(-\frac{u'Q'Qu}{2\sigma^2}\right) \quad (66)$$

or, by the definition of Q and summarizing eqn (60) as $y^{(\lambda)} = X^{(\lambda)}\beta + u$,

$$\mathscr{L}(y^{(\lambda)}; \beta, \sigma^2, \rho_1) = (2\pi\sigma^2)^{-N/2}(1-\rho_1^2)^{1/2} \quad (67)$$

$$\times \exp\left(\frac{-\left[((1-\rho_1^2)^{1/2}(y_1^{(\lambda)} - X_1^{(\lambda)}\beta))^2 + \sum_{t=2}^{N}((y_t^{(\lambda)} - \rho_1 y_{t-1}^{(\lambda)}) - (X_t^{(\lambda)} - \rho_1 X_{t-1}^{(\lambda)})\beta)^2\right]}{2\sigma^2}\right).$$

But, as y is observed, the likelihood of the sample is

$$\mathscr{L}(y; \beta, \sigma^2, \rho_1, \lambda) = (2\pi\sigma^2)^{-N/2}(1-\rho_1^2)^{1/2}. \quad (68)$$

$$\times \exp\left(\frac{-\left[((1-\rho_1^2)^{1/2}(y_1^{(\lambda)} - X_1^{(\lambda)}\beta))^2 + \sum_{t=2}^{N}((y_t^{(\lambda)} - \rho_1 y_{t-1}^{(\lambda)}) - (X_t^{(\lambda)} - \rho_1 X_{t-1}^{(\lambda)})\beta)^2\right]}{2\sigma^2}\right). J,$$

where the Jacobian of the transformation on the dependent variable is

$$J = \det\left[\frac{\partial y_t^{(\lambda)}}{\partial y_{t'}}\right] = \sum_{t=1}^{N} y_t^{\lambda-1}, \tag{69}$$

since the matrix $[\partial y^{(\lambda)}/\partial y]$ is of the form

$$\begin{bmatrix} y_1^{\lambda-1} & & & & \\ & \cdot & & \bigcirc & \\ & & \cdot & & \\ & \neq \bigcirc & & \cdot & \\ & & & & y_N^{\lambda-1} \end{bmatrix} \tag{70}$$

due to the fact that y_t is influenced by $y_{t'}$ only if $t \geq t'$.

The transformation Q can be used to define the transformed variables

$$y^{(\lambda)}\bigstar = Qy^{(\lambda)} \text{ and } X^{(\lambda)}\bigstar = QX^{(\lambda)} \tag{71}$$

and to rewrite the likelihood function as

$$\begin{aligned}\mathscr{L}(y;\beta,\sigma^2,\rho_1,\lambda) &= (2\pi\sigma^2)^{-N/2}(1-\rho_1^2)^{1/2} \\ &\times \exp\left(-\frac{[(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)'(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)]}{2\sigma^2}\right) \\ &\times \prod_{t=1}^{N} y_t^{\lambda-1}.\end{aligned} \tag{72}$$

The log likelihood function is then

$$\begin{aligned}L(y;\beta,\sigma^2,\rho_1,\lambda) &= -\frac{N}{2}\ln(2\pi\sigma^2) + \frac{1}{2}\ln(1-\rho_1^2) \\ &\quad - \left(\frac{[(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)'(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)]}{2\sigma^2}\right) \\ &\quad + (\lambda-1)\sum_{t=1}^{N}\ln y_t.\end{aligned} \tag{73}$$

Maximizing (73) with respect to β and σ^2, given λ and ρ_1, we obtain as in Savin and White (1976), the following estimates of $\beta(\rho_1,\lambda)$ and $\sigma^2(\rho_1,\lambda)$:

$$\hat\beta(\rho_1,\lambda) = (X^{(\lambda)}\bigstar' X^{(\lambda)}\bigstar)^{-1}X^{(\lambda)}\bigstar' y^{(\lambda)}\bigstar, \tag{74}$$

$$\hat\sigma^2(\rho_1,\lambda) = \frac{1}{N}(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\hat\beta(\rho_1,\lambda))'(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\hat\beta(\rho_1,\lambda)). \tag{75}$$

Substituting (74) and (75) into (73), the concentrated log likelihood function is

$$L(y;\rho_1,\lambda) = -\frac{N}{2}\{\ln(2\pi)+1\} - \frac{N}{2}\ln(\hat\sigma^2(\rho_1,\lambda)) + \frac{1}{2}\ln(1-\rho_1^2) + (\lambda-1)\sum_{t=1}^{N}\ln y_t. \tag{76}$$

Savin and White estimated the parameters of likelihood function (73) and performed various likelihood ratio tests for ρ_1 and λ singly and jointly on real and artificial data, in contrast with Zarembka (1968, 1974) who considered briefly the case of first order autoregressive disturbances but without joint hypothesis testing of ρ_1 and λ.

(B) *A simplified formulation in the presence of first order autocorrelation.* Our problem is more complicated than that tackled by Savin and White because we believe that eqn (61) is more appropriate to our problem than eqn (62). Specifically, we are concerned with two demand functions—one for each market under examination—with presumably different autocorrelation structures. This means that the desired equivalent of likelihood function (68) should in principle incorporate terms from the generalized autoregressive schemes. This incorporation would lead to computational difficulties and we will circumvent the problem by dropping as many observations as the highest order of autocorrelation will require.

To see the nature of this simplification, note that neglecting the first observation in the first order case means that the value of the determinant of Q defined in (64) is 1 instead of $(1-\rho_1^2)^{1/2}$, that the term for e_1^2 in eqn (67) is ignored and that the first term in eqn (69) is also disregarded. In consequence the concentrated log likelihood function (76) becomes

$$L(y;\rho_1,\lambda) = -\frac{N-1}{2}\{\ln(2\pi)+1\} - \frac{N-1}{2}\ln(\hat\sigma^2(\rho_1,\lambda)) + (\lambda-1)\sum_{t=2}^{N}\ln y_t \tag{77}$$

in which one should note that the sum in the last term is from the second observation onwards. Of course, ignoring the first observation amounts to assuming that the autoregressive scheme just started instead of assuming that it always existed and makes no difference asymptotically (Goldfeld and Quandt, 1972, p. 185) However, Beach and MacKinnon (1978) have pointed out that ignoring the first observation is less desirable on theoretical grounds than including it because it removes from the likelihood function the built-in constraint that $|\rho_1| < 1$; they also showed in a linear case ($\lambda = 1$) from a Monte Carlo analysis with two explanatory variables that, for some values of ρ_1, some significant differences still persisted with 50 sample points. Because our sample consists of 181 observations, we have good reasons to expect that the standard procedure which neglects observations will be appropriately efficient.

(C) *The simplified formulation with multiple autocorrelation.* This simplification allows us to extend eqn (72) to multiple autocorrelation cases in which, under usual stationarity conditions, the parameter vector is $\rho = (\rho_1, \ldots, \rho_l, \ldots, \rho_R)$:

$$\begin{aligned}\mathscr{L}(y;\beta,\sigma^2,\rho,\lambda) &= (2\pi\sigma^2)^{-((N-R)/2)} \\ &\times \exp\left[-\frac{(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)'(y^{(\lambda)}\bigstar - X^{(\lambda)}\bigstar\beta)}{2\sigma^2}\right]\prod_{t=1+R}^{N} y_t^{\lambda-1},\end{aligned} \tag{78}$$

where, now for $t = 1 + R, \ldots, N$, we have defined

$$y_t^{(\lambda)}\bigstar = y_t^{(\lambda)} - \sum_{l=1}^{R} \rho_l y_{t-l}^{(\lambda)},$$

and

$$X_t^{(\lambda)}\bigstar = X_t^{(\lambda)} - \sum_{l=1}^{R} \rho_l X_{t-l}^{(\lambda)} \tag{79}$$

and where the Jacobian of the transformation of the dependent variable comes from a matrix of the same form as (70) in which the first R rows have been ignored. The concentrated log likelihood function which corresponds to (76) is

$$L(y; \rho, \lambda) = -\frac{N-R}{2}\{\ln(2\pi)+1\} - \frac{N-R}{2}\ln(\hat{\sigma}^2(\rho, \lambda)) + (\lambda - 1)\sum_{t=1+R}^{N} \ln y_t. \tag{80}$$

Equation (80) is the appropriate simplified log concentrated likelihood for model CECM given in (58). For model TLCM defined in (59), the appropriate expression is

$$L(y; \rho, \lambda, \mu_y) = -\frac{N-R}{2}\{\ln(2\pi)+1\} - \frac{N-R}{2}\ln(\hat{\sigma}^2(\rho, \lambda, \mu_y)) + (\lambda - 1)\sum_{t=1+R}^{N} \ln(y_t + \mu_y). \tag{81}$$

3.2.3.2 *Maximization technique.* Likelihood functions (80) or (81) will be maximized by performing a grid search. Note that the equations which describe the general model (53) may be combined to yield

$$(y_t + \mu_y)^{(\lambda_y)} - \sum_{l=1}^{R} \rho_l (y_{t-l} + \mu_y)^{(\lambda_y)} = \sum_{k=1}^{K} \beta_k (X_{k_t} + \mu_k)^{(\lambda_k)} - \sum_{k=1}^{K}\sum_{l=1}^{R} \beta_k \rho_l (X_{k_{t-l}} + \mu_k)^{(\lambda_k)} + e_t, \tag{82}$$

where $t = 1 + R, \ldots, N$ and the constant term has been dropped for expository purposes.

For given values of transformation vectors λ and μ, eqn (82) is a non-linear equation in the coefficients β_k and ρ_l. To estimate the parameters of that equation, we identify the autoregressive scheme by examining, in the spirit of Box and Jenkins (1970), the autocorrelation and partial autocorrelation functions of the unbiased estimates of the residuals $\hat{u}_t$ obtained from applying least-squares to the first equation contained in (53) for $\lambda = 1$ and $\mu = 0$. This provides an initial specification of the autoregressive scheme: it is this specification which will be tested jointly with the functional form parameters. The particular non-linear equation associated with the autoregressive scheme specified in this manner is then solved by the maximum likelihood Cochrane–Orcutt (1949) iterative procedure using 15 iterations to insure a high degree of accuracy. This technique converges to at least a local minimum (Sargan, 1964; Fair, 1971) and no case of multiple minima has been reported (Johnston, 1972). Values of the likelihood functions (80) and (81) are obtained for different values of the transformation vectors λ and μ without changing the specification of the autoregressive process. This is not because the autoregressive structure is independent of the functional form but because it would make no sense to change the specification of the process continually if we are primarily interested in how the estimates of a given parameter structure are influenced by λ and μ. A different process is specified for each of the market equations we consider. In both cases the highest order of autocorrelation considered is the 12th: eqn (82) is therefore estimated with $t = 13, \ldots, 181$ in each transit demand function.

3.3 *Results for the adult market equation*

3.3.1 *Estimation and tests of the functional form.* We will report the results of the CECM and TLCM models at the same time and use conditional and unconditional tests of significance as we did for the cross-sectional models.

The likelihood function is plotted for various values of λ and μ in Fig. 9. Although generally steeper for $\lambda < 1$ than for $\lambda > 1$, it gets progressively flatter as μ_y increases from $-9{,}477{,}108$, the limit value if the nonnegativity constraint defined in (3) is to be respected, to 0.

The maximum values of the function corresponding to different values of μ_y are very close to one another as is more clearly shown on a portion of the likelihood surface near the optimum in Fig. 10 where the dotted ridge slowly rises to a maximum at $\hat{\lambda} = 0.95$ and $\mu_y = -9{,}477{,}108$. Although one could climb further if the nonnegativity constraint were not binding at that point, the rate of change of the likelihood function with respect to μ_y suggests that an unconstrained maximum would not be very different.

The first three tests listed in Table 10 are used to test for the significance of the functional form parameter estimates. In Table 11, it is shown that, irrespective of the conditional value of μ_y, one is justified to reject the hypothesis that λ is smaller that 0.85 or larger than 1.1 at the 99% level of significance (the percentage points of the χ^2 distribution with one degree of freedom are supplied in Table 2). Consequently, a significantly non-linear form cannot be accepted.

The conditional test for μ_y, given λ, presented in Table 12 shows that μ_y is never significantly different from zero, whatever value λ might have.

This is confirmed by the unconditional test provided in Table 13 which shows that the linear form cannot be rejected with μ_y equal to zero or to its minimum value $-9{,}477{,}108$. There is no way of choosing between the linear CECM model and the linear TLCM model: Ockham's razor would make one select the linear CECM model, which was the original specification of Gaudry's model.

3.3.2 *Effects of the functional form on the distribution of the residuals.* The last two tests shown in Table 10 are used to study the significance of the autoregressive

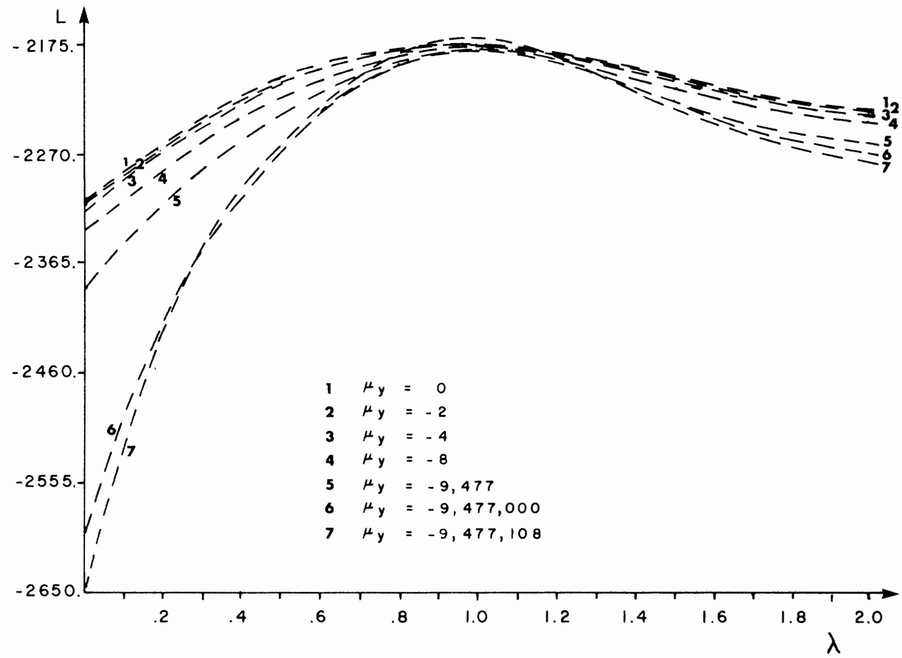

Fig. 9. Adult market equation; log likelihood function.

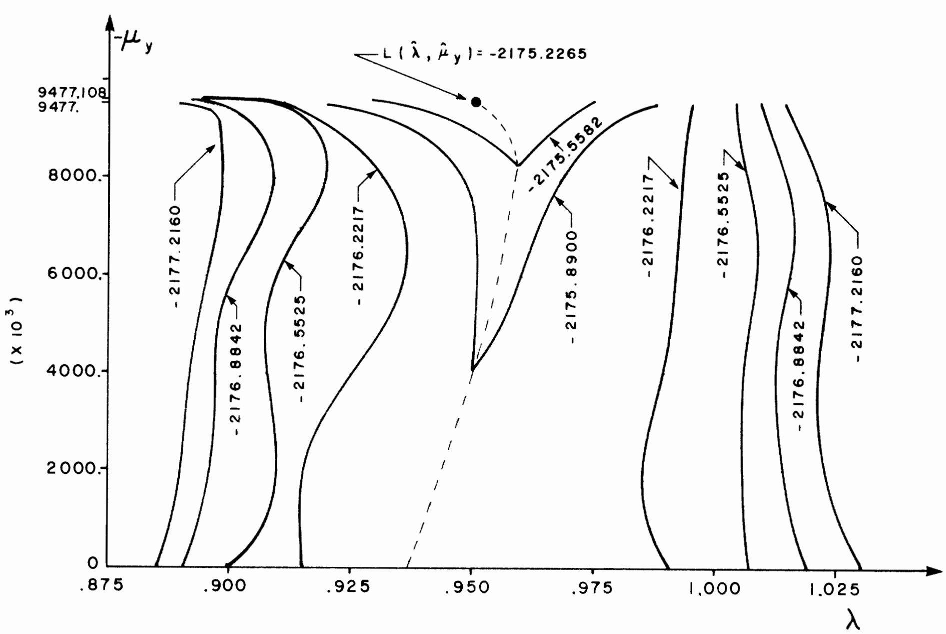

Fig. 10. Adult market equation; log likelihood surface near optimum λ and μ_y.

structure. In the first unconditional test of ρ, the difference between the value of the likelihood at the optimum and its value for $(\hat{\lambda}, \hat{\mu}_y, \rho = 0)$ is 18.5089: with the two degrees of freedom required by the non-zero elements of ρ, the null hypothesis that $\rho = 0$ is rejected with a high degree of confidence. In the joint unconditional test for (λ, μ_y, ρ), the appropriate difference in log likelihood values between the value at the optimum and the value for $(\lambda = 1, \mu_y = 0, \rho = 0)$ is 31.2462: with four degrees of freedom, the null hypothesis is also confidently rejected.

These tests indicate that the orders of the autoregressive scheme postulated for the $\lambda = 1$, $\mu_y = 0$, case are not significantly different at the optimum point $(\hat{\lambda}, \hat{\mu}_y)$. This is confirmed by an analysis of the autocorrelation and partial autocorrelation functions of least-squares residuals presented in Table 14: as one passes from a linear model in A to the optimal functional form in B, there is no significant change in structure; taking serial correlation into accounts yields, in C, approximately "white noise" residuals. These results are not surprising because the values of λ and μ_y which maximize the likelihood are. not significantly different from those of the original linear model.

Table 10. Adult market equation; likelihood ratio tests for functional form and autoregressive structure

Test	Hypothesis	Alternative	Test statistic	Critical region
$C(\lambda)$	H_0: $\lambda = \tilde{\lambda}$ given $\mu_y=\tilde{\mu}_y$	H_1: $\lambda \neq \tilde{\lambda}$ given $\mu_y=\tilde{\mu}_y$	$L(\hat{\lambda},\mu_y=\tilde{\mu}_y,\hat{\rho}) - L(\lambda=\tilde{\lambda},\mu_y=\tilde{\mu}_y,\hat{\rho})$	$C(\lambda) > \frac{1}{2}\chi^2_\alpha$ (1)
$C(\mu_y)$	H_0: $\mu_y = \tilde{\mu}_y$ given $\lambda=\tilde{\lambda}$	H_1: $\mu_y \neq \tilde{\mu}_y$ given $\lambda=\tilde{\lambda}$	$L(\hat{\mu}_y,\lambda=\tilde{\lambda},\hat{\rho}) - L(\mu_y=\tilde{\mu}_y,\lambda=\tilde{\lambda},\hat{\rho})$	$C(\mu_y) > \frac{1}{2}\chi^2_\alpha$ (1)
$U(\lambda,\mu_y)$	H_0: $\lambda=\tilde{\lambda},\mu_y=\tilde{\mu}_y$	H_1: $\lambda\neq\tilde{\lambda}$ $\mu_y\neq\tilde{\mu}_y$	$L(\hat{\lambda},\hat{\mu}_y,\hat{\rho}) - L(\lambda=\tilde{\lambda},\mu_y=\tilde{\mu}_y,\hat{\rho})$	$U(\lambda,\mu_y) > \frac{1}{2}\chi^2_\alpha$ (2)
$U(\rho)$	H_0: $\rho = \tilde{\rho}$	H_1: $\rho \neq \tilde{\rho}$	$L(\hat{\lambda},\hat{\mu}_y,\hat{\rho}) - L(\hat{\lambda},\hat{\mu}_y,\rho=\tilde{\rho})$	$U(\rho) > \frac{1}{2}\chi^2_\alpha$ (NR)*
$U(\lambda,\mu_y,\rho)$	H_0: $\lambda=\tilde{\lambda},\mu_y=\tilde{\mu}_y,\rho=\tilde{\rho}$	H_1: $\lambda\neq\tilde{\lambda},\mu_y\neq\tilde{\mu}_y,\rho\neq\tilde{\rho}$	$L(\hat{\lambda},\hat{\mu}_y,\hat{\rho}) - L(\lambda=\tilde{\lambda},\mu_y=\tilde{\mu}_y,\rho=\tilde{\rho})$	$U(\lambda,\mu_y,\rho) > \frac{1}{2}\chi^2_\alpha$ (2+NR)*

*NR = number of non-zero autoregressive parameters in vector ρ.

Table 11. Adult market equation; conditional test for λ

$C(\lambda)$	$\mu_y=0$		$\mu_y=-4{,}000{,}000$		$\mu_y=-8{,}000{,}000$		$\mu_y=-9{,}477{,}108$	
λ	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.
0	-2329.8700	153.9713	-2340.1060	164.2160	-2380.3711	204.5107	-2646.9820	471.7555
.2	-2276.7477	100.8490	-2284.9908	109.1008	-2314.8804	139.0200	-2429.2153	253.9888
.4	-2236.6176	60.7189	-2242.9842	67.0942	-2262.8926	87.0322	-2306.4734	131.2469
.6	-2204.9141	29.0154	-2209.2264	33.3364	-2221.5790	45.7186	-2237.0892	61.8627
.8	-2182.0037	6.1050	-2183.2269	7.3369	-2187.0110	11.1506	-2190.2050	14.9785
.9	-2176.5625	0.6638	-2176.7008	0.8108	-2177.1248	1.2644	-2176.2023	0.9758
$\hat{\lambda}$	-2175.8987	0	-2175.8900	0	-2175.8604	0	-2175.2265	0
1.0	-2176.3832	0.4845	-2176.3832	0.4932	-2176.3832	0.5228	-2176.3832	0.5228
1.1	-2179.8889	3.9902	-2180.2088	4.3188	-2181.0761	5.2157	-2183.3801	8.1536
1.2	-2185.5454	9.6467	-2186.4145	10.5245	-2188.7246	12.8642	-2193.7894	18.5629
1.25	-2188.8200	12.9213	-2189.9959	14.1059	-2193.0736	17.2132	-2199.4788	24.2523
1.5	-2206.3763	30.4776	-2208.9592	33.0692	-2215.2559	39.3955	-2227.3103	52.0838
2.0	-2234.4793	58.5806	-2237.8936	62.0036	-2246.6877	70.8273	-2267.2660	92.0395

Table 12. Adult market equation; conditional test for μ_y

$C(\mu_y)$	$\lambda=0$		$\lambda=.95$		$\lambda=1.0$		$\lambda=2.0$	
μ_y	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.	$L(\lambda,\mu_y,\hat{\rho})$	Stat.
0	-2329.8700	0	-2175.8987	0.6722	-2176.3832	0	-2234.4793	0
-2,000,000	-2333.7531	3.8831	-2175.8942	0.6677	-2176.3832	0	-2235.8371	1.3578
-4,000,000	-2340.1060	10.2360	-2175.8900	0.6635	-2176.3832	0	-2237.8936	3.4143
-6,000,000	-2351.8862	22.0162	-2175.8870	0.6605	-2176.3832	0	-2241.1332	6.6539
-8,000,000	-2380.3711	50.5011	-2175.8832	0.6567	-2176.3832	0	-2246.6877	12.2084
-9,477,000	-2601.0233	271.1533	-2175.4072	0.1807	-2176.3832	0	-2263.6533	29.1740
-9,477,108	-2646.9820	317.1120	-2175.2265	0	-2176.3832	0	-2267.2660	32.7867

Table 13. Adult market equation; unconditional test for λ and μ_y

λ	μ_y	$L(\lambda,\mu_y)$	Stat.
0	0	-2329.8700	154.6435
.2	0	-2276.7477	101.5212
.4	0	-2236.6176	61.3911
.6	0	-2204.9141	29.6876
.8	0	-2182.0037	6.7772
.95	0	-2175.8987	0.6722
1.0	0	-2176.3832	1.05
1.5	0	-2206.3763	31.1498
2.0	0	-2234.4793	59.2528
0	-9,477,108	-2646.9820	471.7555
.2	-9,477,108	-2429.2153	253.9888
.4	-9,477,108	-2306.4734	131.2469
.6	-9,477,108	-2237.0892	61.8627
.8	-9,477,108	-2190.2050	14.9785
.95	-9,477,108	-2175.2265	0
1.0	-9,477,108	-2176.3832	1.05
1.5	-9,477,108	-2227.3103	52.0838
2.0	-9,477,108	-2267.2660	92.0395

In our cross-sectional analysis, we noted how the transformation parameters reduced the skewness and kurtosis of the residuals. Although we are dealing here with time-series data in which heteroskedasticity is less likely to occur than in cross-sectional data, it is interesting to note in passing the extent to which the transformation is robust to heteroskedasticity using Zarembka's (1974) approximate procedure. First, let the parameter h present the relationship of the square root of the variance of the dependent variable y_t to its expectation: $[\mathrm{var}(y_t)]^{1/2} = \sqrt{c}[E(y_t)]^h$, where c is some constant. Then, the procedure consists of finding that value of the transformation λ, conditional upon a belief about the degree of heteroskedasticity expressed by h, such that the slope of the likelihood function (80) or (81) at, say, $\hat{\lambda}(h)$ is given by

$$(1-\lambda-h)N\frac{\mathrm{var}(\ln y)}{\frac{1}{2}+(1-\lambda-h)^2\,\mathrm{var}(\ln y)}. \tag{83}$$

If we make the relatively strong assumption that $h = 1$, the new estimate of λ computed from (81) for the TLCM model is approximately 1.3 instead of 0.95 ± 0.10 (with a 95% confidence interval) which implies that our original estimates are sensitive to heteroskedasticity and, as in our cross-sectional study, may reflect its estimation.

Table 15 provides a measure of the coefficient of multiple correlation R^2. The R^2 in terms of doubly transformed variables used here, DT, is not the same as the R^2 in terms of the original variables, OV, which could have been used to provide an intuitive measure of the quality of the adjustment obtained from the model. They

Table 14. Adult market equation; autocorrelation and partial autocorrelation functions of residuals

Lags													Estimated standard error for row
	A. $\lambda = 1$, $\mu_y = 0$, $\rho = 0$; 181 observations; least-squares residuals												
	AUTOCORRELATIONS												
1-12	.30	.07	.10	.05	.04	.03	.03	-.03	.00	-.07	.06	.36	(.07)
13-24	.01	-.19	-.09	-.06	-.06	-.08	-.06	-.16	-.14	-.23	-.09	.10	(.08)
25-36	-.06	-.31	-.14	-.07	-.07	-.05	-.06	-.12	-.09	-.07	.05	.24	(.10)
37-48	.17	-.12	.09	.04	-.04	-.07	-.05	-.01	-.04	.06	.15	.34	(.11)
49-60	.22	-.04	.04	.02	-.03	-.04	-.06	-.04	-.14	-.04	-.04	-.13	(.12)
	PARTIAL AUTOCORRELATIONS												
1-12	.30	-.03	.10	-.01	.03	.00	.02	-.06	.03	-.09	.13	.36	(.07)
	B. $\hat{\lambda} = .95$, $\hat{\mu}_y = 9,477,108$, $\rho = 0$; 181 observations; least-squares residuals												
	AUTOCORRELATIONS												
1-12	.29	.07	.10	.05	.05	.03	.03	-.03	-.01	-.07	.07	.37	(.07)
13-24	.02	-.18	-.09	-.07	-.04	-.10	-.07	-.16	-.15	-.22	-.09	.10	(.08)
25-36	-.06	-.31	-.13	-.09	-.07	-.06	-.06	-.12	-.09	-.06	.06	.23	(.10)
37-48	.17	-.12	.09	.03	-.03	-.06	-.04	-.01	-.04	.07	.16	.33	(.11)
49-60	.22	-.03	.04	.02	-.02	-.04	-.06	-.06	-.14	-.02	-.04	.12	(.12)
	PARTIAL AUTOCORRELATIONS												
1-12	.29	-.02	.09	-.00	.04	.00	.01	-.05	.01	-.08	.13	.35	(.07)
	C. $\hat{\lambda} = .95$, $\hat{\mu}_y = -9,477,108$, $\hat{\rho} = \hat{\rho}_1, \hat{\rho}_{12}$; 169 observations, 15th iteration Cochrane-Orcutt residuals												
	AUTOCORRELATIONS												
1-12	.15	.05	.12	.12	-.01	-.00	.02	.05	.08	.02	.06	-.07	(.08)
13-24	-.14	-.03	-.10	-.01	-.03	-.12	.02	-.02	-.07	-.14	.02	-.11	(.08)
25-36	-.08	-.19	-.02	.02	-.01	.01	-.01	-.10	-.02	-.10	-.12	-.00	(.09)
37-48	.07	-.10	.08	.01	.01	-.06	-.04	.06	.03	.09	.14	.10	(.09)
49-60	.01	.05	.02	-.06	-.04	-.02	-.02	.05	.09	.07	-.02	.10	(.10)
	PARTIAL AUTOCORRELATIONS												
1-12	.15	.03	.11	.09	-.05	-.01	-.00	.04	.08	-.00	.04	-.11	(.08)

Table 15. Adult market equation; R^2

λ	R^2 in terms of $y^{(\lambda)}\star$	
	$\mu_y = 0$	$\mu_y = \hat{\mu}_y$
0	.8610043	.7149096
.2	.9214585	.8699967
.4	.9429024	.9363636
.6	.9583589	.9526755
.8	.9730985	.9709826
.9	.9761472	.9786077
$\hat{\lambda}$	.9759198	.9772867
1.0	.9750439	.9750439
1.1	.9721290	.9698936
1.25	.9659502	.9606491
1.5	.9524869	.9417343
1.75	.9341717	.9167930
2.0	.9198391	.9117633

differ as follows:

$$DT = 1 - \frac{\sum (y_t^{(\lambda)}\star - \hat{y}_t^{(\lambda)}\star)^2}{\sum (y_t^{(\lambda)}\star - \bar{y}_t^{(\lambda)}\star)^2}, \qquad OV = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y}_t)^2}, \tag{84}$$

where all sums are from $1+R$ to N and where the double transformation $(\lambda)\star$ is the generalization to multiple orders of autocorrelation of (71) and, in DT, y_t stands for y_t, or for $(y_t + \mu_y)$, according to the case.

3.3.3 *Effects of the functional form on the significance of the variables.* Although we have not been able to reject Gaudry's current model ($\lambda = 1$, $\mu_y = 0$), it is still interesting to note the impact which an improperly specified functional form would have had on the significance of the variables. As we noted in Section 2.3.3. above, t-statistics are biased at points other than those associated with optimal functional form parameter values so that our comments cannot be rigorous. However, they may still be indicative of the difficulties which can arise in interpreting results obtained from a model in which the functional form has been improperly specified. Because of the large number of variables in the model, we shall illustrate this impact with four explanatory variables (fare, waiting time, travel time, income) and with the autoregressive parameters. Figures 11 and 12 successively show the t-statistics for these variables as they vary with λ for $\mu_y = 0$ and $\mu_y = \hat{\mu}_y$. In both cases the statistical significance of the variable is in general the greatest around the estimated optimal or "true" linear model. Moreover, at $\lambda = 1$ and around that point, the service variables have appropriately negative signs, the income variable has a positive sign and, as one expects with monthly time-series data, both first and twelfth order autocorrelation coefficients are positive. Had the logarithmic form ($\lambda = 0$) been used, some of these variables would have changed signs, the worst results being obtained for $\mu_y = \hat{\mu}_y$ where 4 signs change as compared to 1 with $\mu_y = 0$: although it is interesting to note that service level variable waiting-time has a positive sign, the most interesting of these sign changes is that of the income variable which implies that public transit is an inferior good with a logarithmic form and a superior good with a linear form for $\mu_y = \hat{\mu}_y$. The results calculated at $\hat{\lambda}$ or around that point appear to be more

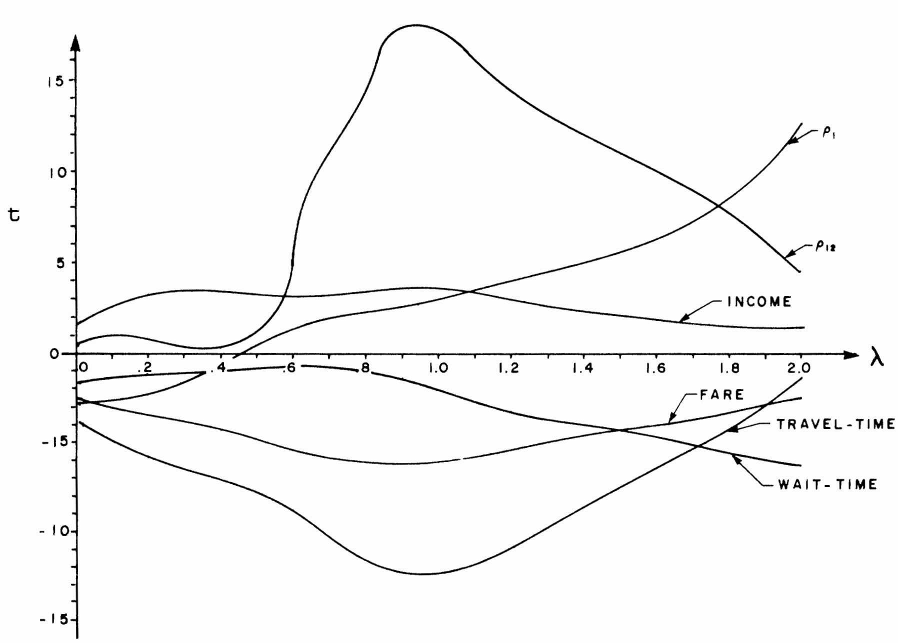

Fig. 11. Adult market equation; t-statistics for selected variables and the autoregressive parameters, $\mu_y = 0$.

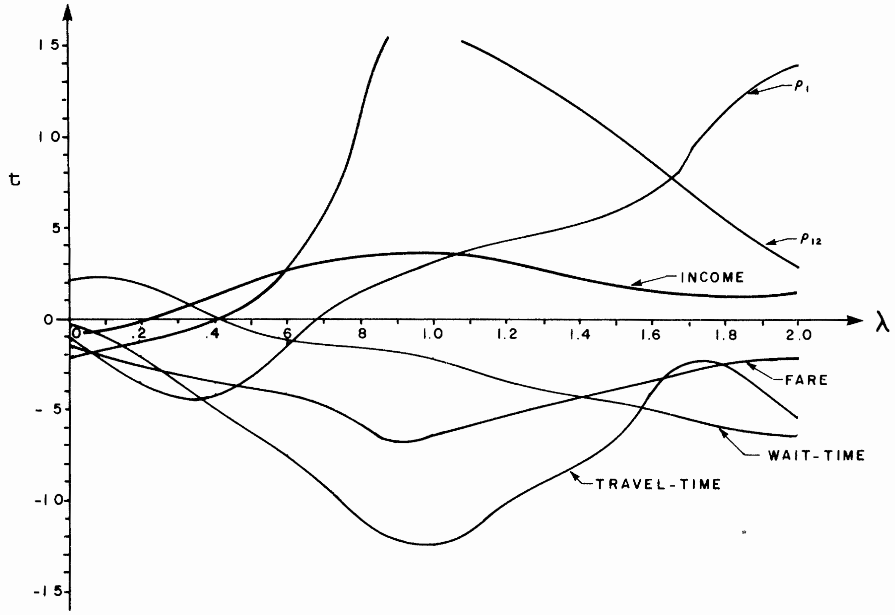

Fig. 12. Adult market equation; t-statistics for selected variables and the autoregressive parameters, $\mu_y = \hat{\mu}_y$.

satisfactory and to have more convincing signs than elsewhere.

3.3.4 *Effects of the functional form on elasticities and autoregressive coefficients.* Similar results occur when the demand elasticities shown in Figs. 13 and 14 are examined. They are computed at the mean value of the observations using Baumol's (1965) convention which reverses the estimated sign of the variable. Although demand elasticities are reasonable independently of λ when $\mu_y = 0$, some of those which have a reasonable sign in both figures, such as the fare and income, take unreasonable values when $\mu_y = \hat{\mu}_y$: with the logarithmic form, the income elasticity of demand is 1.6 and the price elasticity of demand is equal to 1. Nothing of interest can be said about the numerical values of the autoregressive parameters except that the value of ρ_{12} peaks around $\lambda = 1$ in both figures: we are of course quite far from a model of first differences and it would have been improper to formulate it in those terms.

Although the impact of making incorrect assumptions about the functional form is not as considerable on demand elasticities as it appears to be on the sign and level of the t-statistics, the results in both cases in general support the optimal linear model. Further details on the "circumstantial evidence" supplied by t-statistics, demand elasticities and autoregressive coefficients are

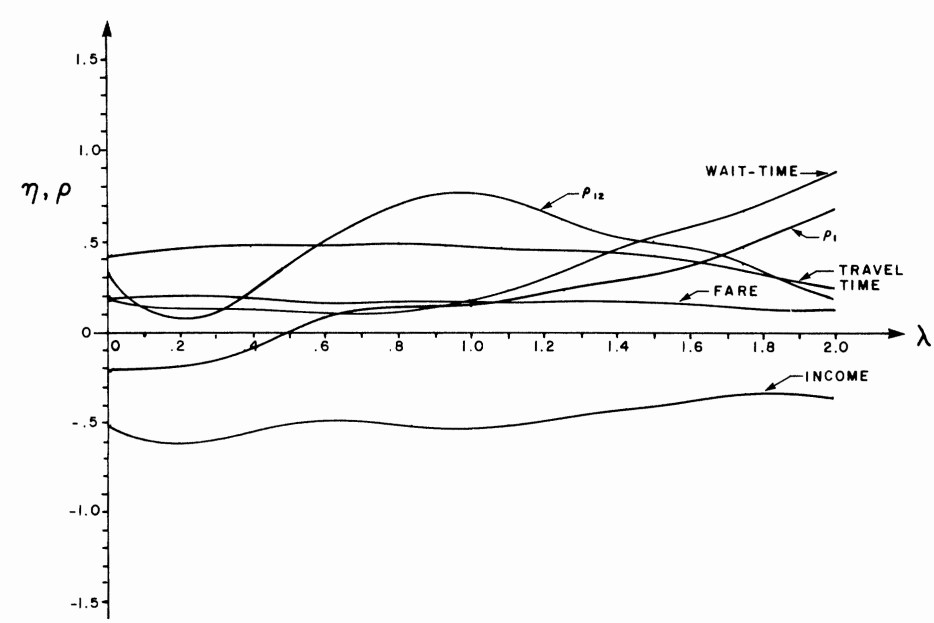

Fig. 13. Adult market equation; selected elasticities and values of autoregressive coefficients, $\mu_y = 0$.

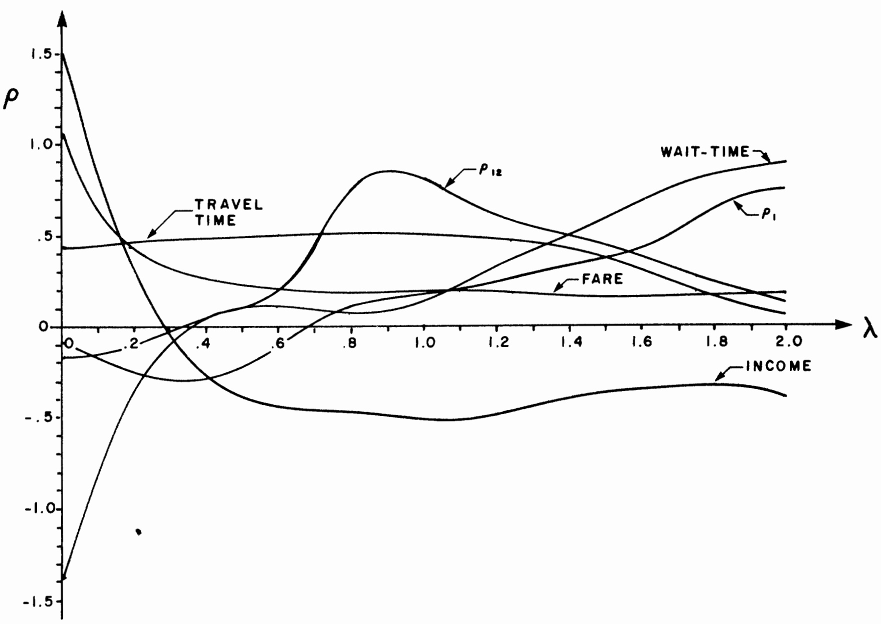

Fig. 14. Adult market equation; selected elasticities and values of autoregressive coefficients, $\mu_y = \hat{\mu}_y$.

Table 16. Adult market equation; t-statistics, elasticities and autoregressive coefficients

	Fare		Wait-time		Travel-time		Income		ρ_1		ρ_{12}	
λ	η	t	η	t	η	t	η	t	value	t	value	t
					$\mu_y = 0$							
0	.189	-2.49	.199	-1.82	.428	4.01	-.503	1.90	-.22	-2.92	.35	0.47
.2	.202	-3.66	.112	-1.33	.464	5.96	-.610	3.03	-.19	-2.53	.07	0.89
.4	.194	-4.35	.124	-1.44	.484	7.43	-.539	3.09	-.06	-0.78	.21	0.29
.6	.189	-5.20	.104	-0.97	.494	8.90	-.479	2.98	.08	1.24	.45	6.82
.8	.186	-6.36	.098	-1.19	.495	-11.83	-.506	3.43	.12	2.21	.71	13.40
.840	.185	-6.57	.099	-1.31	.494	-12.43	-.513	3.56	.11	2.18	.75	15.56
.8521	.185	-6.62	.100	-1.35	.494	-12.56	-.515	3.59	.11	2.18	.76	16.10
.877	.184	-6.67	.105	-1.44	.494	-12.76	-.519	3.65	.10	2.20	.77	16.94
.90	.183	-6.67	.111	-1.55	.494	-12.87	-.521	3.69	.10	2.26	.77	17.37
.9521	.182	-6.59	.134	-1.83	.494	-12.89	-.523	3.71	.11	2.48	.76	17.50
1.00	.180	-6.42	.161	-2.15	.493	-12.70	-.520	3.67	.13	2.77	.74	17.05
1.041	.179	-6.25	.187	-2.41	.492	-12.43	-.515	3.60	.14	3.02	.72	16.53
1.0521	.178	-6.20	.195	-2.48	.491	-12.35	-.514	3.57	.15	3.08	.71	16.38
1.098	.177	-5.99	.226	-2.75	.489	-11.96	-.505	3.45	.16	3.33	.69	15.75
1.1	.177	-5.98	.227	-2.76	.488	-11.94	-.505	3.45	.16	3.34	.69	15.73
1.25	.172	-5.26	.336	-3.52	.472	-10.43	-.466	2.96	.21	4.05	.61	13.82
1.5	.162	-4.20	.504	-4.46	.425	- 7.96	-.397	2.19	.30	5.23	.50	11.14
1.75	.150	-3.21	.677	-5.40	.327	- 5.13	-.323	1.51	.45	7.58	.37	8.26
2.0	.142	-2.45	.870	-6.68	.127	- 1.71	-.371	1.44	.67	12.44	.19	4.34
					$\mu_y = \hat{\mu}_y$							
0	1.043	-1.65	-1.372	1.84	.427	- 0.51	1.562	-0.81	-.09	-1.15	-.16	-2.10
.2	.424	-2.87	-.377	1.98	.440	- 2.19	.227	-0.45	-.27	-3.66	-.10	-1.42
.4	.253	-3.81	.030	-0.33	.486	- 5.35	-.321	1.37	-.31	-4.24	-.03	-0.39
.6	.189	-4.12	.108	-1.32	.500	- 7.67	-.437	2.51	-.10	-1.36	.17	2.21
.8	.184	-5.82	.047	-0.46	.510	-10.97	-.476	3.07	.09	1.54	.66	11.99
.873	.184	-6.83	.039	-0.56	.494	-13.11	-.514	3.63	.05	1.17	.83	18.44
.9	.184	-6.93	.052	-0.77	.493	-13.38	-.524	3.76	.05	1.27	.83	19.57
.95	.182	-6.73	.105	-1.49	.495	-13.22	-.527	3.78	.09	2.01	.79	18.50
1.00	.180	-6.42	.161	-2.15	.493	-12.70	-.520	3.67	.13	2.77	.74	17.05
1.051	.179	-6.12	.213	-2.66	.489	-12.07	-.511	3.52	.16	3.32	.69	16.00
1.1	.178	-5.85	.259	-3.03	.482	-11.43	-.501	3.35	.18	3.68	.66	15.20
1.25	.175	-5.08	.382	-3.78	.457	9.54	-.467	2.83	.24	4.42	.58	13.24
1.5	.170	-3.99	.549	-4.49	.394	6.70	-.396	2.01	.34	5.83	.47	10.37
1.75	.164	-2.88	.789	-5.84	.175	2.34	-.302	1.19	.60	10.97	.25	6.01
2.0	.155	-2.37	.894	-6.43	.048	5.47	-.408	1.44	.72	13.89	.13	2.83

*By definition, elasticities are computed with the opposite of the sign of the estimated coefficient or t-statistic.

provided in Table 16. These practical criteria give further support to the results of the tests of functional form: the results are in general more reasonable at the point of optimal value of functional form parameters than elsewhere.

3.4 *Results for the schoolchildren market equation*

The remainder of this section is devoted to a summary of the results obtained for the second market. In order not to prolong our paper unduly, we shall summarize the results rapidly and note a few differences with the adult market results. Because the CECM model ($\mu_y = 0$) is a special case of the TLCM model ($\mu_y \neq 0$), we will again examine them jointly.

3.4.1 *Estimation and tests of the functional form.* As shown in Figs. 15 and 16, the likelihood function closely resembles that of the adult market with a maximum at $\hat{\lambda} = 0.86$ and $\hat{\mu} = -1{,}194{,}609$, its maximum value under the binding constraint $y + \mu_y > 0$ defined in (3).

As was the case for the adult market equation, the conditional test for λ, found in Table 17, shows that one could not reject the hypothesis that $\hat{\lambda} = 0.86$ is equal to 1, for $\mu_y = -1{,}194{,}609$ at the 95% level; in contrast with adults, the conditional test for μ_y, found in Table 18, shows that one could reject the hypothesis that μ_y is equal to zero, but only for $\lambda = 0.86$. However the unconditional test, found in Table 19, implies that the original linear model with $\lambda = 1$ and $\mu_y = 0$ cannot be

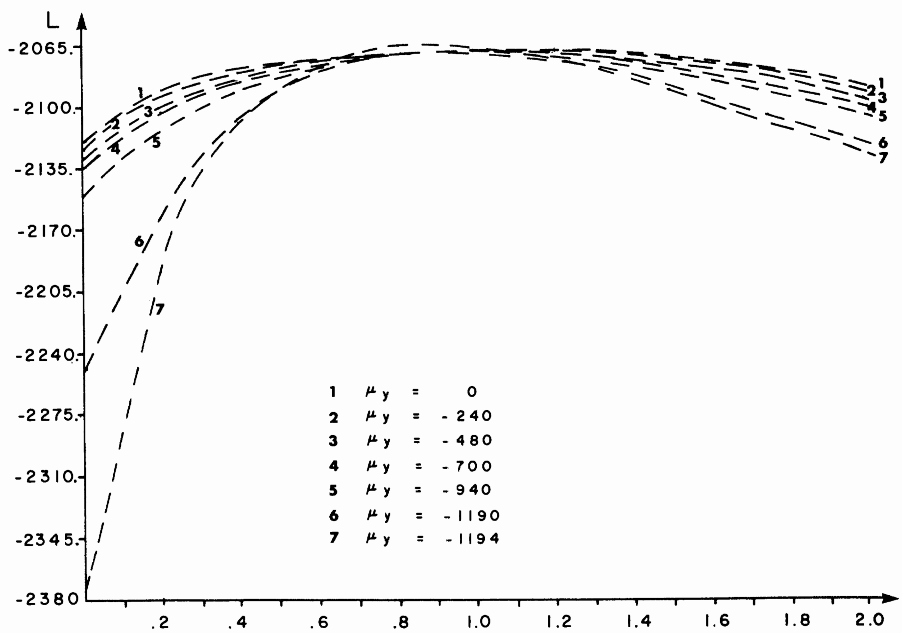

Fig. 15. Schoolchildren market equation; log likelihood function.

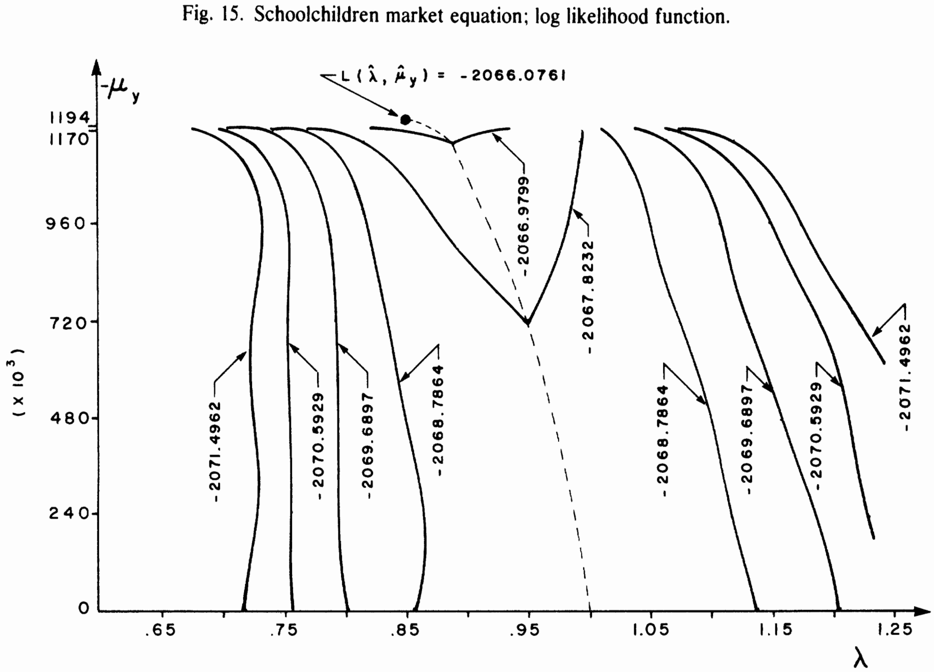

Fig. 16. Schoolchildren market equation; log likelihood surface near optimum λ and μ_y.

Table 17. Schoolchildren market equation; conditional test for λ

C(λ)	μ_y=0		μ_y= -480,000		μ_y= -940,000		μ_y= -1,194,609	
λ	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.
0	-2118.5233	50.5243	-2127.7623	59.8047	-2151.4300	83.7305	-2378.1581	312.0814
.2	-2095.1598	27.1608	-2100.1012	32.1436	-2113.8296	46.1301	-2197.8264	131.7497
.4	-2083.2925	15.2935	-2085.1341	17.1765	-2090.5840	22.8845	-2111.2009	45.1242
.6	-2074.9831	6.9841	-2075.2122	7.2546	-2076.1858	8,4863	-2076.9154	10.8387
.8	-2069.7286	1.7296	-2069.4496	1.4920	-2068.9570	1.2575	-2066.5853	0.5086
$\hat\lambda$	-2067.9990	0	-2067.9576	0	-2067.6995	0	-2066.0767	0
1.0	-2067.9990*	0	-2067.9990	0.0414	-2067.9990	0.2995	-2067.9990	1.9223
1.1	-2068.3751	0.3761	-2068.7228	0.7652	-2069.4362	1.7367	-2071.3946	5.3179
1.2	-2069.4354	1.4364	-2070.2264	2.2688	-2071.8456	4.1461	-2075.9798	9.9031
1.5	-2075.1233	7.1243	-2071.6118	9.6542	-2082.5820	14.8825	-2093.7210	27.6443
2.0	-2088.0513	20.0523	-2094.1017	26.1441	-2105.6332	37.9337	-2128.9198	62.8431

*For μ_y = 0, $\hat\lambda$ = 1.0.

Table 18. Schoolchildren market equation; conditional test for μ_y

	λ = 0		λ = .86		λ= 1.0		λ = 2.0	
μ_y	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.	L(λ,μ_y,$\hat\rho$)	Stat.
0	-2118.5233	0	-2068.8396	2.7629	-2067.9990	0	-2088.0513	0
-240,000	-2122.2734	3.7501	-2068.7253	2.6486	-2067.9990	0	-2090.6330	2.5817
-480,000	-2127.7623	9.2390	-2068.5746	2.4979	-2067.9990	0	-2094.1017	6.0504
-700,000	-2135.6297	17.1064	-2068.3839	2.3072	-2067.9990	0	-2098.5002	10.4469
-940,000	-2151.4300	32.9067	-2068.0654	1.9887	-2067.9990	0	-2105.6332	17.5819
-1,190,000	-2249.4399	130.9166	-2067.1071	1.0304	-2067.9990	0	-2121.2743	33.2230
-1,194,609	-2378.1581	259.6348	-2066.0767	0	-2067.9990	0	-2128.9198	40.8685

Table 19. Schoolchildren market equation; unconditional test for λ and μ_y

λ	μ_y	L(λ,μ_y)	Stat.
0	0	-2118.5233	52.4466
.2	0	-2095.1598	29.0831
.4	0	-2083.2925	17.2158
.6	0	-2074.9831	8.9064
.8	0	-2069.7286	3.6519
.86	0	-2068.8596	2.7629
1.0	0	-2067.9990	1.9233
1.5	0	-2075.1233	9.0466
2.0	0	-2088.0513	21.9746
0	-1,194,609	-2378.1581	312,0814
.2	-1.194,609	-2197.8264	131.7497
.4	-1,194,609	-2111.2009	45.1242
.6	-1,194,609	-2076.9154	10.8387
.8	-1,194,609	-2066.5853	0.5086
.86	-1,194,609	-2066.0767	0
1.0	-1,194,609	-2067.9990	1.9223
1.5	-1,194,609	-2093.7210	27.6443
2.0	-1,194,609	-2128.9198	62,8431

rejected. The statistical grounds for departing from the original model are therefore stronger than they were for the adult equation but not strong enough to do it with a high degree of confidence: the TLCM model with $\hat\lambda$ = 0.86 and $\hat\mu_y = -1{,}194{,}609$ appears on the whole only marginally preferable to the linear model, if at all.

3.4.2 *Effects of the functional form on the distribution of residuals.* For the unconditional test $U(\rho)$ of Table 10, the nest statistic was 25.1849; this value implied, with 4 degrees of freedom required by the four non-zero elements of vector ρ, that the null hypothesis $\rho = 0$ could be confidently rejected (percentage points given with Table 5). For the joint unconditional test $U(\lambda, \mu_y, \rho)$, the appropriate statistic was 25.8834 implying, with 6 degrees of freedom (percentage points table not given) that the null hypothesis ($\lambda = 1$, $\mu_y = 0$, $\rho = 0$) could also be rejected with confidence. These results of more convincing unconditional tests lead one to believe that, although we may be using too complex an original specification of the autoregressive structure $\rho = (\rho_1, \rho_3, \rho_4, \rho_{12})$, so that weaker lags are affected by the functional form parameters, as shown in Table 20, the strength of that structure is not fundamently weakened by the estimation of the functional form parameters. This effect of the functional form on the estimates of ρ is stronger than it was for the adult equation, presumably because the optimal λ is further away from 1 than it was in the adult market equation. In neither case have we been able to

Table 20. Schoolchildren market equation; autocorrelation and partial autocorrelation functions of residuals

Lags													Estimated standard error for row
	A. $\lambda = 1$, $\mu_y = 0$, $\rho = 0$; 181 observations												
	AUTOCORRELATIONS												
1-12	.16	-.02	-.22	-.20	-.05	.09	-.05	-.08	-.17	-.01	.02	.40	(.07)
13-24	.07	-.01	-.14	-.10	-.04	.04	-.10	-.04	-.20	-.06	.04	.24	(.09)
25-36	.03	-.04	-.11	-.02	.03	.09	.02	-.05	-.10	-.06	.02	.19	(.10)
37-48	.01	.01	-.07	.06	.07	.01	-.05	-.17	-.09	-.02	.05	.20	(.11)
49-60	.01	-.01	-.05	-.03	.05	-.01	-.14	-.15	-.06	.03	.11	.20	(.11)
	PARTIAL AUTOCORRELATIONS												
1-12	.16	-.04	-.22	-.15	-.01	.05	-.15	-.11	-.15	.01	-.06	.34	(.07)
	B. $\hat{\lambda} = .86$, $\hat{\mu}_y = -1{,}194{,}609$, $\rho = 0$; 181 observations; least-squares residuals												
	AUTOCORRELATIONS												
1-12	.15	-.02	-.24	-.19	-.07	.09	-.08	-.08	-.17	-.01	.02	.42	(.07)
13-24	.08	.02	-.13	-.09	-.06	.06	-.08	-.06	-.19	-.07	.01	.25	(.09)
25-36	.04	.00	-.11	-.04	-.00	.09	.02	-.04	-.11	-.07	.00	.20	(.10)
37-48	.03	.04	-.08	.04	.06	.01	-.05	-.14	-.10	-.01	.03	.20	(.11)
49-60	.03	.02	-.06	-.05	.05	.01	-.12	-.15	-.07	.02	.11	.20	(.11)
	PARTIAL AUTOCORRELATIONS												
1-12	.15	-.04	-.23	-.13	-.03	.06	-.19	-.11	-.15	-.02	-.07	.35	(.07)
	C. $\hat{\lambda} = .86$, $\hat{\mu}_y = -1{,}194{,}609$, $\hat{\rho} = \hat{\rho}_1, \hat{\rho}_3, \hat{\rho}_4, \hat{\rho}_{12}$; 169 observations; 15th iteration Cochrane-Orcutt residuals												
	AUTOCORRELATIONS												
1-12	.04	-.01	.01	-.16	-.01	.02	-.13	.04	-.10	.08	-.08	-.02	(.08)
13-24	.01	.03	.03	.06	-.07	-.05	-.09	-.04	-.17	-.02	-.01	.04	(.08)
25-36	-.01	-.03	.02	-.02	-.02	.17	.04	.05	-.03	-.08	.04	.02	(.09)
37-48	.01	.03	-.13	.03	.04	-.00	.02	-.11	-.01	.01	-.01	.04	(.09)
49-60	-.01	.04	-.04	-.10	-.01	.00	-.09	-.08	.02	.01	.08	.09	(.09)
	PARTIAL AUTOCORRELATIONS												
1-12	.04	-.01	.01	-.16	.01	.02	-.13	.03	-.11	.11	-.14	.01	(.08)

isolate a dramatic change of the autoregressive structure because our optimal forms have been close to the original linear form.

By contrast, reestimation of λ under the assumption that $h = 1$ in (83) yielded approximately 1.95 instead of 0.86 ± 0.14 (with a 95% interval) which implies that our original estimate of λ in the schoolchildren market equation may reflect an estimation of heteroskecasticity to a much greater extent than did our estimates of λ in the adult market equation.

Table 21 presents the R^2 coefficient in terms of doubly transformed variables and computed as in (84).

3.4.3 *Effect of the functional form on the significance of variables.* Subject to the warnings made in Sections 2.3.3. and 3.3.3. above, Figs. 17 and 18 give an idea of the impact of assuming an incorrect functional form on the statistical significance of variables. As for the adult equation, assuming that the form is logarithmic instead of approximately linear changes the sign of a number of variables (travel time, for $\mu_y = 0$; travel time and waiting time, for $\mu_y = \hat{\mu}_y$) in an unreasonable direction and also changes the sign of other variables (ρ_4 and income, for $\mu_y = 0$; ρ_4 and income, for $\mu_y = \hat{\mu}_y$), the true signs of which are unknown; as in the adult equation, the results for $\mu_y = \hat{\mu}_y$ are worse than those for $\mu_y = 0$ and they provide support for the linear or approximately linear form established by the functional form tests.

3.4.4 *Effects of the functional form on elasticities and autoregressive coefficients.* If, in Figs. 19 and 20, one retains the variables which have a reasonable sign for the logarithmic form, the income elasticity is the only one which is absurdly high; other absurdly high elasticities,

Table 21. Schoolchildren market equation; R^2

λ	R^2 in terms of $(\lambda)_*$: $\mu_y = 0$	R^2 in terms of $(\lambda)_*$: $\mu_y = \hat{\mu}_y$
0	.8307513	.3330711
.2	.9548996	.6893671
.4	.8602800	.8531399
.6	.8607228	.8867933
.8	.8571847	.8728477
$\hat{\lambda}$	.8489854	.8660312
1.0	.8489854	.8489868
1.1	.8420659	.8355506
1.3	.8204326	.8060134
1.5	.8029628	.7864064
2.0	.7846937	.7789682

notably for $\mu_y = \hat{\mu}_y$, belong to the waiting time and travel time variables which have unreasonable signs for the logarithmic form.

Therefore, as we found for adult equation, the linear model cannot be rejected with confidence and the circumstancial evidence yielded by the behavior of t-statistics and of demand elasticities, provided in details in Table 22 for other models such as the log linear one, emphasizes the importance of using the correct functional form.

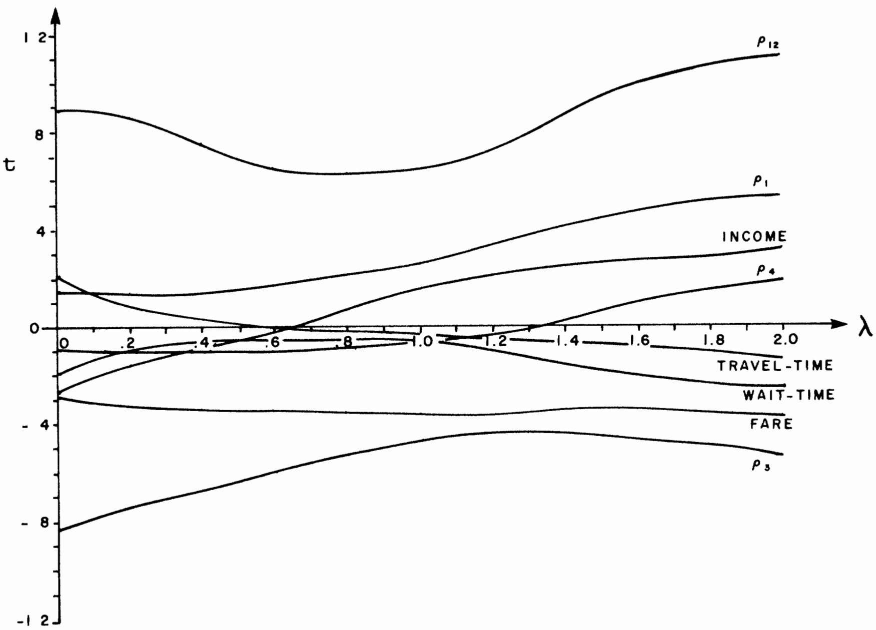

Fig. 17. Schoolchildren market equation; t-statistics for selected variables and the autoregressive parameters, $\mu_y = 0$.

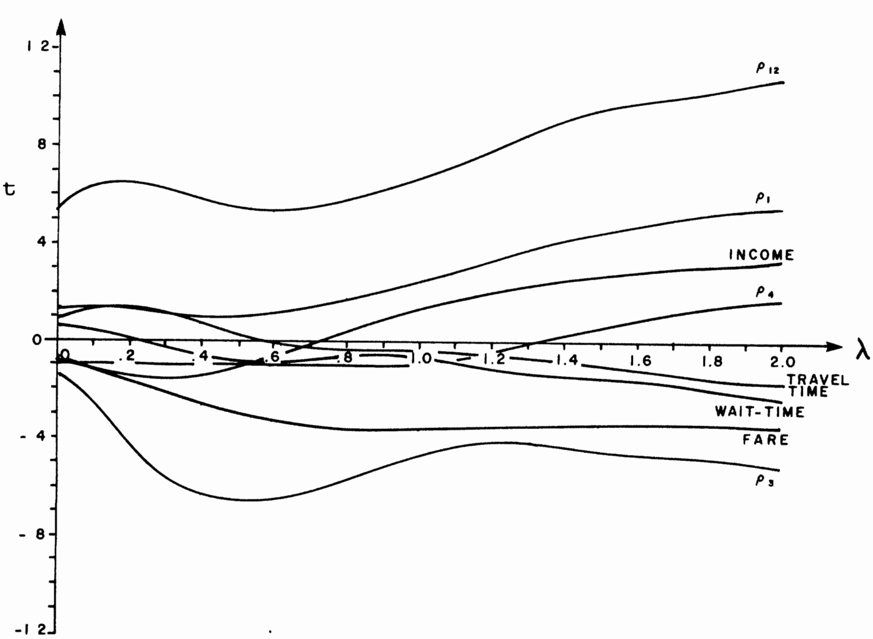

Fig. 18. Schoolchildren market equation; t-statistics for selected variables and the autoregressive parameters, $\mu_y = \hat{\mu}_y$.

4. CONCLUSION

We have shown how particular models found in the literature can be seen as special cases of a broad family of models in which transformation parameters used on the fixed part, or parameters which specify the stochastic part, of such models have been constrained in many ways: the most current of these models are the linear or log-linear formats for the fixed part and the use of first differences of residuals for the stochastic part of time-series models; it is the standard logit model in market share models. We have estimated aggregate models which used fewer restrictions and shown from these particular examples that weaker restrictions were preferable not only because statistical tests might show that they dominate models which use stronger restrictions but also because the sign, significance and elasticities associated with parameters of explanatory variables determined jointly with functional form parameters might be

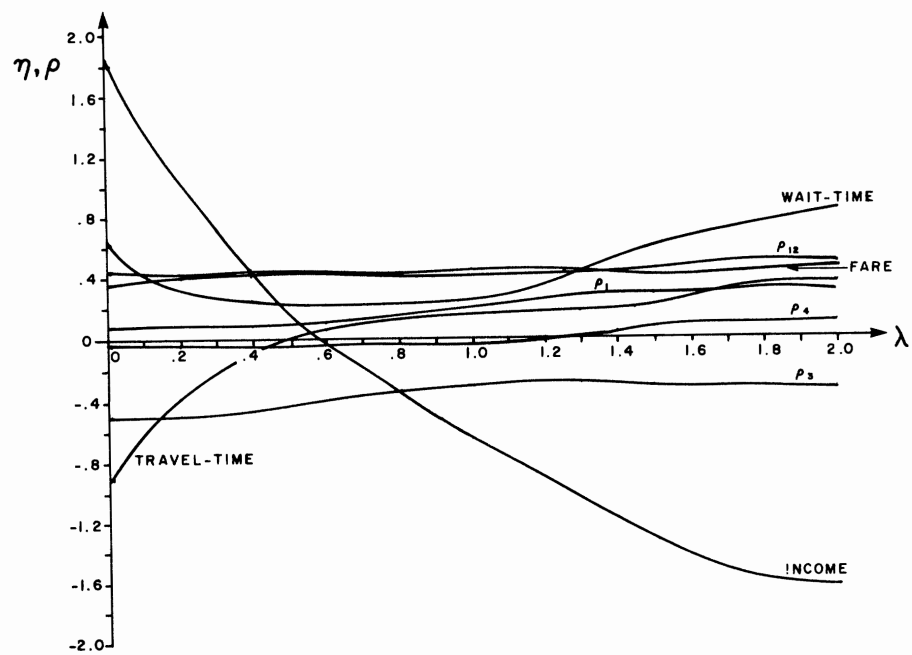

Fig. 19. Schoolchildren market equation; selected elasticities and values of autoregressive coefficients, $\mu_y = 0$.

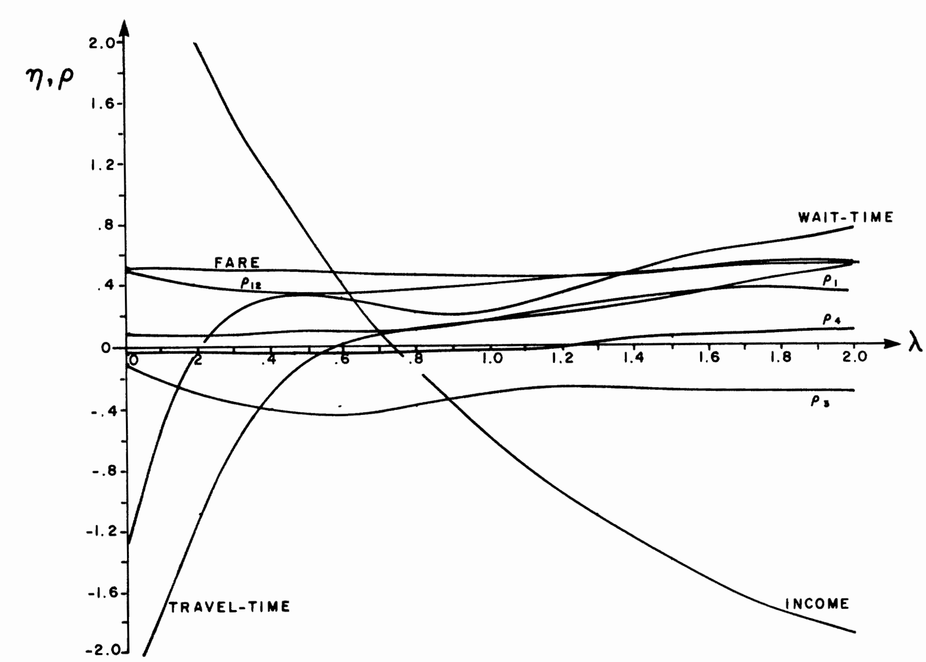

Fig. 20. Schoolchildren market equation; selected elasticities and values of autoregressive coefficients, $\mu_y = \hat{\mu}_y$.

more convincing than results obtained otherwise. We have shown in particular that a standard logit mode share model and logarithmic total demand models were clearly inferior to flexible form models. The fact that optimal values of the functional form parameters seem to be associated with the most convincing signs and demand elasticities apparently provides an *ex post* justification of the frequently used approach of trying a few forms and keeping that which yields "good" results. The considerable impact of functional form parameters on all other parameters of the models we have considered suggests that this approach can go a long way towards improving a number of models in current use, such as the standard logit model. Whether any gains to be made in this way are more important than those made by using simultaneous equations estimation techniques should be the object of further research in which these problems would be considered jointly.

Table 22. Schoolchildren market equation; t-statistics, elasticities and autoregressive coefficients

	Fare		Wait-time		Travel-time		Income		ρ_1		ρ_3		ρ_4		ρ_{12}	
	η	t	η	t	η	t	η	t	value	t	value	t	value	t	value	t
							$\mu_y = 0$									
.0	.463	-2.96	0.637	-1.96	-0.909	2.07	1.879	-2.76	.10	1.54	-.46	-8.24	-.03	-0.43	.36	8.89
.2	.457	-3.44	0.326	-1.00	-0.395	0.99	0.995	-1.70	.10	1.46	-.43	-7.67	-.06	-0.87	.40	8.57
.4	.442	-3.53	0.258	-0.82	-0.105	0.29	0.459	-0.87	.10	1.58	-.41	-6.90	-.06	-0.90	.41	7.44
.6	.431	-3.59	0.224	-0.76	0.057	-0.17	0.044	-0.09	.12	1.77	-.38	-6.03	-.07	-0.94	.41	6.51
.653	.430	-3.61	0.217	-0.75	0.084	-0.25	-0.058	0.12	.13	1.85	-.37	-5.82	-.07	-0.95	.41	6.37
.791	.431	-3.67	0.205	-0.74	0.131	-0.41	-0.310	0.65	.14	2.08	-.35	-5.36	-.07	-0.99	.41	6.20
.8	.431	-3.68	0.204	-0.74	0.133	-0.42	-0.325	0.69	.14	2.09	-.35	-5.33	-.07	-0.99	.41	6.20
.9	.434	-3.72	0.203	-0.75	0.150	-0.48	-0.493	1.05	.16	2.29	-.33	-5.08	-.07	-1.00	.41	6.24
1.00	.438	-3.75	0.213	-0.80	0.158	-0.52	-0.650	1.39	.17	2.51	-.32	-4.87	-.07	-0.97	.41	6.37
1.1	.441	-3.76	0.241	-0.89	0.161	-0.54	-0.798	1.69	.19	2.78	-.31	-4.70	-.06	-0.87	.41	6.65
1.231	.443	-3.71	0.334	-1.17	0.164	-0.57	-0.980	2.03	.22	3.30	-.29	-4.55	-.04	-0.51	.43	7.43
1.3	.443	-3.68	0.410	-1.38	0.170	-0.60	-1.067	2.18	.24	3.63	-.29	-4.54	-.02	-0.23	.43	8.00
1.428	.447	-3.64	0.549	-1.76	0.195	-0.71	-1.206	2.39	.27	4.17	-.30	-4.63	.02	0.31	.45	8.93
1.5	.450	-3.64	0.614	-1.94	0.216	-0.79	-1.274	2.50	.28	4.41	-.29	-4.73	.04	0.60	.46	9.30
2.0	.485	-3.79	0.853	-2.63	0.376	-1.45	-1.656	3.04	.32	5.24	-.31	-5.30	.11	1.85	.51	11.03
							$\mu_y = \hat{\mu}_y$									
0	.536	-0.62	-1.345	0.57	-2.533	0.98	2.734	-0.66	.10	1.34	-.11	-1.58	-.07	-0.97	.53	5.26
.2	.541	-1.89	-0.021	0.03	-1.157	1.33	2.033	-1.56	.09	1.29	-.28	-4.14	-.07	-1.04	.44	6.57
.4	.467	-2.85	0.345	-0.88	-0.353	0.72	1.197	-1.74	.07	0.98	-.41	-6.46	-.07	-0.99	.37	5.66
.6	.432	-3.37	0.290	-1.06	0.039	-0.10	0.461	-0.89	.07	1.05	-.43	-6.75	-.08	-1.09	.35	5.26
.664	.431	-3.49	0.252	-0.96	0.087	-0.25	0.239	-0.48	.09	1.26	-.42	-6.48	-.08	-1.07	.36	5.38
.8	.433	-3.65	0.205	-0.79	0.134	-0.41	-0.174	0.37	.12	1.79	-.38	-5.76	-.07	-1.02	.39	5.76
.86	.434	-3.70	0.198	-0.76	0.144	-0.45	-0.332	0.71	.14	2.01	-.36	-5.45	-.07	-1.01	.39	5.94
.9	.435	-3.72	0.199	-0.76	0.149	-0.48	-0.429	0.92	.15	2.16	-.35	-5.27	-.07	-1.00	.40	6.06
1.00	.438	-3.75	0.213	-0.80	0.158	-0.52	-0.650	1.39	.17	2.51	-.32	-4.87	-.07	-0.97	.41	6.37
1.10	.440	-3.74	0.250	-0.90	0.168	-0.57	-0.846	1.79	.19	2.88	-.30	-4.59	-.06	-0.86	.42	6.80
1.121	.441	-3.73	0.262	-0.94	0.170	-0.58	-0.884	1.86	.20	2.97	-.30	-4.54	-.06	-0.82	.42	6.91
1.30	.444	-3.64	0.416	-1.35	0.206	-0.72	-1.177	2.36	.25	3.81	-.28	-4.39	-.01	-0.22	.44	8.23
1.50	.455	-3.59	0.578	-1.76	0.284	-1.02	-1.434	2.73	.29	4.56	-.29	-4.61	.04	0.56	.46	9.35
2.00	.504	-3.65	0.765	-2.20	0.496	-1.80	-1.958	3.33	.33	5.46	-.31	-5.10	.11	1.76	.50	10.81

*By definition, elasticities are computed with the opposite of the sign of the estimated coefficient or t-statistic.

Acknowledgements—The views presented in this paper are the authors' and in no way engage the institutions with which they are affiliated. The authors are jointly responsible for the whole of the analysis but Marc Gaudry, of the *Département des sciences économiques de l'Université de Montréal*, assumes primary responsibility for the time-series application and Michael Wills, of the Strategic Planning Group of Transport Canada, Ottawa, for the cross-sectional application. Much of the research was carried out during the summer of 1976 at the *Centre de recherche sur les transports de l'Université de Montréal* and supported by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'éducation du Québec*. Marcel Dagenais and Claude Montmarquette provided valuable advice; Danielle Bilodeau and Claude Achim supplied computing assistance for the time-series model. Partial results were presented at the June 1977 meeting of the Canadian Economics Association.

REFERENCES

Anscombe F. J. and Tukey J. W. (1954) *The criticism of transformations.* Paper presented before the American Statistical Association and Biometric Society, Montreal.

Baumol W. J. (1965) *Economic Theory and Operations Analysis*, Second Edn, Chap. 9. Prentice-Hall, Englewood Cliffs, New Jersey.

Beach C. M. and MacKinnon J. G. (1978) A maximum likelihood procedure for regression with autocorrelated errors. *Econometrica*, **46**, 51–58.

Box G. E. P. and Cox D. R. (1964) An analysis of transformations. *J. Royal Statistical Soc.* Series B, **26**, 211–243.

Box G. E. P. and Jenkins G. M. (1970) *Time-Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

Box G. E. P. and Tidwell P. W. (1962) Transformation of the independent variables. *Technometrics* **4**, 531–550.

Canadian Transport Commission [Rea J. C., Dahl M. D., Hooper

T. J., Ku K. D., Platts J. B., Ranger L. J. and Wills M. J.] (1976) *Intercity passenger transport in Canada: analysis of the consequences of alternative pricing and network strategies.* Research Report 254.

Cochrane D. and Orcutt G. H. (1949) Application of least-squares regression to relationships containing autocorrelated error terms. *J. Am. Statistical Ass.* **44**, 32–61.

Crow R. T., Young K. H. and Cooley T. (1973) Alternative demand functions for abstract transportation modes. *Transpn Res.* **7**, 335–354.

Domencich T. A., Kraft G. and Valette J. P. (1968) Estimation of urban passenger travel behavior: an economic demand model. *Highway Research Record* 238, 64–78.

Fair R. C. (1971) *A comparison of alternative estimators of macroeconomic models.* Research Memorandum No. 121. Econometric Research Program, Princeton University.

Fairhurst M. H. and Morris P. J. (1975) *Variations in the demand for bus and rail travel up to* 1974. Economic Research Report R 210, London Transport Executive.

Fisher F. M. (1962) *A Priori Information and Time-Series Analysis: Essays in Economic Theory and Measurement,* Chap. 6. North-Holland, Amsterdam.

Gaudry M. (1973) The demand for public transit in Montreal and its implications for transportation planning and cost-benefit analysis. Ph.D. Dissertation, Princeton University.

Gaudry M. (1975) An aggregate time-series analysis of urban transit demand: the Montreal case. *Transpn Res.* **9**, 249–258.

Gaudry M. (1977) Seemingly unrelated static and dynamic urban travel demands. Publication Nol 61, Centre de recherche sur les transports, Université de Montréal. *Transpn Res.* **12**, 195–212.

Goldfeld S. M. and Quandt R. E. (1972) *Nonlinear Methods in Econometrics,* Chap. 7. North-Holland, Amsterdam.

Huang T. J. (1973) *An econometric analysis of the determinants of demand for local transit in Edmonton,* 1961–70. Economic Studies Report No. 1, Finance Department. Edmonton, Canada.

Johnston J. (1972) *Econometric Methods.* 2nd Edn, p. 262.

Kadiyala K. R. (1969) A transformation used to circumvent the problem of autocorrelation. *Econometrica* **36**, 93–96.

Kau J. B. and Sirmans C. F. (1976) *The functional form of the gravity model: a new technique with empirical results.* Working paper. Department of Finance, University of Illinois, presented at the North American Meeting of the Regional Science Association in Toronto.

Kraft G. (1963) *Demand for intercity passenger travel in the Washington-Boston corridor*—V. Northeast Corridor Project, U.S. Department of Transportation.

Lancaster K. J. (1966) A new approach to consumer theory. *J. Political Economy* **LXXXLV**, 132–157.

McLynn J. M., Goldman A. J., Meyers P. R. and Watkins R. H. (1968) Analysis of a market split model. *J. Res. National Bureau of Standards* **72-B**, No. 1.

McLynn J. M. and Woronka T. (1969) Passenger demand and modal split models. Northeast Corridor Project, U.S. Department of Transportation, Report No. 230.

Meyer J. R. and Straszheim M. H. (1971) *Pricing and Project Evaluation.* Brookings Institution, Washington.

Monsod S. (1967) Further developments in the cross-sectional rail passenger demand model. Studies in Travel Demand. *Mathematica* Vol. III.

Monsod S. (1969) Relative shares model. Studies in Travel Demand. *Mathematica* Vol. V.

Quandt R. E. and Baumol W. J. (1966) The demand for abstract transport modes: theory and measurement. *J. Regional Sci.* **6**(2), 13–26.

Sargan J. D. (1964) Wages and prices in the United Kingdom: a case study in econometric methodology. *Econometric Analysis for National Economic Planning.* (Edited by P. E. Hart *et al.*) Butterworth, London.

Savin N. E. and White K. J. (1976) *Estimation and testing for functional form and autocorrelation: a simultaneous approach.* Discussion Paper 76–26, Department of Economics, University of British Columbia.

Schlesselman J. (1971) Power families: a note on the Box and Cox transformation. *J. Royal Statistical Soc.* Series B., **33**.

Spitzer J. J. (1977) A simultaneous equations system of money demand and supply using generalized functional forms. *J. Econometrics* **5**, 117–128.

Straszheim M. J. (1969) *The International Airline Industry.* The Brookings Institution, Washington.

Tukey J. W. (1957) On the comparative anatomy of transformations. *Annals of Mathematical Statistics* **28**, 602–632.

U.S. Civil Aeronautics Board (1966) *Measuring the elasticity of air passenger demand; a study of changes over time from* 1953–64. Washington.

Wills M. J. (1974) *Estimating the demand for travel using abstract mode methodology.* Research Report. Ministry of State for Urban Affairs, Ottawa.

Young K. H. (1969) An abstract mode approach to the demand for travel. *Transpn Res.* **3**, 443–461.

Zarembka P. (1968) Functional form in the demand for money. *J. Am. Statistical Assoc.* 502–511.

Zarembka P. (1974) Transformation of variables in econometrics. *Frontiers in Econometrics* (Edited by P. Zarembka) Chap. 3.