Title

Review and Evaluation of Transport-Related Fuel-Use Forecasting Models

Authors
Golob, Thomas F.
Kitamura, Ryuichi
Brownstone, David
et al.

Publication Date
1991-04-01

Review and Evaluation of Transport-Related Fuel-Use Forecasting Models

UCI-ITS-WP-91-2

Thomas F. Golob [1]
Ryuichi Kitamura [2]
David Brownstone [3]
Trucy Cameron [4]
Will Recker [1]

[1] Institute of Transportation Studies
University of California, Irvine

[2] Department of Civil Engineering
University of California, Davis

[3] Department of Economics
University of California, Irvine

[4] Department of Economics
University of California, Los Angeles

April 1991

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

Prepared for the California Energy Commission

Table of Contents

Financial Disclosure Statement

This research was funded by the California Energy Commission. The total ceiling amount for this contract (Contract No. 500-88-015) was $92,030.00. The research is reported in the final report: Review and Evaluation of Transport-Related Fuel-Use Forecasting Models.

Implementation Statement

Copies of this report will be distributed to selected individuals at the Sacramento office of the California Energy Commission. The procedures and results described in this report may be used to evaluate models designated to project fuel consumption in the State of California.

Disclaimer

Acknowledgements

The authors wish to express their thanks to several individuals whose valuable assistance made this report possible. Ram M. Pendyala and Konstadinos G. Goulias at the University of California, Davis provided suggestions for the improvement of freight energy demand models. Saney Morrison, Barbara Neenan and Mandar Khanal at the University of California, Irvine helped in the compilation of the reviews on disaggregate and aggregate forecasting models. The editorial assistance of Elizabeth Morrison is also gratefully acknowledged.

I. EXECUTIVE SUMMARY

This report provides an exhaustive review of the state-of-the-art in transportation-related energy demand forecasting models. Models used for forecasting fuel demand and consumption in the passenger vehicle, transit and freight sectors of the transportation economy are described and discussed. Based on the review, recommendations are made on how the accuracy and robustness of the forecasts can be improved while explicitly incorporating behavioral relationships among various users of transportation services.

The private vehicle model presently used by the California Energy Commission suffers from the following main draw backs:

1) The model is estimated on cross sectional data sets and thus provides forecasts based on cross sectional variations across the population.

2) The model does not incorporate a demographic forecasting tool. Socio-demographic and economic characteristics are considered exogenous to the model system and therefore cannot be used for various scenario analyses.

3) The model does not adequately model the vehicle transaction behavior and vehicle usage behavior of households, which are highly inter-related.

Aggregate time series models will provide the most accurate forecasts if there are no major policy shifts. If the forecasting model has to be sensitive to policies such as vehicle emission taxes or mandatory use of alternative fuels, then a more complex disaggregate model is required. The proposed model will incorporate a detailed vehicle transaction model which is in turn estimated from models of fuel type choice, vehicle utilization and vehicle demand. This model will be able to predict all types of transactions; acquiring a new or used car, replacing a currently owned car, disposing of a currently owned car and preserving the present fleet. Various econometric techniques are now available for developing dynamic models of travel behavior that will relate these choices in a simultaneous equation framework. The choice to purchase alternatively fueled vehicles will be modeled from stated preference data until they become widespread. Another major component of the proposed model system would be a demographic forecasting system where different scenarios can be developed and analyzed. This demographic system would simulate different scenarios of employment levels and household evolution.

The development of such a model system would require conducting a panel survey of about 2500 households, bi-annually for the first few years, after which it can be done annually. It is suggested that a telephone modified random digit dialing survey be performed. The households will have to be replenished over the waves of the survey to reflect the changing demographics of the region and to compensate for the attrition that is associated with a panel survey. The data to be collected in the survey will include socio-demographic, economic, vehicle transaction, and vehicle utilization information of the household.

Forecasting at the market level will be done using a simultaneous equation system to allow vehicle prices, fuel prices and usage to be treated as endogeneous to the system. Unfortunately, fuel and vehicle prices are primarily determined by national and international market forces. Therefore implementation of endogenous fuel and vehicle prices will require building a national model or assuming that the California market is representative of the national market. The latter assumption is clearly suspect for forecasting the effects of California-specific policies such as mandatory alternative fuel use.

Dynamic vehicle transaction and utilization models combined with demographic simulators are a new approach to transportation demand forecasting that will account for the rapidly changing urban structure of our society.

Existing aggregate models relating freight transport energy demand to regional economic activity perform quite well. Although these models could be improved by formulating detailed disaggregate models of shippers' and receivers' behavior, the benefits probably do not justify the high cost of data collection. Since shippers and receivers attempt to minimize total costs, attention should be focused on collecting and forecasting cost components for different modes and fuel types. One neglected cost component is reliability, including on-time performance and damage/loss frequencies.

Transit competes with private vehicle utilization, so transit demand should be integrated and jointly estimated with the private vehicle model. Since there are so few transit users in California, this integration will require a special panel survey of transit users in addition to the panel survey needed for calibrating the personal vehicle model. Until transit utilization increases by an order of magnitude from its current levels, the costs of this additional survey and modeling clearly exceeds the benefits of better forecasts.

This report is organized as follows. The first section contains a summary of the recommendations on energy demand forecasting models. The second section discusses the aspects of scenario analysis while the third section is a review of Train's vehicle choice model which has been widely used. The fourth section reviews aggregate forecasting models, disaggregate forecasting tools, and transit and freight energy demand forecasting models .

II. RECOMMENDATIONS

A. Private Vehicle Model for Fuel Use Forecasting

Development of a new CEC forecasting capability, based on a new personal vehicle usage model (PVM), new data set, and new user software, is highly recommended. No existing model, including the present CEC PVM, meets CEC current and projected needs. The time is ripe for development of a new capability: methodological breakthroughs during the last ten years make possible and practical a new joint vehicle choice and usage choice model to drive fuel use forecasting. The new model would borrow from several existing models, including the current CEC PVM model. However, the model must have unique features to meet specific CEC needs.

The recommended new forecasting tool is comprised of a new dynamic personal vehicle ownership and usage model system, supported by a new panel data set and new simulation software for forecasting and scenario evaluation. The data set would have three uses: (1) for initial model calibration, (2) for model updating (especially with the possible appearance of clean fuel and ultra clean fuel vehicles), and (3) as a continuous basis for forecasting. Due to its size and complexity, the proposed model will not perform as well as simple aggregate time series models when there are no major policy changes. The new model is needed to realistically forecast the effects of proposed major policy shifts.

B. The Proposed Model

a. Basic Structure

The model must be fully dynamic to meet CEC needs and to take advantage of state-of-the-art forecasting techniques. That is, the model must be capable of: (1) describing the vehicle holdings and usage levels for each vehicle in a household at a given point in time, (2) capturing household transaction decisions as to whether or not to change their vehicle holdings, and (3) capturing household usage decisions as to whether or not to change levels of usage (with and without ownership transactions). The explanatory variables must include, but are not limited to: aging of the vehicles owned by the household, changes in the characteristics of available new vehicles (including new fuels), changes in environmental conditions (such as travel times affected by congestion levels), and changes in household characteristics (such as labor force participation, residential and employment locations, and disposable income).

The first capability of the model is similar to that demonstrated by models such as the current CEC PVM (Train, 1986; see also Mannering and Winston, 1985; Berkovec and Rust, 1985; Hensher et al., 1987; de Jong, 1989b). The second and third (dynamic) capabilities are critical for the reasons discussed in Clark et al. (1982), Davies and Pickles (1985), Goodwin and Mogridge (1981), Hensher and Wrigley (1986), and Kitamura (1990). Specifically, these capabilities are critical in the present context for the following reasons: Vehicle purchase and disposal is a process, and the temporal dimensions of this process are important to accurate forecasting because they are related to critical explanatory variables. Static models, by definition, are incapable of including the timing of household vehicle purchases without introducing so- called "fudge factors", such as pseudo transaction costs and dummy variables for current vehicle holdings. These fudge factors invariably introduce parameter estimation biases because the error terms and the pseudo-dynamic variables can never be independent (endogeneity bias). Such biases are not generally detectable in static models. Unfortunately, they are also not negligible, as demonstrated by the result that pseudo transaction cost or holdings variables are typically the most important explanatory variables in static vehicle choice models.

A static holding model cannot, by definition, include transaction costs which are involved in the purchase or disposal of a vehicle. A demand model that neglects to include these costs will introduce a significant error unless the costs are very small. However, purchase of a vehicle is a significant expenditure for most households, and households are thus willing to allocate non-negligible transaction costs. Those costs include time and money devoted to acquiring information about various car makes and models, search time, and time devoted to negotiation and the actual transaction. The omission of transaction costs leads to biased and inconsistent estimators (Hocherman et al., 1987). For these reasons, it is strongly recommended that the "vehicle type" choice component of the proposed modeling system be of the transactions type, covering household acquisition, replacement, and disposal behavior.

Static representations of both choice of vehicle and choice of usage level cannot account for effects that are relatively uniform over the cross-sectional (single point in time) sample on which such models are calibrated. However, cross-sectionally uniform effects are likely to vary over time, and CEC forecasting must be sensitive to them. Such effects include, but are not limited to: (1) macroeconomic trends in costs versus disposable incomes (e.g., fuel prices as distinct from vehicle fuel economy levels), (2) trends in tastes (especially those concerning vehicle attributes), and (3) attitudes (such as those concerning the environment). These influences on personal vehicle fuel usage have to date been captured only in aggregate time-series models; they are absent from disaggregate choice models. They, can however, be incorporated in a dynamic disaggregate choice model.

The elasticities calculated from all static models for forecasting purposes are accurate estimates of change only if the attributes of the alternatives are restricted to incremental changes and if the behavior under study has no response time lags, no anticipated effects, and is otherwise independent of past history. It is particularly unlikely that household vehicle holdings and usage behavior conforms to these very restrictive conditions. Brand loyalty is one example of the effects of past decisions. Another example is anticipated changes in household incomes, labor force participation, and drivers license holdings. In any event, there is no reason to take this modeling risk, as dynamic models can capture the expected non-stationary processes.

Dynamic models can represent asymmetry in choice, while static models cannot (Kitamura, 1987). The factors influencing household decisions to acquire an additional car are logically quite different than those affecting decisions to dispose of a vehicle (Manski and Goldin, 1983). Also, responses to increases in costs (and other factors) are likely to be different from responses to decreases in costs (Goodwin, 1989).

All models are subject to specification errors due to missing explanatory variables and the lumping together of households with different unobserved variable levels. This is known as heterogeneity bias. With static models it is extremely difficult to estimate the extent of these heterogeneity errors and it is almost impossible to correct for them. Dynamic models offer the potential of compensation for heterogeneity errors by taking advantage of repeated measurements on the same households. Models incorporating individual-specific effects are variously known as "error components" or "random effects" models in econometrics. Dynamic discrete choice models of car ownership and usage incorporating such individual-specific effects have been recently developed (Hensher and

Smith, 1989; Kitamura and Bunch, 1990; Meurs, 1989, 1990; van Wissen and Golob, 1990) and should be incorporated into the proposed modeling system.

Finally, the utility maximization principle on which existing static vehicle choice models is based is fundamentally flawed for multi-car households, unless either: (1) households replace all vehicles simultaneously, or (2) there are no changes in the explanatory (observed and missing) variables over time; both of these conditions are highly unlikely events. The static principle is that households optimize their vehicle fleet. Dynamic models can be based on a less restrictive and more logical principle that households optimize the decision to replace (or acquire or dispose of) any vehicle conditional on the characteristics of the other vehicles in their fleet. This is a far more realistic treatment of the process of change in vehicle holdings that can be incorporated easily in a dynamic utility maximizing model. It is not necessary to abandon utility maximization, only the simplifications used to fit models to static data need be eliminated.

The proposed dynamic model is similar to static models, such as the present CEC PVM, in its specification of jointly dependent variables that represent discrete vehicle choice and continuous vehicle usage levels. In a dynamic model, however, these mutually dependent variables are additional functions of the same variables measured at a previous point in time. This incorporates state dependence into the model, allowing for the explicit modeling of the transaction decision as to whether or not to change vehicle holdings. It also allows time lags and anticipated effects from explanatory variables (e.g., decisions to replace or acquire might be a lagged function of an increase in income). Moreover, usage and the number and type of vehicles can be dynamically related; e.g., an increase in usage on an existing number of vehicles (due to, say, an increase in the number of workers in the household or a longer commute) might lead to an increase in the number of vehicles later in time (a non-instantaneous response to "need").

Dynamic transaction models offer at least two significant advantages in forecasting future vehicle ownership as well. First, a dynamic transaction model is capable of predicting the frequency and nature of vehicle transactions, and therefore is capable of forecasting the turnover of the vehicle fleet in the region. This capability is crucial in forecasting fuel consumption and air quality impact. Second, use of transaction models allows more flexible and detailed scenario formulation, extending the scope of forecasting. For example, forecasting with a transaction model would distinguish the effect of a gradual income increase of 20% over a 10-year period, and the effect of a stagnant economy for the first five years (0% increase) then a rapid growth (20% increase) over the second five years. This ability, in turn, facilitates the analysis of a wider range of policy options. For example, various pricing policies for clean vehicles and fuels can be examined in connection with their market penetration, which can expand over time. A dynamic model is capable of representing the time dimension accurately; therefore, such a model can be used to examine policies that are implemented in stages, as is often the case in reality.

b. Explanatory Variables

The model must be sensitive to those variables that are either policy sensitive or potentially important predictors of future personal vehicle fuel usage. These include:

- fuel prices
- vehicle prices (new and used)
- disposable incomes (including gross incomes, tax rates, and cost of living)
- attributes of available vehicles (including fuel economy levels, fuel type, seating capacity, performance, price for new and used models)
- work trip length (spatial distributions of residences and jobs)
- congestion levels (travel speeds)

- labor force participation (including full/part-time employment by position in the household)
- flexible work hours and telecommuting
- household socio-demographics (including household size by age group, single-parent households);
- ridesharing incentives and transit/carpool service levels (including carpool lane availability, transit and parking costs)
- characteristics of the existing household fleet, including which household member is the primary driver of each vehicle.

In addition to these variables, the explanatory power of the model is expected to be influenced significantly by its ability to synthesize several complex issues. Specifically, the age of the vehicles owned by a household is anticipated to be one of the principal determinants of vehicle transaction behavior. If the model is to be applicable to non-conventional vehicles (e.g., clean fuel and ultra clean fuel vehicles), it is required that it be sensitive to market penetration and fuel distribution networks. This will require information on the number of fueling stations near the individuals' home and work as well as measures of general availability of alternative fuel throughout the state.

Another important determinant of vehicle holdings is household formation and dissolution (Manski and Sherman, 1981; Town, 1983). Although this process will be modeled separately, it is important that we collect enough information about effective ownership and control of vehicles in multi-vehicle households to effectively track the effects of household change.

c. *Model Specification*

A principal advantage of disaggregate models is that the modeling unit is also the elementary decision making unit. In the proposed model, as in the present CEC PVM, the most appropriate modeling unit is a household (for obvious economic and social reasons). However, the effects of actions taken by a single elementary unit have negligible effect on the system in which it operates. For example, the equilibrium price of a specific type of vehicle is not effected by the single transaction performed by a single household. In a similar way the congestion effects of a single car traveling on a highway are negligible. This fact simplifies tremendously the estimation of disaggregate demand functions. In the proposed model, when estimating demand functions variables, such as vehicle and fuel prices, travel times and congestion levels can be treated as exogenous variables. This is not the case when estimating aggregate demand functions where the dependent variable is, for example, the total demand for vehicles of a specific type. When total demand is estimated, vehicle prices can no longer be considered as exogenous variables. In such a case a complete system of simultaneous equations which define both demand and supply must be defined and estimated in order to obtain consistent estimators.

The recommended model structure can be broken down into five interrelated (non-hierarchical) components:

- Vehicle transaction choice, given vehicle utilization, and the fuel type, and make/model/vintage composition for the household fleet.

- Fuel type choice, given a transaction (acquire or replace) and utilization.

- Vehicle type (make/model/vintage) choice, given fuel type, and utilization.

- Vehicle utilization, given fuel type, and make/model/vintage composition for the household fleet.

- Vehicle and fuel supply by type, given demand.

Vehicle transaction behavior can be described as a function of the characteristics of the vehicles in the household fleet, attributes of commuting trips, attributes of available vehicles (new and used), vehicle and fuel prices, taxation, and other cost elements. Also important are changes in employment, license holdings, and other pertinent household attributes. It is required that the model be able to predict all types of transactions that could be realized: acquire an additional (new or used) car, replace a car currently owned, dispose of a car that is currently owned, and preserve the present fleet (no transaction). Because new cars are produced with annual cycles, it is logical to formulate a model to predict vehicle transactions within the same annual intervals. Household vehicle holdings (number of vehicles owned) at the beginning of any particular year is then determined once vehicle transactions are determined.

A theoretical foundation of such a transaction model must be developed, based on recent research developments published in the literature, as extended and focused to CEC needs in a proposed future study. In particular, the specification of vehicle ownership costs needs to be carefully developed. Prorated annual vehicle ownership costs vary depending on how long the vehicle is kept. Therefore, vehicle ownership costs are not entirely exogenous, but are partially determined by the household's holding duration choice. The endogeneity of vehicle ownership costs is not a problem unique to transaction models, but universally applies to all car ownership and vehicle type choice models. It, however, has not been addressed in the literature.

A critical element in formulating such a dynamic transaction model is the representation of heterogeneity (unobserved effects that vary across households, but remain stable over time for each household; see Heckman, 1981). As noted earlier, cross-sectional analysis is incapable of properly accounting for heterogeneity, possibly leading to erroneous model formulations and forecasts. Recently proposed generalized beta-logistic models (Hensher and Wrigley, 1986; Smith et al., 1989) and error-component ordered-response probit models (Kitamura and Bunch, 1990) offer a potential methodology to formulate a vehicle transaction model that accounts for both heterogeneity and dynamic effects (e.g., dependence on the past behavior). Use of general multinomial probit models (Daganzo, 1979) should also be examined. While earlier classes of multinomial probit models severely limit the number of tractable alternatives, new estimation procedures and algorithms proposed recently (McFadden, 1989; Bunch and Kitamura, 1990) may facilitate the estimation of a model involving multiple time points and multiple alternatives with temporally correlated errors. Bootstrapping techniques (Efron, 1982) should be used to get consistent standard error estimates for model parameters and forecasts.

At the conceptual level, "fuel type" choice and "vehicle type" choice can be integrated as a single model. However, separate formulation is recommended here because (at least in the initial stages of model development) the former will be based on stated-preference (SP) data while the latter is to be constructed on available revealed-preference (RP) data. Estimating a single model jointly based on these two types of data is probably infeasible because RP data are unlikely to include alternative-fuel vehicles as alternatives. If RP data which include alternative-fuel vehicles become available, then these data can be combined with SP data for more efficient model estimation (Ben-Akiva and Morikawa, 1989). Once alternative-fueled vehicles are available in the market, then a special choice-based sample of alternative-vehicle purchasers should be added to the proposed panel study. Reliable techniques are available for analyzing merged random and choice-based cross-sectional samples (Manski and McFadden 1981), and these techniques have been recently extended to panel data (Lancaster and Imbens, 1988; Wurzel, 1988).

Existing "vehicle type" choice models, such as the one in the current CEC model system, can be combined with the vehicle transaction model system. A dynamic model system estimated with a longitudinal data set has the clear advantage that the effects of brand loyalty, transaction costs, etc., can be adequately evaluated and incorporated into forecasting while properly accounting for their endogeneity. The presentation of costs contains the same problem as discussed above. It is recommended that the vehicle type choice model be formulated such that it will reflect the endogenous nature of vehicle ownership costs.

Similar to the "vehicle type" choice model component, existing vehicle utilization models can be combined with the proposed model system. Again, estimation of a dynamic utilization model with longitudinal data with appropriate estimation procedures will facilitate the identification of endogenous effects, such as the effect of the utilization in the previous period (see Meurs, 1989, 1990; Hensher and Smith, 1990). Roy's identity, (Clarke et.al., 1982) often used as the theoretical link between a vehicle type choice model and a utilization model, must be reformulated in a dynamic context if such a link is desired; an initial effort can be found in Hensher et al. (1987).

Close attention must be given to the relationships among these model components. It is strongly advised that a causal analysis be performed to determine the most plausible structure that ties the components together in a dynamic context. Statistical methods suitable for such effort exist. For example, structural equations models have been applied recently to probe mutual causal relationships among changes over time in car ownership, trip generation, mode use, and income and other household attributes (Golob, 1987, 1989;

Golob and van Wissen, 1989). The above four model components representing vehicle transaction, fuel type, vehicle type, and utilization, must be linked to form a dynamic structure that is consistent with the results of such modeling efforts.

C. The Data

The data required can be collected efficiently by a panel survey of a random sample of California households. The minimum sample size, which would be accurately determined as a first step in implementation, is on the order of 2,500 households; sample refreshment would be required at all waves of the panel to maintain the original size by replacing households that drop out during the course of the survey. Panel data are more cost effective in measuring population changes than cross-sectional surveys (e.g., Cochran, 1977).

Biannual panel waves would be most effective for the first two or three waves (12 to 18 months), with subsequent conversion to annual waves. The sample recruitment and the survey itself can be most effectively conducted by telephone, possibly with mail follow-up contacts. It is important to follow all individuals as they leave sample households and establish new households, as is currently done in the University of Michigan PSID survey (Panel Survey of Income Dynamics, described in Survey Research Center, 1972 and Duncan et al., 1987).

The first wave of the panel would include retrospective questions, allowing model estimation to begin immediately upon completion of the first wave. The model estimation and forecasts would be continually updated as later waves become available. Forecasting could commence after model estimation with two full waves (plus retrospective data), but accuracies would continue to improve as long as the panel is maintained. The first three to four waves are most important for the original development and testing. Beyond this point, the major benefit in continuing the panel would be in a continuous updating of parameter estimates and forecasts in response to general trends, and for the ongoing validation of forecasts.

a. Survey Method and Sampling

A telephone survey is recommended, with possible mail follow-ups to track panel households over time. Telephone surveying has become the method of choice in most fields in political and social science, as well as in marketing research. There is compelling evidence to indicate that the data needed to support CEC PVM needs can be collected most effectively by telephone survey using the recommended modified random digit dialing sampling scheme and well trained interviewers and supervisors (Bishop, et al., 1988; Groves and Kahn, 1979; de Leeuw and van der Zouwen, 1988).

It is strongly recommended that the panel be a random sample of all California households, including households currently without vehicles. The coverage should be statewide. Sample selection can be accomplished by implementing a Mitofsky-Waksburg telephone survey design (Waksburg, 1978), which is a two-stage random digit dialing design. This is similar to pure random digit dialing in maintaining an equal probability of residential selection, but at lower costs. The cost savings come from an efficient elimination of commercial and non-working numbers (Burkheimer and Levinsohn, 1988).

The result is a sample of households with telephones clustered by telephone central office code (the first six digits of the ten digit telephone number). The parameters of the sampling frame (the number of clusters and their size, and the number of callbacks for non-answers) should be established at the time of survey design. Analytical procedures are available to minimize costs while satisfying variance requirements on key variables (e.g., Mason and Immerman, 1988). The final sample size should be adjusted based on the detailed sample design process and statistics on car ownership and transaction frequencies accumulated during the interviewing process.

b. Survey Content

The following information would be collected in the first wave of the panel survey:

- Household characteristics
    - tenure (own/rent)
    - socio-demographics
    - income
    - labor force participation

- Car ownership
    - make/model/vintage for each vehicle
    - who owns and maintains each vehicle
    - primary driver, usual driving patterns

- Car transactions
    - date of purchase for each vehicle
    - whether each vehicle purchased new or used
    - primary reasons for each transaction (e.g., acquired for new driver or worker, replacement of car at set period, old vehicle unreliable, new model desired, more seating space needed)

- Car usage
  - estimated miles over last year for each vehicle
  - odometer readings for each vehicle
  - usual commuting uses of each vehicle

- Commuting behavior
  - work locations, distance and time of travel for each worker
  - work schedules for each worker (including flexible schedules and telecommuting for any worker)
  - ridesharing behavior of any worker
  - availability of flexible schedules, ridesharing, etc. and parking costs at workplace

- Housing characteristics
  - residential location
  - housing type
  - length of time at current residence
  - anticipated subsequent tenure

- Sampling control
  - Number of household telephone lines (including modem and FAX)

Follow-up waves would use CATI (Computer Aided Telephone Interview) techniques so that households can be reminded of answers from previous waves and thus reduce measurement errors. Each successive wave of the panel following the initial one would then collect:

- Changes in household characteristics since last wave
  - demographics
  - income
  - labor force participation
  - expected changes in household characteristics

- Car transactions since last wave
  - date of each acquisition or disposal
  - acquisitions: make/model/vintage, new/used, odometer reading
  - disposals: identification of which household vehicle, how disposed of
  - primary reasons for each transaction (e.g., acquired for new driver or worker, replacement of car at set period, new model desired, old vehicle unreliable, more seating space needed, fuel, maintenance, insurance too costly)

- Car usage
  - odometer readings for each vehicle
  - any changes in usual commuting uses of each vehicle

- Changes in commuting behavior
  - work locations, distance and time of travel for each worker
  - work schedules for each worker (including flexible schedules and telecommuting for any worker)
  - ridesharing behavior of any worker
  - changes in availability of work schedule, ridesharing options and costs

- Changes in housing characteristics
    - residential location
    - housing type
    - length of time at current residence

c. *Survey Biases*

Survey biases can be divided into four types: (1) sample biases because certain households are excluded from the sample, (2) non-response bias due to refusal to participate in the panel, (3) attrition bias resulting from nonparticipation in later waves, and (4) non-random missing and inaccurate data. Techniques exist for minimizing each of these biases.

Sample biases arise in the recommended survey approach from three sources: First, only households with telephones are included. While there is a relatively small number of households without telephones (probably less than 5 percent in California), certain minority groups have disproportionally high proportions. Fortunately, this can be controlled for in the weighting factors used to expand model estimates and forecasts to population totals, using external knowledge of telephone coverage by population subgroup and geographical area (e.g., Thornberry and Massey, 1988; Groves, 1987). The 1990 census is particularly timely in this regard. Second, some households have more than one phone line, affecting their probability of being selected. Again this can be taken into account in both recruitment and in weighting results, by using both external and internal data (see survey content, below). Finally, there will be no response at some households. This is minimized by making many callbacks to no-answer numbers at various times, and data are available on which to select the optimal number of callbacks (e.g., Sebold, 1988; Massey, et al., 1981).

Non-response is a problem to the extent that participating households are different from nonparticipating households. Information can be gained on this by asking questions on variables anticipated to be related to nonresponse at the beginning of the survey, so that some data will be recorded if a household prematurely terminates an interview. Also, the sample must be compared to external population data (utilizing the 1990 Census when it becomes available), and these comparisons are valid for total nonresponse as well as partial nonresponse. The comparisons are used to develop weighting factors. There is substantial evidence to guide nonresponse considerations in sample design and weighting (e.g., Groves and Lyberg, 1988), and experienced professionals need to participate in the wording of the survey instrument and in the training of interviewers to maximize recruitment. If possible, respondents should be paid (or given a lottery ticket) for their participation, particularly in the first few waves.

Attrition bias is unique to panel surveys (compared to cross-sectional surveys) and can be controlled for by modeling the probability of dropping out as a function of key model variables (Kitamura and Bovy, 1987). It is also possible to include compensation for attrition bias directly in simultaneous equation model systems with pooled samples (and three or more panel waves) by including variables measuring the number of panel waves in which each household has participated (Golob, 1989).

Finally, both attrition bias and missing data can be handled using multiple imputation techniques (Rubin, 1987). These techniques involve creating multiple imputations for each missing data point (or observation in the case of attrition bias). Since panel data generate a wealth of data on each respondent, it is generally possible to use these data to build very accurate imputation models for panel studies. Little and Su (1989) and Brownstone (1990) show that these techniques can greatly reduce the impact of missing data in household panel studies.

d. Forecasting

The process of demand forecasting using such a dynamic model system is fundamentally different from the conventional (either aggregate or disaggregate) forecasting procedure using cross-sectional models. In the conventional approach, values of the explanatory variables are first forecast for the horizon year, which are then plugged into the model to yield a forecast. Forecasting, therefore, is a single shot toward a distant future. Various developments that may take place between the base year and the horizon year do not at all influence the forecast.

Alternatively, the proposed dynamic system can be used to simulate the continuous evolution of vehicle holdings from the base year to the horizon year; gradual changes in the size and composition of the vehicle fleet in the study area can be obtained for various scenarios. This stochastic simulation will utilize the sample households in the data set, and it requires forecasts for the explanatory variables for each year in the simulation period. Using demographic models, the base set of sample households is aged, households are dissolved, and new simulated households are created. For each of these households in the simulation base, the vehicle transaction model is used to forecast the progression of vehicle ownership and use. A number of good demographic models for California are available (Center for Study of California Economy and/or UCLA) which can be used to generate the necessary demographic and socioeconomic forecasts. Similar methods have already been developed and applied by Millar et al. (1987) and Kitamura and Goulias (1989). Scenario analysis is carried out by generating forecasts for two or more different values of the explanatory variables, and then comparing these forecasts to the baseline forecast. Bootstrap techniques can be used to generate consistent prediction errors for these forecasts (Brownstone, 1990).

One problem that arises with scenario forecasts is that, without additional constraints, the vehicle transaction model may predict infeasible behavior. For example, a rapid drop in the price (either purchase and/or operating) of alternative-fuel vehicles may lead to a forecast that a large number of households sell their current vehicles and purchase new alternative-fuel vehicles. The predicted number of purchases may exceed reasonable production capacity, and the sudden disposal of a large number of vehicles may significantly affect used car prices. One way to solve these problems is to add a supply model for new cars and build an endogenous used-car market to dynamically forecast car prices (as in Berkovec and Rust, 1985). An alternative approach is to build a simple ad hoc model which changes new and used car prices to limit unreasonably large forecast quantity changes. The prices resulting from these models are also of interest, since policies which imply large shifts in new and/or used car prices over short time periods may be politically infeasible.

Such a forecasting methodology is very powerful. A question remains as to whether such simulation can be reasonably performed within computational resources typically available to policy evaluation. Recent experience indicates such simulation forecasting is unlikely to be a computational burden. For example, 15-year simulation runs with about 1,500 households, with internal demographic, car ownership and mobility simulation, were conveniently performed on a minicomputer and also on a small mainframe (Vax 750). (Kitamura and Goulias, 1989). It is anticipated that a 386-based PC can reasonably handle simulations of this size. It is also anticipated that repeated simulation runs need not be a computational burden with a household sample size that reasonably represents the California population, a sample size of, say, 2,500.

In the forecasting procedure special treatment must be given to variables which are not exogenous to the system when the total demand is considered. In reality, prices of vehicles and fuel types are determined by equilibrium between demand and supply for cars and fuel, and with regard to fuel prices and congestion on highways. In the estimation phase those variables can be treated as exogenous. However, in the forecasting phase the behavior of the total market is simulated, and those variables must be treated as endogenous variables and their value should be determined internally within the forecasting procedure. The practical implication of this approach is that a complete set of simultaneous equations which define supply and demand functions of the vehicle and fuel markets needs to be solved at each simulation point in the forecasting process. As a first approximation of market equilibrium it can be assumed that prices of new cars are determined in an efficient market by marginal production costs and only the prices of used cars are determined internally as a function of demand and supply. Similar simplifying assumptions will be necessary regarding fuel prices, taking into account production constraints only.

In order to implement the proposed dynamic equilibrium forecasting procedure, it will also be necessary to define a process by which cars are scrapped. Such a process can be defined either as a normative model defining a price level at which a vehicle is junked; or preferably, a disaggregate scrappage model can be estimated. To predict and evaluate long term processes in the vehicle market, the equilibrium procedure is a necessary tool. To test short term effects on the market due to various price policies, taxation or other administrative actions the use of demand models only will suffice.

Dynamic vehicle transaction models as well as demographic simulators are novel concepts in transportation demand forecasting. Plausibility and accuracy of the conventional approach to forecasting using cross-sectional models is strongly suspect, however, when the socio-demographic structure of an urban society changes rapidly and planning policies mandate the incorporation of nonexistent vehicles and fuels. Meeting this mandate seems to necessitate an approach that emphasizes changes over time, such as the one recommended in this report.

D. Transit and Freight Energy Demand Forecasting Models

Transit and freight energy demand forecasting models tend to be aggregate fuel consumption models based on regional population and economic characteristics. It is possible to develop more realistic disaggregate models, but the costs of data collection exceed the probable benefits. Since transit currently uses less than 1.5% of surface transportation energy (excluding pipelines), even gross forecast errors will have little effect. Total freight energy use is primarily determined by regional economic activity, which must be forecast by other methods.

Transit energy demand forecasting models tend to first estimate the demand for public transit in terms of number of public transit trips and ridership. This is done by modeling trip generation and car ownership as a function of various socio-demographic and economic characteristics. Trip generation and mode used are modeled by trip purpose and distance traveled. The market shares of public transit are determined and energy demand is estimated for various scenarios of fuel prices, fuel availability and changes in transit fares and operating costs.

The models have been estimated on cross sectional data sets without regard for the dynamic components of travel behavior. Elasticities of fuel consumption, transit use and car ownership are estimated from cross sectional models to forecast future transit ridership and fares. As described in the previous section, this may not be entirely appropriate as this would assume stability in the behavioral relationships over time. Longitudinal studies need to be performed to see how transit use is causally related to vehicle ownership and other socio-demographic and economic characteristics in a dynamic framework. Moreover, transit use may be influenced by changes in the vehicle fleet as alternatively fueled vehicles become more widespread. These changes can also be tracked using longitudinal studies so that the rate at which alternatively fueled vehicles penetrate the market need not be assumed.

Another area in which improvement can be made is with regard to the true cost of transit use. While transit fares and government subsidies paid by the taxpayer provide accurate estimates of the money cost of transit, the true "full" cost to the individual should incorporate the waiting time cost and increased travel time cost. Efforts at modeling the cost of time using wage rates may not be appropriate as the cost of time differs depending on the type of trip being made.

It is well known that land use development patterns have influenced transit ridership widely in the last few decades. Rapid suburbanization and decentralization of economic and commercial activities have seen the decline of transit. However, existing transit ridership estimation methods do not currently incorporate land use models. The causal relationships existing among transit ridership, land use development, residential and job locations and mode choice for different types of trips need to be determined if accurate forecasting models are desired. It is in this area that a potential exists for improvement of the currently available models because recent advances in the causal modeling of simultaneous relationships have made this possible.

Freight energy demand models estimate freight fuel consumption based on demands and supplies of commodities in the economy. These models first estimate the demand and supply of different goods in various areas of the country or region under study. The different shippers and carriers are assumed to operate under equilibrium conditions and the flows of goods on different links of the network are determined to find the volume of freight flow. This volume determines the equilibrium total vehicle miles and the necessary fuel consumption.

Improvements can be made to these models in several respects. First, the models assume specific rates of regional growth and specific changes in demands and supplies of goods. These growth rates are assumed to be exogenous to the model being estimated. However, it is important to consider these rates of regional growth and demand and supply relationships as being influenced by transportation costs and changes in socio-demographic characteristics. Therefore, models should be developed in which these factors are determined and treated endogenous to the system. This can be done only by estimating models on longitudinal data sets.

Although it is reasonable to assume that shippers, receivers, and carriers act to minimize expected costs, current models do not reflect all cost components. One important problem is modeling and measuring reliability, including carriers' on-time performance and fleet reliability as a function of vehicle maintenance and replacement policies. Calibrating such a model would require access to detailed sensitive operational data from carriers and shippers. Although these data have been obtained from individual companies for special studies, it would be extremely difficult to collect for a representative sample of carriers and shippers..

III. SCENARIO ANALYSIS

Scenario analysis consists of comparing two forecasts from the demand model described in the previous section, where each forecast is based on different values of the exogenous variables. Typically one of these forecasts, called the baseline forecast, is designed to be the best forecast of behavior under the status quo while the other forecast is designed to reflect some proposed policy. Thus scenario analysis is subject to the same problems as the forecasting methodology described in the previous section. There are a number of special features of scenario analysis which are described below.

a. Specification of scenarios

The key requirement is that the underlying model be sensitive to the factors defining the possible scenarios. Therefore it is important that the full set of possible scenarios be identified in advance so that the model can be properly specified. It is also important to consider the range of possible scenarios to insure that the model is capable of forecasting accurately over that range. For example, a model based on a flexible translog utility specification will yield good forecasts for exogenous variables in the range of the calibration data set, but may fail to satisfy the basic properties of a utility function for values far away from the calibration data. In this case it would be better to specify a more restricted model, such as a restricted Translog with constant elasticities, which has better global forecasting properties. The more flexible models are useful in this setting as a benchmark to judge the adequacy of more restrictive models.

Another problem is the need to translate policy scenarios into precise forecasted values for all of the exogenous variables in the model. For example, a policy to increase the number of alternative-fuel vehicles must specify the cost and operating characteristics of these vehicles as well as fuel costs. If the policy is expected to work through the free market (e.g. manufacturers are forced to sell a minimum number of vehicles and would set prices accordingly), then it may be desirable for the model to be able to predict what prices will be necessary in order for manufacturers to meet the quota. This can be accomplished by iterating the demand model with different prices until the model predicts just the desired minimum sales level for the particular vehicle type. (Assuming that the quota is "effective"). Brownstone, Englund, and Persson (1988) and Anas (1987) have both demonstrated the computational feasibility of this technique applied to discrete choice models of home ownership. Due to the complexity of this procedure, it is probably worthwhile to produce specialized software to utilize the values of the exogenous variables dictated by the required policy and produce the time path of outputs needed to support the policy goals.

b. Out-of-sample forecasting errors

Scenario analysis frequently involves forecasting values of the exogenous variables far outside the range of values in the original calibration dataset. Since any model is an approximation, forecasts based on values far from their calibration values are subject to larger errors. This problem can be mitigated somewhat if the model is continuously recalibrated from an ongoing panel study, since natural changes in the population will keep the range of the calibration variables closer to current values.

c. Sensitivity to baseline forecast errors

One advantage of scenario analysis is that it is not crucial that the baseline forecast be correct, only that the change between the baseline and scenario forecast correspond to the correct change in the real world. However, since the scenario analysis is only concerned with differences, it is more sensitive to random measurement error in the endogenous variables. This situation is identical to that encountered in time series or panel data models for first differences.

d. Error analysis

There are three types of errors which must be accounted for: parameter estimation error, model uncertainty, and uncertainty in baseline and alternative scenario exogenous variable values. Using traditional mathematical statistical techniques it is only possible to estimate model uncertainty in simple time series models. The nonlinear, dynamic forecasting methods proposed in the previous section require different, more computationally intensive techniques. Two generally useful approaches are parametric bootstrapping (Efron, 1982) and multiple imputations (Rubin, 1987). Both of these methods involve drawing parameter values and/or forecasts from their assumed distributions and then combining the results of a number of iterations of this process to generate estimated errors.

Most forecasting models do not allow for uncertainty in future values, so the only way to represent this type of uncertainty is to run many different types of scenarios across likely ranges of uncertain exogenous values. In many cases these exogenous variables are the forecasts from some independent model. The same bootstrapping and multiple imputation can also be used to incorporate this information into a total forecast error measure.

Since the forecasts from dynamic models are time paths for the forecasted endogenous variables, the forecast errors must be presented as confidence regions for these time paths. Software will have to be designed to conveniently display the results of these error analyses as confidence bands around time paths.

IV. TRAIN'S VEHICLE HOLDINGS, VEHICLE TYPE, AND UTILIZATION MODEL

The model developed by Train and his colleagues is important because it is the first comprehensive system to treat all relevant aspects of household car ownership and use, i.e., number of vehicles, class (defined in terms of body type and make), age (or "vintage"), and utilization (in terms of vehicle-miles traveled). It is one of the first models attempting to determine changes in behavior over time, to contain "transaction" variables (other examples include Manski and Sherman, 1980), and has strongly influenced subsequent model development efforts (e.g., Mannering and Winston, 1985). The model system consists of nested logit models for the number of vehicles and vehicle type choice, linear models of vehicle utilization, and logit models that assign vehicle miles traveled (VMT) to different trip types. The system is microeconomically derived with Roy's identity used to relate the vehicle choice and utilization (the exact utilization formula is replaced by a linear formulation as an approximation; the use of exact formulation can be found in a more recent study by de Jong (1989a).

The "vehicle quantity" submodel predicts the number of vehicles a household holds. The alternatives are limited to:

- no car
- one car
- exactly two cars

The explanatory variables include

- household income
- number of workers in the household
- number of household members

In addition, the annual average number of transit trips per capita in the area of residence is used to represent the transit service level. Although this vehicle quantity submodel is not formulated as a dynamic model, dynamic quality enters into it indirectly through the constructed by-product variables representing the "average utility in class/vintage choice," or so-called "inclusive value" or "log-sum" variables representing the expected maximum utility of a car (or a pair of cars) chosen. These inclusive value variables are derived from the vehicle "class/vintage" choice submodels (which are pseudo-dynamic as described below) and reflect the prices and other attributes of vehicles available in the market.

The class/vintage submodels, (defined separately for one-vehicle households and two-vehicle households), describe the probability that a particular class-vintage combination will be chosen by a household. Train's system assumes 12 classes defined in terms of body type and make (subsubcompact domestic, subcompact domestic, etc.) and 10 vintage categories (pre-1970, and the years 1970 through 1978), resulting in 120 class/vintage combinations as the alternatives of choice by one-vehicle households. Each combination is characterized in terms of

- typical purchase price
- typical operating costs
- shoulder room
- luggage space
- horsepower
- several dummy variables

These typical characteristics were obtained by calculating their means over all makes and models in each class and vintage. Also included is the logarithm of "the number of makes and models in the class/vintage" to represent the availability of close substitutes in the form of makes and models in the class/vintage".

The "transaction cost dummy" in Train's model system "takes the value of one for vehicles in a class/vintage that the household did not own in the previous year and zero for a class/vintage of the vehicle that the household did own in the previous year." Therefore, "Assuming that no household sells a vehicle within a class/vintage and buys another one in the same class/vintage, this variable represents the psychic, search, and other transaction costs associated with buying a new vehicle" (Train, 1986, p. 155). Being defined according to the class/vintage of the previous year, this variable is implicitly a lagged, endogenous variable, making the model dynamic. The endogeneity of the variable, however, creates "econometric questions concerning the consistency and efficiency" of model estimation (op. cit., p. 240).

Train notes the importance in his model of the logarithm of the number of makes and models in the class/vintage, and the importance of the variance of the front and rear shoulder room across the makes and models within the class/vintage. After estimating a model formulated without these two variables, Train notes considerable changes in the estimated coefficient values. Clearly the size of a class/vintage combination and the variation in vehicle attributes across makes and models are important considerations in vehicle type choice modeling.

Two-vehicle households are assumed to choose from among pairs of class/vintage combinations. This, however, is a very unrealistic assumption; in most cases a household replaces one of its vehicles at a time (or over a relatively short period of, say, one year).

There are two likely reasons for this:

1. A household's decision process is so structured as to update its vehicle fleet by replacing one vehicle at a time, possibly because of the decision-making complexity of simultaneous replacement of the entire fleet (the number of alternatives in the two-vehicle class/vintage choice in Train's system can be as large as 14,400 ($=120^2$)).

2. The transaction cost involved in replacing the entire household fleet is so excessive as to make replacement of the entire fleet practically impossible.

From the first viewpoint, a vehicle is chosen given the set of vehicles the household continued to hold. The choice, then, is a conditional optimization, which may not lead to the same household fleet as would be chosen if the entire fleet were replaced simultaneously. The former may present an optimal solution, which is suboptimal in the latter unconditional, simultaneous choice problem. For example, consider a two-vehicle household which last purchased a vehicle several years ago. Suppose gasoline prices have increased sharply since the last purchase. The fleet the household owns is no longer optimal. However, the household does not plan to replace both vehicles at the same time, but plans to replace the oldest vehicle it owns. The household's fleet, with one vehicle replaced, may then be suboptimal and inferior to the one it would own had it replaced the entire fleet. Similar situations are conceivable after changes in household characteristics such as .

- income
- number of drivers
- number of workers
- number of children.

One could argue that the transaction variable, included also in the two-vehicle class/vintage choice submodel, would account for this apparent discrepancy (the transaction variable is now defined as the number of transactions involved between the previous year's pair and the alternative class/vintage pair). The above suboptimal fleet may in fact be optimal due to the large transaction cost involved in replacing the entire fleet. Unfortunately, the representation of transaction costs is rather problematic in Train's model system with the consequence that it may predict unrealistic purchase behavior. This is illustrated using the following examples:

> Consider a household with a 1978 annual income of less than \$12,000. This household has two pre-1972 vehicles whose "purchase prices" would be \$1,200 and \$800. The contribution of the purchase price to the utility is -0.000531 x (1200 + 800) = -1.062, where -0.000531 is the coefficient of the "purchase price of both vehicles, summed in dollars" (op. cit., Table 8.4). Suppose this household intends to replace one of the cars (valued at \$800) with a vintage 1976-1978 vehicle, whose characteristics are exactly the same as the one being replaced, except for the purchase price and vintage. Suppose the purchase price is \$3,000. In Train's model system, a household's transaction decision is viewed as a choice from among a set of vehicle pairs; one of the pairs coincides with the pair the household currently owns, and all the other pairs have appropriately defined transaction costs assigned to them. Then the difference in the utility value between the new pair and the old pair is

$$[-0.000531 \times (3000 + 1200) - 4.48 + 0.155]$$
$$-[-0.000531 \times (1200 + 800)] = -5.49$$

where -4.48 is the transaction cost coefficient and 0.155 is added due to the vintage 1976-1978 car. The decline in utility is phenomenal. The difference, -5.49, is equivalent to a purchase price of $10,340 (= 5.49/0.000531) for this household. In other words, the expected utility of the new vehicle pair would not be greater than that of the pre-1972 vehicles the household currently holds, *unless* the purchase price of the new vehicle is -$7,340 (= $3,000 - $12,340), or the household receives $7,340 for acquiring the new car!

Likely improvements in vehicle operating costs (reliability and/or reduced maintenance costs) offered by the new vehicle would offset the reduction in utility only partially (the operating cost variable in the model system apparently represents fuel costs only). Suppose that gasoline is 80 cents per gallon, the pre-1972 vehicles get 10 miles per gallon, and the vintage 1976-1978 car is twice as fuel efficient with 20 miles per gallon. The unit operating cost is 8 cents per mile for the older vehicles and 4 cents per mile for the new vehicle. Therefore, the utility difference between the vehicle pairs is

$$[-0.000531 \times (3000 + 1200) - 4.48 + 0.155 - 0.441\,(4 + 8)]$$
$$- [-0.000531 \times (1200 + 800) - 0.441(8 + 8)] = -3.73$$

where -0.441 is the coefficient of operating cost of both vehicles summed in cents per mile. The difference in expected utility is equivalent to a purchase price of $7,020.

These rather unrealistic results are partly due to the "transaction" coefficient of 4.48. This value implies that, for this household, the cost per transaction is equivalent to 4.48/0.000531, or the purchase price of $8,440! In 1978, this was perhaps greater than the prices of many brand-new cars that were in the market. Using the purchase price coefficient of 0.000383 for households with incomes between $12,000 and $20,000, the cost per transaction is equivalent to the purchase price of $11,700, and for households in the highest income category (greater than $20,000), the transaction cost is estimated at $26,150 per transaction! (The last figure, however, involves a large standard error.) Using the estimated coefficients of the class/vintage submodel for one-vehicle households, transaction costs are estimated at $9,600 for households with incomes less than or equal to $12,000 , and $12,800 for those with incomes exceeding $12,000.

Underlying these results is the fact that, although a transaction variable is included in the utility function, the class/vintage submodel is not a transaction model and does not necessarily capture households' transaction behavior properly. The class/vintage submodels attempt to capture both the transaction decision and the vehicle type decision within one specification. The model's emphasis is clearly on the latter decision and the former is treated only superficially. For example, information about the current household vehicle fleet (or, the one from the last period) does not properly enter the model, despite the fact that the remaining life of the vehicle and how long it has been held by the household are likely to be the determinants of vehicle transaction decision.

This superficial treatment possibly leads to another problem, that is, that brand loyalty or switching between body types may not be properly represented by the class/vintage submodels; the transaction dummy signifies the particular class/vintage combinations that are of the same body type or make (domestic vs. foreign). Vehicle transaction behavior as depicted by these submodels is essentially memoryless, except that the current class/vintage combination has a much higher chance of being chosen again (no transaction) because of the heavy penalty for transactions in the model. Although these submodels may replicate the frequency of transactions, they do not explain types of transactions as a feature of transaction behavior. Their validity in forecasting is therefore questionable.

The working of Train's model system can be visualized as follows. (It is not clear how the class/vintage submodels apply when the number of vehicles held by a household changes, and the following discussions assume that the vehicle quantity does not change). At the end of each time interval, say a year, every household considers replacing its fleet, and therefore, compares the utility of the fleet it currently holds against those of available alternative class/vintage combinations. If the household chooses not to transact, the utility difference is zero (assuming no depreciation for simplicity). The purchase price variable is not differentiated between the fleet held by the household and an alternative fleet. This is convenient when utility differences are being evaluated; the resale price of vehicles currently held acts as their "purchase price", (the linear utility formulation in the model implicitly assumes that the marginal utility of the net profit from a transaction is constant and is identical to the marginal cost of purchase. It is also assumed that the resale price of a given class/vintage vehicle equals its purchase price).

This depiction without distinction between a vehicle held (which is an asset) and a vehicle to be purchased (as an expenditure item), however, produces somewhat counterintuitive behavior, the validity of which does not appear to have been tested. To illustrate the point, consider the following utility difference:

(Utility of alternative fleet) - (Utility of current fleet)

= b_1(purchase price of alternative fleet) + ...

+ b_t(number of transactions needed) + ...

- [b_1(purchase price of current fleet) + ...]

Because b_1 is negative, this utility difference increases as the purchase price of the currently held vehicles increases. The difference is positive when a net monetary gain is achieved by "trading down" the vehicles a household holds; therefore the incentive to trade down increases as the value of the household's fleet increases. According to the model, then, the more expensive are the vehicles the household holds, the more attractive other fleets become, *ceteris paribus*. This implies that the probability of transaction, or trading down in this case, increases as the purchase price of the currently held vehicles increases (holding income constant); households with more expensive cars are more likely to trade them down!

These counterintuitive examples appear to have their roots in the difficulty of representing household vehicle ownership behavior as a series of choices from among a choice set of the currently held fleet and alternative fleets. Had the class/vintage submodels been specified without the transaction dummy variable, they would have depicted household vehicle ownership as independent multinomial choices repeated over time. A choice of a fleet in one period would be conditionally independent of the fleet the household owned in the previous period, given the explanatory variable values. Consequently the household fleet would tend to change randomly from period to period. This would not be realistic. The transaction variable avoids this depiction of vehicle ownership behavior as a highly transient process by anchoring a household's fleet to what it has. This, however, has resulted in rather questionable estimates of transaction costs as pointed out above.

More logical, behaviorally consistent formulations are possible by modeling vehicle ownership decisions as a transaction process. Indeed Train's class/vintage submodels would work well if they were applied conditionally, given that the household engages in a transaction (therefore the current class/vintage combination does *not* enter the choice set and the transaction dummy is no longer needed). The difficulty of incorporating the ownership of more than two vehicles would be no longer overwhelming. Differentiating among (i) cost of vehicle acquisition, (ii) fixed costs of vehicle ownership, and (iii) variable costs of maintenance and operation, would become a more amiable task in such a transaction model.

During the years after Train's novel model was introduced, advances were made in both data collection and model estimation that enable the development of dynamic model systems. Statistical hurdles in the formulation of a dynamic transaction model system of household vehicle ownership appear to have been largely removed.

V. REVIEW OF FORECASTING MODELS

The next four sections describe and review various energy demand forecasting models that are applied in the transportation sector. Energy demand forecasting models are used for estimating future fuel consumption in primarily three areas of the transportation sector. These are passenger travel energy demand, transit energy demand and freight energy demand. The passenger travel energy demand models can be further subdivided into aggregate and disaggregate level forecasting models. The transit and freight energy use models tend to be aggregate in nature where demand and supply equilibrium relations are used to estimate flows of goods and people and the aggregate fuel consumption.

The passenger travel fuel demand models, at the aggregate level, often follow the same methodology as that applied in transit and freight models. However, at the disaggregate level, models of car ownership and fuel consumption are estimated at the person or household level. The next two sections review these aggregate and disaggregate forecasting models to see how the state-of-the-art in forecasting can be advanced so as to obtain more reliable and robust predictions. These sections are followed by two more sections reviewing energy demand models for transit use and freight transport.

A. Disaggregate Forecasting Models

In this section, forecasting models that are calibrated at the disaggregate level are reviewed. In contrast to the aggregate models, the disaggregate models use household or person-based information as input and provide output at an equally disaggregate level. Thus most of the variation in the population is retained when applying these techniques. However, the application of these models requires the collection of data through household-based surveys, which can involve considerable time and expense. Most of the models attempt to predict car ownership, car scrappage, and car type choice using the multinomial logit approach. These models hold promise in the area of forecasting and could be made more valuable by incorporating dynamic aspects of travel behavior and by testing the validity of the independence of irrelevant alternatives (IIA) assumption that is inherent in most of them.

CARO Model

CARO (Mogridge, 1989) is an accounting procedure that depicts the evolution of a "car park" while replicating the supply-demand interface in the car market. This model was originally developed for the U.K. and was re-calibrated for the Netherlands. The model consists of four components:

1) Supply Model - Provides an estimate of the number of cars of each age and price in the car park as a function of the depreciation rate in average car prices on a yearly basis, the distribution of car prices by age, and the distribution of the "prices" of cars scrapped.

2) Demand Model - Allocates the car park, as defined by the supply model, to households defined by their income and car ownership levels.

3) Interlock Mechanism - Distributes the cars in the supply model to households in the demand model through the operation of a bias function.

4) Equilibrium Process - Balances supply and demand.

The demand component is a heuristic accounting system. It consists of a relationship expressing the proportion of cars of a given age held by a household in a given income group as a function of:

- the median age of the cars in the park
- median income level
- the proportion of the total car park in a given income group
- the proportion of the total car park of a given age
- "Bates" bias constant

Given the distribution of cars by age in each income group, the distribution of cars by price is then determined assuming cars of a given age are chosen randomly from the price distribution for that age. The ratios of the mean incomes of the distributions of income by car ownership level, and the ratios of the spread values are held constant in the model.

The CARO model is driven by a model of new car purchases. This is motivated by the observation that as incomes increase, expenditure on car purchase increases. All such expenditure eventually ends up as expenditure on new cars. But new cars represent only one-quarter of total stock value, so we would expect the total "car park" value to rise at one-quarter the income increase rate if prices do not change. The park-income relation thus captures the growth in park accurately, in conditions of steady growth. Thus the relationship between car purchase expenditure and disposable income is the key relation in this model.

This depiction is based on the assumption that the proportion of the total household expenditure spent on car transport by car-owning households remains constant to a first approximation. Therefore, an increase in expenditure on fuel is matched by a decrease in expenditure on car purchase. If gasoline prices increase, households are assumed to travel less, use the car less, purchase more fuel efficient cars, etc., but continue to expend the same amount on car transport. Households are thus assumed to be budget maintainers rather than rational actors. Given estimated total expenditure for car purchase for each year, it can then be translated using estimates of average new car prices, into a total number of new cars purchased each year.

CARO is an equilibrating supply-demand model that identifies a strong time-trend component. Exogenous factors entering into the equilibrium process are:

- the period between purchase and resale
- the average price of cars by age

Likewise, inputs to the demand portion of the model include:

- household income distribution by car ownership
- proportion of households by level of car ownership
- price/expenditure ratio per car at a given income level

Inputs to the supply portion of the model include:

- new car growth rate
- distribution of car price by age

The output from the two previous components feeds into the demand/supply "reconciliation" component, whose output, in turn, is:

- car ownership by income group
- price distribution of cars owned at a given level of household income
- age distribution of cars at each income level
- level of car ownership

The equilibration, or reconciliation, process is rather convoluted, and its underlying optimization principles and mechanisms are not evident. Although demand and supply are matched in the process, it is doubtful whether its outcome represents an "equilibrium" in the economic sense. In particular, the internal mechanism through which the distribution of car prices is determined is at best opaque.

The main deficiency of the model is that it is unable to reflect ongoing demographic trends, most notably the increasing driver's license ownership, especially among younger individuals. In its current form, CARO does not incorporate any network level-of-service or accessibility measures, and does not reflect possible effects on car ownership of highway investment, traffic congestion, public transit service levels, and urbanization.

A strength of CARO is its sensitivity to income changes. However, the hypothesis of constancy, or proportionality, in car expenditure needs to be statistically tested using longitudinal observation, and an empirically supported model must be developed to apportion the total expenditure to car purchase and car use. Note that the latter concerns the model's sensitivity to operation costs, car prices, and taxation and other policy schemes that are reflected in the costs of car ownership and operation.

CARO utilizes a large number of parameters that are estimated or adjusted to fit the data. Running the model involves iterations over which parameter values are adjusted by hand to attain stability (or equilibrium). Thus the final results reflect investigator judgement and may not be robust across different investigators.

Van den Broecke Cohort Model

Central to the car ownership forecasting procedure presented by Van den Broecke (1988) is the concept of cohort, i.e., a group of individuals who share some attribute, such as year of birth, in common. In this model, total population is disaggregated according to:

- Birth year (20 five-year periods)
- Sex (2)
- Education (4)
- Partnership status (2)
- Employment status (2)

Each variable has two to twenty categories (indicated by the number in parentheses). This leads to a total of 640 combinations, or a five-dimensional matrix with 640 cells. The model forecasts the population size for each cell for a horizon year, and at the same time estimates the probability that an individual in each cell will hold a driver's license or will own a car in the horizon year. A car ownership forecast can be immediately obtained based on these two factors.

Forecasted driver's license holding is based on holding rates observed in 1985 for age cohorts by education. The forecasting procedure appears to rest on the premise that an individual in a younger cohort who acquires a driver's license will usually do so within a few years after having become 18, that older individuals who haven't acquired a driver's license will probably never do so, and that cohorts of older individuals gradually diminish.

Another variable, income, is introduced to determine the level of car ownership. The average personal income for each matrix cell in the 1980 matrix could be derived from available detailed statistics from several sources. From statistics on the distribution of income changes for population subgroups in the period 1980-1985 the income levels for the cells of the matrix for the base year 1985 could be calculated from the 1980 matrix and these changes. These 1980 cell-average incomes, and calculated 1985 average incomes, are then used to estimate an income elasticity function which relates the change in cell income to the change in cell car ownership. The estimation is also based on cell car ownership probabilities which are calculated based on survey results and other data for 1980, and estimated for 1985. These car ownership probabilities are based on inadequate data bases. The saturation level of car ownership within each cell, another variable, is assumed to be identical to the saturation level for driver's license holding. The final model incorporates the actual characteristics of the total car park as boundary conditions.

The model is capable of accounting for the impact of changes in age distribution and takes into consideration the difference in license holdings and car ownership across population cohorts. Because the model incorporates sex, employment, and partnership status, it can also account for the effect of increasing labor force participation by women, or increasing single-person households. In addition, the model has been extended to make car ownership forecasts sensitive to income changes. If car ownership at the national level is determined primarily by the size of the driver population, then this cohort approach should provide reasonably accurate forecasts.

The procedure adopted in this model rests on several ad hoc assumptions and relationships. For example, the equation relating the cell car ownership probability between the start and end of a five year period's income growth appears purely heuristic. The assumption that the saturation level of cell car ownership is identical to the saturation level of driver's license holdings also appears ad hoc and the license saturation level lacks formal justification.

How future license holding probabilities are inferred for each cohort is not documented and no formal statistical procedure is apparent to distinguish cohort, age and period effects. The model was used to predict 1975 car ownership and the results were compared with observed data. The model parameters were then re-estimated, implying that the model was unable to replicate the observed 1975 car ownership.

There is a lack of clarity regarding the treatment of the effects of two correlated factors -- income and education. Also, licensing levels may vary depending on the degree of urbanization, which is not incorporated in the model.

Another deficiency of the model is that the individual, not the household, is the basic analytical unit. This is incompatible with the notion that the household is a unit where the pooling of resources and allocation of tasks take place. Very probably, the acquisition or disposal of an automobile is a household decision, and automobiles are often shared by driving members of a household. This raises a fundamental question as to whether the probability of car ownership can be assigned to the individual irrespective of the characteristics of the household to which he or she belongs.

The HCG Model

In the HCG models (Geinzer and Daly, 1981; Bradley, 1983; Smyth and Daly, 1985; Daly and Gunn, 1985; Gunn and Pol, 1986; Gunn, van de Hoorn and Daly, 1987; HCG, 1989), car ownership is forecast conditional on household driving license status. In the original study (Geinzer and Daly, 1981), household license status was specified as a function of:

- Sex
- Age
- Employment status
- Education
- Household income
- Number of children
- Population density

(The first four attributes are those of the adults in the household.)

A four-alternative logit model was used. The alternatives were specified as:

- Neither adult licensed
- First adult only licensed
- Second adult only licensed
- Both adults licensed

Population density was replaced in a later version of the model with two variables, representing "regional effect" (or Randstad dummy) and "low urban effect" (or urban/rural dummy). The model was also extended to include households with three or more licensed drivers.

The HCG models incorporate an explicit treatment of license holding and a range of other explanatory variables, primarily socioeconomic in nature.

The car ownership model system of the original study comprises two models, "0/1 car model" and "1/2 car model." The former is applied to one-license households, and the latter to two-license households. Households with no licensed driver are assumed to have no car, while households with two or more drivers are assumed to have at least one car available. The models are specified as functions of

- Personal business ("logsums") accessibility (0/1, 1/2)
- Disposable income (0/1, 1/2)
- Sex of driver (0/1)
- Age of driver (0/1)
- Multi-worker dummy (1/2)
- Number of children by age (1/2)

where "0/1" and "1/2" refer to the "0/1 car model" and "1/2 car model", respectively.

The fit of these models is poor. Although the full set of variables are jointly highly significant, much of the variation is explained by the alternative specific constant terms with limited contributions from the other explanatory variables. The overall χ^2 statistic of the 0/1 car model is 368.7 with degrees of freedom (df) 6, indicating the model as a whole is highly significant. However, the χ^2 statistic associated with the explanatory variables is 34.0 (df=5), significant still, but to a much lesser extent. The same observation can be made of the estimation results of the latter model where the "logsums" accessibility measures of the original models are replaced by the "regional effect" and "low urban effect" dummy variables. Importantly, another newly introduced variable, "trend effect," is significant, and indicates a reduction in car ownership among single-license households.

The formulation of car ownership models conditional on household license holdings is superior to virtually all discrete choice car ownership models in which variables representing household license status are introduced into the "utility" function as additive terms. The HCG formulation is capable of capturing interaction effects involving license ownership and the other explanatory variables. The model also treats household license holdings as an endogenous variable. This structure allows the formulation of car ownership models as separate binary models, entirely bypassing the IIA issue.

The HCG driver's license holdings model fits the data well and appears to be stable across regions, but it has potential limitations. Because the decision to acquire a driver's license tends to be made in an early stage of a person's life, and the person would tend to maintain a driver's license once he acquires one, those factors that contributed to a person's earlier decision to obtain a license may not be present in a cross-sectional data set used to estimate a license holdings model. In this sense, the HCG license holdings model is a descriptive model, as opposed to a causal model, that evidently well captures the variation in license holdings observed across households in a cross-section. The model, therefore, is susceptible to structural changes affecting license holdings. In addition, the model is subject to the recognized limitation that it contains no mechanism to represent the ongoing trends in license holdings.

The car ownership models account for much smaller fractions of the variation in the data than does the license holdings model, implying that license holding is the major determinant of household car ownership, and once the former is accounted for, the latter is mostly explained. Nonetheless, the car ownership models indicate that income, age, and sex are important determinants of 0-car vs. 1-car ownership for one-driver households, and that multi-car ownership varies meaningfully depending on accessibility, income, number of workers, and presence of children in the household.

A shortcoming of the HCG model is that it does not capture the observed time-trends in license holdings and car ownership. The limitation arises from the fact that the model was estimated using cross-sectional data sets. Although geographical transferability has been examined, no mechanism to reflect a trend is present in these models. In particular, the license holding model does not capture the cohort effect that is evident in the national statistics compiled by van den Broecke.

Car ownership and maintenance costs enter the HCG models throughout a disposable income variable, which is based on car ownership costs estimated for the Citroen 2CV, a common inexpensive automobile. Car ownership and operation costs are, in reality, endogenous variables largely determined by the household's choice of:

- Car type (make, model, vintage)
- Holding duration
- Utilization

The disposable income variable in the HCG model serves only as a crude proxy. A car type choice model would allow precise specification of car ownership cost, which in turn would lead to the model's enhanced sensitivity to policy measures that differentially influence the ownership and operation costs of different types of cars.

Berkovec's Model

Berkovec (1985) describes a model wherein vehicle stock is forecasted by estimating vehicle characteristics, scrappage, used car prices, vehicle demand, and number of vehicles demanded. This is accomplished through the use of three submodels: 1) Scrappage submodel, 2) Auto type choice submodel, and 3) Auto ownership submodel.

The scrappage submodel is an exponential model which was calibrated by a log log regression. It utilizes:

- Vehicle price variables (endogenous)

- Dummy variables for vehicle year
- Dummy variables for vehicle class
- Vehicle weight variable (exogenous)
- Dummy variable for old trucks

A GLS transformation was done in the scrappage model to correct for heteroskedasticity caused by the substitution of log actual scrappage rates for the preferred log scrappage probabilities.

Auto demand was estimated using a two-level nested logit model which was calibrated by maximum likelihood methods with cross sectional data. The auto type choice submodel utilizes:

- Capital cost variables (endogenous)
- Operating cost variable (exogenous)
- Seat space variables (exogenous)
- Dummy variable for trucks
- Dummy variable for vans
- Dummy variable representing vehicles
- Dummy variable for old cars
- Age variable (exogenous)
- New truck variable (exogenous)
- Foreign make variable (exogenous)
- Sports model variable (exogenous)

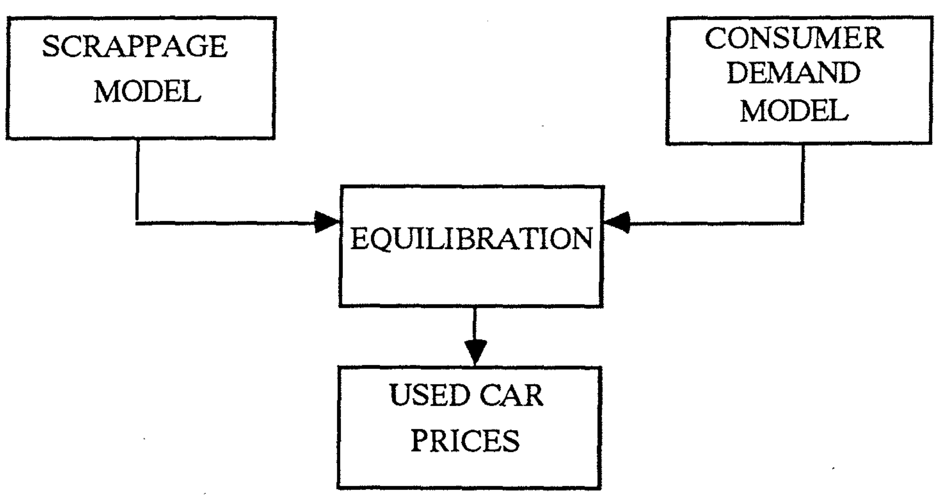
The auto ownership submodel utilizes values for one, two and three vehicles.

Data sources included:

- Auto ownership data from the National Transportation Survey done by Cambridge Systematics (CSI) for NSF in 1978
- Registration data compiled by R.L. Polk Company for 1977 and 1978
- Automobile characteristics data from a modified version of the CSI auto attributes file
- Actual auto scrappage rates for 1978

The three models are static and were calibrated simultaneously. Equilibration was used to obtain equilibrium price vectors for used cars (see following figure). Prices were calculated by Newton's Method.

The model is estimated for 12 consumer groups consisting of three income groups and four household size categories. It is aggregated by vehicle type (13 possible) and year of model (for 10 years) for a total of 131 vehicle types. The 131st vehicle type is for model years earlier than 1969. No distinction is made between fuel types, only vehicle size.



A constraint of the model is that the quantity of new and used cars at a given time must be equal to the total demand for vehicles, used and new, at that time. Also, at a given time, the total new car supply for a given type of car at given prices, plus the quantity of that type of car existing in stock, must be equal to the aggregate demand for that type at the given prices.plus the number of that type of car scrapped during the period.

It is further assumed that all transactions occur simultaneously at equilibrium prices and transaction costs are minimal. New car attributes are fixed exogenously.

The following external estimates are needed for each of the forecast years:

- Population characteristics
- Vehicle characteristics
- Gasoline prices

It would be necessary to recalibrate the model using updated information on population, vehicle characteristics, and gasoline prices in order to utilize the model. In addition, a personal computer as well as a statistical package capable of computing logs and exponentials would be required.

Berkowitz's Model

In a study by Berkowitz, et.al. (1985), gasoline demand was derived from disaggregated vehicle holdings and usage decisions. The model considered three decisions made by the household:

- Vehicle holdings
- Mode to work
- Non-work vehicle usage

where the choices made are the outcomes of a utility-maximization problem. From these choices of vehicle type and desired usage, household fuel consumption was derived.

The model framework is of the discrete/continuous type. A direct random utility household function is assumed for multi-vehicle households due to the complications of dealing with two nonsymmetric usage demand equations when pre-specified indirect utility functions are used to derive the demand functions. The non-work usage choice model is obtained as a solution to a constrained optimization problem in which direct utility obtained from a given choice is maximized subject to the budget constraint faced by the traveller.

An indirect utility function is used for the discrete choice decisions regarding mode to work and choice of vehicle holdings.

The model has three steps corresponding to three sub-models:

- Non-work vehicle usage submodel
- Work trip mode choice submodel
- Vehicle holdings choice submodel

The three decisions made by the household are assumed to be nested. Each decision is taken hierarchically with reference to the utility expected at subsequent stages of the decision process. Vehicle holdings are decided at the beginning of each period based on expectations regarding work and non-work vehicle usage. Conditional on the holdings, mode to work is decided given the expectation of the non-work vehicle usage. In the last stage, non-work usage is determined conditional on vehicle holdings and mode to work choice made in the previous stages.

In the vehicle holdings choice model, an inclusive value term in the utility of alternative vehicle holdings provides a measure of the expected utility of optimal mode choice for a given vehicle bundle. The utility function of the vehicle holdings choice model also contains a complex composite term which is a function of:

- Household income
- Cost of holding the vehicle bundle
- Mean price per kilometer of the bundle

The coefficients of the variables making up this term are determined by the expected non-work utility elasticities and thus a link is established between the continuous usage and the discrete components of the model system.

There is also an endogenously determined cost variable in the utility function of the mode-to-work choice model. This cost variable is constructed from, among other things, the expected non-work utility elasticities. Thus another endogenous link is established between the continuous and the discrete components of the model system.

The non-work vehicle usage model utilized the following exogenous variables:

- Rural households
- Urban area (population equal to or greater than 500,000)
- Number of non-working drivers in the household
- Comparative size of the vehicles in a two-vehicle household
- Vehicle type
- Region of the country (study was conducted in Canada)

Households were stratified with the following for the work trip mode choice model:

- One worker households, worker drives alone

- Two-worker households, workers share a vehicle
- Two-worker households, one worker takes transit while other worker drives
- Two-worker households, each worker drives a vehicle
- Number of non-working drivers in the household, for transit alternatives
- Urban area (population equal to or greater than 500,000), for transit alternatives
- Travel time to work

The work trip mode choice model also utilized an endogenous cost variable that was a function of the total household cost as related to mode choice, household income, and expected non-work utility elasticities.

The vehicle holdings model looked at households as follows:

- No vehicles
- No vehicles, household in an urban area
- No vehicles, household has access to transit
- Two vehicles, no workers
- Two vehicles, one worker
- Two vehicles, two workers
- Two vehicles, number of licensed drivers
- Vehicle holdings include a truck, household in a rural area
- Holdings included a "large" vehicle, household in a rural area
- Age of vehicles
- Vehicle horsepower
- Origin of vehicle (foreign or domestic)
- Vehicle classification (e.g., sports, luxury)

The vehicle holdings model also utilized an "inclusive value" endogenous variable which represents the expected maximum utility associated with the lower level mode choice decision, given a choice of vehicle holdings and a complex composite term which is a function of household income, the cost of holding a vehicle bundle, and the price per kilometer of operating the vehicles in the bundle. The coefficients of the variables making up this term are determined by the expected non-work utility elasticities.

Vehicles were grouped into 18 distinct categories based upon a four-way classification scheme:

- Length
- Vintage
- Type (standard sedan, luxury station wagon, or truck)
- Origin of manufacture (domestic or foreign)

There were 171 possible combinations or bundles of vehicles for two-vehicle households, 18 for one vehicle households, and one for no-vehicle households.

Single equations were estimated for each of the three models. Selectivity bias correction was not applied in the continuous non-work usage model. An inclusive variable provided the linkage between the work mode choice model and the vehicle holdings choice model. Additionally there was one variable each in the vehicle holdings and in the work mode choice models that provided the linkage between the continuous non-work usage model and the two discrete choice models.

Dynamic factors entered the model system in the vehicle holdings choice model.

The non-work usage model was estimated using the SUR (Seemingly Unrelated Regression) method due to the likelihood of correlation in the errors for two cars in a multi-vehicle household. The two discrete choice models were estimated separately using the maximum likelihood estimation method. The following assumptions/simplifications were made:

1) The non-linearities in the random variables in the indirect utility function were eliminated through a Taylor Series Expansion of the non-linear terms, allowing an approximately linear model for estimation.

2) While separating the representative utility in the indirect utility function into auto-specific and mode-specific components, approximations such as the expression of some nonlinear terms into appropriate Taylor Series expansions are adopted.

Constraints imposed by the model include:

1) The number of auto holdings is restricted to two or less.

2) The number of workers per household is restricted to two or less.

3) The mode to work is restricted to the following: drive alone, use public transportation, or carpool with other household members.

4) The coefficient of the composite variable in the vehicle holdings choice utility is restricted to one.

Household data were collected using a questionnaire designed to elicit:

- Household socioeconomic demographics
- Vehicle descriptions for currently and previously held vehicles
- Mode of travel to work
- Data linking individual household members with specific household vehicles

A vehicle-use log was also included in the questionnaire to collect information on the use of each vehicle by the household over the survey week. The survey was administered to a nationwide panel of Canadian households in Fall 1982.

Vehicle characteristic data were collected for each vehicle owned by the household or disposed of by the household during the preceding year. Used car prices were derived from Ward's Canadian Black Book, for 1977 and newer vehicles. Pre-1977 prices were obtained from MacLean Hunter's Red Book with a corresponding adjustment for comparison with the Black Book. Fuel efficiency figures were derived from Transport Canada Fuel Economy Guides. Interior volume characteristics were gathered from various issues of Consumer Reports while the Automotive Service Data Books provided information on cubic inch displacement, compression ratios and vehicle horsepower.

The forecasting procedure adopted in this study was sample enumeration within a simple Monte Carlo simulation, designed to generate household vehicle holdings and usage decisions in each year of the forecast period. The procedure began with the November 1982 sample of Canadian households gathered during the household survey. This sample of households was taken as being representative of the Canadian population of households in "year zero" of the simulation. Given the policy scenario under consideration, changes in the "system state" for "year one" were then specified. Given these changes, and the known characteristics of the households in the sample (including their current vehicle holdings), the probability of each household selecting any given feasible vehicle bundle to hold in year one could be calculated using the model. The actual bundle which each household chose to hold in year one was then predicted by drawing a uniformly distributed random number defined over the range zero to one, and selecting the vehicle bundle within whose probability range the random number happened to fall.

Once the household vehicle holdings for year one were determined, household vehicle usage and gasoline consumption were then determined. The work trip mode choice component of these calculations was determined using the same Monte Carlo procedure described for the vehicle holdings choice. The non-work vehicle usage for each household was then determined given the auto and mode choices. Finally, household weekly gasoline consumption was calculated. The procedure was then repeated for successive years in the forecast period.

A five-year forecast period was chosen as being sufficiently long to permit households to adjust to the policy initiatives under consideration, but not so long as to significantly violate the "short-run" assumptions built into the model. One-year time increments were chosen based on the assumption that households rarely engage in "multiple transactions" with respect to their vehicle holdings within a single year.

The forecasts made by this model are all point estimates and as such are of unknown precision. Ideally, the simulation should have been replicated many times to trace the distribution of possible outcomes. The authors state that this was not done because of the complexity of the model and the cost involved in running the simulation model through many replications for each policy alternative examined. The implications of different system states on gasoline consumption can be explored while forecasting. Different system states can be obtained by varying the policy scenario. Recalibration of the prediction model does not appear to be needed as the model is used only for short term prediction. Changes in other correlated explanatory variables are taken into account during forecasting.

To utilize this model, a household survey would be necessary to collect the data described above. Vehicle characteristics data should be readily available from published sources. Any one of the many widely available standard statistical packages should suffice for the type and extent of computations that are needed for studies such as this one.

Electric Vehicle Disaggregate Models

It is difficult to anticipate the impact of new transportation technologies on the demand for energy for transportation purposes. If the technology in question is not yet widely available, we certainly cannot observe quantities demanded at different prices for different configurations for enough different consumers to produce reliable forecasts of ultimate adoption and utilization. The only really tractable method for determining probable consumer responses to the availability of new technology is to elicit information by the use of "contingent" surveys. Research by Calfee (1980, 1985) on electric cars and Hill (1987) on electric utility vehicles anticipates this type of approach.

Survey methods are now widely used for the valuation of non-market environmental resources. If there is no traditional market for access to an environmental resource (in the sense of a specific per-unit price for consumption) it is very difficult to come up with a dollar figure for the social value of that resource. (Cummings, Brookshire and Schultze, 1986, and Mitchell and Carson, 1989 provide detailed assessments of "contingent valuation methods" for non-market valuation.) For new transportation technologies, we have a non-market situation not because of the public goods problem, but simply because the good in question is not yet available. In either case, respondents have little or no experience with placing a value upon the good, so special elicitation methods are required.

A very preliminary description of how market research can be conducted using contingent valuation methods is contained in Cameron and James (1987). The interviewer must describe to the respondent in as much realistic detail as possible the characteristics of the new transportation technology (for example, electric cars). The respondent can then be asked whether or not they would purchase one of these cars if the capital and operating expenses were as specified in the hypothetical scenario. (The caveats associated with the hypothetical nature of contingent valuation surveys have been discussed and evaluated at length in the environmental valuation literature, so much is known about the framing of the survey questions to minimize biases.) In conjunction with the socioeconomic characteristics of the respondents and the usual assumptions regarding some commonality of preferences in the population,the researcher can then derive the "demand function" for the new technology implied by these responses.

Using a fully disaggregated probabilistic choice analysis model, Calfee (1980, 1985) estimates the individual's choice of electric vs. conventional automobiles. Preference data were collected from a selected group of 51 individuals by giving each individual thirty choices, each of which had three alternatives. This provided a direct sample distribution of preferences, with no a priori restriction on the distributions of coefficients. The potential market shares were then predicted by determining the choice of each individual, and aggregating those choices to produce the predicted market share. The functional form for each utility function was taken to be the same. Individual choices were assumed to be stochastic. That is, individuals were assumed to make random trade-offs between vehicle attributes. The following hypotheses were adopted:

1) Individual choices are random in the sense that trade-offs fluctuate within limits. To some degree this randomness will occur even within a single experiment.

2) Variation between individuals is much greater than within individuals. Thus one individual may fluctuate between Toyota and Datsun, and another between Cadillac and Buick, but neither is expected to choose from the other's preference set.

Measurements of intra-individual variations were suppressed while seeking to estimate inter-individual variations. This was accomplished by estimating utility functions for each of a sample of consumers. Individual choices are assumed to be governed by a probabilistic process which may correspond to something along the lines of random trade-offs among attributes.

Because of constraints in data collection, only five attributes were used to describe each vehicle:

- Price
- Capacity (number of adult passengers)
- Top speed
- Range (in miles per day)
- Operating cost (including the periodic costs of battery replacement, but excluding depreciation and interest charges, which were subsumed into price)

Multinomial logit models were then estimated for each of the individuals in the data set using the maximum likelihood estimation method.

In the utility function, physical attributes (speed, range, and capacity) were entered as natural logs in order to allow for declining marginal utility. Price and operating costs were entered linearly. Even though marginal disutility was suspected to increase with those attributes, it was believed that this specification allowed for simple, unambiguous calculations of such items as a single, constant revealed value for each attribute for each respondent, as well as the subjective discount rates of the subjects. The variables used in the utility function were:

- Dummy variable for electric or conventional vehicle
- Purchase price
- Capacity
- Top speed (assumed to be 65 for conventional vehicles)
- Daily range in miles between fill-ups or recharges (daily range for conventional vehicles was assumed to be 400 miles)
- Operating costs
- Vector of parameters unique to each individual

This study considered a relatively small subsection of the potential market for electric vehicles. Only private demand for automobiles was considered, ignoring demand from business and demand for trucks and vans. All demand was assumed to be for second cars in multi-vehicle households, and data were gathered from only a restricted socioeconomic segment of the automobile consuming public - one that may be more likely than the average population to choose electric vehicles. In addition, the study assumed the existence of a fully mature market of electric vehicles, including repair facilities that do not currently exist. Also, the study excluded the effects of supply-side adjustments such as changes in neighborhood layout or commuting habits, which may effect demand for electric vehicles.

Market shares for electric vehicles were computed for various scenarios of electric vehicle technology advancement. To utilize this model, new region-specific surveys would be needed to estimate consumer trade-offs correctly. Additionally, a random sampling methodology, or some other unbiased sampling method, would be necessary in order to apply the results to the general population.

Manski and Sherman Model

The model described by Manski and Sherman (1980) estimates the magnitude and mix of new car sales, used car scrappage, and composition of used car holdings using household level discrete choice models. Given assumptions on the design and prices of future new vehicles and fuel prices, the behavior of a sample of households is simulated and vehicle market equilibrium conditions for each year of the forecast period are solved.

In contrast to the conventional aggregate stock adjustment models, where new vehicle sales are explained as the fractional reduction of discrepancies between desired and actual vehicle stocks, this model treats new vehicle sales, used vehicle scrappage, and used vehicle prices as jointly endogenous variables that solve a set of market equilibrium conditions.

To model the equilibrium in the vehicle market, two sets of modeling efforts are undertaken: one for the demand side, and one for the supply side. The demand side consists of a household's choice of:

- Quantity of vehicles to own
- Type of vehicles to own
- Subsequent use of the vehicle(s)

Vehicle use was not modeled. The size of vehicle holdings and its composition are not modeled simultaneously. The vehicle holdings are modeled first and then the choice of the type conditional on the quantity is modeled.

The decision between selling and scrapping a vehicle is said to depend on the price each transaction is expected to yield. It is assumed that there is no linkage between a household's decision regarding vehicle holdings and the sales-scrappage price the household can realize from disposal of currently held vehicles.

The new vehicle market is portrayed by assuming that manufacturers set new vehicle designs and pricing decisions at the beginning of each sales year and these are fixed for the year. New vehicle supplies are assumed to be perfectly elastic until the end of the sales year, when production ceases.

In this simulation approach to forecasting, the vehicle quantity model predicts the number of vehicles each household in the sample would own. Given a set of prices, the vehicle choice model then predicts the composition of holdings for each household. The scrap-sell model determines how vehicles are disposed. In the final step of a given year's simulation, a set of vehicle prices is searched to generate desired holdings, sales, and scrappage, and to solve a feasible set of equilibrium conditions.

The projected vehicle designs for each future year were represented as revisions to the designs of a base year. A Corporate Tech Planning, Inc. report is used to formulate a most probable scenario for vehicle weight changes expected to be made by domestic auto manufactures. To predict new auto prices, the cost consequences of projected design changes are considered and a two percent per year real increase in the price of domestic automobiles is used.

a. Summary of Models

Mogridge, 1989

- Depicts the evolution of a "car park" through a car accounting procedure.
- Based on the observation that as incomes increase the expenditure on car purchase increases and the proportion of the household total expenditure spent on car transport remains constant to a first approximation.

Van den Broecke, 1988

- Forecasts population by cohort along with probability of individuals holding driver's licenses and owning automobiles in the forecast year.
- Model is formulated with individuals as the basic analytical unit, not households.

Geinzer and Daly, 1981; Bradley, 1983; Smyth and Daly, 1985; Daly and Gunn, 1985; Gunn and Pol, 1986; Gun, van de Hoorn and Daly, 1987; HCG, 1989

- Car ownership is forecast conditional on household driver's license status. Driver's license holdings are forecast based on a host of socioeconomic variables.
- Trends in license holdings and car ownership are not captured.

Berkovec, 1985

- Vehicle stock is derived by estimating scrappage, used car prices, vehicle demand, and number of vehicles demanded.
- Households are categorized by three income levels and four household sizes.

Berkowitz, et al., 1985

- Disaggregated vehicle holdings and usage decisions are estimated through a utility maximization problem. Household fuel consumption is derived from these choices.
- The implications of different policy scenarios on gasoline consumption can be explored while forecasting.
- A household survey is required to gather socioeconomic, vehicle holdings, and driving habit data.

Calfee, 1980; 1985

- Estimates the individual's choice of electric vs. conventional automobiles.
- Surveys were used to gather data for each subject in the sample to estimate individual utility functions.
- Only private demand for automobiles was considered, ignoring demand from business or for trucks or vans.

Manski and Sherman, 1980

- The magnitude and mix of new car sales, used car scrappage, and composition of used car holdings are estimated using household level discrete choice models. They are then treated as jointly endogenous variables that solve a set of market equilibrium conditions.

b. Critical Factors

Model	Vehicle Holdings Described for each Household at Given Point	Capture Household Decisions: - to change vehicle holdings - to change usage levels - aging vehicles - external changes - HH characteristics
Mogridge, 1989	Holdings based on income.	No
Van den Broecke, 1988	Holdings estimated for individuals, not households.	No
Geinzer and Daly, 1981 Bradley, 1983; Smyth and Daly 1985; Daly and Gunn 1985; Gunn and Pol, 1986; Gun, van der Hoorn, and Daly, 1987; and HGC, 1989	Estimates vehicle holdings based on household driver's license status.	No
Berkovec, 1985	Holdings are estimated for twelve household types. Type is determined by income and size.	Auto demand and scrappage decisions are based on costs, size, age, make, and model.
Berkowitz, 1985	Holdings are a function of income and vehicle costs.	Vehicle holdings, mode to work, and non-work vehicle usage decisions are modeled for the household.
Calfee, 1980; 1985	No	No
Manski and Sherman, 1985	Households are characterized by income, size, residential location, number of workers, age, and previous ownership to estimate vehicle holdings.	Vehicle usage is not modeled.

B. Aggregate Forecasting Models

This section provides an overview of aggregate forecasting models developed and used in the last few decades. Most of these models consist of three fundamental steps. The first step involves the estimation of vehicle miles traveled by different fleet: passenger vehicles, trucks, and airlines. The second step involves estimating the characteristics and number of future fleets of vehicles using methods ranging from simple linear projections to more complex econometric techniques. The third step involves using the estimates from the first two steps to estimate the future fuel and energy demand for the transportation sector. The aggregate nature of the modeling technique results in a loss in disaggregate level information. This is the main disadvantage of these models. However, data sets to apply them are readily available from various Federal and State agencies, and thus may prove useful if quick and initial rough estimates are required before performing a detailed analysis.

<u>Erlbaum's Model</u>

Erlbaum, et.al. (1977) studied the impact of price, availability, and efficiency on automotive energy forecasts. Aggregate automotive fuel demand models were estimated for each of New York State's SMSA's and the rest of the state, using monthly data for a 48-month period from 1972 to 1975.

The impact of changes in policy-related variables on gasoline consumption were examined through the following approach:

1) Vehicle Miles of Travel (VMT) were estimated for the base year, 1975, by expanding the known 1972 VMT by the ratio of the annual regional traffic count for 1975 to the annual regional traffic count for 1972.

2) Future year VMT was estimated by expanding the 1975 VMT by a) the ratio of population, and b) the ratio of propensity to travel, for the forecast and base years. Regional traffic counts were taken to be a measure of "propensity to travel."

3) Regional energy consumption was estimated from the regional VMT figure and the regional average fleet efficiency.

Regression analysis was used to estimate the elasticity of Monthly Traffic Counts (MTC's) using:

- Gas Price/Consumer Price Index
- (Average gallons of gas sold in two contiguous months) x (average fleet efficiency)
- Labor Force Participation Rate
- Unemployment Rate
- Autos/Population Ratio
- Retail Sales Index
- Transit Ridership (NYC only)
- Seasonal Driving Index

Demand and supply equilibration was achieved through summation of gas demand (the model output) over all regions and equating it with the annual gas supply (the model input), over the state. However, this equilibration process is not built into the model and could only be done externally in an iterative fashion.

The temporal aggregation of this model is at two levels: 1)the estimation of elasticities of monthly traffic counts with respect to the variables listed previously was done with monthly data, and 2) the estimated elasticities were used in forecasting yearly traffic count and VMT variables utilizing annual data and thus, also creating an endogenous link.

This model assumes that annual fluctuations in regional VMT will correspond to annual fluctuations in the desire to travel, as measured by traffic volume. Future year VMT is also assumed to be dependent on growth in population and growth in propensity to travel, which is given as a function of the economic climate, gas price, and gas availability. The model also makes assumptions as to the growth in variables used to estimate traffic counts, including population, unemployment, labor force participation, and the consumer price index. The traffic count elasticities can be updated, however, by recalibrating the regression equation with revised data.

This model was used to test various policy scenarios related to transportation energy analysis and forecasting. In particular, the impact of the national vehicle fleet turnover requirement and the supply and price of gas, was explored. Data was obtained from: the New York State DOT, Economic Development Board, Department of Labor, Data Services, City Transit Authority; U.S. Department of Commerce, FHWA Monthly Motor Gasoline Reported by States, Oil and Gas Journal, and the Society of Automotive Engineers Transactions.

Once the data are collected, minimal effort is required to run the model. Any standard statistical package will suffice and run time is relatively short.

Millar's Model

The model presented by Millar, et.al. (1982) projects total transportation energy consumption through energy demand projections for each of the major transportation modes and submodels. The model incorporates categorical analysis, regression, multinomial logit analysis, a direct demand model, and iterative proportional fitting procedures. In all, ten modeling steps are utilized, some requiring manual iteration.

The demand for personal vehicles was estimated with an incremental logit equation using:

- Vehicle initial cost/household income class
- Vehicle operating cost per mile
- Vehicle capacity weighted by household size
- Household VMT/month
- (Vehicle weight) x (head of household age)
- (Vehicle weight) x (head of household education)
- (Vehicle performance index) x (head of household age)

Information on demographics, travel patterns, and characteristics of motor vehicles was obtained from the Nationwide Personal Transportation Study (NPTS) conducted in 1977 by the Bureau of the Census. Population is aggregated at the household level and personal vehicles are aggregated into nine categories with only gasoline, diesel, and electric vehicles considered.

Regression analysis was used to forecast the total number of vehicles in commercial/municipal automotive fleets. The variables used in this analysis included:

- GNP (1972 dollars)
- Number of autos in fleets of four or more

- Projection year

The FRATE (Freight Responsive Accounting for Transportation Energy) model was used to project the number of commercial vehicles comprising the commercial truck fleet as a function of economic activity. Local and intercity truck stocks are projected independently using the following variables:

- Truck characteristics
- Truck stock
- Ton-miles travelled
- Average annual miles/vehicle
- Vehicle fuel type
- Future fuel efficiencies

Buses were categorized by school, transit, and intercity use. Bus stocks and activities were projected as a function of:

- Bus stock
- Projected population growth
- Average vehicle utilization
- Fuel types
- Fuel efficiencies

Intercity passenger travel demand was estimated through the use of direct demand models as found in POINTS, the CTR Passenger Oriented Intercity Network Travel Simulation model. The calibration of the POINTS travel demand model attempts to replicate the base year origin-destination matrix. Input to the model included:

- Intercity passenger travel by:
  - Origin-Destination pair

- Trip purpose
- Mode choice
- Travel time
- Travel cost
- Distance

- Travel growth rates between nodes
- VMT or PMT based energy intensities
- Nodal population
- Government employment
- Hotel service receipts

Non-local travel patterns were obtained from the National Travel Survey (NTS), a household survey performed at three month intervals by the Bureau of the Census in 1977.

Total trip number, distribution, and modal split were estimated simultaneously by the model. The model's four calibration coefficients were not automatically revised until a pre-specified threshold was met. Therefore, it was necessary to manually update these coefficients until they converged.

International air total revenue passenger-miles (RPMs) and energy usage were projected using:

- Annual estimates of RPMs
- Revenue per RPMs
- Annual GNP of the U.S., U.K., West Germany, and Japan

General aviation activity was projected with the following input variables:

- Rate of change of forecast year VMTs to base year VMTs
- Rate of change of population
- Rate of change of real general aviation cost
- Rate of change of real per-capita disposable income
- Rate of change in commercial airline fares

There is no stratification applied in this model; however, the categorical analysis and iterative proportional fitting do effectively break the sample into homogeneous groups.

To recalibrate the model, the external input variables mentioned previously must be updated. Thus, a sizable data collection effort would be required. The required intercity and freight software packages, POINTS and FRATE, and the iterative proportional fitting, are also time consuming. Additional required software includes a statistical package capable of handling regression analysis and maximum likelihood estimation. A spreadsheet would also be helpful to keep the data in an easily readable and usable form.

Kulash's Model

Kulash (1976) developed a model for forecasting long-run automobile demand. The objective of this model is to estimate the effects of alternate policies for improving automobile fuel economy on future gasoline consumption, vehicle miles traveled, new car sales, fleet size, fleet composition, stock of cars in use, new car prices, and fuel economy. The model projects total new automobile sales and details the composition of total sales by vehicle size class.

New automobile sales are assumed to be related to the gap between a "target" stock and the existing stock of automobiles at the beginning of the current period, less those autos that will be scrapped during the year. A stock adjustment approach is used to forecast this demand for new automobiles using:

- Household income
- Auto stock on hand at the beginning of the forecast year
- Number of autos scrapped during the forecast year
- Number of autos per household by income group
- Household totals by income group

As automobile ownership by income group is relatively stable over time, it is assumed that the relationship between target ownership and household income can be estimated cross-sectionally with 1970 data. The proportions of miles traveled by each age group of autos are also based on cross-sectional data and remain constant. VMT per household is estimated utilizing the following forecast year variables:

- Total VMT
- Total real disposable income
- Real gasoline costs per mile

Fuel economies and prices of three size classes of autos are calculated through a procedure which attempts to minimize the sum of:

- Automobile cost to the consumer
- Automobile travel cost to the consumer (inversely proportional to fuel economy)
- Consumer tax/rebates based on fuel economies
- Industry civil penalties based on fuel economy standards

The shares of the new auto stock that are produced by the various automobile manufacturers are also assumed constant.

Data required to run the model include:

- Gasoline prices
- New car fuel economy policies

Data already built into the model include:

- Income
- Population
- Historical stock of autos and fuel economies
- Unemployment rate
- Target auto ownership
- Technological costs of achieving future fuel economy ratings
- Projected fuel economy

This model has been used in policy analysis by the Federal Energy Administration, the Information Administration, the White House, The Transportation Systems Center, the Congressional Budget Office and the National Highway Traffic Safety Administration.

<u>Immers' Model</u>

In a paper on determinants of car mobility in the Netherlands, Immers, et.al. (1988) present conclusions regarding travel and other transport-related behavior drawn from a time series analysis of trends in travel related variables, and two regression analyses that estimate daily mileage per person and percentage car ownership by homogeneous population groups.

The first regression equation estimates daily per person mileage utilizing income and fuel price. The second equation estimates the percentage car ownership utilizing income level, age, and sex of the respondent.

Results of the analysis indicate that:

- Fuel prices only marginally influence mobility
- Reduction in automobile fuel consumption is mainly due to manufacturers introducing more fuel efficient vehicles
- Personal net income is the most important variable influencing the mobility of car drivers

Data were drawn from the Netherlands Travel Survey for the period 1978-1982. The survey contains relevant information about approximately 1,000,000 trips and the respondent.

Gur's Model

Gur, et.al. (1987) analyze differences in travel and fuel purchase patterns by demographically aggregated household groups. The methodology adopted for this analysis was the multiple classification analysis (MCA). MCA was used to relate observed differences in fuel purchase frequency, tank inventories, vehicle miles traveled, and fuel expenditures to such index-type household characteristics as income and residence location, as well as sex, race, and age of household head.

Monthly fuel purchase logs from the Residential Energy Consumption Survey's Household Transportation Panel (TP) served as the data source for this analysis. The data set contains detailed fuel purchase, demographic, and socioeconomic information for a monthly sample of approximately 1,000 households that used vehicles for personal transportation, along with weighting factors for expanding the estimates to national monthly totals. The data spanned the 28-month period from June 1979 through September 1981.

It is not clear whether the type of analysis presented in this paper can be used to forecast future energy usage because forecasting the demographic and socioeconomic information utilized in the analysis is, in itself, a non-trivial task.

Manski's Model

Manski (1983) contends that disaggregate models of household vehicle choice behavior are inadequate for forecasting purposes because prices and designs of vehicles are assumed as given, while in reality, they are market-determined. Therefore, a short run equilibrium model is presented in this study. The "short run" is defined as a period in which new car designs and prices are fixed but used car prices adjust competitively to market forces.

Used car prices and scrappage rates and new car sales are determined for a given short run period. The core of the work is an aggregate random utility model ("aggregate" because it is not conditioned explicitly on household attributes). Therefore, holdings of segments of the population can not be forecasted, only forecasts of the population as a whole can be obtained.

It is assumed that in each market period, consumers evaluate their current holdings of vehicles and update them as desired. It is also assumed that an individual holds at most one car. The utility of holding is based on:

- Net value of transportation services of the vehicle type
- Present market value of the vehicle type
- Expected resale value of the vehicle type

The choice set is the set of distinct vehicle types in existence in addition to the alternative of owning no car. A car is sold in the used car market only if the difference between the price fetched by the car and the cost of non-routine maintenance on the car in that market period is greater than the scrap value of the car. Three conditions must be satisfied to obtain equilibrium:

1) The price of a used car cannot be lower than the scrap value and expected discounted depreciation cannot be negative.

2) There exists a marketable set of used vintages among which consumers are indifferent and the other used vintages are dominated by the marketable group.

3) Vehicle demand equals vehicle supply.

As a result of this analysis, the qualitative effects of changes in tax-subsidy policy on the size and composition of aggregate automobile holdings can be predicted. Changes in tax-subsidy policy may include:

- Customs duties and value added tax
- Annual registration fees
- Gasoline tax
- Taxes on replacement parts
- Operating cost subsidies

Quantitative forecasts are obtained through econometric analysis of observed market equilibria. To accomplish this it is assumed that:

- Taste variation terms are independent and identically distributed with an extreme value distribution.
- The quality of an auto type can be derived through observed attributes and vintage.
- Expected resale value is dependent on auto qualities and repair costs.

A model explaining vehicle demands was estimated using regression analysis with observations on the Israeli automobile market in three successive years - 1978 through 1980. Automobile qualities included:

- Fuel consumption
- Power
- Transmission
- Height and width
- Trunk space
- Wagon space
- Number of doors
- Years of availability in Israeli market

The number of vehicles of each model was obtained from the Israeli Central Bureau of Statistics. The vehicle aging process was analyzed and a model estimating vehicle quantity was derived using empirical data for 1300 model-vintage pairs. The empirical data included continent of origin and engine size. Finally, automobile holdings by model and vintage are forecasted for a given market period based on:

- Population size
- Per capita income
- Policy variables

- Vehicle characteristics
- New car prices
- Used car stocks

Edwards' Model

In a study of urban transportation energy, Edwards (1977) attempts to explain the fundamental relationships between urban spatial structure and energy consumed in urban passenger travel. The research is premised on the hypothesis that the structure of urban land-use and transport networks can have a very significant effect on the consumption of energy in urban passenger travel.

A Lowry-type model was used together with simple models of trip generation, mode choice and trip assignment to produce realistic land-use patterns and characteristic travel behavior for a given city. Interzonal trip matrices were computed for each of five trip purposes: work, two types of shopping or service, non-home based, and social. Following the assignment of trips to the network, energy models for automobile as well as transit were utilized to assess the transportation energy consumed over all trip purposes in each of several different urban spatial structures.

. The principal components found to be important in explaining differences in energy use were: 1) the urban form, 2) the overall level of service provided by the transportation system (as measured by overall travel speed), and 3) the extent to which public transit plays a role in the transport system (mode split).

With respect to urban form, four distinct dimensions were separately identified as contributing subcomponents:

- urban shape
- geographic extent (developed land area in square miles)
- the degree of concentration of population about the city centroid (measured by the second moment)
- the degree of employment concentration around the city centroid (also measured by the second moment)

The results presented are conditional upon assumptions relating to preferences of workers for various trip lengths that are inherent in the specification of work trips in the Lowry-type model. These assumptions can be inaccurate and thus the resulting residential location process can be highly erroneous. Thus, the modeling technique presented here may only be useful for spatial planning and spatial structure analysis, and not for energy consumption.

<u>Cardell's Model</u>

The CRA Hedonic Demand Model (Cardell, et.al., 1976) estimates the impact of changes in vehicle attributes on the market shares of the particular vehicles that are offered by manufacturers. The model assumes that automobile purchases maximize consumers' utility, subject to their individual preferences for vehicle attributes and for the vehicle offerings produced by manufacturers.

The distribution of consumers' tastes for each of up to ten (of over fifty possible) attributes of vehicles, is modeled based on: 1) Actual vehicle market shares, 2) A measurement of each attribute for each vehicle, and 3) A random probability of the importance consumers attached to each of the attributes.

The model seeks to find the distribution of consumers' utility functions that most closely reproduces the market shares actually observed for the individual models of vehicles. Algorithms are provided that use the taste distribution parameter estimates to simulate the market's response to changes in the attributes of the vehicles offered. Various applications of the model have involved using these algorithms on alternative data bases. Different parameter estimates occurred when different vehicle attributes were studied. Thus, the algorithms are used to build the specific version of the model through the estimation of the parameters as well as to simulate alternate scenarios.

For the market share predictions to be reasonable, the set of automobile models included for consideration must cover nearly the entire market. In order to predict new-car sales, estimates of the total price elasticity of new-car demand must be obtained from other sources. Thus, a relatively detailed vehicle data base is required to run the model. The data base would include:

- New product offerings
- Vehicle attributes
    - prices
    - acceleration
    - frequency of repair,
    - volume
    - passenger area
    - weight
    - fuel efficiency, etc.

- Projected policies
    - tariff levels
    - vehicle taxes, etc.

The model requires running a time-consuming computer program consisting primarily of an algorithm for the numerical evaluation of probability integrals using the Monte-Carlo method.

Reza-Spiro Model

Reza and Spiro (1979) developed a model which describes the demand for passenger autos, for miles traveled, and for the attributes or qualities of autos. In this case quality is measured by the average weight of autos. The demand for gasoline is then estimated with the average efficiency of the fleet of autos. This econometric model was estimated with a linear specification utilizing:

- Income
- Automobile stock
- Automobile prices
- Miles driven
- Fuel .

The model's equations were estimated using either the cost per mile of driving the existing stock, or of driving the new stock, or the price of gasoline.

Estimations of the demand for gasoline were then derived from these equations, and short-run and long-run effects of changes in the price of gasoline on miles traveled and gasoline consumption were determined. The short-run effect of a change in the price of gasoline on gasoline consumption results only from a change in miles traveled by the existing stock of cars. The long-run effect results from a combination of changes in miles traveled, the size of the stock of cars, and the composition of the stock of cars.

The model was estimated with quarterly data from 1969 through 1976. Data sources included: Survey of Current Business for income, interest rate, gasoline demand, car price, and the consumer price index; Bureau of Labor Statistics for price of gas; Ward's Automotive Yearbook for stock of cars. Cars were divided into eight size classes and the car model with the greatest number of registrations in the class in each model year was used to represent the class. The overall efficiency of new cars is a function of the weighted average of registrations over the entire sample period. Depreciation is an average equal to 0.0234. Fuel efficiency data for new cars came from the Automobile Club or Italy's World Cars, then these were averaged with the efficiency of the existing stock, after deflating 20% to account for the difference in new car estimates and actual car performance.

Cervero's Model

An autoregressive-integrated-moving average (ARIMA) model (Cervero, 1985) for forecasting monthly highway energy consumption attempts to replicate the past behavior of a time series by decomposing data into interpretable secular, seasonal and stochastic components rather than investigating any direct causal relationships. This approach is based on the observation that gasoline consumption has historically proven to be both price and income inelastic, thus, the exclusion of these variables in the interest of explicitly modeling seasonal, secular, and stochastic influences would be appropriate.

The model utilizes monthly U.S. gasoline consumption data from the 1977-1983 period compiled by the Federal Highway Administration from state motor fuel tax receipts.

Gasoline consumption is modeled as a weighted sum of current and past errors of the lagged values of the series itself. This model applies to a stationary series. Stationarity is achieved by differencing.

Distinct secular and seasonal factors, revolving around 1-month and 6-month cycles, emerged from the time series analysis that offer a practical basis for short-term energy forecasting. The model was used to forecast gasoline consumption for the 12-month period of 1983. For most months, forecasts were accurate to within 1% of actual consumption. Although there seemed to be some slight systematic underestimation of consumption levels, the model correctly forecasted the latter increases in gasoline usage.

The advantage of ARIMA modeling for short-term forecasting of highway gasoline usage seems to lie primarily in the simplicity of its structure and its relatively light data requirements.

Adler's Model

Adler, et.al. (1980) designed a computer-based dynamic simulation model that analyzes the interactions between energy supply and transportation-related energy use to be used to analyze policies such as transit service improvements, auto fuel efficiency mandates, gasoline taxes, gasoline price changes, and carpooling incentive programs. Scenarios such as different population growth rates, economic growth rates, and petroleum pricing and supply policies, may also be tested. The model projects to the year 2020.

The model has seven sectors, each performing a specific function:

1) Computes mode-specific travel demand and fuel use

2) Represents industry/consumer response to gasoline prices and government policies

3) Represents the transit sector response to changes in ridership and to policy

4) Represents carpool-specific levels of service

5) Determines the effect of highway condition and congestion on average network speed

6) Projects economic and population growth

7) Converts crude oil prices to equivalent gasoline prices

The decisions of the auto industry were also modeled. First, the costs associated with auto production were computed and prices for each of five auto size classes were estimated. Those costs influencing price include:

- Technology costs of fuel efficiency
- Gasoline costs
- Penalties for not meeting government mandated fuel economy standards
- Government imposed excise taxes based on auto fuel economy

Second, the fuel efficiency within each class that minimizes lifecycle costs of auto ownership was found using the derivatives of the four variables listed above.

A consumer choice model was used to determine the market share of each auto class. The factors involved include:

- Household income
- Number of licensed drivers per household
- New car price by class
- Number of autos per household

Total vehicle travel distance was used to compute congestion and road deterioration effects. The transit, carpool, and highway sectors compute levels of service by mode, given modal characteristics and travel volumes. These levels of service were used by the travel sector to determine travel patterns.

The demographic sector computes household characteristics such as number of households, mean household income, number of licensed drivers per household, and number of autos per household. Inputs include the total number of autos from the auto sector, and the distribution of autos across income classes, from the travel sector.

The cost sector used a wellhead crude oil cost to compute the price of gasoline, including the average state fuel tax computed in the highway revenues subsector. Values for fuel use and fuel availability from the travel sector were used to compute a price multiplier from fuel shortfall.

Several significant general scenario assumptions were made in this model. These are that; the internal combustion engine will be used in automobiles over the model's lifetime but diesel engines will be also widely used, that standards of the 1977 Clean Air Act amendments will be met, and that motor vehicle safety standards will be met without adverse impact on automobile fuel efficiencies.

Economic inputs were projected by Data Resources, Inc. Oil supply and price were forecast by the Department of Energy's FOSSIL79 and NEP2000 models. This model incorporates concepts from the following: Faucett Automobile Sector Forecasting Model, Sweeney Vintage Capital Model, and the MIT-TRANS model of travel demand at the urban area level. The consumer choice sector is borrowed from the disaggregate individual-level auto type choice model developed by Lave and Train.

The technology cost curve parameters, relating fuel economy rating increments to manufacturing costs, were estimated using data compiled for the Wharton EFA Automobile Demand Model.

Mullen-White Model

The model developed by Mullen and White (1977) utilizes a logistic function to estimate car ownership by person type using:

- Public transport quality
- Personal income
- Head of Household income or spouses income if person is head of household
- Motoring costs

Aggregation of the model is at the individual (person-type) level. Person-types for which a model was calibrated included:

- Male head of household
- Female head of household
- Male retired
- Female housewife

A saturation level of car ownership was specified for each person-type. Because little evidence was available to support the estimation of saturation levels, these levels were assumed subjectively.

Because each of these groups is assumed to follow a logistic car ownership curve, total car ownership, represented by the sum of these curves, does not follow a logistic curve even approximately. Examining the pattern of current car ownership, it was concluded that total car ownership will follow a path significantly different from the logistic: the car ownership curve was not symmetrical about its middle point as is the logistic, but demonstrated a very rapid growth in its lower part leading to a much earlier deceleration of growth rate. The growth in cars amongst low-income groups such as pensioners will occur over a long period, implying that car ownership will continue to grow slowly until it ultimately reaches saturation at a much later date than in the current official forecasts.

This model was used to forecast car-ownership and traffic growth in Telford, England. However, it does not forecast vehicle miles traveled and subsequently does not project energy consumption.

Wadhwa Model

The objective of a study by Wadhwa (1980) was to present a comprehensive analysis of energy use in passenger transportation in Australia in order to aid the process of energy policy design. A simulation model was developed to provide a framework for determining future passenger fuel demand in response to a variety of assumptions about controlled and uncontrolled social, economic, technological, and institutional factors.

The fuel demand for automobiles is assumed to be a function of the passenger task to be undertaken by automobiles and the efficiency with which it is accomplished. The determinants of the passenger transportation task by automobiles are:

- Population
- Number of cars per capita
- Passenger kilometers per car per year

Car ownership levels and its intensity of use are dependent on a host of socio-economic variables which includes:

- Incomes
- Motoring costs
- Car prices
- Residential density
- Age and sex
- Household composition
- Public transport quality
- Transport policy

In this model, car ownership is estimated with a simple logistic model with data on car ownership level and growth rate from 1965 through 1975. The car saturation level is one of the most important parameters in the logistic growth curve method. Here, a value of 0.65 was taken for this parameter. Key factors that appear to affect car usage are:

- Incomes
- Gasoline price and locational patterns

In the passenger transport task, travel was disaggregated into journey to work and other trip purposes. This was because the elasticity of travel with respect to incomes and prices vary widely with the trip purpose. Work trips have been found to be fairly inelastic whereas non-work journeys have significantly higher elasticities.

The efficiency with which the passenger task is accomplished is represented by energy intensiveness. Some important vehicle characteristics affecting energy intensiveness are weight of the car, the engine size and the auto-fuel efficiency.

The simulation model developed to determine the future oil demand for automobiles in Australia was based on the demand determinants listed above. Speed, weight, auto fuel efficiency and load factor on fuel demand were introduced in the model.

a. Summary of Models

Erlbaum, et.al., 1977

- Estimates fuel usage only, not auto demand.
- Can test effect of various policy scenarios (e.g., fleet turnover requirement, supply and price of gas).
- Model was estimated for New York State only.

Millar, et.al., 1982

- Estimates total energy consumption through energy demand projections for each of the major transportation modes and submodes.
- Sizable data collection (e.g., uses Nationwide Personal Transportation Survey from 1977, National Travel Survey from 1977, data regarding all commercial and recreational transportation modes).

Kulash, 1976

- Projects total new automobile sales and details the composition of total sales by vehicle class size.
- Can be used to analyze the effects of alternate fuel conservation policies.

- Requires that actual and future year vehicle populations be maintained by vehicle class and by model year detail and that scrappage and fleet size be computed on this detailed basis.

Immers, et.al., 1988

- Estimates daily per person mileage and percentage car ownership by homogeneous population groups.
- Data was drawn from the Netherlands Travel Survey (includes information regarding approximately 1,000,000 trips and the respondent). This effectively precludes a duplication of this study.

Gur, et.al., 1987

- Summarizes travel and fuel purchase patterns by demographically aggregated household groups.
- Requires extensive forecasting of demographic and socioeconomic information.

Manski, 1983

- Forecasts automobile holdings for the population as a whole.
- Only models a period in which new car designs and prices remain unchanged but used car prices adjust to competitive market forces.
- Allows for qualitative analysis of the effects of policy changes on automobile holdings.

Edwards, 1977

- Models the relationship between urban spatial structure and energy consumed in urban passenger travel.
- Applicable for understanding the variation in energy requirements under various scenarios of future urban development, not for forecasting energy consumption.

Cardell, et.al., 1976

- Estimates the impact of changes in vehicle attributes on the market shares of the particular vehicles that are offered by manufacturers.
- A relatively detailed vehicle data base is required to run the model.

Reza and Spiro, 1979

- Models the demand for passenger cars, for miles traveled, and for the class of the car by weight. Gasoline consumption is estimated based on the average efficiency of the fleet of cars.

Cervero, 1985

- Model replicates the past behavior of a time series to forecast monthly highway energy consumption.
- Does not attempt to determine a causal explanation for the time series' behavior.
- Suited for short-term forecasting.

Adler, et.al., 1980

- A computer-based dynamic simulation model that analyzes the interactions between energy supply and transportation-related energy use.
- Designed to be used to analyze policies such as transit service improvements, auto fuel efficiency mandates, gasoline taxes, gasoline price changes, and carpooling incentive programs.
- Projects to the year 2020.

Mullen and White, 1977

- Estimates car ownership by person-type.

- Does not estimate future mileage or energy consumption.

Wadhwa, 1980

- A fuel demand simulation model which provides a framework for determining future passenger fuel demand in a variety of scenarios.
- Applicable for analysis of various policy and economic scenarios. May be of limited use in forecasting energy demand.

b. Critical Factors

Model	Vehicle Holdings Described for each Household at Given Point	Capture Household Decisions: - to change vehicle holdings - to change usage levels - aging vehicles - external changes - HH characteristics
Erlbaum et.al., 1977	No	No
Millar et.al.,1982	Uses 1977 Nationwide Personal Transportation Study and modified Lave-Train model.	Demographic and economic variables assumed to determine total vehicle holdings. These patterns were assumed to be stable. Lave-Train utility function was used in order to identify attractiveness of vehicles and thus vehicle types were distributed.
Kulash et.al., 1976	Holdings based on income	Automobiles classified by weight. Turnover based on economic factors (e.g., unemployment, new auto prices).
Immers, et.al., 1988	Percentage of car ownership by income level, age & sex.	No
Gur, et.al., 1987	No	No
Manski, 1983	Only forecasts automobile holdings for the population as a whole.	For population as a whole only. Considers vehicle age and attributes only.
Edwards, 1977	No	No
Cardell, et.al., 19	No	No
Reza and Spiro, 1979	No	No
Cervero, 1985	No	No
Adler, 1980	Yes	Borrows from the Lave-Train model.
Mullen and White, 1977	Estimates car ownership by person-type.	No
Wadhwa, 1980	No	No

C. Freight Models

Most existing freight energy demand models relate freight volume and spatial distribution to the level and distribution of economic activity. These models perform well when the underlying economic and demographic forecasts are accurate, and poorly otherwise. Even short-run regional economic and demographic forecasting is difficult, and evaluating the accuracy of these models is beyond the scope of this report.

There are several classes of models that are available for freight forecasting: UTPS-type macro models, shipper/supplier decision models, and equilibrium models. These classes are not entirely exclusive. For example, mode choice models which describe the shipper's choice of transportation mode, are also an important component of macro models in the first group. Equilibrium models are based on shipper/supplier decision models. Selected models in these three groups are reviewed later in this chapter. In addition, several studies exist on more specific aspects of freight transportation, e.g., tariff determination and vehicle energy efficiencies. Some of these studies are also summarized in this report.

Improvements can be made in several areas. The time dimension has not been explicitly included in shippers' mode choice models. Mode choice are volume are influenced by seasonal variations (Winston, 1981) and/or the production cycles of firms. Dynamic models which account for these variations should be developed.

Another area where further research is needed is in the development of fuel switching models for carriers. The introduction of alternatively fueled vehicles in carriers' fleets, triggered by Federal and State regulations, is likely to affect carriers' freight transport charges and consequently shippers' mode choice decisions. Cost models describing the cost of owning and operating alternatively fueled vehicles need to be developed in order to determine penetration rates of different alternative fuels. Production and marketing of alternative fuels are still in their infancy and the future demand for each fuel type can be better assessed through a scenario analysis where various rates of penetration are analyzed. The development of fleet purchase behavior by carriers is therefore an extremely important research area. Hill (1987) is a pioneering study of carriers' demand for electric light utility vehicles.

a. Proceedings

Several consolidated publications of articles on freight demand and energy forecasting are available. In particular, two special issues of the *Transportation Research Journal* have been published on this subject area. The first special issue, entitled "Intercity Freight Modelling" and edited by T.L. Friesz and R.L. Tobin (1983) contain papers that address the issues of more realistic models of network equilibrium, transportation demand models reflecting known behavioral aspects, more realistic prediction of cost and rates, and better operations models. All these elements are crucial for energy use estimates. The second issue, "Transportation Systems and Logistics" edited by L. Bianco and A. la Bella (1988) represents the state of the art in logistics and freight supply.

Prior to these special issues, the National Research Council organized a workshop with the intent of providing a forum for discussing the importance of accurate freight demand forecasting and the problems associated with present forecasting procedures. Its participants were planners, engineers and economists who shared an active interest in freight movement and forecasting. The papers contained in the proceedings (NRC, 1978) address a broad range of issues related to freight transportation and forecasting. Their brief summary is given below.

Klein and Loxley (1978) use a two stage model structure to estimate freight demand. In the first stage, an input-output analysis is performed to obtain economic, demographic and market projections. These are then used to estimate the total ton-miles for each commodity. A modal split model is developed based on various mode characteristics such as tariffs.

Wild and Bentz (1978) propose a more elaborate model structure in which shipments at the national level are first estimated, which are then broken down to regional level forecasts. A detailed listing of the input data and output projections produced by the models is offered in the paper. Herwald (NRC, 1978) discusses how improvements in transportation technology influence firms' decision-making processes. He studies the case of Westinghouse Electric Corporation and observes the manner in which it grew and spread out. He also observes the modal split for freight movement over the years and notes how improvements in road transportation pushed freight off the railroad. Lawrence (NRC, 1978) uses econometric models to forecast long-range demand for motor freight carriers. He provides some estimates of modal shares and freight revenue, compiled by the IU International Management Corporation in 1978.

The proceedings also contains a few presentations on future transportation demand for nondurable goods, bulk commodities and coal. Also presented are panel discussions on data requirements for forecasting freight transportation demand. The panelists were in favor of using disaggregate data. The data used in the models presented in the proceedings are economic, demographic and transportation related. Population growth by region, employment growth, production and demand of commodities, vehicle stock, fuel economy and level-of-service attributes of different transport modes are among the major variables of the models. Most of these data are available from the Bureau of Census. The sources included the Census of Transportation, the Census of Retail and Wholesale Trade, Census of Manufactures, and Census of Population and Housing.

Papers in Freight Transport Planning and Logistics

Bianco and la Bella (1988) offer a collection of papers presented in Bressanone, Italy, in July 1987, at the International Seminar on Freight Transport Planning and Logistics. The papers in this volume span from simple trend analyses to more complex simulations of freight flows to logistics. Some of the papers are similar in their contents to those included in our review. The volume contains a summary of research in the United States, Canada and Europe, thus constituting a good depiction of the state of research in freight transport.

b. Freight Transportation Demand Forecasting Models

Several large-scale freight transportation demand forecasting models have been developed. These are similar in their structures to commonly used urban passenger travel demand forecasting systems. Trip distribution (determination of the volume between each pair of origin and destination zones) and modal split (assignment of shipments among different modes of transportation, e.g., air, truck and railroad) are often their key components.

Suits' Model

One of the first motor-truck demand models was estimated by Suits (1956) for a report to the Ford Division. The demand for three size-classes of trucks is modeled-light trucks with gross vehicle weight of 10000 lbs and under, medium trucks of 10001 to 16000 lbs, and heavy trucks over 16000 lbs. The model is used to calculate the expected demand for each type of truck in 1956. This is a historical example of an industry-sponsored, publicly available research effort in econometric motor vehicle demand modeling. Least squares regression equations were estimated for each class of trucks. First differences, the increase or decrease in a variable value over the preceding year, are used to center the analysis on changes in demand rather than on absolute levels. The period analyzed was 1935-1955, because of data availability. The period of the second world war (1942-45) was ignored. The dependent variable in the regression equation was number of new registrations of trucks while the explanatory variables included gross national product, existing stock of trucks per billion dollars of goods moved, and income of farmers.

Vinod's Model

The study reported by Vinod (1969) is probably one of the first studies in freight transport demand in which disaggregate choice models are applied. Model components are sequentially organized as follows. First national growth is established and outputs of different commodities are computed by simple projection. Second, the national growth is distributed across regions. Third, commodity flows among regions are computed, and finally modal split models are used to determine freight flows by mode between regions. The variables used in the model system include: shipment weight, value of shipment, tariff rate, shipment distance, employment rate, per capita income, and retail sales.

Two linear modal split models are used. The first model is based on "typical" commodity weight and shipment distance. The second model considers the proportions of shipments falling in categories defined by cost and weight. The transportation modes considered are rail, motor carrier, private truck, air, and water. Sixteen composite shipper groups (commodity types) and the 25 SMSA regions are used in the study. The data used are from the 1963 Census of Transportation.

Carlton's Model

The work by Carlton et al. (1975) aims at developing an approach to intercity freight modal split analysis. The motivation underlying the analysis is to develop the ability to evaluate proposed innovations in freight systems. The modeling effort is based on the "abstract-commodity" concept in which each commodity to be shipped is represented in terms of generic measurements of its attributes, e.g., shipment weight, distance shipped, value per ton and tariff charged. Two types of modal split models are estimated. The first is the linear regression model applied to explain the proportion of shipment made by each mode. The second is the logit model where the probability that a particular mode is chosen is expressed as a function of attractiveness measures (or "utilities") of all available modes. It is assumed that the shipper uses that mode which has a maximum utility.

The models are applied to the Boston production area, and the proportions of each commodity shipped by truck and rail are estimated. The variables used in these models can be grouped into two. The first group consists of commodity characteristics, including: value per pound, density, shelf life, method of inventory, annual use volume and need for special services. The second comprises transportation mode characteristics: travel time, waiting time, time reliability, loss and damage, special services, minimum shipment size and transport tariff. The 1972 Census of Transportation is the major source of data for the models estimated in this work. The U.S. Interstate Commerce Commission provided data on the commodity characteristics and the regional demand and supply of commodities.

Bronzini-Miller Model

The report by Bronzini and Miller (1979), part of a series prepared for the U.S. Department of Transportation, describes the development of a demand forecasting model and the prediction of the impact of energy conservation strategies on intercity freight transportation. This report is the documentation of a transportation network model, which is a component of the Transportation Systems Center (TSC) Freight Energy Model. The transportation network model assigns origin-destination commodity flows to modes and routes. It also determines traffic load on each network link, transportation cost, transit time, and energy usage. Examples on input-output methods of economic analysis, computational methods used, and programmer instructions are provided in the report.

The network model is defined for the 173 economic regions defined by the Bureau of Economic Analysis (BEA). Commodity flows among the regions are exogenous (pre-determined) to the model. Network link flows are estimated using a shortest path algorithm. The model system includes cost functions, time functions, energy functions, capacity constraints, mode constraints, and inertia effects.

Kim-Hinkle Model

Kim and Hinkle (1982) advocate the use of UTPS-style approaches to freight demand modeling and forecasting ("UTPS" stands for "Urban Transportation Planning System," which is a passenger demand forecasting package that consists of four main components, trip generation, trip distribution, modal split, and network assignment). The authors claim that small modifications to the existing UTPS package can produce acceptable statewide freight transport estimates. The issues they analyze include impacts of deregulation, rail mergers, areawide socioeconomic changes, rate changes, energy availability and level of service changes. The model system proposed is made of five components: network analysis models, freight transport demand analysis models, vehicle requirements models, assignment model, and evaluation models. The application of the model system with real data is not presented (the authors reference a practical application of their methodology in Korea).

Southworth's Model

The freight planning model by Southworth, et al. (1982) was developed to aid the Chicago Area Transportation Study (CATS) in devising a Year 2000 Motor Freight Development Plan. The purpose of the study was to improve the efficiency of the existing roadway network for freight and passenger transport and to model the location of truck terminal activity and its impacts.

The model systems consist of total truck trip generation, distribution and assignment models. A multiple linear regression model is used to estimate daily truck trip generation. The distribution is achieved by modeling the number of trips between origin-destination pairs as a function of travel time. This trip matrix is then input to the mixed person-freight assignment model to estimate link volumes. The assignment model estimates peak period travel times and peak period travel volumes. A non-linear optimization procedure is used in the assignment to obtain equilibrium network flow using Wardrop's principle that each user uses that route which minimizes his origin-to-destination travel time.

The terminal location analysis is based on the premise that the most influential factors are the proximity to demand, proximity to other carriers, proximity to major interstate and urban highways, and land rents and prices. The accessibility of each location is evaluated by calculating accessibility indices which are assumed to be functions of trip generation, trip distribution and travel time. These spatial accessibility indices are used to construct contour maps where locations with equal accessibilities are joined together. The final step involves maximizing system access to service demands to determine the number of trucks that would be served at the terminal.

The models were applied to the Chicago area using the CATS planning network to simulate total vehicle miles, average speed, and air quality standards under various policies. All network data consisting of traffic volumes, travel times, cost functions, node and link characteristics, land use information, etc. were obtained from the CATS. Commodity and commercial vehicle data were also available from the CATS. The commodity data required included the perishability, price and demand. Commercial vehicle characteristics included the miles traveled, age, range of operation, and fuel use.

California Freight Energy Demand Model

The California Freight Energy Demand Model (Jack Faucett Assoc., 1983) was developed to assist the California Energy Commission (CEC) in forecasting rail-freight and truck fuel consumption in California and in estimating future fuel consumption under alternative scenarios. Projections are developed for 17 activities/commodities with the state divided into five regions. The truck fleet projections are developed for 40 categories of size and body type.

The freight model consists of three programs: a Growth Program, the Main Program and a third program which performs some preliminary data aggregation. The Growth Program develops projections of commercial-use automobile stock (for use in the Personal Vehicle Model), ton-miles of rail/truck competitive freight, and VMT of commercial-use trucks. Commercial-use automobiles are distinguished on the basis of size, fuel type and place of manufacturing. There are 13 categories in all. Changes in the number of vehicles in each of these categories are projected based on projected changes in employment in those economic sectors which are presumed to be the major users of commercial automobiles, i.e., services, retail and wholesale trade. Intrastate freight movement is distinguished from interstate movement, and trucks are assigned between the two by age and by whether they are generally used for local or non-local operations.

Projections of changes in rail/truck freight traffic and commercial-use truck activity are assumed to be proportional to changes in a set of economic indicators which correspond to the 17 commercial use activities. All indicators are constructed from economic projections for California. Projections of total competitive freight traffic are obtained by commodity, by applying the corresponding ratios of economic indicators to base-year competitive freight traffic. For each commodity, the resulting projections are assumed to be distributed over all the regions and origin-destination regions in the same way as they were in the base year. Preliminary projections of truck ton-miles are obtained in the same way. Preliminary VMT estimates for heavy vehicles are obtained by dividing the ton-mile projections by the effective average payload of the vehicles, while those for light trucks are obtained by applying the economic-indicator ratios directly to the corresponding estimates of base-year VMT.

The main program performs a series of calculations:

Projection of fuel efficiency by truck type and railroad car type: The fuel efficiency of trucks is assumed to be a function of vehicle type, fuel type, range of operation (local or non-local) and load carried. The model assumes that any reduction in efficiency due to aging is compensated by improved technology and driving practices.

Estimation of changes in truck and rail transport costs for competitive freight: Changes in modal costs are calculated on the basis of changes in fuel efficiency and exogenously specified changes in various factors that affect transport costs such as diesel prices, labor costs, capital costs and federal and local taxes.

Modal diversion: The Modal diversion is then estimated based on the relative costs between the two modes (truck and rail) of freight movement.

Development of truck vehicle stock: Vehicle stock projections are developed in a series of steps. First, survival rates are applied to the previous years vehicle stock and the VMT provided by these vehicles is estimated. This gives the remaining VMT required. The number of new LPG and Methanol vehicles and the VMT they provide is estimated. The split between new gasoline and diesel vehicles is determined and a preliminary estimate of the number of such vehicles is obtained. These estimates are adjusted to be consistent with the desired local/non-local split.

Allocation of trucks to regions: LPG vehicles and their VMT are distributed over the five California regions in proportion to the base-year distribution of these vehicles. Methanol vehicles and their VMT are distributed in proportion as externally specified. The remaining VMT is allocated to gasoline and diesel vehicles.

Projection of truck and rail fuel consumption: Fuel consumption is estimated by multiplying ton-miles by the corresponding fuel-consumption coefficients and summing over commodities and modes.

The model system was applied to obtain projections in the state of California. A comparison of model estimates with previous year's data shows that the estimates are quite accurate. The data required are extensive, including the vehicle stock and level of transport activity, e.g., VMT by each category of truck and ton-miles of commodities transported by truck and rail. Average VMT, average truck payload, truck survival rates (the fraction of the fleet that remains in service in the following year), etc. are estimated as functions of exogenous variables. Additional information required includes projections of economic activities, fuel prices and other transport costs, vehicle fuel efficiency, and personal-use trucking. These projections were made using data available from the CEC and other agencies and applying various techniques, e.g., linear regression and exponential extrapolation.

Loxley's Model

A preliminary model of U.S. light-duty vehicle (<10000 lbs) demand and fuel consumption is presented by Loxley, et.al. (1981). The model is a long-run, annual, econometric model providing both forecasts and policy analyses and is an improvement over previous Wharton Models. Light trucks are divided into personal-use and commercial use vehicles. The personal-use light truck energy demand is estimated similar to automobile demand while commercial-use vehicle demand is estimated separately. Variables included in the model formulations are sales of vehicles, new registrations, scrappage, purchase prices, fuel economy, vehicle miles traveled, gasoline consumption, other operating costs, taxes and the influence of demographics.

Roberts-Greene Model

Roberts and Greene (1983) performed an analysis of recent historical heavy truck fuel data and projected future trends in heavy truck fuel economy and fuel consumption. The unique feature of this study is the in-depth consideration of the various technological improvements which would lead to improvements in fuel economy. Literature estimates are used to arrive at fuel economy improvement figures for various technological improvements. Market penetrations of various technologies are projected based on the Weibull distribution. Fuel economy improvements possible from technological advancements are compared with those possible with current technologies only. It is estimated that savings from advanced technologies could come to 4.1 billion gallons of fuel per year by the year 2000. The main serious drawback of this study is that they do not account for the increasing volume of commodities shipped by heavy trucks and the resulting increase in heavy truck VMT. They also do not account for federal mandates on fuel economy. For example, in 1982, the U.S.D.O.T. set rules that the average fuel economy standards for two-wheel drive and four-wheel drive light trucks should be at least 18.0 mpg and 16.0 mpg respectively (NHTSA, 1980).

Oum's Model

A slightly more disaggregate model was developed by Oum (1979) where the demand for the three modes of freight transport was formulated and estimated as the intermediate inputs to the production and distribution sectors of the economy. One of the main problems with this research was the non-availability of adequate data on freight rates. Despite the data limitations, his model estimated modal shares quite well. The study found that there is significant competition between the three modes of freight transport and that there is a lag structure in the shippers' response to changes in freight rates. The application of this kind of derived demand model at a disaggregate level should be useful for analyzing multi-modal demands for freight transport, especially inter-modal price competition.

Winston's Model

A fully disaggregate model of shipper and receiver mode choice behavior based on the principle of utility-maximization was developed by Winston (1981). This model was estimated using data of 1975-76 from a large number of shipments covering a wide range of commodities, length of haul and origin-destination pairs. The samples were choice-based and weights were applied to account for this. Variables included in the model system are the quantity shipped, value of the commodity, freight and refrigeration charges, loss and damage rates, and mean and variance of transit time. The disaggregate model structure resulted in an improved understanding of the shippers' and receivers' decision making processes. This study found that freight charges are the most important determinant of mode choice. The other variables were found to be highly commodity specific in their influence of mode choice.

This study was applied later (Winston, et al., 1990) to determine the effects of deregulation on surface freight. Shippers' and carriers' profitability from deregulation was determined by comparing their costs and benefits under regulated and deregulated conditions. It was found that deregulation substantially improved the benefits accruing to users of all modes of freight transport.

Energy and Environmental Analysis, Inc. Model

In 1980, Energy and Environmental Analysis, Inc. developed a computer model called Technology/Cost Segment Model (TCSM) to forecast the consumer cost of improving domestic light duty truck fuel economy (EEAI, 1980a, 1980b). It uses an accounting methodology to forecast technology penetration based on disaggregate estimates of cost and fuel economy improvements for individual technologies. The model was used to predict the baseline fuel economies of the vehicles expected to be available in 1985. The dependent variable of the model is fuel consumption per mile, while the predictor variables included test weight of vehicle, horsepower of the vehicle, engine volume, rear axle ratio and transmission efficiency.

Later the same year, Energy and Environmental Analysis, Inc. developed a Medium and Heavy Duty Truck Fuel Demand Module in their Highway Fuel Consumption Model (HFCM) into which the TCSM was also integrated (EEAI, 1980c). The model structure is the same in this module as the one in the TCSM. The model is not able to account for improvements in the fuel economy of diesel trucks. Fixed survival curves and fixed annual VMT per truck vintage curves were used. These can not reflect changes in behavioral or economic parameters, however, which truck demand would be sensitive to.

Thus, it is probably understating the trend in fleet average fuel economy and is thereby overstating fuel consumption. The model outputs are average new vehicle fuel economy, new vehicle registrations, average fleet fuel economy, total vehicle miles traveled and fuel consumption disaggregated by gasoline and diesel fuel. Model inputs included historical data of annual new vehicle registrations, survival of new vehicles, annual VMT per vintage vehicle, annual diesel penetration of new vehicle registration and their diesel fuel economy (EEAI, 1980).

Cantwell's Data Description

The Transportation Systems Center maintains a database on light-duty trucks. Two reports prepared for the U.S.D.O.T. by Cantwell (1979, 1980) describe the scope, definition, collection, and recording of data on domestic and imported light-duty trucks sold in the U.S. during the model years 1955 through 1977. Some 80 attributes are used to describe each model, including name, transmission type, engine size, wheelbase, weight, capacity, fuel economy, performance, production volumes, price, options and cost of options.

Transportation Energy and Emissions Modeling System

The Argonne National Laboratory developed a broader Transportation Energy and Emissions Modeling System (TEEMS) which has a freight component called FRATE3 (Freight Responsive Accounting for Transportation Energy). FRATE3 uses an econometric approach to forecast the volume of goods movement and applies utility maximizing procedures to estimate probable mode shares for commodity and shipment groups with competing modal alternatives. A user manual for this model has been written by Vyas and Millar (1986).

c. Modal Choice and Other Shipper/Supplier Decision Models

Hartwig-Linton Model

A 1974 study by Hartwig and Linton attempts to identify an appropriate model structure and significant variables in freight demand modeling (Hartwig and Linton, 1974). Shippers' modal choice between truck and rail is the focus of the study. The authors develop discrete choice models including logit, probit, and discriminant analysis models. The explanatory variables used in the analysis are transit time, freight cost, reliability, and value of the commodity. The data base was created using freight bills containing information on origin-destination, freight charge, freight rate, shipment weight, routing, type and number of commodities, date shipped, and date received. The three types of models are found to perform equally well. The study concludes that freight cost, reliability, and value of the commodity shipped are the most significant factors affecting shippers' mode choice.

Oum's Model

A theoretical derivation of a freight transportation demand model is given in a 1980 Canadian study (Oum, 1980). The analysis is based on the neoclassical economic theory, i.e., minimization of a cost function given a performance function representing the production technology available to the shipper. The shippers' transportation sectoral (modal) unit cost is expressed as a function of freight rates, quality attributes of service and haul distance. Modal revenue share functions are derived from four alternative formulations of the transportation sectoral unit cost. Four "nested" models are considered: a general model, a model strictly independent of distance, a mode-specific with hedonic aggregator (quality-adjusted price function) model, and a model with identical hedonic aggregators. Each of the models is presented in versions including both speed and reliability, speed alone, and reliability alone.

Argonne National Laboratory Model

Another freight network model was developed by Argonne National Laboratory (1984) based on the equilibrium concept. The primary goal of this work was to develop and apply an accurate and efficient network model that can be used to predict the performance of a highly complex multicommodity, multimodal freight transportation system in a large region. In this model system, the freight transportation is characterized by a dichotomy assumed in the decision-making process. Some decisions are made by shippers, while others are made by carriers. Shippers decide the origin, destination and size of shipment, while carriers decide the location of the site of shipment, the mode, scheduling and routing.

First, the behavior of shippers is modeled under the following principles:

1) The shipper will select an origin/path combination that has the minimum delivered price.
2) The greater the delivered price, the smaller the shipper's demand.
3) The greater the demand for a commodity, the greater its price.
4) The choice of origin/path combination made by shippers can be specified by a decreasing function of its own delivered price.
5) Between any origin-destination pair, the delivered prices of all used paths are equal and less than that of any unused path.

The carrier's decisions are modeled in two stages. The first is inter-carrier behavior to account for the possibility of a shipment being made by two different modes. The principles of the behavioral model are:

1) The shipper determines the location of inter-modal trans-shipment sites.

2) The shipper determines the location of railroad inter-lining sites.

3) The only routes used will be those that have the minimum number of inter-lining routes.

4) Each carrier will use the shortest distance path.

5) Of the eligible routes, the one that maximizes the profit will be used.

6) If there is more than one potential originating carrier, the shipment is divided among all potential carriers.

The second stage addresses intra-carrier behavior to account for the scheduling and routing decisions made by carriers. Carriers are assumed to route shipments over their network to minimize total cost. A system optimization network problem is set up using Wardrop's second principle and solved using the Kuhn-Tucker conditions. Network link flows are determined.

At the end of the study, a few insights into how the rates charged by carriers could be modeled are given. The three basic approaches are: empirical (observing past behavior of carriers), simulation (simulate the competitive environment to obtain a rate function) and theoretical (make realistic assumptions regarding the economic environment and set up rate determining rules under which carriers will operate).

The data required include commodity flow data, i.e., regional demand and supply for each commodity to be shipped. This is obtained from the Bureau of Economic Analysis. Also needed are detailed network data such as nodes, links, traffic volumes, congestion delays, and cost functions on each link. This information can be obtained from the Transportation Systems Center of the U.S. Department of Transportation.

The model was applied in three different scenarios. The first was to study the northeastern region of the nation in depth. The model predicted almost half of the link loadings correctly, which was found to be better than the prediction by the Multimodal Network Model of Bronzini and Miller (1979). The model was then applied to the national rail network, and results were found to be better than the earlier application. There was an increase in the accuracy of prediction. Finally, the model was applied to the nation as a whole, and it was found that the model predicted 63% of the network link volumes accurately.

A multivariate system of costs and shares is estimated in the study (Argonne National Laboratory, 1984) using as explanatory variables, freight rates, commodity flows, link distances, and origin-destination transit time (time elapsed between the request for service and delivery of the shipment). The data source is the Canadian Freight Transportation Model (CFTM) data base, which contains 78 commodity groups and 69 geographic regions. The author selected eight commodity groups and two modes (trucks and rail) for the models estimated. Each commodity group is modeled separately. The results indicate that the price and quality elasticities vary from commodity to commodity and from link to link. Trucks dominate high-value and short-distance markets. Competition is observed in the low-value and short-distance market. Long-haul markets are largely dominated by rail.

Hassan's Model

Hassan (1981), with the intent of identifying causal relations in the decision structure of the shippers and provide guidelines for the carriers in improving their level of service, attempts to determine factors that affect shippers' modal choice. The two modes considered are truck and air. The analysis is confined to small shipments of manufactured goods. The study can be classified as an attitudinal analysis. Through the use of a logit model the importance of mode attributes on the shippers' choice was assessed. Mode attributes used are: origin-destination travel time, shipping rates, weight of shipment, and dummy variables for mode and type of commodity. The data were obtained from traffic flow data from the U.S. Census of Transportation Commodity Transportation Survey, and data on Commodity Attributes from the MIT Center of Transportation Studies Commodity Attribute Data File. Shipping rates were estimated with linear regression.

Chang's Model

Developing a tool for analyzing the impact of alternative policies on freight transportation is the objective of the study by Chang, Roberts and Ben-Akiva (1981). Disaggregate models are used to model firms' decision making behavior assuming shippers and carriers minimize costs and maximize profits. A firm is considered to have a hierarchical decision structure. First, it has to determine its location in the long run. Second, it has to decide on the level of production, technology to adopt, etc. Finally, it makes short-term logistic decisions.

In the development of the short-run freight demand model, the study states that a firm's choice on logistics can be modeled using either a deterministic cost function or a random cost function. In the first case, the cost function is assumed to be fully observable, while in the second case, it is assumed that there is an unobservable random component in the cost function. The deterministic cost model involves the setting up of a cost function to be minimized under certain constraints and linear programming solution procedures are adopted.

The disaggregate data base used in the study involves the following four broad categories of variables:

1) Level-of-service attributes for each mode, shipment size, and shipment origin and destination. These include waiting time, transit time, reliability, tariff, loss and damage, and delay time. These are available from the Census of Transportation and the USDOT.

2) Commodity attributes which describe the characteristics of the commodity being transported. These include variables such as commodity value, density, shelf life and perishability. These attributes are available from the disaggregate commodity attributes file assembled by the Freight Study Group at MIT.

3) Receiver attributes describing the characteristics of the receiving firm include the annual use rate of the commodity being ordered, the variability in use rate, the risk of stockout, the type of inventory system used, and the storage and capital costs. Receiver attributes are available from the Census of Wholesale and Retail Trade.

4) The fourth category of variables describes the market for the commodity being shipped. Variables of this type include price, supply availability, etc. and data is taken from the Census of Manufacturers.

The model is applied to estimate modal split and shipment size for transporting freight between Houston and Chicago, New York and Pittsburgh and between Los Angeles and Boston. The model reasonably replicated the observation in the data. However, the model tended to predict a distribution of demand spread across modes even when observed demand concentrated on one mode.

d. Equilibrium Models

More rigorous treatments of freight transportation that embodies the concept of economic demand-supply equilibrium can be found in several studies.

Friedlander's Model

Friedlaender (1976) probes into market and network equilibrium conditions for freight transportation and defines a theoretical framework of a comprehensive model system for demand forecasting. The proposed model system consists of four interactive components: a regional transportation model, a national input-output model, a national macro model, and a regional income model. The model is formulated at an aggregate level, thus making the model simpler to use. The study, however, involves no data analysis and no model components are statistically estimated.

Within the theoretical framework, however, direct and indirect effects of a policy action on each of the four modeled sectors can be assessed. This analytical work thus attempts to gain a comprehensive solution to the problem of transportation policy evaluation. Although the author presents some discussion on possible use of the model system in a changing environment, the model system is conceived primarily for static analyses.

Harker's Model

Another theoretical probe into demand-supply relationship under equilibrium conditions can be found in Harker and Friesz (1984), where a theoretical freight market prediction model is developed based on behavioral discussions. The basis of this network equilibrium model is neoclassical economic theory. The model system is formulated as a short-run simultaneous freight trip generation, distribution, modal split, and network assignment model. It is made of three components: demand, supply, and an integrating economic mechanism. The carriers (supply) are assumed to operate independently while maximizing their respective profits. The demand-side is represented by a spatial equilibrium model. In this study, however, no specific model forms are assigned to most of the model components, and no empirical analysis is performed. The integrating mechanism is also defined only in general terms. The variables considered in the theoretical framework are costs and other network attributes.

This work is further extended in Harker (1987) in an attempt at providing a method for predicting the quantity of freight shipments between cities, the costs associated with these movements, and the volume of freight flow on the network by mode. Harker claims that econometric models do not consider the characteristics of transportation networks in detail, and thus ignore the complexities of a transportation system. Harker further notes that spatial price equilibrium models consider transportation networks while representing the interactions among producers, consumers and shippers. However, the interactions involving carriers are left out of the analysis in this type of approach. Transportation firms are characterized by network cost functions, but their decision processes are not explicitly dealt with. He also notes that freight network equilibrium models focus on the relationship between shippers and carriers, but without paying due attention to the importance of the role of producers and consumers in the transport of goods.

The model developed by Harker is a synthesis of the freight network and spatial price equilibrium approaches. He calls it the Generalized Spatial Price Equilibrium Model, or GSPEM, which attempts to overcome the limitations of earlier research by incorporating behavioral models of the producers, consumers, shippers and carriers into a single mathematical framework. The GSPEM consists of sub-models for each agent in the system.

Carriers are assumed to operate under the following conditions:

1) Carriers serve their supplies on a fixed network. Hence the economic time frame considered in this analysis is the short-run.
2) Each carrier is a profit maximizing firm.
3) Carriers do not collude when setting supply levels.
4) Each carrier attempts to minimize the cost of producing a given set of outputs.

Each carrier's network consists of a subset of geographic regions linked together by a specified level of freight transportation service. Alternative service levels as well as alternative modes are represented using multiple arcs between the same region pairs.

Shippers decide the quantity to be shipped, origin and destination of the shipment, and a carrier for the shipment. Shippers are assumed to operate under the following conditions:

a) Shippers possess no market power in the market for transportation services. Hence, they take the price of this service as given.
b) Shippers compete noncooperatively for the services of the carriers
c) Shippers attempt to minimize the cost of shipping goods.

The supply and demand models are then integrated in order to determine the equilibrium freight flows in the carrier's and shipper's networks as well as the prices in the economy associated with these flows.

The GSPEM conceptualizes freight movement using the following equilibrium conditions:

a) Carriers individually minimize the cost of transporting goods over their networks.

b) Carriers price according to demand (actually, a "rate function" is used to describe the way carriers fix the price of services under different scenarios).

c) Supply equals the demand in each transportation market.

d) Shippers individually minimize the cost of shipping goods over their network.

e) Shippers send goods between regions only if it is economically attractive to do so (i.e. shippers act according to a spatial price equilibrium).

f) There is a conservation of freight flows in every region.

Under these conditions, finding an equilibrium freight flow involves the solving of a constrained optimization problem where the costs of shippers and carriers are minimized. This optimization problem can be solved using the Frank-Wolfe algorithm.

This model is applied to the U.S. coal economy of 1980, and predictions of freight flow for coal are made. Harker first estimated the national supply function using three-stage least squares regression estimation and then apportioned this to different regions. Regional demand functions were developed similarly. The application included 16 rail carriers and 2 water carriers in this analysis. It was assumed that these carriers used marginal cost pricing in determining their freight rates. The carrier network had 7688 arcs, 2577 nodes and 4245 origin-destination pairs while the shipper network had 6993 arcs, 960 nodes and 1238 origin-destination pairs. The Frank-Wolfe algorithm was applied and the base year (1980) predictions were obtained. It was found that the predicted supply and demand for coal in each region fitted the actual values quite closely. However, there were marked differences between the predicted and actual coal movements along individual links in the network. The author suggests that carriers were probably making suboptimal routing decisions in 1980 when this model was not available.

The data requirements for the model are rather extensive. First, information is needed on the regions from, to, or through which freight will flow. Second, regional supply and demand functions are needed. If these are not available, they can be estimated from the national demand and supply figures using a disaggregation procedure. Additionally, network data are needed for each carrier and shipper, including the number of arcs, nodes, arc attributes, and origin-destination pairs. Also needed is a marginal cost function for each carrier arc, which was estimated in the study from the average cost functions developed as part of the National Energy Transportation Study (Bronzini, 1979).

A few limitations are pointed out by the author suggesting the need for further research. First, the simple supply and demand functions used in GSPEM may not be capable of capturing complex market behaviors. Second, the fixed capital and administrative costs need to be better understood in order to determine how these costs are allocated to specific moves. Third, the freight rate-setting behavior of carriers needs to be understood more clearly. Finally, an algorithm which makes computations easier and faster could make it possible to lift some of the assumptions that are required as the model is currently formulated.

D. Transit Forecasting Models

Transit energy forecasting models are reviewed in this section. These models estimate transit fuel consumption by predicting the market share of transit and the number of transit trips. These estimates are obtained from aggregate socio-demographic and economic characteristics of the population by modeling trip generation and car ownership as a function of these characteristics. The estimates are obtained for various scenarios of fuel prices and availability. The percentage share of transit trips is assumed to be determined by distance and time of trip. No behavioral relations are incorporated into predicting the share of transit and number of transit trips. This void could be filled by incorporating transit into the personal vehicle model proposed earlier.

Bruck's Model

A study by Bruck, Garrison, Reinhardt, and Skiles (1981) was undertaken as a strategic analysis of the petroleum situation and how it would affect travel behavior and public transportation. They developed five scenarios reflecting the manner in which the petroleum situation was most likely to evolve and affect urban public transportation during the 1980-2000 period. Analysis of the consequences of the situations described in each scenario pointed to the likelihood of reduced travel in the remainder of the century, and gradual increase in mass transportation use except in supply disruption, or major price shock situations. The study pointed to likely increases in transit use, offset to an extent with rising transit fares, if these eventualities were to occur.

The five scenarios, three general and two special ones, were as follows:

1) Scenario 1: Optimum Response. In this scenario, all instruments available to the public sector are utilized. Implementation of substitute technologies was accelerated

where feasible. Petroleum price increased gradually. Petroleum imports were reduced to approximately 3.5 million barrels per day in 1990 and 1 million barrel per day in 2000.

2) Scenario 2: Reliance on Market Forces. Here, major energy policies of the early 1980's remained unaltered, with the exception of a more gradual approach to synthetic fuels production than envisioned in the 1980 legislation. Market forces were depended upon to induce continued reduction of petroleum use and technology substitution. Petroleum price increased gradually. Petroleum imports were reduced to approximately 4 million barrels per day in 1990 and 1.5 million barrels per day in 2000.

3) Scenario 3: Reversal. The basic conditions were the same as in Scenario 2, except that there was a period of petroleum price stability in the early 1980's. The response was identical to that in the 1974-78 period. Petroleum use increased, and conservation oriented behavior was reduced. The Organization of Petroleum Exporting Countries (OPEC) responded by limiting production and sharply increasing price. The resulting oil shock was more severe than the previous ones and resulted in a prolonged recession and high unemployment. Much time was required to reinstate policies for reducing petroleum use.

4) Special Scenario A: Supply Disruption. A supply disruption occurred in the Persian Gulf causing a shortfall of approximately twenty percent. Because of relatively high levels of supply in oil company storage facilities and availability of petroleum from the SPR, a six-to nine-month time cushion was available to prepare for the shortage situation. If contingency plans, including policies such as pricing, supply allocation measures, rationing, etc. had been developed and could rapidly be made operational,

dislocation resulting from the supply disruption could be limited greatly. The principal points made were that a future supply disruption was unlikely to have the sudden character of the 1973-74 disruption and could be dealt with by careful contingency planning.

5) Special Scenario B: Transit Prices. The basic conditions in this scenario were the same as those in Scenario 2: gradually increasing petroleum prices and prevalence of market forces. Transit fares doubled over a period of 4-5 years. The situation was one of rising gasoline prices and more rapidly rising transit prices.

The travel behavior implications of the conditions described in the scenarios were then examined. The most general implication of rising petroleum prices postulated in all scenarios was reduced mobility during the remainder of the century. Utilization of public transportation would increase under the conditions described in all scenarios, gradually in the situations of Scenarios 1 and 2, greatly in the price chock and supply disruption situations, and very gradually in the greatly increased transit fares scenario.

Among the scenarios, the highest probability was assigned to Scenarios 2 and B. These are identical except for the added factor of greatly increased transit costs in Special Scenario B.

The RMC Model

The RMC demand model developed by Roche and Lago is a micro cross-section national intercity passenger transportation model. It predicts household trip rates by mode and purpose of travel for five different income groups, four different distance markets, and three different groupings for the number of persons on a trip (party size). Modes of travel examined included auto, bus, rail and air. The model was developed to investigate energy consumption rates for intercity passenger transportation.

Trip purposes investigated include: 1) Business and convention travel, 2) Visit friends, relatives and family business travel, and 3) Travel for recreation, sightseeing, entertainment, and other.

Four distance markets were distinguished:

- 100 to 250 miles
- 250 to 600 miles
- 600 to 1,250 miles
- over 1,250 miles

Three groups of party size were distinguished:

- One person trip
- Two person trip
- More than two persons on the trip

Demand equations were estimated for each mode and travel purpose through regression analysis. Independent variables used in the models include:

- Per person fare
- Travel time
- Hourly wage corresponding to a given income group
- Time valuation as proportion of hourly wages for a given trip purpose

The income elasticities reported were close to 1.0 for business travel and ranged from 0.20 to 0.50 for family and recreation travel. The income elasticities were always higher for air and auto and very low for bus. Distance elasticities were positive for air (denoting the competitiveness of air travel with increasing trip distance) and negative for other modes. The reasonableness of these estimates strengthens the overall reliability of the results of the subsequent simulation of energy conservation policies.

The demand model was used to obtain appropriate estimates of value of time, depending on the significance of the cost, minimum cost ratio, and the multiple correlation coefficient for each equation. The model is limited by data quality and statistical reliability. Data was obtained from the National Travel Survey (NTS) of the Census Bureau, which restricts the intercity passenger transportation market to trips exceeding 100 miles. The NTS was based on a sample of 24,000 households making 75,101 intercity trips. The intercity trips excluded military travel, student travel to and from school, community to work trips, and round trips of less than 200 miles.

In addition, adjustment factors were used in conjunction with projections obtained from the NTS data base to correct for sizable differences found between intercity passenger miles and vehicle miles reported in the NTS and those reported by other official sources. Only for air passenger transportation were both sources in agreement. Figures from AMTRAK and the Federal Highway Administration's Nationwide Personal Transportation Study (NPTS) substantially exceeded those reported to the National Travel Survey.

A problem of multicollinearity among the independent variables (particularly the correlation of trip distances with fares, costs, and travel times) -- was so pervasive that it was necessary to substitute modal costs and travel time ratios in lieu of each mode's travel time and costs. Thus, the estimated price elasticities refer to ratios of each mode's costs to the lowest costs exhibited by any mode. The use of costs and time ratios rather than the actual costs and travel times thus limits the usefulness of the model.

The main strengths of the model are its simplicity, the additive nature of its errors, and the reasonableness of the coefficient estimates. Most multi-modal demand models are of the chain model variety, in which total travel is forecasted first, with modal share forecasts being applied next to the total travel forecast. In these chain models, the errors have a multiplicative nature. The RMC model is a modal trip model (with one equation for each mode and travel purpose), and the errors are additive rather than multiplicative.

The main weaknesses of the RMC model concern: 1) its lack of separate own-price elasticity for each mode due to the multicollinearity between fares and distances; and 2) the incompleteness of the National Travel Survey data base, which makes it necessary to apply correction factors to derive comparable national travel estimates.

Hartgen-Erlbaum Model

Hartgen and Erlbaum (1980) developed a model in which long-range baseline forecasts of energy use in each of the subsectors of the transportation sector in New York State were prepared using a variety of techniques relating energy use to:

- Projections of energy supply
- Real Gross State Product
- Disposable income
- Inflation
- Population
- Employment
- Fuel price

Projections of these variables were provided by the New York State Energy Office. The forecasting methods are generally not computerized and rely on simple curve fitting procedures.

Transit energy was defined as consisting of several major parts:

- Subway
- Public bus
- Commuter rail systems
- School buses
- Intercity

The study assumes an annual 1% increase in VMT and energy use in NYC area bus and rapid transit and an annual 0.5% increase in VMT and energy use of commuter rail and upstate bus.

School bus VMT was constant for a few years preceding the study, and while population in New York State was expected to rise, the general tendency toward smaller families suggested the school age population would rise at a slower rate than overall population. Therefore, a 0.5% annual rate of growth for school bus VMT was assumed for New York State. An annual growth rate of 1.0% in intercity bus VMT was assumed.

A straightforward extrapolation of known energy requirements produced five-year forecasts for a twenty-year period. Intercity passenger rail travel in New York State came directly from earlier research by the authors. Forecasts were made for rail passenger travel in the NYC-Buffalo corridor for the years 1975-1980. These forecasts accounted for the train, track, and service improvements planned by AMTRAK. Using common conversion factors, the energy use was converted into equivalent gallons of gasoline and equivalent gallons of diesel fuel. Rail Division staff at NYSDOT indicated that 1995 rail passenger volumes would be three times that of 1980 levels, increasing from 980,000 to approximately 3 million passengers. Based on that assumption, the energy used for rail passenger travel in New York State was estimated to increase 100% by 1995, from 652 x 109 BTU in 1980 to approximately 1300 x 109 BTU.

Air passenger aviation fuel constitutes a substantial share of New York transportation energy because of the major national and international services in the NYC region. Forecasts were made relating current year fuel use to seat-miles, revenue passenger miles per capita, projected population, load factor, and an index relating future travel to the ratio of future business activity to the 1975 business activity. Air carriers were projected to increase 142% overall or 4.5% annually.

These long-range curve-fitting projection techniques lend themselves best to providing baseline projections of energy consumption for quick responses to a variety of policy questions.

Obeng's Model

The hypothesis that transit inputs have inelastic demand since transit demand itself is inelastic was tested by Obeng (1984) using a sample system, the City Transit Division (CDT) of the Southeastern Pennsylvania Transportation Authority (SEPTA) in Philadelphia. A homogeneous transcendental logarithmic (translog) cost function was developed from which demand surfaces for the various inputs were derived. Direct and cross-elasticities of demand were calculated from the various demand surfaces.

It was concluded that labor demand is inelastic. However, electricity demand was found to be elastic in the relevant range and diesel fuel demand was found to be inelastic when diesel fuel's share of total operating cost was less than seven percent.

A translog cost model was developed because it was found that translog functions provide a flexible approach to the estimation of cost models which permits input demand to derived and, in turn, allows price elasticities to be determined. They also lend themselves well to allowing various economic hypotheses, such as economies of scale, to be tested.

Transit operating cost was defined as a function of:

- output in passengers
- price per kilowatt-hour of electricity
- average wage rate
- price per gallon of diesel fuel
- total number of passenger vehicles considered fixed

This function was expanded using Taylor series.

For transit systems, government supported services and most input prices are exogenously determined. There are also institutional constraints on deficits which require adoption of cost minimization policies. This implied that a linear homogeneity assumption could be placed on the coefficients of the translog cost function.

A constrained form of the cost equation and two share equations were estimated using Joint Generalized Least Square Estimation techniques and Zellner's method of handling seemingly unrelated equations. The constraints were handled by expressing the coefficients of terms involving electricity in terms of the other coefficients and substituting the result in the cost and share equations.

Demand elasticities were determined from the values of the coefficients obtained from the model. The elasticities show that diesel fuel demand is inelastic when diesel fuel's share of operating cost is less than seven percent. However, when diesel fuel's share of operating cost is seven percent or more, its demand is elastic. This implies running fewer buses, operating fewer miles of service than before, or, on multimodal systems, substituting rail for bus service where both use the same corridor. On single mode systems, modal substitution is impossible and diesel fuel price increases in the elastic range could have severe consequences on service levels unless a service subsidy is provided.

Elasticity of electricity demand decreases with increases in electricity's share of operating cost in the range considered. The absolute values of these elasticities are higher than those of diesel fuel. It was also seen that electricity demand is inelastic when its share of operating cost is less than 26.42%.

Kain's Model

Kain, et.al. (1977) propose a model which estimates auto ownership and mode choices based on household and urban structure characteristics. Two stages of the model are defined. One predicts the probabilities that a household will own zero, one, two or more automobiles; the other predicts the probabilities that employed household members will commute to work as auto drivers, auto passengers, bus passengers, or rail passengers, or by walking. Probabilities are based on a variety of household, transit availability, and urban structure characteristics.

Linear, recursive equations were estimated, consisting of disaggregated probabilities that, when added, result in the probability that a particular household will choose a level of auto ownership or a travel-to-work mode. The number of cars owned is based on:

- family type and composition
- income
- residential location

- employment location of workers
- density and age of housing stock in the region
- levels of available highway and transit service

These variables and the number of cars owned are used to explain the mode choice. There are thirty auto ownership and 120 mode choice equations, since the total sample of households was stratified by:

- race
- number of employed members
- work place location (CBD, central city, suburban ring)
- household type (husband-wife, single head, primary individual)
- age of head of household
- number of drivers
- number of children
- annual household income
- dwelling type and age

Urban structure variables included:

- road route miles per capita
- annual bus seat-miles per capita
- annual rail seat-miles per capita
- fraction of housing in region built before 1920
- fraction of housing that is single family

Mean values for the variables were used in making projections for aggregate regions.

Samples of data on 346,893 households and 407,731 workers in the largest 125 SMSA's in the U.S. from the 1970 Census were used. A projection for the year 1990 was done. It was predicted that there would be a shift towards more auto ownership by households, more commuting by auto driving, and a lower percentage of commuting for each of the other modes. With regard to the urban structure explanatory variables, a comparison was made with the model's predictions and the actual data for Boston and Phoenix. The model predicted well.

By looking at different cities rather than a single metropolitan area, it can be seen how different urban spatial structures and the levels of highway and transit investment affect household decisions. Also, the distinction between races illustrates how housing discrimination causes blacks and whites to respond differently to changes in land use patterns and transit supply.

VI. REFERENCES

Adler, T.J., J.W. Ison and J.C. Geinzer (1980) Interactions Between Energy Supply and Transportation Related Energy Use. Final Report. Resource Policy Center, Thayer School of Engineering, Dartmouth College.

Anas, A. (1987). *Modeling in Urban and Regional Economics*. Harwood Academic Publishers, New York.

Argonne National Laboratory (1984) Freight Network Modeling System. Freight Network Equilibrium Model-Theory and Validation. Prepared by Environmental Research Division. Technical Memorandum, Vol. I.

Ben-Akiva, M. and T. Morikawa (1989). Estimation of mode switching modes from revealed preferences and stated intentions. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Kyoto, Japan, July. *Transportation Research* (forthcoming).

Berkovec, J. (1985). Forecasting Automobile Demand Using Disaggregate Choice Models. *Transportation Research. B*, 19B(4): 315-29.

Berkovec, J. and J. Rust (1985). A Nested Logit Model of Automobile Holdings for One Vehicle Households. *Transportation Research B*, 19B(4): 275-85.

Berkowitz, **et.al.**(1985). A Dissaggregate Analysis of Multi-Vehicle Holdings, Usage, Gasoline Demand. Department of Economics and Institute for Policy Analysis. Working Papers Series. University of Toronto, No. 8508.

Bianco, L. and A. La Bella, Eds. (1988) *Freight Transport Planing and Logistics*. Lecture Notes in Economics and Mathematical Systems. Springer Verlag.

Bishop, G.F., H-J. Hippler, N. Schwarz and F. Strack (1988). A comparison of response effects in self-administered and telephone surveys. In R.M. Groves, et al., eds., *Telephone Survey Methodology*. New York: Wiley.

Bradley, M. (1983) Transferability of the ZVL Car Ownership Models. Memorandum, Cambridge Systematics Europe, the Hague, The Netherlands, 10 p.

Broecke, van den (1988) Long term forecasting of car ownership with the cohort processing model. Paper presented at the Fifth International Conference on Travel Behavior Analysis, Aix-en-Provence, France, October.

Bronzini. M.S. and R.C. Miller (1979) Freight Transportation Energy Use. Volume II-Methodology and Program Documentation. Final Report prepared on behalf of CACI, Inc. for the USDOT.

Brownstone, D. (1990). Multiple Imputations for Panel Studies. Working Paper, Department of Economics, University of California, Irvine.

Brownstone, D., P. Englund, and M. Persson (1988). A Microsimulation Model of the Swedish Housing Market. *Journal of Urban Economics* 23.

Bruck, H.W., W.L. Garrison, N. Reinhardt, and M. Skiles (1981). A Strategic Perspective on Public Transportation in a Time of Uncertain Petroleum Supply. Final Report No. UCB-ITS-RR-81-6. Prepared for the U.S. Department of Transportation.

Bunch, D.S. and R. Kitamura (1990). Multinomial Probit Model Estimation Revisited: Testing of New Algorithms and Evaluation of Alternative Model Specification for Trinomial Models of Household Car Ownership. Research Report UCD-TRG-RR-4, Transportation Research Group, University of California, Davis, CA.

Burkheimer, G.J. and J.R. Levinsohn (1988). Implementing the Mitofsky-Waksberg sampling design with accelerated sequential replacement. In R.M. Groves, et al., eds., *Telephone Survey Methodology*. New York: Wiley

Calfee, J.E. (1980). Potential demand for electric automobiles. Ph.D. thesis. University of California, Berkeley.

Calfee, J.E. (1985). Estimating the demand for electric automobiles using fully disaggregated probabilistic choice analysis. *Transportation Research B*, 19B(4).

Cameron, T.A. and M.D. James (1987). Estimating Willingness to Pay from Survey Data: An Alternative Pre-Test-Market Evaluation Procedure. Journal of Marketing Research, Vol. 24, November: 389-95.

Cantwell, C. (1979). Light Duty Truck Characteristics. Historical Data Base, U.S. Department of Transportation, Report No. DOT-TSC-NHTSA-79-37 (December).

Cantwell, C. (1980). Light Duty Truck Characteristics, Historical Data Base, U.S. Department of Transportation, Report No. DOT-TSC-NHTSA-80-22, November.

Cardell, Boyd, Mellman, and Dunbar (1974) The CRA Hedonic Demand Model, Charles River Associates, Inc., Boston, MA.

Carlton, R., B. Morgan, M. Costello, R. Tucker, and G. Wang (1975) Intercity Freight Mode Choice. Working paper. Transportation Systems Center, USDOT.

Cervero, R. (1985) Short-run Forecasting of Highway Gasoline Consumption. *Transportation Research A*, 19A(4).

Cervero, R. (1986) *Suburban Gridlock*, Center for Urban Policy and Research, New Brunswick, NJ.

Chang, Y.S., P.O. Roberts and M. Ben-Akiva (1981) Development of a Policy Sensitive Model for Forecasting Freight Demand. Final report prepared for USDOT.

Clarke, M., M. Dix, and P.B. Goodwin (1982). Some issues of dynamics in forecasting travel behaviour: A discussion paper. *Transportation*, 11(2), 153-172.

Cochran, W.G. (1977). *Sampling Techniques*. 3rd Ed. New York: Wiley

Cummings, R.G., D.S. Brookshire, and W.D. Schultze (1986). Valuing Environmental Goods: An Assessment of the Contingent Valuation Method, Rowman and Allanheld, Totowa, New Jersey.

Daganzo, C.F. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forecasting.* New York: Academic Press.

Daly, A. and H. Gunn (1986) Cost-Effective Methods for National Level Demand Forecasting. In *Behavioral Research for Transport Policy.* VNU Science Press. Utrecht, Netherlands.

Davies, R.B. and A.R. Pickles, (1985). Longitudinal vs. cross-sectional methods for behavioural research: A first-round knockout. *Environment & Planning A,* 17:1315-1329.

Duncan G.J., F.T. Juster, and J.N. Morgan (1987). The role of panel studies in research on economic behavior. *Transportation Research* 21A: 249-264.

Edwards (1977) Use of a Lowry-Type Spatial Allocation Model in an Urban Transportation Energy Study.

Efron, B. (1982). *The Jackknife, Bootstrap and Other Resampling Plans.* Philadelphia: SIAM.

Energy and Environmental Analysis, Inc. (1980a). Technical Analysis of the Light Duty Truck Fuel Economy Standards for 1985 Model Year. Department of Energy Report, May.

Energy and Environmental Analysis, Inc. (1980b). Technology Cost Segment Model for Light Duty Trucks. Prepared for U.S. Department of Energy, December.

Energy and Environmental Analysis, Inc. (1980c). Medium and Heavy-Duty Truck Fuel Demand Module. Final Report. Department of Energy, May.

Erlbaum, N.S., D. Hartgen, and G.S. Cohen (1977) Automotive Energy Forecasts: impact of price, availability and efficiency. Research Report No. 133. New York State Department of Transportation, Planning Division.

Friedlaender, A.F. (1976) An Integrated Policy Model for the Surface Freight Transportation Industries. Final Report prepared by MIT Center of Transportation Studies. DOT University Research Program.

Friesz, T.L. and R.L. Tobin, eds. (1983) Intercity Freight Modelling. Special Issue of *Transportation Research A,* 17A(4).

Geinzer, J. and A. Daly (1981). Zuidvleugel Study - Report 9, Models of Car Ownership and License Holding. Cambridge Systematics Europe, The Hague, The Netherlands.

Golob, T. (1987). The causal influences of income and car ownership on trip generation by mode. *Journal of Transport Economics and Policy,* 23:141-162.

Golob, T.F. (1989). The Dynamics of Household Travel Time Expenditures and Car Ownership Decisions. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Japan, July. *Transportation Research* (forthcoming).

Golob T.F. and L.J. van Wissen (1989). A Joint Household Travel Distance Generation and Car Ownership Model. *Transportation Research,* 23B: 471-491.

Goodwin P.B. and M.J.H. Mogridge (1981). Hypotheses for a fully dynamic model of car ownership. *International Journal of Transport Economics,* 8(3), 313-26.

Goodwin, P.B. (1989) Family Changes and Public Transport Use 1984-1987: A Dynamic Analysis using Panel Data. In *Transportation,* 16(2): 121-54.

Groves, R.M. (1987). Research on survey data quality. *Public Opinion Quarterly,* 50th Anniversary Issue, S156-S172.

Groves, R.M. and L.E. Lyberg (1988). An overview of nonresponse issues in telephone surveys. In R.M. Groves, et al., eds., *Telephone Survey Methodology.* New York: Wiley.

Groves, R.M. and R.L. Kahn (1979). *Surveys by telephone: A national comparison with personal interviews.* New York: Academic Press.

Gunn, H. and H. Pol (1986). The OVD Regional Travel Demand Forecasting System. In Planning and Transport Research and Computation (International) Co. Meeting. Proceedings of Seminar M held at the 14th PTRC Annual Summer Meeting.

Gunn, H.T., van der Hoorn, and A. Daly (1987). Long-range, country-wide travel demand forecasts from models of individual choice. Paper presented at the Fifth International Conference on Travel Behavior Analysis, Aix-en-Provence, France, October.

Gur, Y., M.M. Mintz, and R. Morrison (1987). Demographic Influences on Household Travel and Fuel Purchase Behavior. Transportation Research Record 1155. National Research Council, Washington D.C.

Harker, P.T. (1987) Predicting Intercity Freight Flows. Department of Decision Sciences, The Wharton School, University of Pennsylvania. Ph.D. Thesis.

Harker, P.T. and T.L. Friesz (1984) Prediction of Intercity Freight Flows-I: Theory. Preliminary report. Department of Decision Sciences, The Wharton School, University of Pennsylvania.

Hartgen, D. and N.S. Erlbaum (1980) Long-range Forecasts of Transportation Energy Consumption in New York State. Preliminary Research Report no. 158. New York State Department of Transportation, Planning Division.

Hartwig, J.C. and W.E. Linton (1974) Disaggregate Mode Choice Models of Intercity Freight Movement. Ph.D. Thesis. Northwestern University.

Hassan, H. (1981) A Behavioral Mode Choice Model Comparing Air and Truck Freight Transport. Ph.D. Dissertation. Institute of Transportation Studies, University of California, Berkeley.

Hague Consulting Group (1989). Creation of the Landelijk Model, National Traffic Forecasting System. The Hague, The Netherlands, April.

Heckman, J.J. (1981). Heterogeneity and state dependence. In S. Rosen, ed., *Studies in Labor Markets.*, Chicago, IL: University of Chicago Press:. 91-139.

Hensher, D.A. and N. Wrigley (1986). Statistical modelling of discrete choices in discrete time with panel data. In *Behavioral Research for Transport Policy*, VNU Science Press, Utrecht. The Netherlands:. 97-116.

Hensher, D.A. and N.C. Smith (1989). Dynamic random effects continuous choice modeling: estimating automobile utilization with panel data. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Kyoto, Japan, July. *Transportation Research* (forthcoming).

Hensher, D.A. and N.C. Smith (1990). Estimating Automobile Utilization with Panel Data: An Investigation of Alternative Assumptions for the Initial Conditions and Error Covariances. *Transportation Research,* 24A(6).

Hensher, D.A., P.O. Barnard, N.C. Smith and F.W. Milthorpe (1987). Dynamic random effects continuous choice modeling: estimating automobile utilization with panel data. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Kyoto, Japan, July. *Transportation Research* (forthcoming).

Hill, D.H. (1987). Derived Demand Estimation with Survey Experiments: Commercial Electric Vehicles. *Review of Economics and Statistics*, 69(2): 277-85.

Hocherman, I., J.N. Prashker and M. Ben-Akiva (1987). The transaction cost approach to modeling the demand for automobiles. *Journal of Advanced Transportation* 21(2): pp 173-193.

Immers, Scheltes, and Hamerslag (1988). Research into factors influencing car mobility in the Netherlands. Netherlands Center for Transportation Research Training and Consultancy Services. Rijswijk, Netherlands.

Jack Faucett Associates (1983). California Freight Energy Demand Model. Consultant Report prepared for the California Energy Commission.

Jong, G.C. de (1989a). Some Joint Models of Car Ownership and Car Use. Unpublished Ph.D. thesis, Faculteit der Economishe Wetenschappen en Econometrie, Universiteit van Amsterdam, The Netherlands.

Jong, G.C. de (1989b). Simulating Car Costs Changes Using an Indirect Utility Model of Car Ownership and Annual Mileage. Presented at the PTRC Summer Annual Meeting, Brighton, September.

Kain, J.F. and G.R. Fauth (1976) The effects of Urban Structure on Household Auto Ownership Decisions and Journey to Work Mode Choices. Urban Planning Policy Analysis and Administration Research Paper R 76-1. Department of City and Regional Planning, Harvard University, Cambridge, MA.

Kain, J.F., G.R. Fauth, and J. Zax (1977). Forecasting Auto Ownership and Mode Choice for U.S. Metropolitan Areas. Research Report R77-4. Department of City and Regional Planning, Harvard University, Cambridge, MA.

Kim, T.J. and J.J. Hinkle (1982) A Model for Statewide Freight Transportation Planning. *Transportation Research Record* No. 889. National Research Council, Washington D.C.

Kitamura, R. (1987). The asymmetry of the change in household car ownership and utilization: A panel analysis. In the proceedings of the Fifth International Conference on Travel Behavior, Gower Publishing, Aldershot, England (forthcoming).

Kitamura, R. (1990). Panel analysis in transportation planning: An overview. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Kyoto, Japan, July. *Transportation Research* (forthcoming).

Kitamura, R. and P. Bovy (1987). Analysis of attrition biases and trip reporting errors for panel data. *Transportation Research* 21A: pp 287-302.

Kitamura, R. and D.S. Bunch (1990). Heterogeneity and state dependence in household car ownership: A Panel Analysis Using Ordered-Response Probit Models with Error Components. Research Report UCD-TRG-RR-89-6, Transportation Research Group, University of California, Davis, CA.

Kitamura, R. and Goulias K.G.(1989) MIDAS: A Travel Demand Forecasting Tool Based on a Dynamic Model System of Household Car Ownership and Mobility. Final report for Projectbureau Integrale Verkeer- en Vervoerstudies, Ministerie van Verkeer en Waterstaat. The Netherlands.

Klein, L.R. and C.J. Loxley (1978). Transportation Demand-Aggregate and Major Freight Category Demand Estimation. In Forecasts of Freight System Demand and Related Research Needs. Proceedings of a Workshop. The National Research Council, Washington D.C.

Kulash (1976). Forecasting Long-Run Automobile Demand. Strategies for Reducing Gasoline Consumption Through Improved Motor Vehicle Efficiency. TRB Special Report No. 169:14-19. National Research Council, Washington, D.C.

Lancaster, T. and G. Imbens (1988). Choice-based sampling of dynamic populations. Presented at International Symposium on Panel Data and Labor Market Studies, Amsterdam, Dec. 15-17.

Leeuw, E. de and J. van der Zouwen (1988). Data quality in telephone and face to face surveys: a comparative meta-analysis. In R.M. Groves, et al., eds., *Telephone Survey Methodology*. New York: Wiley.

Little, R.J.A. and H-L. Su (1989). Item nonresponse in panel surveys. In D. Kasprzyk, G.J. Duncan, G.Kalton and M.P. Singh, eds., *Panel Surveys*, 400-425. New York: Wiley.

Loxley, C., P. Luce, T. Osiecki, and K. Rodenrys (1981). The Demand for Light Duty Trucks, The Wharton EFA Motor Vehicle Demand Model (Mark II). Philadelphia, PA: Wharton EFA, Inc. Report No. DOT-HS-805 667, January.

Mannering, F. and C. Winston (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics*, 16(2), 215-36.

Manski, C.F. (1983). Analysis of Equilibrium automobile holdings in Israel with Aggregate discrete choice models. *Transportation Research B*, 17B(5).

Manski, C.F. and E. Goldin (1983). An econometric analysis of automobile scrappage. *Transportation Science*, 17(4):. 365-75.

Manski, C.F. and D. McFadden (1981). Alternative estimators and sample designs for discrete choice analysis. In C. Manski and D. McFadden (eds). *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: M.I.T. Press.

Manski, C.F. and L. Sherman (1980). Forecasting Equilibrium Motor Vehicle Holdings by Means of Disaggregate Models. *Transportation Research Record* no. 764: 96-103.

Manski, C.F. and L. Sherman (1981). Forecasting equilibrium motor vehicle holdings by means of disaggregate models. *Transportation Research Record* 764: 96-103.

Mason, R.E. and F.W. Immerman (1988). Minimum cost sample allocation for Mitofsky-Waksberg random digit dialing. In R.M. Groves, et al., eds., *Telephone Survey Methodology.* New York: Wiley.

Massey,J.T., P.R. Barker, and S. Hsiung (1981). An investigation of response in a telephone survey, *Proceedings of the Section on Survey Research Methods, American Statistical Association*:. 426-431.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica,* 57(5): 995-1026.

Meurs, H.J. (1989). Dynamic analysis of trip generation. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto University, Kyoto, Japan, July. *Transportation Research* (forthcoming).

Meurs, H.J. (1990). Trip generation models with permanent unobserved effects. *Transportation Research* (forthcoming).

Millar, M., J. Bunch, A.D. Vyas, M. Kaplan, R. Knorr, V. Mendiratta, and C. Saricks (1982). Baseline Projections of Transportation Energy Consumption by Mode: 1981 Update. Argonne, IL. Argonne National Labs. Report No. ANL/CN SV-28.

Millar, E.J., P.J. Noehammer, and D.R. Ross (1987). A micro-simulation model of residential mobility. In Proceedings of the International Symposium on Transport, Communication and Urban Form:2, *Analytical Techniques and Case Studies* (edited by W. Young), Monash University, Clayton, Victoria:. 217-234.

Mitchell, R.C., and R.T. Carson (1989) Using Surveys to Value Public Goods: The Contingent Valuation Method, Resources for the Future, Washington D.C.

Mogridge, M.J.H. (1989). The Prediction of Car Ownership and Use Revisited. *Journal of Transport Economics and Policy,* 23: 55-74.

Mullen and White (1977). Forecasting Car Ownership: A New Approach. Parts I and II. *Traffic Engineering and Control,* vol. 18, part 1: pp 354-61, part II: pp. 422-26.

National Highway Traffic Safety Administration (1980). Final Regulatory Analysis of the Model Year 1982 Light Truck Fuel Economy Standards. Washington D.C.: U.S. Department of Transportation. Report No. DOT/HS-805291, March.

National Research Council (1978). Forecasts of Freight System Demand and Related Research Needs. Proceedings of a Workshop. Washington D.C. July.

Obeng, K. (1984). The Economics of Bus Transit Operation. In *The Logistics and Transportation Review* 20(1).

Oum, T.H. (1980). Demand for Freight Transportation with a Special Emphasis on Mode Choice in Canada. Ph.D. Thesis. Centre for Transportation Studies, University of British Columbia.

Oum, T.H (1979). Derived Demand for Freight Transport and Inter-Modal Competition in Canada. *Journal of Transport Economics and Policy*. Vol. XIII, No. 1:. 149-168.

Reza and Spiro (1979). The Demand for Passenger Car Transport Services and for Gasoline. *Journal of Transportation Economic Policy* 13(3):. 304-319.

Roberts, G.F. and D.L. Greene (1983). Trends in Heavy Truck Energy Use and Efficiency. Oak Ridge, TN: Oak Ridge National Laboratory. Report No. ORNL/TM-8843.

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Sebold, J. (1988). Survey period length, unanswered numbers, and nonresponse in telephone surveys. In R.M. Groves, et al., eds., *Telephone Survey Methodology*. New York: Wiley.

Smith, N.C., D.A. Hensher and N. Wrigley (1989). A Dynamic Discrete Choice Sequence Model: Method and an Illustrative Application to Automobile Transactions. Unpublished manuscript, March.

Smyth, A.W. and A.J. Daly (1985) Extensions and Improvements to Models of Driving License Holdings. Cambridge Systematics Europe, the Hague, the Netherlands. Prepared for Dienst Verkeerskune, Rijkswaterstaat.

Southworth, F., Y. J. Lee and D. Zavattero (1982) A Motor Freight Planning Model for Chicago. Research Report, publication no.9. University of Illinois at Urbana-Champaign.

Suits, D.B. (1956). Demand for New Motor Trucks in the United States 1935-1955. Ford Motor Company.

Survey Research Center, University of Michigan, (1972). *A panel study of income dynamics: Study design, procedures, available data*. Ann Arbor, Michigan, Institute for Social Research, University of Michigan.

Thornberry, O.T. and J.T. Massey (1988). Trends in United States telephone coverage across time and subgroups. In R.M. Groves, et al., eds., *Telephone Survey Methodology*. New York: Wiley.

Town, S.W. (1983). Reasons for car acquisition. In *Research for Transport Policies in a Changing World: Proceedings of World Conference on Transport Research*, Hamburg, Vol. 1, SNV Studiengesellschaft Nahverkehr, Hamburg:. 321-30.

Train (1986). *Qualitative choice analysis: theory, econometrics, and applications to auto demand*. MIT Press Series in Transportation Studies 10.

..od, H.D. (1969) Studies on the Demand for Freight Transportation, volume II: Forecasting the Freight Demand by Stages. Final report prepared by Mathematica for The Northeast Corridor Transportation Project, USDOT.

Vyas, A.D. and M. Millar (1986). User Manual for Freight Responsive Accounting for Transportation Energy (FRATE3) Model. Argonne, IL: Argonne National Laboratory, June.

Wadhwa (1980). Simulation and analysis of energy demand for passenger automobiles in Australia. *Transportation Research* 14A: 235-240.

Waksburg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association* 73: 40-46.

Wild, T.E. and Bentz (1978) Freight Travel Demand. In Forecasts of Freight System Demand and Related Research Needs. Proceedings of a Workshop. The National Research Council, Washington D.C.

Winston, C. (1981) A Disaggregate Model of the Demand for Intercity Freight Transportation. *Econometrica,* 49(4). 981-1006.

Winston, C., T.M.Corsi, C.M. Grimm, and C.A. Evans (1990) *The Economic Effects of Surface Freight Deregulation.* The Brookings Institution, Washington D.C.

Wissen, L.G. van and T.F. Golob (1990). A dynamic model of car fuel type choice and mobility. Research Report UCI-ITS-WP-90-3, Institute of Transportation Studies, University of California, Irvine. Prepared for Traffic Engineering Division, Public Works Administration, Ministry of Transport and Public Works, the Netherlands.

Wurzel, E. (1988). Distribution of single-spell durations in sample designs with time aggregated data. Discussion Paper A-162, Sfb 303, Universitat Bonn.