ENERGY 2020 Documentation

Volume 4

Theoretical Derivation

Last Revision: June 2017

Systematic Solutions, Inc.
Beavercreek, Ohio 45435

www.ENERGY 2020.com

Table of Contents

Theoretical Derivation

Theoretical Derivation

ENERGY 2020: CONSUMER CHOICE THEORY

This chapter provides a basis for using the multinomial logit in the consumer choice components of the ENERGY 2020 model. This discussion will primarily rest on extensive quotations and the research of others. Rather than reinvent the wheel, the developers of ENERGY 2020 took advantage of the well-supported research related to consumer choice as it applies to energy simulation.

CONSUMER CHOICE SIMULATION

The socioeconomic environment, of which energy is a component, is the consequence of people making choices. They choose to build a house, store, or factory. They decide to emphasize capital, operating, or energy efficiency in the process of providing goods and services. They choose the fuel used to heat their homes; they choose the efficiency of the furnace and other energy using equipment; and they decide how to operate their furnaces and equipment. The basic characteristic of consumers is that they make choices: choices to acquire, specify, and use. Therefore, a proper representation of energy use must be a proper representation of how choices are made and the energy impact of those choices.

Typically a choice can be portrayed as a selection among a spectrum of alternatives. Faced with the selection options, a particular or discrete choice is made based on the preference of the consumer. The mathematical characterization of this choice process is called discrete choice analysis. The preferences are a function of observable quantities such as price and unobservable quantities such as style or taste. Additionally, consumer uncertainty in both the observable and unobservable portions of the individual's preference function means that the mathematical formulation of the choice process must be based on an estimation process, as are those estimations performed for more common econometric representations.

Consumer Utility and the Multinomial Logit

The utility, **U**, of a preference can be defined as $\mathbf{U_{in}=V_{in}+\varepsilon_{in}}$ where **V** is a dependent term and **ε** is an error term. **V** may depend on any number of characteristics **x** of a choice **i** and has an arbitrary functional form. "The generality and limits of this form deserve emphasis. A variable [utility] may be a component of **x**, a function specifying a nonlinear transformation, or interacting components of **x**, or a function specifying an interaction between **x** [choice attributes] and **s** [decision-maker attributes] variables."[1]

---

[1]McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics,* Ed. P. Zarembka, New York, Academic Press, 1974, page 114.

Using these preferences, for a given situation **n**, the probability $\mathbf{P_n}$ of a consumer making a particular choice **i** can be determined with the use of a multinomial logit (MNL).

"The multinomial logit is expressed as:

$$P_n(i) = \frac{e^{V_{in}}}{\sum_{j=1}^{N} e^{V_{jn}}} \tag{1}$$

This model (or equivalent variants of it) can be derived in a great number of ways. Its original formulation is due to Luce (1959), a mathematical psychologist. He derived the form of the [above] equation by making assumptions about the choice probabilities rather than the disturbances."[2]

The EPRI REEPS model uses the MNL formulation. The EPRI report starts the discussion as follows:

"Our choice of functional forms for the choice probabilities has been guided by several considerations. First, the functions must be computationally tractable, so that calibration and simulation on relatively large populations is possible. Second, the forms must be sufficiently flexible to adapt to the patterns of substitution and complementarity found in the data, without restrictive a priori assumptions. Third, households are assumed to be motivated to minimize the lifecycle cost of achieving specified levels of service, and more generally to weigh the desirability of energy-consuming services against other commodities in allocating their incomes. The functional forms for the choice probabilities should be consistent with such behavior. A family of functional forms for choice probabilities which meet these criteria and are therefore selected for our analysis are termed *nested logit* models... A nested logit model is a generalization of [the] form called the multinomial logit."[3]

MNL Characteristics

If the utility $\mathbf{V_i}$ of a choice i greatly exceeds that of any other option, then the probability that the choice **i**, as shown in equation (1) will be actually chosen approaches unity. If all the choices, as perceived by the consumer, have the same utility, then all the choices will have an equal probability of occurrence or **1/N** where **N** is the total number of choice options available. That is, if the utilities are the same, then consumers cannot tell the difference between the options and they are just as likely to pick any option. Relative to a large population, this implies that equal proportions of each option will be selected. If the choice were only between

[2]Ben-Akiva, M., *Discrete Choice Analysis: Theory and Applications*, MIT Press, Cambridge, MA, 1985, page 103.

[3]Cambridge Systematics, Inc., *Residential End-Use Energy Planning Model System (REEPS)*, Electric Power Research Institute, Report EA-2512, Palo Alto, California, July 1982, pages 3-9.

two options, then the probability would be 50/50 or ½ that either of the options would be chosen. This phenomena is a natural and reasonable consequence of both equation (1) and consumer choice theory. This does not mean a choice is not being made; it simply means that a consumer has no basis for a particular choice if all the options have equal utility. Note that in reality there is only an infinitesimal chance that all the options have the same utility.

Equation (1) has the feature that it allows the full range of utilities. If the utility function is a function of price, the price can range between 0 and infinity. This use of infinitesimally small or large values is not a problem from an empirical perspective. "Since empirically, a zero probability is indistinguishable from one that is extremely small, there is little loss of generality in assuming that the selection probabilities are all possible for the positive alternative sets in the experiment."[4] Moreover, more conventional econometric methods using elasticities have the identical theoretical considerations in that infinitesimally small prices would lead to infinite demands and infinitely large prices would lead to infinitesimally small demands and imply infinite energy efficiency.

Independence From Irrelevant Alternatives

"Three properties of logit probabilities [have been discussed], namely that they (1) range from zero to one, (2) sum to one over alternatives, and (3) are a sigmoid or S-shaped [cumulative distribution shaped] function of representative utility. Each of these properties is quite reasonable, and in fact, the first two are logically necessary. Logit probabilities also exhibit a property, however, that, at least in some contexts, is not desirable. This is called the independence from irrelevant alternatives property or the IIA property for short.

"The IIA property has been the focus of considerable discussion in the literature and not a small amount of confusion."[5]

"Generally the attributes entering the [MNL] for a specific alternative **j** depend solely on features of this specific alternative, and not on features of other alternatives. In this case the multinomial model is said to have the property of independence from irrelevant alternatives (IIA).

The term 'independence from irrelevant alternatives' refers to the property that the relative odds of two alternatives are independent of the availability and attributes of other alternatives.

However, it is possible for $\mathbf{V_j}$ to depend on iterations between features of alternative **j** and other alternatives, in which case the MNL model does *not* have the IIA property."[6]

---

[4]McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics,* Ed. P. Zarembka, New York, Academic Press, 1974, page 109.

[5]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986, page 18.

"Despite its practical advantages, the IIA property is a restriction that is not realistic in many situations. Recent work has indicated, however, that the IIA property in logit models is not as restrictive as it might at first seem...

McFadden shows that any model that specifies choice probabilities, including models that do not exhibit IIA, can be expressed in the *form* of a logit model [emphasis from original text]. That is, it is possible to express any choice probability [in the MNL form.] [The proof follows in the K. Train text. See footnote seven].

This shows that the logit probabilities, with the appropriate specification [of parameters] equal the true probabilities. Stated another way, any choice model can, with an appropriate choice [of estimated linear parameters], be put into the logit form. This concept gives rise to the term 'mother logit'"[7]

"What this discussion implies is that the logit specification can be used in situations for which IIA does not hold. All that is required is that additional variables be added to the representative utility, in particular, variables that relate to alternatives other than the one for which the representative utility is designated."[8]

Train also says that these extra variables are constant terms simply added to the utility function for each choice alternative prior to estimation of **V**. The modeler "estimates the model with all ... alternatives in the choice set and includes a constant term in the specification of the representative utility of the ... alternatives ..."[9]

McFadden performed other tests to show that:

"In particular, if the desirability of different alternatives tends to be fairly sharply differentiated for most households, which is the case unless the weights in [the MNL equation] are small in magnitude, the market cross elasticities are primarily determined by the distribution of households and are virtually independent of whether the household choice probabilities have the IIA property or not. Furthermore, the MNL functional form is rather robust empirically in that it will often describe observed choice behavior adequately even when the forces underlying that behavior are theoretically inconsistent with the IIA property."[10]

---

[6]Cambridge Systematics, Inc., *Residential End-Use Energy Planning Model System (REEPS)*, Electric Power Research Institute, Report EA-2512, Palo Alto, California, July 1982, pages 3-10.

[7]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986, page 21.

[8]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986, page 22.

[9]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986, page 23.

[10]Cambridge Systematics, Inc., *Residential End-Use Energy Planning Model System (REEPS)*, Electric Power Research Institute, Report EA-2512, Palo Alto, California, July 1982, pages 3-11.

In an earlier work, McFadden explains the concept further:

"Nevertheless, empirical evidence is that the MNL model is relatively robust, as measured by goodness of fit or prediction accuracy, in many cases in which the IIA [independence of irrelevant alternatives] property is theoretically implausible.

The restrictive IIA feature of the MNL model is present only when the vector $\mathbf{x_{it}}$ for alternative **i** is independent of the attributes of alternatives other than **i**. When this restriction is dropped, the MNL form is sufficiently flexible to approximate any continuous positive choice probability model on a compact [limited and defined] set of explanatory variables. Specifically if **P(i|xt)** is continuous, then it can be approximated globally to any desired degree of accuracy by the [standard] MNL model ..."[11]

Distributional Basis

"If we assume that the $\mathbf{U_{in}=V_{in}+\varepsilon_{in}}$ for all **i** ... and that all the disturbances in (1) are independently distributed, (2) identically distributed, and (3) Gumbel-distributed with a location parameter $\tilde{n}$ and a scale parameter $\mu > 0$, then

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j=1}^{N} e^{\mu V_{jn}}} \tag{2}$$

Say ε is Gumbel-distributed. Then [the cumulative form is:]

$$F(\varepsilon) = \exp(-\exp^{-\mu*(-\tilde{n})}) \ldots \tag{3}$$

As in the case of binary logit, the assumption of a constant $\tilde{n}$ for all alternatives, or $\tilde{n}=0$, is not in any sense restrictive as long as each systematic utility has a constant term. Similarly, the assumption that the disturbances are Gumbel-distributed can be defended as an approximation to the normal density. It is also used only for reasons of analytical convenience."[12]

If $\tilde{n}=0$, then the distribution is called the Weibul distribution. The Weibul distribution is more commonly cited than the Gumbel because it is the form actually used in practice.

"Thus, the probability distribution function on the generic technology price can be derived from distributions on the specific technology costs. It can be shown that the distribution of the

[11]McFadden, D., "Qualitative Response Models," in *Advances in Econometrics,* Ed.. Werner Hildenbrand, Cambridge University Press, New York, 1982, p.10.

[12]Ben-Akiva, M., *Discrete Choice Analysis: Theory and Applications*, MIT Press, Cambridge, MA, 1985, p.104.

least-cost from a sample of independently-distributed costs approaches the Weibul distribution."[13]

"This model, or a derivative, has been used in a variety of energy modeling applications."[14]

McFadden chooses the Weibul distribution a priori:

"Suppose each member of a population of utility-maximizing consumers has a utility function ... [whose error terms are] distributed with the Weibul (Gnedenko, extreme value) distribution."[15]

Despite the support for the multinomial logit, alternative distributions have been studied to achieve a more theoretical avoidance of IIA problems. First, McFadden's experience:

"If [the error term] is assumed to be multivariate normal, the resulting discrete response model is termed the multinomial probit (MNP) model ... when correlation is permitted between alternatives, so that the [covariance of the error terms] is not diagonal, the MNP model does not have the IIA or related restrictive properties ... However, for [more than five choice options], the computational time required for [estimation] ... is excessive."[16]

Next, Ben Akiva reviewed the topic:

"Recent works [using the Probit] have resolved some of the computational problems. However, only a few, very limited applications have appeared in [the] literature, and there is still no evidence to suggest in which situations the greater generality of multinomial probit is worth the additional computational problems resulting from its use."[17]

Ben-Akiva spent some time on the problem as noted in his earlier work:

---

[13]Boyd, D.W., et.al., *Abbreviated R&D Program Portfolio Selection Workbook: Market Share Model Appendix,* Decision Focus Incorporated, Palo Alto, California, U.S. Department of Energy contract DE-AC05-7BET05474, 1979, p. 6.

[14]Boyd, D.W., et.al., *Abbreviated R&D Program Portfolio Selection Workbook: Market Share Model Appendix,* Decision Focus Incorporated, Palo Alto, California, U.S. Department of Energy contract DE-AC05-7BET05474, 1979, p. 11. See, for instance, Cazalet, E.G., *General Equilibrium Modeling: The Methodology of the SRI-Gulf Model,* Final Report prepared by Decision Focus, Inc., for the Federal Energy Administration, Stanford Research Institute, Menlo Park, California, May 1977. See also A. Masevice, *A Review and Assessment of the Fossil1 Supply Structures.* Thayer School of Engineering, Dartmouth College, Master of Science Thesis [Advisor - George Backus], September, 1978.

[15]McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics,* Ed. P. Zarembka, New York, Academic Press, 1974, page 111.

[16]McFadden, D., "Qualitative Response Models," in *Advances in Econometrics,* Ed.. Werner Hildenbrand, Cambridge University Press, New York, 1982, p.18.

[17]Ben-Akiva, M., *Discrete Choice Analysis: Theory and Applications*, MIT Press, Cambridge, MA, 1985, p.128.

"The basic choice model that is used in this study for all alternative models is the multinomial logit model. Other choice models that might be considered to be superior from a theoretical point of view, such as the multiple probit model, are more complicated. It is not evident, however, that the added expense for more sophisticated choice models is worthwhile."[18]

Charles River Associates (CRA) also addressed the problem:

"Three commonly used models, the probit and logit, and a third known as the Cauchy probability model, give ogives ... and are virtually indistinguishable except at probabilities close to zero or one, where the probit model approaches the limiting values most rapidly, the Cauchy model the least rapidly. Within the range of most data, these models provide essentially equivalent probability functions and except for computational reasons, there is little to choose [statistically] among them. The logit model has computational advantages since it is a closed (explicit) functional form. The probit model, on the other hand, has an argument as the limit of an integral which cannot be expressed in closed form."[19]

Although CRA brings up the Cauchy distribution in this entry it is never brought up anywhere else in their discussions. One possible reason for the omission is McFadden's concern for positive finite moments which the Cauchy distribution does not have.[20]

Near the end of its review, CRA is down to only two approaches, the linear probability model and the conditional logit model (a form of MNL):

"The two models for multiple choice developed above, the multiple choice linear probability model and the conditional logit model, prove to be the most useful for demand analysis ... These models provide the advantage of practical empirical implementability along with a satisfactory theoretical justification in terms of the underlying behavior of individual decision makers.[21]

By the end of their review, however, CRA's last alternative to the multinomial logit is rejected.

[18]Ben-Akiva, M., *Structure of Passenger Travel Demand Models,* MIT, Department of Civil Engineering, Ph.D. Thesis, June, 1973, p. 171.

[19]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-11.

[20]McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics,* Ed. P. Zarembka, New York, Academic Press, 1974, page 111, footnote 4.

[21]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-28.

"We conclude that the linear probability model as formulated ... does not yield a practical estimation procedure with satisfactory statistical properties."[22]

The references above claim to include all distributions that could be justified for use in choice analysis. Other distributions have characteristics which violate the necessary requirements of consumer choice theory or provide currently untenable mathematical difficulties. All competent research to date indicates that the multinomial logit, although it has limitations just like any other approach, provides the most acceptable means to simulate consumer choice.

Market Share MNL in ENERGY 2020

Consumers, as simulated in ENERGY 2020, make choices relative to fuel selection for each energy end-use. These choices are simulated in ENERGY 2020 using the multinomial logit. The use of detailed multinomial logit formulations in energy demand has already been noted in the reference to the EPRI REEPS residential energy model developed by D. McFadden. It is also used in the EPRI COMMEND commercial model[23] by incorporating the multinomial logit work of Cohen and Baughman as the market share simulation.[24] The original Oak Ridge Residential Model developed by Eric Hirst also uses the multinomial logit for the market share calculation.[25]

Utility Function Form

The utility function is often written clearly, for example, as a simple function of price **(P_i)** with the constant (non-price) term noted above by Train. [26]

$$V_i = A_i + B*P_i \tag{4}$$

[22]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-47.

[23]Jackson, J.R., et. al., "Conservation Policy Analysis and End-Use Models: A Commercial Sector Example" in *Proceedings: End-Use Models and Conservation Analysis*, Electric Power Research Institute, Report EPRI EA 2509, Palo Alto, CA, July 1982, page 13.

[24]Cohn, S., *Fuel Choice and Aggregate Energy Demand in the Commercial Sector,* Oak Ridge National Laboratory, ORNL/CON-27, December, 1978 and Baughman, M.L. and Joskow, P.L., "Energy Consumption and Fuel Choice by Residential and Commercial Consumers in the United States" in *Energy Systems and Policy,* Volume 1, No. 4, 1974.

[25]Hirst, E., et. al., *An Improved Engineering Model of Residential Energy Use,* Oak Ridge National Laboratory, ORNL-CON-8, April 1977, page 18.

[26]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986.

in ENERGY 2020, the log-linear form is used:

$$V_i = a_i + b*\ln(P_i) \tag{5}$$

An implication of this form is that the consumers are more sensitive to the proportional (percent) differences in costs than in absolute ($) differences. This means a one dollar difference is less important in a thousand dollar furnace decision than it is in a three dollar light-bulb decision.

There is substantial support for this formulation. The derivation of the market share function based purely on technological cost distributions leads directly to the form (with **a=0** - no non-price component) as shown in the work of Decision Focus, Inc. and the Institute for Economic Analysis.[27]

With this formulation, equation (1) becomes

$$MS(i) = m_i \frac{P_i^b}{\sum_{j=1}^{N} m_j P_j^b} \tag{6}$$

where **MS**$_i$ replaces the probability usage in equation (1) to avoid confusion with the use of "P" for prices. Also,

$$m_i = \exp(a_i) \tag{7}$$

This m term is called the market share multiplier in ENERGY 2020 but it is just the constant required to avoid IIA concerns. When m is defined to be 1.0, as is common in technology assessment analysis, the equation becomes:

$$MS_i = \frac{P_i^{-b}}{\sum_{j=1}^{N} P_j^{-b}} \tag{8}$$

This simpler form is the most commonly used form of market share calculation as noted in the SRI/Gulf Model, the GEMS model, the DRI Energy Model and LMSTM along with the ORIM model noted above.[28] It is also the method taught by EEI and EPRI for DSM analysis.[29]

[27]Work of Decision Focus: Boyd, D.W., et.al., *Abbreviated R&D Program Portfolio Selection Workbook: Market Share Model Appendix,* Decision Focus Incorporated, Palo Alto, California, U.S. Department of Energy contract DE-AC05-7BET05474, 1979; and for the Institute for Economic Analysis: Reister, D, et. al., "The Oak Ridge Industrial Model: An Introduction," in *Proceedings: End-Use Models and Conservation Analysis*, Electric Power Research Institute, Report EPRI EA 2509, Palo Alto, CA, July 1982, pages 6-14.

[28]**SRI/Gulf Model:** Electric Power Research Institute, *Fuel and Energy Price Forecasts,* Volume 2, Report EPRI EA-433, Palo Alto, CA, 1977, p. 6-7:

Nonetheless, note that the price (P) can be any complicated function (including the real price) necessary to specify the perceived value of the commodity or service.

Basic research in choice analysis also tends to favor the log-linear approach:

"The formulation employed by the McLynn and Woronk model (1969) is equivalent to (the MNL) equation if all the variables $\mathbf{X_{itk}}$ are replaced by their logs ... This formulation can be written as:" [30]

$$P(i,A) = \frac{X_{itk}^{\theta k}}{\sum_{j=1}^{N} X_{jtk}^{\theta K}} \tag{9}$$

"Specification of explicit probability functions for the 'strict utility' specification in the [MNL] equation can be completed by specifying parametric forms for the function **V(x,s)**. We shall consider several cases. First suppose this function is log-linear in unknown parameters ..."[31]

Parameter Specification

The parameters (b) associated with choice variables are generally the same for all choice options, consistent with the derivation of the MNL-form above. The usage stems from the concept that all choices have equal uncertainty relative to the consumer.[32] The b parameters can be allowed to vary by alternative, however, provided the data truly supports the assertion

GEMS Model: Cazalet, E.G., *General Equilibrium Modeling: The Methodology of the SRI-Gulf Model,* Final Report prepared by Decision Focus, Inc., for the Federal Energy Administration, Stanford Research Institute, Menlo Park, California, May 1977, p. 4-6;
DRI Energy Model: Data Resources, Inc., *DRI Energy Modeling System Documentation,* Data Resources, Inc., Cambridge, MA, 1984, p.11;
LMSTM: Decision Focus, Incorporated, *User's Guide to the Load Management Strategy Testing Model,* Electric Power Research Institute, EPRI EA-3653-CCM, August 1984, p. C-2;
ORIM Model: Reister, D, et. al., "The Oak Ridge Industrial Model: An Introduction," in *Proceedings: End-Use Models and Conservation Analysis*, Electric Power Research Institute, Report EPRI EA 2509, Palo Alto, CA, July 1982, p. 6-14.

[29]Battelle Columbus Laboratory and Synergetic Resource Corporation, *Demand-Side Management,* Edison Electric Institute and Electric Power Research Institute, EPRI EA/EM-3597, Volume 2, December 1984, p. 32.

[30]Ben-Akiva, M., *Structure of Passenger Travel Demand Models,* MIT, Department of Civil Engineering, Ph.D. Thesis, June, 1973, p. 177.

[31]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-26.

[32]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986, pp. 37-40; Ben-Akiva, M., *Discrete Choice Analysis: Theory and Applications*, MIT Press, Cambridge, MA, 1985, p.111; and McFadden, D., "Qualitative Response Models," in *Advances in Econometrics,* Ed.. Werner Hildenbrand, Cambridge University Press, New York, 1982, p.4.

that the choices are naturally indexed [unique onto themselves]. [33] When the microcomputer-based maximum-likelihood procedure described below is fully functional, nonconventional analysis assuming varying b parameters can be performed.

Thus the multinomial logit (based on the Weibul distribution) used in ENERGY 2020 is the only form supported in the literature (other than a theoretical effort to advance the potential use of the probit model - based on the normal distribution).

ESTIMATION OF MNL PARAMETERS

The estimation of MNL parameters is abundantly discussed in the literature. The functional form of the MNL equation causes the ordinary least square estimation process to be biased. Therefore the method of maximum likelihood estimation is used. [34]

"In many multiple choice applications using available data, regression methods are not applicable and the maximum-likelihood method is the only practical procedure available."[35]

ENERGY 2020 Estimation

In ENERGY 2020, a non-price and a price related parameter are estimated for each fuel by end-use and economic category. These parameters were originally estimated in the DEMAND81 model using national data and non-linear least-squares. At the time, maximum-likelihood estimation packages were not commercially available. However, as McFadden notes "an alternative to maximum-likelihood estimation is to use non-linear least squares ...."[36] Nonlinear least-square estimation is a computer intensive operation. Therefore, re-estimation of the price response portion of the function was not routinely performed. It was assumed that the price response behavior would not be locally variable. Local tastes and socioeconomic environment (the non-price) were however assumed to be local. The non-price parameter was then re-estimated by ordinary least-squares for each implementation of ENERGY 2020. Studies show that "least-squares estimation leads to substantial overestimates of the price

[33]McFadden, D., "Qualitative Response Models," in *Advances in Econometrics,* Ed. Werner Hildenbrand, Cambridge University Press, New York, 1982, p.5.

[34]Fomby, T., et.al, *Advanced Econometric Methods*, Springer Verlag, New York, 1984, Section 16.4.

[35]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-49.

[36]McFadden, D., "Qualitative Response Models," in *Advances in Econometrics,* Ed. Werner Hildenbrand, Cambridge University Press, New York, 1982, p.7.

sensitivity..."[37] Therefore, this process should overestimate the conservation associated with market shifts and thus be less controversial from a regulatory perspective.

Nonetheless, recent computer hardware advances now allow maximum-likelihood estimation to be performed routinely on microcomputers, the platform for ENERGY 2020. Further, as ENERGY 2020 is used for analyses where there are limited historical data, maximum-likelihood estimation becomes more important because "limited Monte Carlo studies and analytical solutions suggest the maximum-likelihood estimators are also satisfactory in small samples."[38] Limited testing of the components of a maximum-likelihood estimation routine for the ENERGY 2020 calibration has been completed. This routine will provide the statistical reporting unique to MNL estimation. The ENERGY 2020 maximum-likelihood routine is based on the work of S. Cosslett which focuses on the use of aggregate data for efficient MNL estimation.[39] This is the type of data most readily available to energy modelers.

Data Sources

Historical data, applicable to the service area, for the estimation of the MNL are obtained using published data from several sources corrected (scaled) to be self-consistent. The energy use data by economic sector (residential, commercial, industrial) at the state level are available from the U.S. Department of Energy.[40] These data are scaled to the service area based on historical utility sales by economic sector. Industrial energy use is further disaggregated into SIC (Standard Industrial Category) designations by utility billing data or the Annual Survey of Manufacturers.[41] The Survey of Manufacturers also provides the SIC-specific proportions of fuel use (coal, oil, gas, electricity, cogeneration) for each historical year. End-use information is often available from utility or other institutional surveys.[42] The most appropriate data available are used.

---

[37]Dubin, J., and McFadden, D., "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," in *Proceedings: End-Use Models and Conservation Analysis,* EPRI Report EPRI EA 2509, Palo Alto, CA, July 1982, pages 13-20.

[38]Charles River Associates, *A Disaggregated Behavioral Model of Urban Travel Demand,* U.S. Department of Transportation, Contract No. FH-11-756, Final Report, March, 1972, pages 5-41.

[39]Cosslett, S.R., "Efficient Estimation of Discrete Choice Models", in *Structural Analysis of Discrete Data with Econometric Applications,* ed. C. Manski and D. McFadden, MIT Press, Cambridge, MA, 1986, Chapter 2.

[40]U.S. Department of Energy, *State Energy Data Report,* Energy Information Administration, DOE/EIA-0214, 1978 and later.

[41]U.S. Department of Commerce, *Annual Survey of Manufacturers,* Washington, DC, 1987 and later.

[42]See, for example, American Gas Association, *Gas Facts,* Arlington, VA., 1975 and later; U.S. Department of Commerce, *Census of Housing,* Washington, DC, 1970, 1980; U.S. Department of Energy, *End Use Energy consumption Data Base: Series I Table,* Energy Information Administration, DOE/EIA-0014, June 1978: U.S. Department of Energy, *Residential Energy consumption Survey,* Energy Information Administration, DOE/EIA-0207/5, July 1980 and later; U.S. Department of Energy, *Nonresidential Buildings Energy Consumption Survey,*

The utility sales are assumed to be the only values which are correct in an absolute sense. All other data are only presumed correct in a relative sense. That is, the data can be used for scaling (proportions) when the errors associated with that data can be assumed to cancel-*out in the equation*. (It is generally assumed that the information in any survey data set has the same proportional error for each fuel, SIC, or end-use - all portions of the data are “equally” in error, e.g., 20% overestimated or 50% underestimated.) “Proportional data” is interpolated for missing data.

By using data for historical demands, appliance efficiency and appliance life, additions and retirements to the appliance stock by fuel and end-use can be estimated to derive historical market shares.[43] These historical market shares are then used to estimate the MNL. Price information comes from the utility, state, or the U.S. Department of Energy.[44]

The focus here is to use the best data available categorized in the same manner as the utility uses the data for required regulatory matters. This same data would be used in effectively the same way whether the formal model used were an end-use, econometric, or MNL-based model.

EFFICIENCY TRADE-OFF AS BINOMIAL LOGIT

The decision to invest in higher capital cost (higher energy efficiency) equipment or structures in the face of higher energy prices is a consumer choice. It is a binomial logit choice in that it is the choice between two quantities, capital cost and operating costs (fuel). The result of the choice determines the efficiency of the new equipment or structure. The multinomial logit, equation (1), reduces to a much simpler form when only two choices are involved:

$$P_n(i) = 1/(1+e^{(V1-V2)}) \qquad (10.)$$

or

$$P_n(i) = 1/(1+e^{(V1)}/e^{(V2)}) \qquad (11)$$

Energy Information Administration, DOE/EIA-1278, June 1981 and later; and Electric Power Research Institute, EPRI EM-5126 *Energy Use Patterns and Indicators,* Palo Alto, CA, April 1987.

[43] for appliance efficiency see: Association of Home Appliance Manufacturers, *Energy Efficiency and Consumption Trends,* Chicago, Illinois, July 1, 1984, and Geller, H., *Energy and Economic Savings from National Appliance Efficiency Standards,* American Council for an Energy-Efficient Economy, Washington, D.C., August 1986. For appliance efficiencies and appliance life see U.S. Department of Energy, *Annual Report to Congress,* Energy Information Administration, DOE/EIA-0173(198X)/3, 1981 and later.

[44] See U.S. Department of Energy, *State Energy Prices by Major Economic Sector,* Energy Information Administration, DOE/EIA-0190, 1981 and U.S. Department of Energy, *State Energy Price and Expenditure Report,* Energy Information Administration, DOE/EIA-0376(8X), 1984 and later.

If **V** is log-linear, as used in ENERGY 2020, the “form” becomes:

$$P_n(i) = 1/(1+(V1/V2)) \tag{12}$$

Functional Form Selection

A review of capital-efficiency trade-off literature shows only algebraic variations of the two forms above for determining capital cost versus efficiency. (The function presented in the documentation can always be algebraically transformed to correspond exactly to a binomial logit.) The logit has the necessary functional S-shape. The curve must be asymptotic and reach the maximum (finite) efficiency at infinite costs. The curves estimated here are empirical continuous curves reflecting consumer choice in light of actual technology alternatives.

Least-Cost Curves

Least-cost curves which can also be used in demand analysis, including ENERGY 2020 analyses, are discrete (discontinuous) engineering curves which order a selection of energy efficient technologies based on estimated (engineering-based) energy savings. Least-cost curves are not used to determine the choice consumers make; they are used to determine the impacts of energy programs if consumers chose energy efficiency technologies based on the economic decisions used by the analyst. The ordering of least-cost options on the least-cost curve is still an open issue, hotly debated. Examples of least-cost generation are available from a variety of sources.[45] These curves have the same general shape as the binomial logit and are well approximated by the logit. The primary difference is that the “least-cost logit” is shifted toward the zero axis because it would have consumers investing in higher efficiency equipment at a much lower energy price. That is, it would infer that consumers place much more utility on reducing long-term energy costs than the historical data indicate.

Binomial Logit Basis

The binomial logit curves can also be reconciled as a composite of the market share of all available technologies chosen by consumers as energy prices vary. The resulting binomial logit can then be construed as a “fit” of the average efficiency selected by those “multinomial” choices.

[45]See, for example: Meier, A. *Supply Curves of Conserved Energy,* Lawrence Berkeley Laboratory, May 1988; Krause, F., *Analysis of Michigan’s Demand-Side Electricity Resources in the Residential Sector,* Lawrence Berkeley Laboratory, LBL-23025, February, 1987; Ford, A. and Naill, R., *Conservation Policy in the Pacific Northwest,* Bonneville Power Administration, May 1985; and Synergetic Resource Corporation, *Industrial Electricity Conservation Potential in the Pacific Northwest,* Volumes I and II, report No. 7077-R2, Bala Cywyd, Pennsylvania, March 1983.

Those that use the log linear form of the binomial logit are ENERGY 2020 and the Oak Ridge Residential Model.[46] The linear form is used in the REEPS and COMMEND model and several independent studies.[47]

Binomial Logit Sensitivity

Empirical tests using both forms under worst case conditions (at the center point where the probability is ½ and operating cost utility and capital cost utility are equal) show that a 25% change in capital cost (the independent variable) produces a 2% difference in the model results. A 50% change leads to a 10% difference. During model usage, these curves are only affecting new investments, so their immediate impact on model results is reduced by an additional order of magnitude. Note also, that the recently announced 25% improvement in efficiency standards for refrigeration is expected to produce only a 10% increase in capital costs. The two forms, in this situation, would agree within 0.4%! Thus the sensitivity to the form used in ENERGY 2020 and the only used alternative is indistinguishable.

ESTIMATION OF TRADE-OFF CURVES

The trade-off curves are only estimated once when the raw historical data on historical efficiency, capital cost, and fuel prices are entered into the ENERGY 2020 databases. The binomial logit is a two parameter curve. Therefore, the two (binomial choices) can be thought of as two equations (both a function of energy prices) with two unknowns. These equations are solved by simple point estimates.

Algebraic Solution

Two features can be determined about the choice equation under particular conditions (the year 1972 for ENERGY 2020 calculations.) These are the actual choices of capital cost and efficiency (the first known) and the slope of the curve when the choice was made (the second known). The functional form of the curve has been derived a priori. The solution for the

[46] For ENERGY 2020: Backus, G., and J. Amlin, *ENERGY 2020 Integrated Policy Model Documentation* (three volumes), Policy Assessment Corporation, St. Paul, Minnesota, April 1987.
For Oak Ridge: Hirst, E., et. al., "The Oak Ridge National Laboratory's Residential Energy Use Model: Version 7.1" in *Proceedings: End-Use Models and Conservation Analysis,* Electric Power Research Institute, Report EPRI EA 2509, Palo Alto, CA, July, 1982.

[47]See, for example: Corum, K., et. al., "A Simulation Analysis of Alternative Policies to Simulate Energy conservation in Commercial Buildings," in *Proceedings: End-Use Models and Conservation Analysis,* Electric Power Research Institute, Report EPRI EA 2509, Palo Alto, CA, July 1982; O'Neal, D., and Corum, K., "Investment in Energy Efficient Houses: An Estimate of Discount Rates Implicit in New Home Construction Practices," in *Energy,* Volume 7, No. 4, Pergamon Press Ltd., 1982; Ruderman, H., et.al., "The Behavior of the Market for Energy Efficiency in Residential Appliances Including Heating and Cooling Equipment," in *The Energy Journal,* Volume 8, No. 1, 1987.
For REEPS and COMMEND see Cambridge Systematics, Inc., *Residential End-Use Energy Planning Model System (REEPS)*, Electric Power Research Institute, Report EA-2512, Palo Alto, California, July 1982.

parameters is then simply to find two conditions for which the unknown parameters can be solved. The capital cost and efficiency can be found in readily available historical data. The slope of the curve in an infinitesimal region at the decision-point can be calculated by "perturbing" the solution of the cost function around a point. This calculation provides the ε needed to solve the parameters of the globally-applicable binomial logit. This slope calculation at one point has no other purpose and is unrelated, functionally, to the binomial logit used for all capital cost and efficiency calculations. The ε calculation is just part of a mathematical process to solve the parameters of the binomial (trade-off) logit.

The trade-off curve is only estimated at the "1972 point" because that "point" was prior to any changes in energy prices. The data for that year closely approximates an equilibrium market. This provides an easy basis for data interpretation in that marginal versus average issues need not be addressed. (Alternatively, AHAM, AGA, ASHRAE, or other survey data could be used to perform a complete maximum-likelihood estimation of the trade-off curve, but any biases or incompleteness issues must be reconciled.)

Estimation Confidence

The use of a curve based on only 1972 data provides a significant test of function validity. The curve is applied historically for the years 1975-1992 as well as for the future. Independent, historical estimates of appliance efficiency can be compared to those produced by the curve in the historical simulation. To date those comparisons have been favorable (conversations with J. Davulis, Central Maine Power Company; R. Terrell, Wisconsin Power and Light; M. Jurabchi, when with the Massachusetts Office of Energy Resources). Formal comparisons have not yet been performed because definitive data on historical efficiencies are still lacking. Additional years of self-consistent appliance efficiency surveys should resolve this problem.

DEMAND TRADE-OFF CURVE DERIVATION

This section derives the cost-versus-efficiency trade-off curves used in ENERGY 2020. This derivation was originally developed for the U.S. Department of Energy's DEMAND81 model.[48] This derivation also details how the demand coefficients in the model are estimated.

The demand trade-off curve derivation begins with a generalized cost function:

$$MCO = CCR*CC+OMC+P/N \tag{1}$$

Where

MCO = Marginal cost of output ($/Unit)
CCR = Capital charge rate (($/Yr)/$)
CC = Marginal capital cost ($/(Unit/Yr))
OMC = Marginal operating and maintenance costs ($/Unit)

[48]Backus, George A., *DEMAND81: National Energy Policy Model,* School of Industrial Engineering, Purdue University, Reports AFC-7 through AFC-10, 1981.

P = Marginal price of energy ($/BTU)
N = Marginal efficiency (Unit/BTU)

For the general economy, output is measured in dollars of goods. For an energy conversion process (here converting primary fuel BTUs to useful process BTUs), output is measured in BTUs of useful (process) energy. For a transportation sector, output would be measured in equivalent vehicle-miles.

This functional form is consistent with the classical definition:

$$MCO_j = \sum a_i * (I/O)_i \qquad (2)$$

Where
a_i = cost per unit of input factor "I"
$(I/O)_i$ = units of input factor "I" to produce one unit of output "j"

For the purposes here, only capital and energy are explicitly considered. The OMC term is an aggregate variable representing all other input factors such as labor and materials. Capital costs (CC) are assumed to be a function of technological advance and energy costs only. Operating and maintenance costs are assumed to be proportional to capital cost (and energy costs to the extent that capital costs are a function of energy costs.) As machines become more complicated, higher cost labor and maintenance are required. Empirical studies support this assumption.[49]

$$OMC = OCF * CC \qquad (3)$$

where OCF is the unit operation cost factor ($/Yr)/$

Process efficiency is assumed to be a function of technological advance, capital costs, and energy costs. At the margin, perceptions of the trade-off between cost and efficiency stipulate that:

$$d\text{MCO}/d\text{N} = 0 \qquad (4)$$

Where "*d*" is the ordinary differential operator. (This analysis could proceed using partial derivatives; the results would be the same and the additional mathematical arguments would only detract from the clarity of the derivation.)

Technological advance is exogenous but assumed to be changing over time. Therefore, at any instant, capital cost can be written as a function of process efficiency for small perturbations of N as:

$$CC^*/CC_B = (N^*/N_B)^{\varepsilon} \qquad (5)$$

[49]See Backus, G., *FOSSIL79: National Energy Policy Model*, Resource Policy Center, Thayer School of Engineering, Dartmouth College, Report No. DSD-165 through DSD-168, 1979; and U.S. Department of Energy, *FOSSIL2 Energy Policy Model Documentation.* NTIS Document DOE/70143-02, Washington, D.C., October, 1980.

Where

CC^*, N^* = perturbed values

CC_B, N_B = base values before perturbation

ε = elasticity and derivative of curve at "B"

For algebraic ease:

$$CC^o = CC^*/CC_B \tag{6}$$

$$N^o = N^*/N_B \tag{7}$$

$$P^o = P^*/P_B \tag{8}$$

Using equations 3,5,7 and 8, equation (1) can be rewritten as:

$$MCO = (CCR+OCF)*CC_B * (N^o)^{\varepsilon} + P_B/N_B * P^o/N^o \tag{9}$$

Equation 9 can be used in equation 4:

$$dMCO/dNo = (CCR+OCF) * CC_B * \varepsilon * (N^o)^{\varepsilon-1} - P_B/N_B * P^o/(N^o)^2 = 0 \tag{10}$$

or in the base year when equations 6,7 and 8 equal unity:

$$\varepsilon = [P_B/N_B]/[(CCR+OCF) * CC_B] \tag{11}$$

This equation guarantees that the value added from energy or capital is equal at the margin as required by classical economics. Note that ε is always positive.

To increase the utility of equation 5, there needs to be a function "*f*" such that:

$$CC^*/CCN = f\,(N^*/N_{max}) \tag{12}$$

Where CCN is a normalizing capital cost varying only with technological advance and N_{max} is the maximum obtainable efficiency currently available at any cost.

Now the coordinate systems can be changed by multiplying equation 5 by CCN/CCN and N_{max} / N_{max}:

$$CC^*/CCN * CCN/CC_B = (N^*/N_{max} * (N_{max}/N_B)^{\varepsilon} \tag{13}$$

or

$$CR = (\beta*NR)^{\varepsilon}/\alpha \;\Big|_{B} \tag{14}$$

where:

$CR = CC^*/CCN$ *(15)*

$NR = N^*/N_{max}$ *(16)*

$\alpha = CCN/CC_B$ *(17)*

$\beta = N_{max}/N_B$ *(18)*

Note that the slope of equation 14 in the base year (base values) is:

$$dCR/dNR = \varepsilon * \beta^{\varepsilon} * NR^{\varepsilon-1} / \alpha \,|_B \quad (19)$$

By definition it is assumed here that as:

$$NR \rightarrow 1 \text{ then } CR \rightarrow \infty \quad (20)$$

(i.e. as $N^* \rightarrow N_{max}$)

As implied by a production function with substitution:

$$CR \rightarrow 0 \text{ as } NR \rightarrow 0 \quad (21)$$

This expression assumes that there can be no output without energy, which more strictly assumes that if:

$$CR > 0 \text{ then } NR > 0 \quad (22)$$

It also assumes that capital is required for energy to be useful, i.e., if:

$$CR = 0 \text{ then } NR = 0 \quad (23)$$

The market share function satisfies all these requirements:

$$NR = 1/(1+CR^{\mu}) \quad (24)$$

Equation 24 is the market share function with only two choices - trading energy efficiency (fuel cost) for capital costs. Here, the market share is the share of the maximum efficiency. Note that NR equals 0.5 when CR equals 1.0 (i.e., CC equals CCN) and that μ is always negative. The appearance of the market share makes sense given that it reflects how choices are made with real-world, imperfect information/perceptions.

Equation 11 can be solved using historical data. For use in equation 1, equation 24 would be rearranged to yield:

$$CR = (1/NR-1)^{1/\mu} = \Phi^{h} \quad (25)$$

Note that dCR/dNR must equal the value obtained from equation 19 in the base year. From the chain rule:

$$dCR/dNR = dCR/d\Phi * d\Phi/dNR \quad (26)$$

$$= -h*\Phi^{h-1} * NR^{-2} \quad (27)$$

$$= -1/\mu * (1/NR-1)^{1/\mu-1} *NR^{-2} \quad (28)$$

From equations 28 and 19:

$$\varepsilon*\beta^{\varepsilon} * NR^{\varepsilon-1} / \alpha = -1/\mu * (1/NR-1)^{1/\mu-1} * NR^{-2} \quad (29)$$

In the base year (from equations 17 and 18):

$$\varepsilon*NR^{-\varepsilon} *NR^{\varepsilon-1} *CC_B/CCN = -1/\mu * (1/NR-1)^{1/\mu-1} *NR^{-2} \quad (30)$$

or by noting that (1/NR-1) equals (1-NR)/NR:

$$-\varepsilon*\mu * CC_B/CCN = (1-NR)^{1/\mu-1} * NR^{-1/\mu} \quad (31)$$

In equation 31, ε, CC_B, and NR can be obtained directly form historical data and engineering estimates (i.e., N_{max}). CCN and μ are the only unknowns in equation 31. Equation 25 also defines CCN and μ. Equation 25 can be used to generate an equation with μ as the only unknown. Equation 31 becomes:

$$-\varepsilon*\mu * (1/NR-1)^{1/\mu} = (1-NR)^{1/\mu-1} * NR^{-1/\mu} \quad (32)$$

or

$$\mu = -1/[\varepsilon*(1-NR)] \quad (33)$$

With μ known, CCN can be found by equation 25:

$$CCN =CC_B/(1/NR-1)^{1/\mu} \quad (34)$$

Now equation 1 can be rewritten by using equation 25:

$$MCO = (CCR+OCF)*CCN*(1/NR-1)^{1/\mu} + PN/N_{max} * PR/NR \quad (35)$$

where PN is a normalizing energy price and:

$$PR = P^*/PN \quad (36)$$

On the margin, equation 4 must be valid at all points; therefore:

$$dMCO/dNR =\Omega*h * (1/NR-1)^{h-1} *NR^{-2} - \theta*PR/NR^2 = 0 \quad (37)$$

where:

$$\Omega = (CCR+OCF)*CCN \quad (38)$$
$$\theta = PN/N_{max} \quad (39)$$
$$h = 1/\mu \quad (40)$$

or

$$\theta*PR = -\Omega*h * (1/NR-1)^{h-1} \quad (41)$$

or

$$NR = 1/[1+(-\theta*PR/(\Omega*h))^{1/h-1}] \quad (42)$$

Note that from equations 38, 39 and 40, $-\theta*PR/(\Omega*h)$ equals:

$$(-\mu/Nmax) / [(CCR+OCF)*CCN] * P^{*} \quad (43)$$

Thus PN can be redefined as:

$$PN = (CCR+OCF)*CCN / (-\mu/N_{max}) \quad (44)$$

and

$$\sigma = \mu/(1-\mu) \quad (45)$$

Finally, equation 42 becomes:

$$NR = 1/(1+PR^{\sigma}) \quad (46)$$

Note that the CR and PR equations are functionally consistent as they must be.

In ENERGY 2020, σ is the fuel trade-off coefficient XXFTC (where XX is the end-use or process prefix), μ is the capital trade-off coefficient XXCTC, PN is the fuel price - normal XXFPN and CCN is the capital cost normal XXCCN. N_{max} are the XXEMs in ENERGY 2020 for each end-use or process.

XXFTC, XXCTC, XXFPN, and XXCCN are solved with 1972 historical data. In 1972 the average (recorded) data also approximate the marginal decision data because energy prices had been constant since 1940. This is long enough for the vintaging effects of capital stocks to be minimal.

DERIVATION OF THE CAPITAL CHARGE RATE

The capital charge rate is the annualization of capital expenses to account for taxes, tax credits, return of principal, return on investment, and interest during construction. The "CCR" equation is:

CCR = (1+R)**(C/3)*(1-ITC/(1+NR)-TR*(TL/2)/(TL/2+NR))
*R/(1-(1+R)**(-BL))/(1-TR)

Where:

R = Real Return on Investment
NR = Nominal Return on Investment
C = Construction Time
ITC= Investment Tax Credit
TR = Tax Rate (Federal plus State income tax)
TL = Tax Life
BL = Book Life

NR=(1-TR)*(1-F)*ND+F*NE
R=(1+NR)/(1+INF)-1
ND=(1+D)*(1+INF)-1
NE=(1+E)*(1+INF)-1

Where:

F = Fraction Equity
INF= Inflation Rate
ND = Nominal Return on Debt (Interest Rate)
D = Real Interest Rate
NE = Nominal Return on Equity
E = Real Return on Equity

For small "INF" (less than 10%/yr), a simpler calculation can be used with acceptable error:

ND=D+INF
NE=E+INF
R=(1-TR)*(1-F)*D+F*E
NR=R+INF

Risk can be added to "R" to reflect uncertainty and a higher required return. Energy 2020 includes financial risk concerns by increasing the required rate of return. Typically, a .02 to .05 risk (RISKN) is used for new technologies.[50]

Although the standard version of ENERGY 2020 uses a constant risk adjustment, a dynamic risk adjustment can be easily calculated. As a first approximation, a technology is assumed to be mature when the demand (D) for it is 10% of the total market demand (MPD). The risk can be reduced over time to reflect this phenomenon:

RISK=RISKN*EXP (-D/MPD)
RR=R+RISK

Where "RR" is the risk-adjusted "R" that can be used instead of "R" in all appropriate equations.

The "(1+R)**(C/3)" term in the "CCR" equation represents interest during construction which must be added to the final cost of the facility. During construction, costs accumulate faster near the end of the project than at the beginning. As a good approximation, it can be assumed that all the construction costs occurred two-thirds of the way through the construction program. That means interest charges ® accumulated for a time equaling "C/3".

The "R/(1-(1+R)**(-BL)" term is the classical capital recovery term.[51] The "(1-TR)" term at the end converts the after tax calculation into before tax dollars.

Investment tax credits reduce the cost of the plant by the tax credit after the first year of operation using "original" dollars. Therefore the value of the tax credit is "ITC/(1+NR)".

Depreciation is expensed for tax purposes during each year of the tax life of the plant. With the double-declining balance method (DDB) of computing depreciation, the depreciation (DEP) of the plant for each capital dollar spent in year "t" is:

DEP(t)=2/TL*(1-2/TL)**(t-1)

Depreciation, under existing laws, is a current dollar phenomena which does not account for inflation. Therefore the net present value of the energy is calculated with the nominal rate of return. If the depreciation life is adequately long to neglect end year effects, then the net present value of depreciation expenses is:

(2/TL)/(NR+2/TL)

Because depreciation is a benefit (negative cost) based on the total plant before investment tax credits, it shows up as an additional negative term in the capital cost modifiers of "CCR:"

(1-ITC/(1+NR)-TR*(TL/2)/(TL/2+NR))

---

[50]Backus, G. A., *FOSSIL79 National Energy Policy Model*, Resource Policy Center, Thayer School of Engineering, Dartmouth College, Report No. DSD-165 through DSD-168, 1979.

[51]Smith, Gerald W., *Engineering Economy: Analysis of Capital Expenditures*, Iowa University Press, Ames, Iowa 1973.

The CCR calculation is naturally appropriate to business decisions but its use in the residential sector may appear artificial. When the CCR calculation is used for the residential sector, TL and C are set to zero because the residential sector can neither write off depreciation expenses nor make adjustments for extended construction times. This makes the calculation exactly correct for housing and any long-term investments.

Concerns can occur when the life of the loan is much shorter than the physical life assumed in the CCR calculation. When short-term loans (2-5 years) are used, the home owner still implicitly discounts the equity portion of equipment and depreciates the equipment over its expected life time. (Consumers do not expect a car or stove to fail as soon as the loan is paid-off; they write-off its value over its actual life time.) Therefore, the CCR calculation can only be incorrect for the debt portion of the investment. When a life cycle cost analysis of the actual cash flows is performed, which leveizes the short-term interest payments with the life of the equipment, the results are essentially identical to those obtained with the CCR calculation here.

PROMULA—HOST LANGUAGE OF ENERGY 2020

The ENERGY 2020 model is a large and complex mainframe-size model. Through the use of the PROMULA computer simulation language, ENERGY 2020 is now available for the IBM PC and IBM compatibles. PROMULA is a product of Mindware Corporation, Columbus, Ohio. The name PROMULA comes from: PROcessor of MUltiple Language Applications. The following briefly describes the language:

PROMULA is a high-productivity applications development tool. It is an integrated multi-language compiler which, in the standard version, can process programs written in any of the following four languages: FORTRAN, PASCAL, BASIC, PROMULA.

PROMULA, which is specifically designed to increase programming productivity, is a tool for problem solving. Its ability to process and integrate programs written in a variety of languages sets it apart from all other application development tools. As an applications and information management system, PROMULA manages both applications and the information associated with applications. It is especially useful for implementing serious, mainframe-size applications on a personal computer. PROMULA applications will run either as stand-alone programs or as integrated systems consisting of many multi-language components.

PROMULA, the language, is a decision support system; it lets the user focus information onto a problem, thus allowing him/her to analyze and evaluate alternative decisions about the problem. It integrates the following basic capabilities:

1. Data management (organize and selectively manipulate data)
2. Data analysis (establish relationships in the data)
3. Modeling (simulate a problem and possible solutions to it)

4. "What if" analysis (compare alternative decisions about the problem)
5. Report generation (display data or results in report form)
6. Graphics (display data or results in plotted form)
7. Menu management (prepare and use pick and data menus)
8. Equation solving (solve systems of simultaneous equations)

PROMULA bridges the transition from third- to fourth-generation programming capability in applications development. FORTRAN, PASCAL, and BASIC are third-generation procedural languages used extensively by both systems and application programmers. PROMULA, on the other hand, is a fourth-generation language designed specifically for programming applications.

PROMULA is also a valid programming alternative to spread sheets or database managers (DBMS) in large-scale applications development. It is easier to write PROMULA programs than it is to write unreadable spread sheet macros or constrained DBMS command sequences. In the PROMULA environment, learning a new language, like PROMULA, is an option, not a requirement. The user can continue to write programs in FORTRAN, PASCAL, or BASIC.

PROMULA System Highlights

A. *Total Programming Environment:* Complete turnkey applications can be written with PROMULA. The system is designed to capitalize on existing applications written in a variety of languages and to minimize programming time in developing new applications. PROMULA is largely self-contained with its own screen editor, compilers, and operating system interface.
   All languages supported by PROMULA are fully featured. No tricks are required to do the detailed types of operations needed by all kinds of applications.
B. *Program Editor:* A screen editor allows the writing and editing of source code and data files.
C. *Transportability:* PROMULA is designed to be transportable among various machines and operating systems. It is written in standard, transportable C. A standard FORTRAN version is also available for traditional mainframe environments.
D. *Compatibility:* PROMULA is not an operating system; rather, it operates within the standard PC-DOS environment and is completely compatible with other software.
E. *Extendibility:* PROMULA can be easily extended to include additional compilers or different language dialects.
F. *Language Integration:* PROMULA is an integrated program development environment. The whole of PROMULA is more valuable and more powerful than the sum of its parts. The PROMULA languages work with and enhance one another.
G. While each language has its own separate compiler, the execution of all languages is processed via a single, central program. All PROMULA languages share a large, built-in function library for performing such operations as database management, screen and menu management, and graphics. The library functions are available to

PROMULA via single commands to FORTRAN and PASCAL via procedure calls, and to BASIC via special statements.

H. *Program Integration:* A PROMULA program can consist of segments written in different languages or even segments compiled by compilers other than PROMULA.
I. *Value Added:* PROMULA adds value to existing FORTRAN, PASCAL, or BASIC codes. By themselves, such codes are strictly computational or procedural "boxes." PROMULA allows complete use of the information contained in the boxes, thus adding value to the codes.
J. *Mainframe PROMULA:* A mainframe version of PROMULA allows PROMULA to be used as a distributed decision support system, trading off the power of the mainframe against the ease-of-use and convenience of the personal computer.

PROMULA Language Highlights

A. *Notation:* PROMULA is a structured language especially useful for developing applications quickly. Its elegant notation, structured concepts, and built-in functions minimize the time it takes to develop serious, mainframe-size applications on a desk top computer. The self-documentation notation of PROMULA enhances the readability of programs.
B. *Language Tutorial:* This reference aid is an on-line, menu-driven tutorial that allows the user to obtain information about PROMULA while programming or using an application.
C. *Language Primer:* This learning aid is a series of commented source programs designed to demonstrate the PROMULA language constructs (nouns) and the PROMULA commands (verbs).
D. *Tutorial Writer:* A tutorial writer allows the creation of menu-driven, application-specific tutorials by simply typing them in. It converts whole books or reports into on-line, menu-driven tutorials.
E. *Menu Manager:* PROMULA's menu manager prepares pick and data menus for "user friendly" applications. Menu preparation is as easy as typing the menus on the screen.
F. *Data Editor:* A screen editor allows data entry and update. Using techniques similar to those found in spread sheet programs, PROMULA can browse through the "pages" of multidimensional arrays to view or change their values.
G. *Report Generator:* A general-purpose report generator displays information in flexible tabular or report formats.
H. *Graphics:* PROMULA supports business graphics (point plots, x-y plots, bar plots, etc.) for both monochrome and color display monitors.
I. *Command Mode:* In command or direct mode, PROMULA accepts a statement, converts it to executable instructions, and proceeds to the next statement.
J. *Compilation Mode:* In indirect or compilation mode, PROMULA compiles a group of statements as a procedure or a program that can be run later. A procedure can be run by other procedures, including itself.

K. *Conversational Mode:* The user can interact with a PROMULA program either in command mode or by responding to conversational prompts. Conversational prompts are valuable when designing a program for use by others.
L. *Debugging Mode:* The user can interrupt a program dialogue, perform local operations in command mode, and return to the same place where she/he left the program. This is a very useful debugging feature.
M. *Multidimensional Data Structure:* Unlike the two-dimensional view of spread sheets, PROMULA supports a multidimensional data structure. Data arrays in PROMULA can have ten dimensions, making it easy to define and consolidate highly structured information. The information of a PROMULA program is structured into variables and sets. Variables are multidimensional arrays whose subscripts are sets. Variables store the information and sets classify information. PROMULA variables can have up to ten dimensions and can be as large as the computer allows.
N. *Array or Matrix Equations:* PROMULA equations are written in standard algebraic notation. The equation operands may be scalars, vectors or multidimensional arrays. Implicit and dummy subscripts allow a condensed notation for array equations. Simpler in notation, this feature is comparable to a similar capability of the APL language.
O. *Equation Solver:* PROMULA's equation solver gives solutions to systems of simultaneous equations.
P. *Variable Management System:* In PROMULA, a program is a database as well as computational procedures. The database contains the input and output variables of the program as well as other supporting information. The program database can be used independently of the program code, and the user can even interrupt a running program to work with its database. In addition to sequential access text files and direct access binary files, PROMULA supports a unique variable management system. This is a multidimensional array management system that is ideal in managing the information usually stored in program variables. Unlike other DBMS systems, which have limited command languages, PROMULA is a fully-featured applications programming language, so it offers full flexibility in analyzing and using information retrieved from its database.
Q. *Program Management System:* PROMULA has a program manager to help handle large, mainframe-size programs. If the program code is too large, the program manager can divide the code into manageable segments. If variable arrays are too large or there are too many for the work space, they can be can stored on disk. PROMULA's variable manager brings only what is needed into the computers work space.
R. *Dynamic Simulation:* PROMULA supports dynamic simulation applications. It has all the special functions needed to develop system dynamics models—models of systems whose variables interact with each other continuously as they evolve over time.

PROMULA was recently selected by the American Public Power Association as the host software for POWER MANAGER - a comprehensive library of management, planning, and engineering applications for public power systems.