Procedure for the Calibration of a Semicompensatory Mode Choice Model

EIJI KAWAMOTO AND JOSÉ REYNALDO SETTI

A two-step method for the calibration of semicompensatory models is presented. To demonstrate the use of the method, it is applied to a model that represents the process of choosing modes for work trips. The calibration of semicompensatory models, such as the one presented here, is not a trivial process because it involves finding the best set of parameters for two functions while satisfying a series of inequalities. In the example shown here, the inequalities are used to determine whether the modal choice predicted by the model corresponds to the user's choice. The best set of parameters is that corresponding to the fewest differences between the observed and predicted choices. The first stage in the proposed calibration process is a preliminary fitting, which attempts to find the maximum of a deterministic function using a process that resembles the maximum likelihood calibration method. The second stage uses the first parameters determined in the first stage as an initial solution and then tries to find the best fit through an exhaustive search around the initial guess. The justification of this two-step procedure is that the efficiency of the calibration process will be increased, since the technique used in the first stage is faster than that used in the second stage. The proposed procedure ensures that an accurate answer is obtained in a reasonable time while allowing the user to determine the sensitivity of each calibration parameter. The calibrated model was able to correctly predict more than 85 percent of the modal choices observed.

Semicompensatory models make up a class of disaggregated behavior models that may be used to represent the behavior of trip makers who are choosing travel modes and routes. Two other classes of disaggregated behavior models may be identified: compensatory and noncompensatory models. The main difference among these three types of models is the assumption about whether compensations can be made among the attributes that influence the trip maker's decision. The assumption of compensatoriety implies that a high level of satisfaction with one attribute offsets low levels of satisfaction with others (*1*). For example, some models assume that time and cost are compensatory attributes. In terms of the trip maker's perception of a mode's utility, this could mean that the higher cost of a particular mode may be offset by the reduction in travel time obtained when using that mode.

The logit and probit models are two well-known compensatory models. In these models, some amount of utility is associated with each travel mode. The value of the utility of a particular travel mode may be calculated as a function of variables that characterize the socioeconomic situation of homogeneous groups of users, travel costs of the mode, and the mode's attributes (such as comfort, safety, etc.). In a compensatory model, the probability of a user choosing a given mode increases as the relative utility of that mode increases.

Noncompensatory models assume that choices are made on the basis of attribute-by-attribute comparisons of available alternatives and minimum thresholds of acceptability. Noncompensatory models do not consider trade-offs among attributes (*1*). Examples of noncompensatory models are the lexicographic, the conjunctive, and the disjunctive models (*1–3*), among others. Young has used the elimination-by-aspect technique proposed by Tversky (*3*) in a residential location-choice model, which is a good example of the application of a noncompensatory model (*4*).

Semicompensatory models are based on the assumption that trip makers perceive and distinguish between two categories of utilities: (*a*) an intrinsic utility of a mode and (*b*) the utility of the money spent to use a given mode. The intrinsic utility of a mode is a function of its attributes (such as comfort, safety, travel time, etc.), whereas the utility of the money spent to use this particular mode depends on the trip maker's socioeconomic characteristics. The model also assumes that compensatoriety is only admitted among attributes classified in the same category (such as cost and income, or comfort and travel time) (*5*).

MATHEMATICAL REPRESENTATION OF UTILITIES

In the context where travel is considered an intermediate activity allowing access to other activities, it may be assumed that all trip makers want to minimize travel time, physical effort, and other inherent effects of locomotion. Therefore, the intrinsic utility of a mode increases as its level of comfort and rapidity increase—where rapidity is defined as the ratio between the origin-to-destination straight-line distance, raised to a certain exponent, and the travel time, raised to another exponent.

The semicompensatory structure assumes that an individual's decision about the use of the mode perceived as having the greatest intrinsic utility depends on the individual's perception of the utility of the amount of money required to use that particular mode, which is a function of the out-of-pocket cost associated with the mode and of socioeconomic factors such as income and number of dependents. If the intrinsic utility of a given mode is greater than the utility of its out-of-pocket cost, that mode will be chosen for the trip; otherwise, this model will be considered too expensive, and the

Department of Transport Engineering, University of São Paulo, 13.560–São Carlos, SP, Brazil.

second-best alternative is taken into consideration in a similar way.

The intrinsic utility of a mode is expressed as a function of the following attributes: travel time, amount of physical effort required (a proxy for comfort), and straight-line distance between origin and destination. The utility of the money spent for using a mode is described as a function of out-of-pocket cost, household income, and number of dependents.

These two utility functions have a multiplicative form, because previous studies have shown the adequacy of the multiplicative rule in representing the perception of a multiattribute stimulus (*6*) and human judgment concerning travel behavior (*5*). In other words, the perception of a set of attributes by a certain user may be represented by a multiplicative model in terms of actually measured values and not perceived values. For instance, the model uses "real" data for travel time or distance instead of values obtained from answers to questionnaires—which are affected by the respondent's perception. Thus, the intrinsic utility of Travel Mode m is given by the expression

$$I_m = \alpha_0 \cdot D^{\alpha_1} \cdot T_m^{\alpha_2} \cdot E_m^{\alpha_3} \qquad (1)$$

where

- I_m = intrinsic utility of Mode m;
- D = straight-line distance between origin and destination;
- T_m = travel time by Mode m;
- E_m = physical effort required for traveling by Mode m, defined as the amount of bodily energy spent by the user when traveling by Mode m, given the travel time; and
- α_i = calibration constants, which transform objective measurements into perceived values.

Note that the level of comfort is taken into account by the model insofar as comfort is the inverse of physical energy, E, raised to some power.

The second equation, for the utility of the money required to use Mode m, is given by

$$S_m = \beta_0 \cdot P_m^{\beta_1} \cdot R^{\beta_2} \cdot N^{\beta_3} \qquad (2)$$

where

- S_m = utility of the money required to use Mode m;
- P_m = out-of-pocket cost for using Mode m;
- R = household income;
- N = number of people depending on the household income; and
- β_i = calibration constants.

A trip maker j chooses the mode for a trip by first ranking the available modes according to their intrinsic utilities: $I_p^j > I_q^j > I_r^j > \ldots$. The intrinsic utility for the most preferred mode (Mode p) is then compared with the utility of the money required to use that mode: if $I_p^j > S_p^j$, then Mode p is chosen; otherwise, the second-highest-ranked mode is considered. Therefore, Mode q is chosen if $I_q^j > S_q^j$. If $I_q^j < S_q^j$, the process is repeated until a mode whose intrinsic utility is higher than the money utility is found.

CALIBRATION OF THE SEMICOMPENSATORY MODEL

The calibration of compensatory disaggregated behavior models uses the probability that an individual belonging to a homogeneous group will choose a certain alternative, measured as the frequency of occurrence of each alternative. The main difficulty in calibration of semicompensatory models is the lack of a measurable variable linked directly to the choice of an alternative (e.g., the probability of choosing private car). However, this does not rule out probabilistic approaches to semicompensatory models—Kawamoto has proposed a probabilistic structure for the semicompensatory model (*7*). The calibration of such a model would require observations of the frequency of mode utilization for homogeneous groups of users.

The semicompensatory model, as proposed by Kawamoto (*5*), should be calibrated for each person in the data set through the comparison of observed and predicted choices. This is because it is almost impossible to determine individual propensities of choosing an alternative from observed individual choices. Although this deterministic approach may cause some operational difficulties, it allows for a better understanding of the process of mode selection because the underlying assumptions about the structure of the trip maker's behavior are explicit.

The multiple regression approach for the calibration of the model was discarded because of potential problems in the collection of accurate data. To use a multiple regression model, it would be necessary to know the points of indifference between the two utilities. Therefore, each subject interviewed would be required to state at least one combination of attributes of a mode that would make that mode's intrinsic utility equivalent to the utility of the money required to use it (for instance, the price of fuel that would cause the trip maker to stop using a car, and so on). Responses to this type of question are usually not reliable because the subject must think about hypothetical situations and not about real ones. Furthermore, it would be necessary to assume that these stated combinations of attributes are really representative of the points of indifference between utilities.

Linear programming was also considered for the calibration of the model. The objective function would be some function that would reflect the difference between the predicted and observed choices, subject to the restrictions represented by the inequalities, which would also need to be linearized. The main problem with this approach is that a solution (or solutions) for the problem would have to satisfy all restrictions, a condition that is equivalent to correctly predicting all observed choices and that is very unlikely to occur.

To avoid such pitfalls, Kawamoto has proposed that the best way to calibrate the model would be to use data on choices that people have actually made, given the available travel modes (*8*). Each subject interviewed is asked to rank the available alternatives. It is then possible to find the rank of the mode each person in the sample actually used for his or her trip. For instance, if an individual has three alternative modes available for a trip, the person can rank the modes according to their perceived intrinsic utilities as well as indicating which mode is actually used. Hence, it can be determined whether the mode used is considered best, second-best, or third-best.

If the chosen alternative is the best of the three available, the value of its intrinsic utility (I_1) must not only be the greatest among the three alternatives ($I_1 > I_2 > I_3$, where I_2 and I_3 are the intrinsic utilities of the modes ranked second and third, respectively) but the intrinsic utility of the selected mode (the one ranked best) must also be greater than the utility of the amount corresponding to the out-of-pocket cost of this alternative ($I_1 > S_1$).

If the alternative used is the second-best, the following inequalities are valid:

$$I_1 > I_2 > I_3$$

$$I_1 < S_1$$

$$I_2 > S_2$$

where S_2 is the utility of the amount corresponding to the out-of-pocket cost for the alternative ranked second. Finally, if the individual can only use the third-best alternative, the values of the intrinsic utilities must satisfy the following inequalities:

$$I_1 > I_2 > I_3$$

$$I_1 < S_1$$

$$I_2 < S_2$$

$$I_3 > S_3$$

The number of inequalities that must be verified for a particular trip maker depends on the number of alternatives and the rank of the alternative selected.

The first stage in the two-stage calibration procedure tries to find values for the calibration constants α_i and β_i such that most of the preceding inequalities are satisfied for the largest number of subjects in the sample. The procedure adopted in the first stage resembles the maximum likelihood method, although the utility functions used are deterministic. The second stage uses the results of the first stage as an initial guess and tries, through exhaustive search, to find regions of optimal values around this starting point.

First Stage

The calibration of the semicompensatory model consists of finding a set of parameters that make the previously defined set of inequalities true for the maximum number of individuals in the calibration data set. The first step in the proposed two-stage calibration technique tries to find an initial set of parameters V_0 quickly through a process that resembles the maximum likelihood calibration technique, in spite of the deterministic nature of the functions used.

Kawamoto (*8*) has used a technique for the calibration of semicompensatory models that involves two functions. The first function, $f_{jk}(V_l)$, verifies whether the kth inequality is true for User j, given a parameter vector V_l:

$$f_{jk}(V_l) = \frac{e^{U_x}}{e^{U_x} + e^{U_y}} = \frac{1}{1 + e^{U_y - U_x}} \quad (3)$$

where U_x and U_y are utilities and e is a constant, usually the base of natural logarithms, 2.718 . . . .

This function ranges from 0 to 1: if $f_{jk} > 0.5$, then $U_y < U_x$; if $f_{jk} < 0.5$, then $U_y > U_x$. For each user j there is a corresponding number of inequalities t_j to be checked, which depends on the number of alternatives and on the rank of the selected alternative.

A second function, $g(V_l)$, is defined for a vector of calibration parameters V_l as follows:

$$g(V_l) = \prod_{j=1}^{n} \prod_{k=1}^{t_j} f_{jk} \quad (4)$$

where

f_{jk} = function indicating whether a particular inequality is true (Equation 3) for User j,
n = number of subjects in the sample used for calibration of the model, and
t_j = number of inequalities defined for User j.

This function is submitted to a maximization procedure to find the best set of calibration exponents.

Despite its computational efficiency, three problems are associated with this approach:

1. The function f_{jk} (Equation 3) used to check whether an inequality is true may distort the results because the results of the test are weighted. For instance, consider two situations, one where $f = 0.9$ and another where $f = 0.7$. Both represent situations where the inequalities are true ($f > 0.5$), but higher values of f will generate higher values of g, distorting the results.
2. The maximization of Function g corresponds to the maximization of the number of true inequalities. Unfortunately, the largest number of true inequalities may not correspond to the minimum difference between predicted and observed choices.
3. Although the maximization of Function g produces a vector of calibration parameters V_0, there is no warranty that the minimum difference between predicted and observed choices corresponds to only one vector, V_0. In fact, given the discrete nature of the objective function (number of correctly predicted choices), there may be several vectors that can yield the same degree of precision.

The first stage in the calibration process presented here is largely based on Kawamoto's 1989 procedure. A critical change is that the function f_{jk} is modified to avoid the introduction of distortions because of the weighting of the results of the inequality checks (Item 1). Thus, f_{jk} has been changed to

$$f_{jk} = \begin{cases} 1.0 & \text{if the inequality is true} \\ 0.9 & \text{otherwise} \end{cases} \quad (5)$$

This change eliminates the first of the problems with the former approach. To minimize the influence of the other two problems, the new process includes a second stage, which uses the calibration vector V_0 determined in this first step as a starting point in the search for the best exponents for the utility expressions (described by Equations 1 and 2).

Second Stage

The procedure adopted for the second stage needs an initial "guess" for the calibration parameters—here, the exponents obtained by the first stage. Through an exhaustive search procedure, small variations are introduced in these initial values, and the number of correctly predicted choices is calculated for each variation in each exponent. The number of correctly predicted choices is determined through the computation of the utility functions values for each subject in the sample; if all inequalities for each subject are true, the predicted choice is correct.

This procedure is computationally not efficient. For instance, if the search is carried out for 10 values around the initial guess, there are 10^7 combinations of calibration parameters to be verified, and the number of correctly predicted choices has to be determined for each of these 10^7 vectors. The computational inefficiency of this procedure rules out the possibility of its sole use unless enough computing resources are available.

DATA COLLECTION AND MODEL CALIBRATION RESULTS

Data Collection

The data used to demonstrate the model calibration procedure proposed here were collected in two medium-sized cities in Brazil (São Carlos and Campinas) in May 1989. Both cities are in the state of São Paulo in the southern region of the country. The population of Campinas is roughly 1 million; Campinas is 95 km northwest of the city of São Paulo. São Carlos is about 230 km northwest of São Paulo; the city's population is 160,000. Both Campinas and São Carlos are fairly industrialized and are major urban centers in the state.

The method adopted for the data collection was to interview subjects at their workplaces. In São Carlos, interviews were carried out at the campus of the University of São Paulo (USP). In Campinas, data were collected at the Highway State Department Regional Headquarters (HSD). The choice of sites was based on their availability (the interviewers were known by the workers) and the fact that the reliability of certain responses (such as trip length, travel time, etc.) could be determined.

The inclusion of data from Campinas was meant to avoid calibration based solely on short trips. Travel distances for USP workers range from 0.5 to 5 km, with a mean trip length of 2 km; most trip lengths for HSD workers range from 5 to 10 km, with values as high as 18 km. Although these distances may seem short to the North American reader, any trip longer than 15 km is usually considered to be a long work trip for most Brazilians.

The data collected in the interviews included residential address, workplace address, main mode used in the work trip, family income, work trip length, number of people dependent on the family income, travel time, out-of-pocket cost of the work trip, and how the subject would rank the available modes if no expenses were associated with their use. The interviewees were asked to give their best estimates for travel time, distance and cost—the objective was to find "real" rather than subjective values for these variables. The responses to these items in the questionnaire were later checked against reliably calculated values; whenever any significant inaccuracies were noticed in the subject's answers, the calculated values replaced the subject's estimates. The use of this procedure can be justified by the multiplicative form of the model, which has been proved to be able to transform objective measured values into perceived magnitudes by Stevens (*9*) and Louviere (*6*), among others. The reader is referred to these authors for further details on multiplicative models.

The sample consisted of 95 interviewees, 45 in Campinas and 50 in São Carlos. Data related to modes not actually used by the subjects were determined from other sources of information, such as observed bus and car speeds, bus headways and routes, and so forth. This procedure was adopted to avoid errors introduced from any bias toward a particular mode—subjects may not be able to give an accurate assessment of the attributes of the modes they do not use.

The estimate of the out-of-pocket cost associated with use of a private car was made assuming that (*a*) the only cost actually perceived is the fuel cost, (*b*) the average gas mileage under normal urban traffic conditions is 7 km/L of fuel, and (*c*) the morning warm-up cycle consumes 0.3 L of fuel. Travel time for private car users was estimated considering that (*a*) the average morning warm-up cycle for an average car is 5 min (since a large number of cars are fitted with ethanol-powered engines whose warm-up cycle is longer than that of gas-powered engines), and (*b*) the average speed of a car, under normal traffic conditions, is 30 km/hr.

Travel time for bus transit users was calculated on the basis of the following assumptions: (*a*) the average speed for buses is 15 km/hr under normal traffic conditions and (*b*) the total travel time for bus users is given by the sum of the time to walk from home to the bus stop, the wait at the bus stop (half the average headway), the in-vehicle time, and the time to walk from the bus stop to the workplace. Travel time associated with walking was calculated assuming that the average walking speed is 5 km/hr.

Although there may be some degree of correlation between travel time and out-of-pocket cost for automobile trips of these lengths, there is no such correlation between travel time and travel cost for the other two modes—transit fares are uniform for all routes in both cities, and the out-of-pocket cost for walking is nil. Therefore, it may be assumed that the effects of the correlation between travel time and cost are negligible considering that (*a*) the variable travel time is used in the intrinsic utility model (Equation 1) and the variable cost is used in the monetary utility model (Equation 2), and (*b*) that the data set used includes not only drivers but also walkers and public transit riders.

Finally, Table 1 gives the level of physical effort associated with the use of each travel mode (*10*). The physical effort used during a bus trip was estimated as the weighted average of the energy requirements for walking to and from the bus stop, standing at the stop, and riding a vehicle as a passenger.

TABLE 1 Physical Effort Requirements by Mode (*10*)

Mode	Energy expenditure (kcal/min)
Driving	2.8
Walking	4.5
Riding a bus	2.5

Model Calibration Results

The first stage produced the following calibration parameters:

$$I_m = 100D^{1.03}T_m^{-0.60}E_m^{-1.61} \tag{6}$$

$$S_m = 3{,}680P_m^{1.05}R^{-0.82}N^{0.35} \tag{7}$$

where distance is expressed in kilometers; travel time in minutes; energy consumption in kilocalories per minute; and out-of-pocket cost and household income in American dollars. This model was able to correctly predict the choice of 85.3 percent of the subjects in the data set (81 out of 95 cases). The signs of the calibration parameters obtained are consistent with their expected signs. For instance, the greater the travel distance, the greater the utility of a mode, provided time and physical effort are fixed. If a mode allows a longer distance to be traveled with the same time and energy expenditures as other modes, this mode is clearly superior. Similarly, the utility of a given amount of money, perceived by a person whose family income is fixed, increases as family size increases.

Although it is hard to comment on the absolute magnitude of the exponents, it is possible to verify that the relative magnitude of the calibration parameters is also consistent with the observed behavior. For instance, the interviews indicate that the most important attribute in the perception of a mode's utility is its level of comfort. The calibrated model is consistent with this observation: the variable with the highest exponent is physical effort, a proxy variable for level of comfort. Similarly, in the equation for the perception of the utility of an amount of money, the order of the attributes, in terms of their importance, is the magnitude of the amount itself, family income, and family size. This, also, is consistent with the observations.

The second stage was conceived with the main purpose of improving the initial answer through an exhaustive search procedure. Yet, the number of correctly predicted choices did not increase from the first to the second stage. Instead of increasing the accuracy of forecast, the second step indicated that there are many combinations of exponents that can produce the same number of correctly predicted choices. Table 2 gives exponents of eight models and their averages—the constant α_0 is assumed to equal 100. Any of these eight models, as well as the model with the average exponents, is able to correctly forecast the choices of 81 of the 95 subjects interviewed. If smaller increments were used in the exhaustive search, other models would be found.

TABLE 2 Calibration Parameters

Model	Calibration parameters α_0	β_0	α_1	β_1	α_2	β_2	α_3	β_3
1	100	3400	0.990	1.110	-0.620	-0.820	-1.610	0.340
2	100	3400	0.990	1.110	-0.620	-0.820	-1.610	0.360
3	100	3500	0.990	1.110	-0.600	-0.820	-1.630	0.340
4	100	3500	0.990	1.110	-0.600	-0.820	-1.630	0.360
5	100	3500	0.990	1.110	-0.600	-0.820	-1.630	0.380
6	100	3600	0.990	1.110	-0.620	-0.840	-1.670	0.340
7	100	3600	0.990	1.110	-0.620	-0.840	-1.670	0.360
8	100	3700	0.990	1.110	-0.620	-0.840	-1.630	0.360
mean	–[a]	3525	0.990	1.110	-0.613	-0.827	-1.635	0.355
σ	–	103.510	0.000	0.000	0.010	0.010	0.023	0.014

[a] α_0 was assumed to be a constant.

The existence of multiple solutions able to produce the same number of correctly predicted choices is due to the discrete character of the objective function, the number of correctly forecasted choices. Although small variations in the calibration parameters (as given in Table 2) produce the same number of correct predictions, the set of subjects whose choice was correctly forecast is not the same for all the models. There may be a subset of subjects whose choice is correctly predicted by all models, but there may also be some subjects whose choice is correctly predicted by one model and not by the others. In fact, there is a group of 77 people whose choice is always correctly forecast by the models given in Table 2; the differences found among the results produced by the eight models are due exclusively to the composition of the remaining subset (four people). Therefore, the semicompensatory model's results are stable for the majority of the people in the data set used.

CONCLUSIONS

The two-stage calibration procedure presented here was shown to be a feasible way for calibrating a semicompensatory mode choice model. The calibrated model is able to correctly predict more than 85 percent of the observed choices. A particular characteristic of the proposed calibration procedure is that it is able to come up with many models, each having the same degree of accuracy as measured by the number of correctly predicted choices. This characteristic is due to the discrete nature of the objective function.

Because of the limitations of the data set used, it is not possible to say that the utility functions obtained in the calibration procedure represent the users' perceptions, although the authors believe that the exponents obtained are good approximations to the real ones. Larger data sets would improve the accuracy of the calibration, but larger data sets would also need longer processing times. To analyze the spatial and temporal transferability of the calibrated model, it would be necessary to calibrate the model using data sets collected in different regions and countries.

REFERENCES

1. J. F. Foerster. Mode Choice Decision Process Models: A Comparison of Compensatory and Non-Compensatory Structures. *Transportation Research*, Vol. 13A, 1979, pp. 17–28.
2. T. F. Golob and A. J. Richardson. Noncompensatory and Discontinuous Constructs in Travel Behaviour Models. In *New Horizons in Travel-Behavior Research* (P. R. Stopher, A. H. Meyburg, and W. Brog, eds.), Lexington Books, Lexington, Mass., 1981, pp. 369–384.
3. A. Tversky. Elimination by Aspects: A Theory of Choice. *Psychological Review*, Vol. 79, No. 4, 1972, pp. 281–299.
4. W. Young. A Non-Tradeoff Decision Making Model of Residential Location Choice. *Transportation Research*, Vol. 18A, 1984, pp. 1–11.
5. E. Kawamoto. *A Novel Approach to Modal Choice in Transport Based on Multidimensional Psychophysics* (in Portuguese). Ph.D. thesis. University of São Paulo at São Carlos, São Carlos, São Paulo, Brazil, 1988.

6. J. Louviere. Psychological Measurements of Travel Attributes. In *Determinants of Travel Choice* (D. A. Hensher and Q. Dalvi, eds.), Saxon House, Westmead, Farnborough, Hants., England, 1978, pp. 148–186.
7. E. Kawamoto. Semi-Compensatory Choice Process: A Probabilistic Model (in Portuguese). *Revista de Transporte e Tecnologia*, Vol. 1, No. 2, 1989, pp. 31–39.
8. E. Kawamoto. Calibration of a Semi-Compensatory Model (in Portuguese). Presented at the 1989 Conference of the Brazilian Society for Research in Transportation Engineering (ANPET). Forthcoming in *Revista da ANPET*.
9. S. S. Stevens. *Psychophysics—Introduction to Its Perceptual, Neural, and Social Prospects*. John Wiley & Sons, New York, 1975.
10. *The Demand for Public Transport: Report of the International Collaborative Study of the Factors Affecting Public Transport Patronage*. Transport and Road Research Laboratory, Crowthorne Hill, Berkshire, England, 1980.

Publication of this paper sponsored by Committee on Traveler Behavior and Values.