Household Travel Survey Nonresponse Estimates: The Chicago Experience

ASHISH SEN, SIIM SÖÖT, LIDAN YANG, AND ED CHRISTOPHER

Because response rates vary by household type and by neighborhood, certain groups can be underrepresented. Factoring can rectify this situation somewhat for data used for descriptive purposes, but assumptions underlying many model estimation procedures are violated if factored data are used. Perhaps the only practical solution is to increase the sampling rate of underreporting groups. Because sampling rates should be proportional to reciprocals of response rates, a model to estimate response rates is presented. Such a model could be of value for implementing future surveys. A logit regression model was constructed with the demographic data as independent variables. Observations on the dependent variable, response rate, were obtained from a large-scale household travel survey conducted in the Chicago metropolitan area by the Chicago Area Transportation Study.

Understanding traveler behavior is critical to urban transportation planning and modeling. For this reason, travel surveys are conducted periodically, but their response rates are usually fairly low. Written surveys tend to be cost effective but usually produce low response rates (*1*). Even for oral surveys (including telephone surveys or home interview surveys) where these rates are frequently higher, if one considers people without phones and various prescreening processes, response rates are rarely high (*2*).

A low response rate by itself is not much of a problem. The key difficulty is that these rates vary over different groups. Although in descriptive use of data obtained from such surveys, corrections can be made by means of factoring (*3,4*) this avenue is not always available for modeling uses of these data. For example, consider logit modal split models estimated by maximum likelihood methods. The application of maximum likelihood requires assumptions about the distributions of the number of travelers by mode, and this distribution is usually taken to be multinomial. If one scales up the number of travelers taking each mode by some factor, the resultant products will not have a multinomial distribution. Thus, the use of factored data violates an underlying assumption of the procedure used in estimating the model.

Another example is the usual procedure used for estimating gravity-type trip distribution models, which consists of equating estimated and observed origin trip totals, destination trip totals, and frequencies of trip travel times. This procedure is also a maximum likelihood procedure, and the same comments apply to it as for modal split models. Even for trip generation models, the situation is similar (*4*). Whether one applies a categorical method or a "regression" approach, one is using a linear model, and the assumptions used in estimating it are violated if factored data are used.

Thus, in general, factored data cannot be used for model building and they do not need to be, because assumptions underlying the procedures used to estimate the models do not require the data to have been gathered through random sampling. However, the fact still remains that the travel behavior of groups who are underrepresented in the sample remain ill represented in the final model.

One partial solution to this problem is to increase the sampling rate for such groups—that is, to sample a larger number of households from such groups so that, even if they respond at a lower rate, the group is better represented in the survey. In fact, this is seen as the only solution. Fortunately, because the estimation procedures for the models do not require complete randomness, the level of oversampling can be somewhat rough.

To appropriately oversample from underrepresented groups, one needs to identify groups that have low response rates and to estimate response rates for these groups. Because this information is needed before the survey is conducted it must be obtained from previous surveys. However, because previous surveys would have occurred over a different period and perhaps even a different geographical area, it is likely that the estimates would not be precise. Thus, it is not suggested that an improved sampling scheme would render factoring unnecessary for descriptive (as opposed to model building) uses of travel data. However, selective over- and undersampling will improve the quality of the ultimate data set.

More precise knowledge of nonresponse effects can also lead to other benefits. Special efforts could be made to elicit a higher level of response from underresponding groups. In this latter context, this paper describes in passing one such effort with the Hispanic community and how it paid response dividends. But the key reason for estimating response levels for different groups is for model estimation.

In this study, response rates observed in a large-scale mail survey—the Chicago Area Transportation Study (CATS) Household Travel Survey (HHTS)—were used to estimate a regression model that identifies response rates of key population subgroups. The model is described in the next section. The sections that follow it will be devoted to describing the data on which the model is based and the steps used in constructing the model.

The model was based on Chicago data and no claim is made here about the universality of the findings. Given the importance of the use of judicious over- and undersampling, it is hoped that similar models would be constructed elsewhere. Nevertheless, it is guessed that, in most major U. S. cities, adjustment of sampling rates using the results of this model would be better than no adjustment at all.

A. Sen and S. Sööt, Urban Transportation Center, University of Illinois at Chicago, M/C 357, 1033 West Van Buren Street, Suite 700 South, Chicago, Ill. 60607. L. Yang, Metra, 547 West Jackson Boulevard, Chicago, Ill. 60661. E. Christopher, Chicago Area Transportation Study, 300 West Adams Street, Chicago, Ill. 60606.

RESPONSE RATE MODEL

Model Estimates

The final response rate model is a logit-type model with 12 independent variables. It is of the form

$$\text{RATE}_i = \frac{\exp(Z_i)}{1 + \exp(Z_i)} \tag{1}$$

where RATE_i represents the response rate of zone i, and Z_i is estimated by

$$\begin{aligned}\tilde{Z}_i = 0.38 &- 2.89 PMACH_i - 1.70\text{PLBR}_i - 1.14\text{PTECH}_i \\ &- 1.40\text{PH1}_i - 0.53\ \text{PH4}_i \\ &+ 0.79 PW3_i \\ &- 0.30 S1_i - 0.26 S3_i \\ &+ 2.83\text{PC0}_i - 1.16\ \text{PC02}_i \\ &- 1.09\ \text{PAFRO}_i - 0.58\text{POTHER}_i\end{aligned} \tag{2}$$

The independent variables in the model were obtained from the Census Transportation Planning Package (CTPP) and are represented by boldfacing in Table 1. The table also presents all variables

TABLE 1 Description of Variables in Initial Model

Variable	Description
PEXEC	Proportion of the labor force with executive, administrative and managerial occupations
PPROF *	Proportion with professional occupations
PMACH	Proportion with machine operator occupations
PLBR	Proportion with transportation and material moving, machine handlers, helpers and labors, household service, and service occupations
PTECH	Proportion with technicians, administrative clerical occupations
POTHOCC	Proportion with other occupations, like arm force or farmers
PH1	Proportion of households with 1 member
PH2	Proportion of households with 2 members
PH3 *	Proportion of households with 3 members
PH4	Proportion of households with 4 or more members
PW0	Proportion of households with no worker
PW1	Proportion of households with 1 worker
PW2 *	Proportion of households with 2 workers
PW3	Proportion of households with 3 or more workers
S1	S1 = 1 if mean household income is less than $30,000, S1 = 0 otherwise
S0 *	S0 = 1 if mean household income is between $30,000 and $60,000, S0 = 0 otherwise
S2	S2 = 1 if mean household income is between $60,000 and $100,000, S2 = 0 otherwise
S3	S3 = 1 if mean household income is over $100,000 for this section, S3 = 0 otherwise
PC0	Proportion of households with no vehicle
PC02	Square of PC0
PC1	Proportion of households with 1 vehicle
PC2 *	Proportion of households with 2 vehicles
PC3	Proportion of households with 3 or more vehicles
PWHITE *	Proportion white
PAFRO	Proportion African American
POTHER	Proportion other races
PEMPLY *	Proportion of the employed to the total labor force
PUNEMP	Proportion of the unemployed to the total labor force
PHISPN*	Proportion Hispanic origin
PNHISP	Proportion non-Hispanic origin

1). All observations are based on square mile micro-zones.

2). The boldfaced variables were those included in the final model. See Equation(2).

3). Other variables are base group variables; those variables with * are initially considered to describe the base group.

initially considered in this study. Those variables were chosen on the basis of previous studies, which revealed that nonresponse usually results in underrepresentation of households with low incomes and low education levels (3). Therefore, such variables as occupation, household size, number of workers in the household, household income, vehicle ownership, and race were considered. However, data availability restrained the choice of variables. For example, the CTPP data do not report educational attainment. But because these data often reflect the degree of literacy and civic consciousness, they may have an influence on the response rates. Therefore occupation was considered as a surrogate for education.

Logit Transformation

In the final model, the variables in each class (e.g., occupation) sum to 1. For example, the sum of PEXEC, PPROF, PMACH, PLBR, PTECH, and POTHOCC is 1. If all these variables were left in the model, an overspecified model with consequent acute multicollinearity or singularity would result because the independent variables including the intercept would then be linearly dependent. The remedy used was to exclude one variable from each variable class. Those variables are indicated by an asterisk in the table and constitute what is called a base group.

When every independent variable is zero valued, all households in the zone have the characteristics of the base group (given in Table 1 without boldfacing). The base household group is white, has two or three members, and owns at least one vehicle. It has one or two people working in a professional, managerial, executive or administrative occupation; the annual household income is in the range of $30,000 to $100,000. In this case, because all independent variables are zeros, $\hat{Z}_i$ equals the intercept $\beta_0 = 0.38$ (Equation 2) and the response rate RATE_i could be estimated to be 0.59 [from Equation (1)].

In the process of variable selection, it was found that many coefficient estimates were close to 0, implying that response rates of population subgroups with correspondent characteristics are close to that of the base group. For instance, people with executive, administrative, and managerial occupations had response rates similar to those of the base population, which is professional. The variables corresponding to these coefficients were excluded from the analysis for the sake of parsimony. In the same way variables with like coefficients were candidates for consolidation. The variable selection process and diagnostics will be further discussed later.

The independent variables used in the model were obtained from the CTPP. Therefore, instead of focusing on individuals or households the focus was on areas of residence. This focus would have been problematic if there had been large demographic variations within zones. It was not a problem in this study (3), and since a very good fit was obtained (subject to the caveats mentioned later), the results appear useful.

When the variable being predicted is a proportion—response rate in this case—the logit model is frequently used. Apart from the fact that it makes no predictions that are larger than 1 or less than 0, which is clearly most appropriate in this context, other reasons have been cited about the value of the model for proportions (5). Further discussion of the interpretation of the model appears elsewhere in this paper.

Although in the transportation literature maximum likelihood methods typically are used for estimating logit models, linear least-squares (LS) methods can be used under certain circumstances, after the dependent variable is transformed and appropriate weights are used (as described later). Indeed, LS methods were used in the earliest applications of the logit models (5). The main advantage of linear LS is the wide availability of diagnostics; moreover, the economical usage of computer time makes it possible to experiment with various variable combinations to find the best model fit.

From Equation 1, the following is obtained:

$$\log \{E(\text{RATE}_i/[1 - E(\text{RATE}_i)]\} = Z_i \tag{3}$$

where $\text{RATE}_i = m_i/n_i$, is the ratio of completed surveys to the total number mailed out to zone i. When n_i is a fixed number, and the values of both m_i and (n_i and ($n_i - m_i$) are large enough, the following transformation (6, p. 188) is used:

$$\log(m_i + 0.5) - \log(n_i - m_i + 0.5) = Z_i + \epsilon_i \tag{4}$$

This transformed function can then be estimated by linear LS. Because the variance of the function on the left side is approximately equal to $\{n_i\, E(\text{RATE}_i)\,[1 - E(\text{RATE}_i)]\}^{-1}$, its reciprocal

$$w_i = n_i\, E(\text{RATE}_i)[\,1 - E(\text{RATE}_i)] \tag{5}$$

would be the weights for the linear LS. Because the expected value of response rate, $E(\text{RATE}_i)$, is included in the weight and is unknown, an iteratively reweighted LS estimate is needed. Such a procedure is often carried out using a nonlinear LS program (6, pp. 298–318). In this work the SAS nonlinear LS program, PROC NLIN was used. Once they were computed, weights were inserted into a weighted linear LS procedure to take advantage of the diagnostic methods.

Standard errors and t-values of parameter estimates are given in Table 2. Although the resulting R^2 of 0.38 appears to be low, the fit is good as seen from the following observation. The value of

$$s = \sqrt{\sum_{i=1}^{n} e_i^2 \,/\, (n - k - 1)} \tag{6}$$

called root mean squares error in SAS is 1.2. The n is the number of observations, k is the number of independent variables, and e_i's are residuals. The s^2 is an estimate of the variance of the appropriately

TABLE 2 Results of LS Estimates on Travel Survey Response Rate

Variable	b_j	$s.e.(b_j)$	$t(b_j)$
INTERCEP	0.38	0.1465	2.623
PMACH	-2.89	0.3978	-7.272
PLBR	-1.70	0.2335	-7.290
PTECH	-1.14	0.2069	-5.520
PH1	-1.40	0.2304	-6.076
PH4	-0.53	0.2074	-2.571
PW3	0.79	0.2794	2.820
S1	-0.30	0.0595	-5.063
S3	-0.26	0.0793	-3.273
PC0	2.83	0.3402	8.321
PC02	-1.16	0.5329	-2.180
PAFRO	-1.09	0.0720	-15.097
POTHER	-0.58	0.1879	-3.081

$(R^2 = 0.38,\ R_{adj.}^2 = 0.37,\ s = 1.20)$

weighted residuals. Because the weight is approximately the reciprocal of the standard deviation of Equation 3 as seen earlier, $s^2 \sim 1$ when the model is well specified. For a variety of reasons, the theoretical minimum of 1 is difficult to achieve. Thus, the fit obtained here has to be regarded as excellent.

Model Application

The estimated coefficients of the variables indicate that the independent variables shown in Table 2 have significantly different effects on response rates compared with the base group. As a class of variables, occupation seems to have the greatest effect on response rates. Also household size and vehicle ownership are key variables. However, the coefficients of the variables representing unemployment and households of Hispanic origin are not statistically significant, implying similarity with the base group.

The following example illustrates the use of this model, supposing that there is a diversified midincome zone *i,* in which

- 40 percent of the labor force is employed as professionals and managers and 21 percent is employed as machine operators; 5 percent have transportation, material moving, machine handling, or service occupations; and the remaining 34 percent are technicians or clerks ($\text{PMACH}_i = 0.21$, $\text{PLBR}_i = 0.05$, $\text{PTECH}_i = 0.34$);
- 64 percent of the households have four or more members; another 27 percent of the households have two to three members; and the remaining 9 percent are single-member households ($\text{PH1}_i = 0.09$, $\text{PH4}_i = 0.64$);
- All of the households have one or two workers ($\text{PW3}_i = 0$);
- The mean household income of the zone is \$58,000 ($S1 = 0$, $S2 = 0$);
- 95 percent of the households have at least one vehicle ($\text{PC0}_i = 0.05$, $\text{PCO2}_i = 0.0025$); and
- 12 percent of the households are African-Americans; 10 percent of the households belong to other minority groups; and the rest of them are nonminority whites ($\text{PAFRO}_i = 0.12$, $\text{POTHER}_i = 0.10$);

To estimate the response rate of this zone *i,* Equation 2 may be used obtaining $Z_i \approx -1.21$. The response rate $RATE_{:iti}$ can then be obtained by Equation 1, and its estimate would be 0.23. This information would be helpful in deriving the sampling rate. If the target number of respondents of this zone is 100, then the sample drawn from this zone should be $100/0.23 \approx 435$. More precise knowledge of the nonresponse leads to more effective targeting of the survey distribution and thus helps in obtaining better survey results.

Discussion of Results

This model suggests that individuals with managerial and professional occupations have higher response rates than blue collar workers. Because occupation reflects education, this is not a surprising finding.

Household size is another class of key variables. It is widely believed that larger households are less likely to respond to travel surveys. Because the CATS survey requested that each household member over 14 report all of his or her trips made during a given weekday, large households were candidates for nonresponse. This study suggests that households with four or more members have lower response rates. This low response rate can be partly offset when there are three or more workers in the household. However, it also appears that single-member households have much lower response rates, which may reflect their lifestyle or attitude toward surveying.

Perhaps because of the existence of multicollinearity, household income plays a minor role in this model. The highest response rates come from the middle-income households, namely, \$30,000 to \$100,000, which is specified as the base group. The low response rates for households with lower income might be because those groups are not comfortable with written surveys: lower income usually is associated with lower education. Household members with high incomes might just be too busy to respond.

In contrast to some previous studies (*7*), this model suggests that households without vehicles are more likely to respond. One reason is that the nonresponse effect is represented by the coefficients of occupation, household size, and income. Second, this is not too relevant for suburban households, where vehicle ownership rates reach 100 percent. Still, the higher response rates partly suggest that people without vehicles have a stronger tendency to respond to the survey. Because of their mobility dependence on public transportation, they could be more sensitive to transportation issues.

Chicago is a socially diverse metropolitan area with large minority neighborhoods. It was anticipated that response rates from some communities would be low. Accordingly, CATS made an effort to approach the Hispanic population to improve their response rates. From the beginning of the survey design, special attention was given to neighborhoods with large Hispanic percentages. Survey subjects with Hispanic surnames received a Spanish-language insert, which explained the importance of completing the survey form and provided a toll-free telephone number for assistance. Almost 100 calls were received, and most were given specific instructions on how to complete the survey. These efforts resulted in an improved response, which is also demonstrated by the model estimates. This effort was feasible because the agency could target Hispanic surnames, but was not practical with the African-American community—another area where lower response rates were anticipated. The model suggests that the response rates from African-American communities are lower than they are for the white population. For the "other" minority category, which includes Asian-Americans and Native Americans, the average response rates are also lower.

DATA AND METHODS

Data and Survey Methodology

The CATS HHTS is a travel diary-type survey conducted in a period from 1988 to 1991, using a mail-out/mail-back format (*8*). This survey format proved to be an effective means of collecting travel data.

The survey area encompassed seven counties in the Chicago metropolitan area. A total of 79,346 of the 2.8 million households in the region received the survey instruments. The 19,314 completed and usable questionnaires resulted in an average return rate of 24 percent. However, response rates varied by area. As indicated in Figure 1, they ranged from 13 percent for the city of Chicago outside of the central business district (CBD) to 34 percent for Kendall County in the far southwest suburb. This corresponds to the widespread belief that nonresponse for mail surveys is greater for low-income and less-educated households, many of which reside in

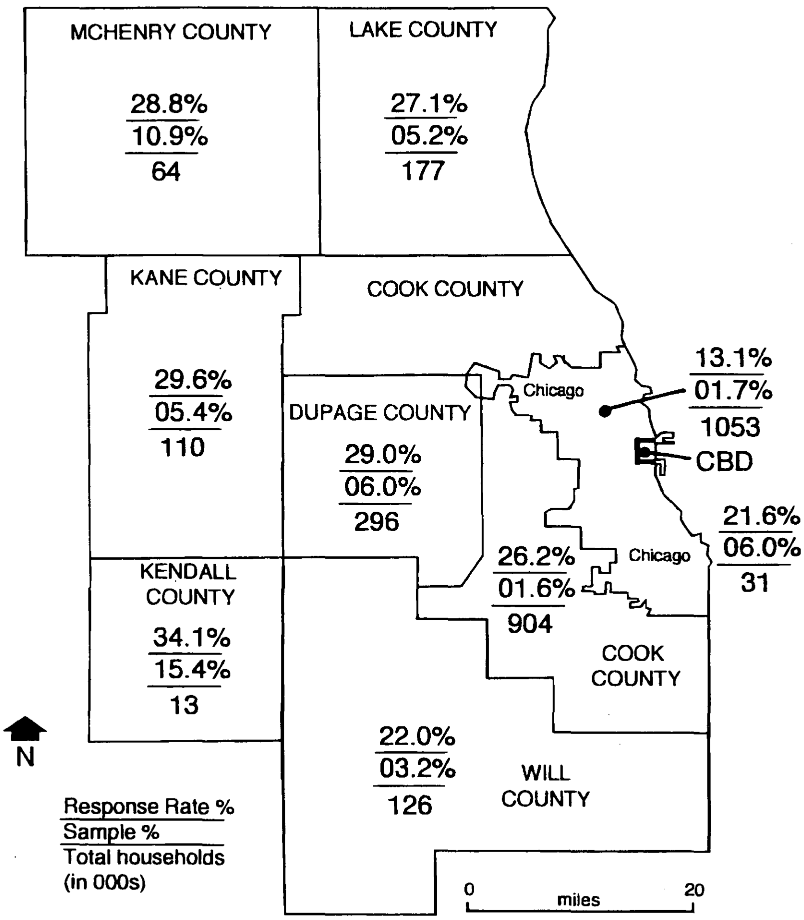

FIGURE 1 CATS household travel survey response rates and sampling sizes.

Chicago. Figure 2 further indicates a wide range of response rates by square-mile zones within the city of Chicago. Response rates were found to be high in the northwest and southwest corners and low west of the CBD.

The sample households were drawn from the local electric company, Commonwealth Edison, records. The number of surveys to be mailed per zone (square-mile area) corresponded to the number of electric meters adjusted by an educated guess of the number of potential respondents. This is the area in which the return rate estimates developed in this paper would have been useful.

Demographic data obtained through the CTPP were used to estimate response rates for population subgroups. These 1990 data were aggregated to a square-mile zone level, defined by the township and range system, to obtain a uniform geographic system with the HHTS. The advantage of the township and range system is that the zones are, for the most part, defined by major arterials resulting in largely homogeneous zones (*9*).

Model Estimate Diagnosis: Outliers and Influential Points

In regression model estimation, it is usual that some observations have large residuals. Sometimes, they occur when some observations reflect conditions or situations different from those under which other observations were obtained. When a few observations with high absolute values of studentized residuals ($|e^*| \geq 3$) were flagged for scrutiny, over half of the zones appeared to belong to Kendall County. This is a suburban county located southwest of Chicago. A rural area recently added to metropolitan Chicago, it consists of two subareas: large residential properties and farms. With a county population of only 38,000, it is vastly different from the neighborhoods throughout the rest of the study area. Therefore, the zones in Kendall County were all excluded.

Two other observations ultimately were deleted. One was a zone in southwest suburban Du Page County. Its response rate (88 percent) was considerably higher than was estimated in the model. A closer study of the zone revealed that this was an area near an employment center. Many residents were professional or skilled workers, either singles or working couples with no children. This group of people was likely to be different from typical urban young professionals, which contributed to a difference in the survey response rate. It was not possible to set a special variable that indicated the difference between urban or suburban young professionals. Besides, the large number of mailed-out surveys resulted in a large weight, which made the point influential. Therefore this data point was excluded from the model estimate. The second case was a zone near downtown Chicago in which there were many large minority households with extremely low incomes. There were also many zero-worker households implying high unemployment rates, which was not revealed in the census

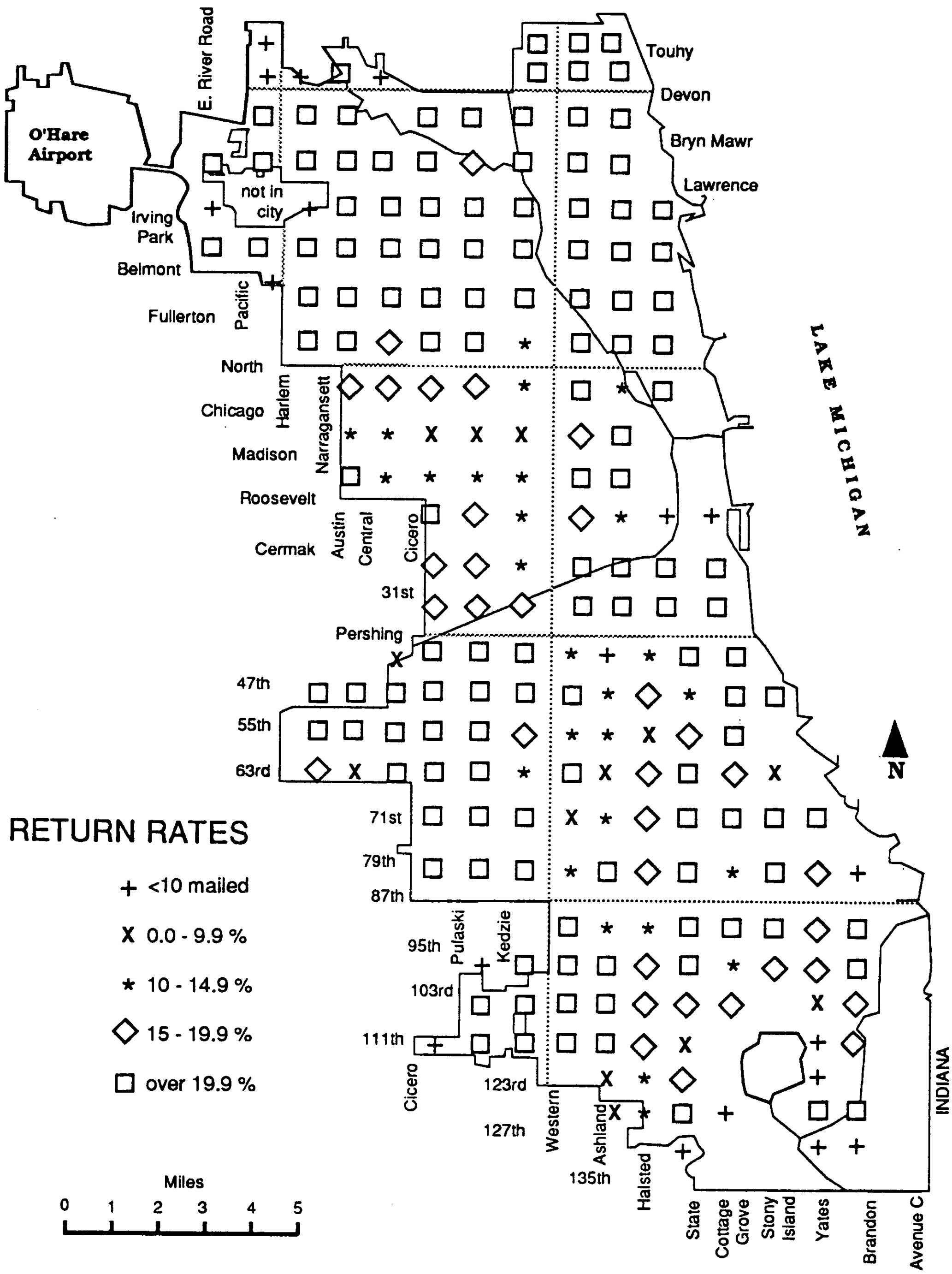

FIGURE 2 City of Chicago travel survey response rates by square-mile zones.

unemployment variable. The large mail-out size gave it a large weight and made it an influential point.

Ideally $|e^*| \geq 2$ occurs for about 5 percent of the observations. In the final model (after discarding Kendall County and the other two observations), there are 77 such data points whose e^* exceeds the critical level of $|e^*| \geq 2$. They are approximately 5.3 percent of the total 1,450 observations, further strengthening the earlier conclusion about a good fit of the model.

Plots of residuals against predicted values and each independent value were also carefully examined to check for the existence of unequal variances and the need for transformations. Figure 3 displays a residual plot. Because in this and other plots no obvious patterns were found, the model appeared acceptable.

Variable Selection and Multicollinearity

Variable selection is a critical process in model building. Given the large number of variables used, multicollinearity is likely present. For instance, low income is usually associated with minority and

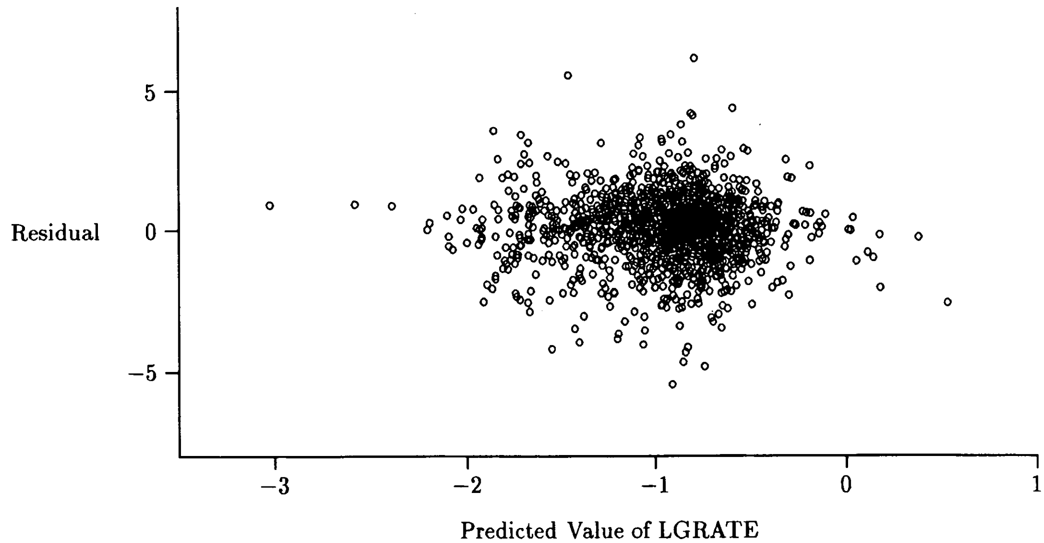

FIGURE 3 Residual versus predicted plot for response rate model.

blue collar occupations; therefore collinearity probably exists among race, income, occupation, and vehicle ownership. In addition, from the collinearity diagnostics, multicollinearity was also found among PH1, PH2, and PH4 (the household-size variables; see Table 1 for definitions). Because the model being constructed is for predictive purposes, multicollinearity is not as much of a problem as for studies in which the significance of variables is the key object—as long as one can reasonably conjecture that the structure of the multicollinearity will be similar for the place and time when the model is applied.

However, parsimonious models are easier to work with and inspire greater confidence because they are easier to interpret. Hence an all-possible-subset-search variable selection procedure was conducted using SAS-weighted PROC RSQUARE. Assuming that the initial model with all variables included is not biased, usually dropping some independent variables will cause bias in the parameter estimates left in the model, except when the values of the deleted variables are orthogonal to those of the remaining variables. To examine the variables, the value of C_p is a good indicator of the presence of noticeable bias. When $C_p \approx p$ (p is the number of independent variables plus one), the bias in the predicteds introduced by dropping variables is usually negligible (*6*, pp. 234–235). Also available are other indicators, such as s^2 and R^2, which estimate the goodness of fit. The combination of those indicators helps in the selection of a concise model with a reasonable level of goodness of fit and low bias. In the all-possible-subsets search, the C_p value began to approach p for $(p - 1) \geq 12$. This indicated that little bias would occur with those sets of suggested 12 independent variables. Their corresponding R^2's were around 0.38. The difference with the highest R^2 occurred only at the third decimal point, which was encouraging.

When variables in one class are similar and their b_j coefficients are alike then consolidation is appropriate. For instance, those occupations requiring analogous skills that also have similar coefficients, such as transportation and material movers, machine handlers, helpers, laborers, and household service and service workers, were combined into one summary variable: PLBR. Also it was seen that response rate estimates for households with one vehicle and households with several vehicles were approximately the same, suggesting that only two variables could be used: households with and without vehicles.

CONCLUSIONS

In a mail-out/mail-back survey, the lower-income, less-educated households are usually underrepresented. Because those population subgroups are usually mobility disadvantaged, it is particularly important to properly estimate and address their needs in transportation planning and policy. However, in many modeling procedures, factoring is not appropriate. Thus, to achieve a desirable number of responses, a carefully designed sampling scheme is necessary in surveys to under- and oversample in subareas.

In this study, such a survey response model was estimated. The model was estimated by linear LS with a logit transformation. Because of the linear LS approach, it was possible to apply a wide range of statistical diagnostic procedures.

The model results do not reveal surprises about response rates. It does, however, provide a means of estimating response rates in urban areas in which considerable demographic variations exist. Although this may not be critical for survey data used for descriptive purposes, it is important if the data are used for modeling.

ACKNOWLEDGMENTS

The authors thank the Illinois Department of Transportation and the Chicago Area Transportation Study for the financial and moral support to undertake this study. In particular the authors express their gratitude to Ed Christopher for his cooperation and assistance in the preparation of this manuscript.

REFERENCES

1. Hassounah, M. I., L. S. Cheah, and G. N. Steuart. Underreporting of Trips in Telephone Interview Travel Surveys. In *Transportation Research Record 1412,* TRB, National Research Council, Washington, D.C., 1993, pp. 90–94.

2. Cohen, B. A., A. P. Lobo, E. L. Fielding, and L. S. Yang. *Analysis of Households Without Phones: Impact for the NPTS 1995.* Office of Highway Information Management, FHWA, U.S. Department of Transportation, 1993.
3. Kim, H. J., J. Li, S. Roodman, A. Sen, and S. Sööt. Factoring Household Travel Surveys. In *Transportation Research Record 1412,* TRB, National Research Council, Washington, D.C., 1993, pp. 17–22.
4. Thakuriah, P., A. Sen, S. Sööt, and E. Christopher. Non-response Bias and Trip Generation Models. In *Transportation Research Record 1412,* TRB, National Research Council, Washington, D. C., 1993, pp. 64–70.
5. Ashton, W. D. *The Logit Transformation with Special Reference to Its Uses in Bioassay.* Hafner Publishing Company, New York, 1972.
6. Sen, A. K., and M. Srivastava. *Regression Analysis, Theory, Methods, and Applications.* Springer-Verlag, New York, 1990.
7. Clark, A. C., and C. Goldstucker. Mail-Out/Mail-Back Travel Survey in Houston, Texas. In *Transportation Research Record 1097,* TRB, National Research Council, Washington, D.C., 1986, pp. 13–19.
8. Ghislandi, A. C., A. R. Fijal, and E. J. Christopher. *CATS 1990 Household Travel Survey: A Methodological Overview.* Working Paper 94-05. Information Services Division, Chicago Area Transportation Study, Illinois, 1994.
9. Christopher, E. J., J. J. Nam, and M. J. Rogus. *CATS Traffic Analysis Zone Structure for the 1990 Census Transportation Planning Package.* Working Paper 94-01. Information Services Division, Chicago Area Transportation Study, Illinois, 1994.

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