Multiply Imputed Sampling Weights for Consistent Inference with Panel Attrition

David Brownstone
Xuehao Chu

UCTC
No 590

The University of California Transportation Center
University of California
Berkeley, CA 94720

The University of California Transportation Center

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, Irvine, and Los Angeles, the Institute of Urban and Regional Development at Berkeley, and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses Faculty and students on other University of California campuses may participate in Center activities Researchers at other universities within the region also have opportunities to collaborate with UC faculty on selected studies.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. It also publishes *Access*, a magazine presenting summaries of selected studies For a list of publications in print, write to the address below

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

University of California
Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel 510/643-7378
FAX: 510/643-5456

Multiply Imputed Sampling Weights for Consistent Inference with Panel Attrition

David Brownstone
Department of Economics
University of California, Irvine CA 92697-5100

Xuehao Chu
Center for Urban Transportation Research
University of South Florida, 4202 E Fowler Avenue
Tampa, FL 33620-5350

Reprinted from
Panels for Transportation Planning Methods and Applications
Chapter 10

UCTC No. 590

The University of California Transportation Center
University of California at Berkeley

P76

Panels for Transportation Planning

Methods and Applications

Edited by

THOMAS F. GOLOB
Institute of Transportation Studies
University of California, Irvine

RYUICHI KITAMURA
Department of Transportation Engineering
Kyoto University

and

LYN LONG
Institute of Transportation Studies
University of California, Irvine

KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON

CHAPTER TEN

MULTIPLY IMPUTED SAMPLING WEIGHTS FOR CONSISTENT INFERENCE WITH PANEL ATTRITION

DAVID BROWNSTONE
Department of Economics, University of California,
Irvine, California 92697-5100, U.S.A.

XUEHAO CHU
Center for Urban Transportation Research,
University of South Florida, 4202 E. Fowler Avenue,
Tampa, FL 33620-5350, U.S.A.

Abstract: This chapter demonstrates a new methodology for correcting panel data models for attrition bias. The method combines Rubin's Multiple Imputations technique with Manski and Lerman's Weighted Exogenous Sample Maximum Likelihood Estimator (WESMLE). Simple Hausman tests for the presence of attrition bias are also derived We demonstrate the technique using a dynamic commute mode choice model estimated from the University of California Transportation Center's Southern California Transportation Panel The methodology is simpler to use than standard maximum likelihood-based procedures. It can be easily modified to use with many panel data estimation and forecasting procedures

INTRODUCTION

Panel studies are often plagued by the attrition of survey respondents. Attrition can bias the sample and limit the usefulness of the panel for long-term dynamic analysis. If the attrition process is correlated with the endogenous variables in the model (called *non-ignorable attrition*), then standard estimation techniques ignoring attrition will yield inconsistent inferences and estimators. Even if the attrition process is independent of the endogenous variables (called *ignorable attrition*), uncorrected attrition may bias forecasts and policy simulations based on the remaining sample. Both of these problems occur in transportation panels. If survey questions are concentrated on the adoption of a new mode or technology, then users of this new mode or technology may be less likely to attrite and the attrition process will be correlated with the endogenous choice variable. Since the main purpose of many analyses of transportation panels is to produce forecasts of the effects of proposed policy changes, it is important to account for the effects of attrition on model forecasts and policy simulations.

This chapter describes a new methodology for obtaining consistent estimates and forecasts from panel models where non-ignorable attrition is present. The basic approach is to use information from early panel waves to fit an attrition model. The inverse of the attrition probabilities resulting from this attrition model are then used as weights in Manski and Lerman's (1977) Weighted Exogenous Maximum Likelihood Estimator (WESMLE). Then, Rubin's (1987) Multiple Imputations technique is used to get consistent standard errors for parameter estimates and model forecasts taking into account uncertainty in the attrition model. This procedure appears to have been first proposed in Brownstone (1991), but it is a simple modification of Rubin (1986).

Relative to joint maximum likelihood estimation of the attrition and choice model, the methodology proposed in this chapter is inefficient. However, this methodology is much easier to calculate than joint maximum likelihood, which is frequently intractable in complex models. The multiple imputations technique proposed here can also be easily combined with closely related methods for handling missing data, and it easily produces consistent forecasts and their standard errors. Simple Hausman (1978) tests can be applied to test for the non-ignorability of the attrition (or missing data) process. Since the WESMLE was originally designed to provide consistent estimates with choice (or response)-based sampling designs, the methodology proposed here can be trivially modified to yield consistent estimates and forecasts for choice-based panels with non-ignorable attrition.

This chapter demonstrates the methodology using a dynamic commute mode choice model calibrated from the University of California Transportation Center's Southern California Transportation Panel. We use the first and fifth waves of the panel (approximately 18 months apart) where there was 40% attrition. Although the attrition process is correlated with the commute mode choice dependent variable and therefore non-ignorable, the magnitude of the resulting biases ignoring attrition is quite small. The next section describes the proposed methodology in more detail. The third section describes the panel data used in the empirical example. The fourth section gives the empirical results and simulations for two policies designed to increase ridesharing in the Greater Los Angeles area.

MULTIPLY-IMPUTED WESMLE

Manski and Lerman (1977) show that a simple modification of the standard Maximum Likelihood estimator for discrete-choice models yields consistent parameter estimates in the presence of choice-based sampling when the proportion of the population choosing each discrete alternative is known. If $L_i(\theta,x_i)$ is the log likelihood function for the i^{th} observation, then Manski and Lerman's WESMLE maximizes:

$$\sum_i \omega_i L_i(\theta, x_i) , \tag{1}$$

where θ is a vector of parameters to be estimated, x_i is the vector of observed characteristics for the i^{th} observation, and the sampling weight, ω_i, is the inverse of the probability that the i^{th} observation (individual) would be chosen from a completely random sample of the population. Of course, if the sampling scheme were completely random, then all of the sampling weights would be equal and the WESMLE would simply be the usual maximum likelihood estimator. Manski and Lerman (1977) show that the WESMLE is consistent and asymptotically normal, but not fully efficient (see Imbens, 1992 for fully efficient alternative estimators). Manski and Lerman's proof actually shows that the WESMLE's properties hold as long as the sampling weights are known with certainty.

A major advantage of the WESMLE is that it can be computed very easily by modifying existing maximum likelihood programs. The WESMLE for both the linear regression model and the multinomial logit model can be computed by appropriately weighting the variables and applying standard maximum likelihood programs. Unfortunately, this procedure yields downward biased standard error estimates, but the consistent estimates given by Manski and Lerman are easy to compute.

A panel survey can always be viewed as the result of the original sampling process and the attrition process. Although the properties of the sampling process are known with certainty in a well-designed panel study, the properties of the attrition process are typically unknown. If they *were* known, then the sampling weights could be easily computed as the inverse of the product of the sampling and attrition probabilities and the WESMLE could be applied to get consistent parameter estimates. Fortunately, there is at least one wave of information about panel attriters, and with some modeling assumptions this information can be used to estimate a model of the attrition process. Unfortunately, the resulting predicted attrition probabilities cannot be used to generate weights for the WESMLE, since this would violate the assumption that the weights are known with certainty.

The simplest way to solve this problem is to use Rubin's (1987) Multiple Imputations technique together with the WESMLE to control for the uncertainty in the weights. This technique, which was used in Brownstone and Golob (1992) to deal with uncertain sampling weights used to generate forecasts from a static commute mode-choice model, uses simulated draws from the distribution of the weights to measure the effects of their uncertainty. Suppose we have a procedure for making independent simulated draws from the sampling distribution of the attrition probabilities (which are given from our estimated attrition model). Conditional on this set of of simulated attrition probabilites, we can compute a vector of sampling weights (as the inverse of the product of the attrition probabilities and the sampling probabilities for the first wave of the panel). This weight vector can in turn be used to get a consistent (conditional on that particular set of weights) estimate of θ and its covariance using the WESMLE. If m independent weight vectors are simulated and m corresponding parameter and covariance estimators, $\tilde{\theta}_j$ and $\tilde{\Sigma}_j$, are computed, then Rubin's Multiple Imputations estimators are given by

$$\hat{\theta} = \sum_{j=1}^{m} \tilde{\theta}_j / m \tag{2}$$

$$\hat{\Sigma} = U + \left(1 + m^{-1}\right) B, \tag{3}$$

where

$$B = \sum_{j=1}^{m} \left(\tilde{\theta}_j - \hat{\theta}\right)\left(\tilde{\theta}_j - \hat{\theta}\right)' \Big/ (m-1) \tag{4}$$

$$U = \sum_{j=1}^{m} \tilde{\Sigma}_j / m. \tag{5}$$

Note that B is an estimate of the covariance among the m parameter estimates for each weight vector, and U is an estimate of the covariance of the estimated parameters given a particular weight vector.

Rubin (1987) shows that for a fixed number of imputed weight vectors, $m \geq 2$, $\hat{\theta}$ is a consistent estimator for θ and $\hat{\Sigma}$ is a consistent estimator of the covariance of $\hat{\theta}$. Of course B will be better estimated if the number of imputed weight vectors is large, and the factor $(1 + m^{-1})$ in equation (3) is to compensate for the effects of small m. Rubin (1987) shows that as m gets large, then the Wald test statistic for the null hypothesis that $\theta = \theta^0$,

$$\left(\theta - \theta^0\right)' \hat{\Sigma}^{-1} \left(\theta - \theta^0\right), \tag{6}$$

is asymptotically distributed according to an F distribution with K (the number of elements in θ) and ν degrees of freedom. ν is given by:

$$\nu = (m - 1)(1 + r_m^{-1})^2 \text{ and} \qquad (7)$$
$$r_m = (1 + m^{-1}) \operatorname{Trace}(BU^{-1})/K .$$

This suggests increasing m until ν is large enough so that the standard Chi-squared distribution of Wald test statistics applies. The results reported in this chapter increased the number of multiple imputations, m, until ν was greater than 100.

Note that the consistency of these multiple imputations estimators crucially requires that the attrition model is correctly specified conditioned on all variables entering the choice model (the x vectors in equation (1)). The multiple vectors of imputed weights must also be drawn independently (conditional on the attrition model) in such a way as to completely reflect all of the uncertainty in the estimated attrition model.

The multiple imputations estimators, $\hat{\theta}$ and $\hat{\Sigma}$, are consistent whether the attrition process is ignorable or not. The standard maximum likelihood estimators, which ignore the sampling and attrition weights, are efficient if both the sampling and attrition processes are ignorable, but inconsistent otherwise. Therefore the statistic:

$$T = \left(\hat{\theta} - \overline{\theta}\right)' \left(\hat{\Sigma} - \overline{\Sigma}\right)^{-1} \left(\hat{\theta} - \overline{\theta}\right), \qquad (8)$$

where $\overline{\theta}$ and $\overline{\Sigma}$ are the maximum likelihood parameter and covariance estimators, is a valid Hausman (1978) test statistic for the null hypothesis that both the sampling and attrition processes are ignorable. Under the null hypothesis, T has a chi-squared distribution with degrees of freedom equal to the rank of $\left(\hat{\Sigma} - \overline{\Sigma}\right)$.

Typically, the most difficult computational step of the multiply-imputed WESMLE described in this section is the repeated computation of the WESMLE for a fixed weight vector. The latter computation can be carried out as easily as standard maximum likelihood for the same choice model, and is generally much simpler than joint maximum likelihood estimation of the choice and attrition model. The more unusual computation is drawing the simulated attrition probabilities from the estimated attrition model. Examples of this are given in Rubin (1987) and Brownstone and Golob (1992). We also give an example of how this might be done in the fourth section of this chapter. We first describe the data used in the empirical example.

DATA DESCRIPTION

These data are from the first wave of a panel study of commute behavior in California's South Coast Air Basin. The study region and survey methodology are more fully described in Uhlaner and Kim (1993). The panel was selected from respondents to a mail survey, and was initiated in February 1990. The first wave of data were drawn from the original sample and from a refreshment sample introduced three months later. The overall response rate for the first-wave mail survey was approximately 50%. The total sample size for the first wave was 2,189 commuters (approximately 1,850 had complete data). Almost all respondents were employed full-time. The fifth wave of the panel was collected beginning in July 1991. The attrition rate (from Wave 1) was 40%, leaving 1,107 respondents whose data were suitable for dynamic analysis.

The panel questionnaires gathered detailed information about each respondent's most recent trip to work, including mode, perceived distance, times of departure and arrival, and number of stops. Respondents were also asked which other commute modes, if any, they used during the previous two weeks. Because of retrospective questioning, the survey provides more information about mode choice than is available from the conventional single-day travel diary typically used in mode choice studies. The first-wave data can be used to fit attrition models, as demonstrated in the next section of this chapter.

Although the initial panel sample was not a probability sample of full-time workers in the South Coast Air Basin, Golob and Brownstone (1992) developed 'sampling' weights by statistically matching the first wave sample to the March 1987 Current Population Survey (CPS) of the U.S. Bureau of the Census. We repeated this process with the 1991 CPS and obtained essentially the same results as did Golob and Brownstone. For the purposes of this chapter, we treat the weights constructed in Golob and Brownstone (1992) as valid sampling weights for the first wave of the panel. These weights are then combined with the attrition weights derived in the next section to obtain valid weights for the panel analysis.

MODELS AND RESULTS

The model we wish to estimate is a simplified but dynamic version of the commute mode-choice model in Golob and Brownstone (1992). We simplify the mode choices to two: 'always drive alone' (during the two-week diary period) and 'rideshare at least once'. Our goal is to model the changes in commute mode choice made between the first and fifth waves of the Southern California Transportation Panel. The simplest such model is a four-alternative multinomial logit model, with the four alternatives being: 'always drive alone in both time periods', 'rideshare in both periods', 'switch from drive alone to rideshare', and 'switch from rideshare to drive-alone'. The choices made by the panel members are shown at the beginning of Table 1 (note that the percentages in Table 1 are not weighted using the sampling weights for the first wave described in the previous section). If we temporarily ignore possible non-ignorable sampling and attrition processes, then we can estimate this multinomial logit model on the panel data by maximum likelihood. The results of this estimation are given in Table 1. Although many of the individual coefficients are significant, it is very difficult to interpret the sign of any single coefficient. Since the ultimate purpose of this type of model is to provide forecasts for the effects of certain policy interventions, we will next calculate the results of two hypothetical policy interventions occurring between the two time periods: giving all commuters access to a guaranteed ride home, and giving all freeway users access to a high-occupancy vehicle lane. The results of these hypothetical simulations, shown in Table 2, are computed exactly as in Golob and Brownstone (1992). The results in Table 2 are given as percentage changes relative to the number choosing the particular alternative in the baseline scenario. For reference purposes, the predicted number in the baseline scenario is given in the third column of the table.

The results in Tables 1 and 2 are only consistent if the attrition and sampling processes are actually ignorable. Brownstone and Golob (1992) investigated the sampling process, and found that it was ignorable for the purposes of model estimation. Therefore, we only investigate the ignorability of the attrition process. Table 3 gives the results from fitting a binomial logit model to the attrition process between Waves 1 and 5 of the panel. Since at least some of the coefficients on the mode choice variables and their interactions are significantly different from zero, the attrition process is not ignorable. The large number of interactions between mode choice and the demographic variables show the complexity of the process. The results in Table 3 also imply that white, middle-aged homeowners with an annual household income of less than $75,000, more education, and more than three vehicles are less likely to attrite from the panel. Those respondents who receive the survey at their work sites (and presumably fill it out during their normal working hours) are also less likely to attrite.

Table 1. Multinomial logit dynamic mode choice ignoring attrition

Dependent Variable	Count	Percent
DA[1] → DA	454	58.28
RS[2] → RS	137	17.59
DA → RS	107	13.74
RS → DA	81	10.40

Independent Variable	Estimated Coefficient	t-Statistic
log[Prob(RS → RS)/Prob(DA → DA)]		
Alternative specific constant	-3.85076	-9.62953
Wave1 reserved parking for rideshare	0.57038	1.80876
Wave1 cost subsidies for rideshare	-0.12829	-0.33768
Wave1 guaranteed ride home for rideshare	1.10713	2.67947
Wave1 other incentive for rideshare	1.44839	3.50626
Wave1 HOV lane available	0.59272	2.32036
Wave1 last commute distance (miles)	3.81516e-02	5.00495
Wave1 household size (number of people)	0.19848	2.33038
Change in reserved parking for rideshare	0.11453	0.35600
Change in cost subsidies for rideshare	0.30644	1.05198
Change in guaranteed ride home for rideshare	0.64529	2.41823
Change in other incentive for rideshare	0.70468	2.09886
Change in HOV lane availability	0.75560	2.59281
Change in last commute distance	-4.52949e-03	-0.79204
Change in household size	0.18792	1.70960
Log[Prob(DA → RS)/Prob(DA → DA)]		
Alternative specific constant	-2.55746	-6.91552
Wave1 reserved parking for rideshare	0.26735	0.77539
Wave1 cost subsidies for rideshare	-0.16491	-0.38833
Wave1 guaranteed ride home for rideshare	1.14798	2.39961
Wave1 other incentive for rideshare	-0.18260	-0.40070
Wave1 HOV lane available	-0.11802	-0.39999
Wave1 last commute distance (miles)	1.96872e-02	2.22477
Wave1 household size (number of people)	0.13687	1.49854
Change in reserved parking for rideshare	0.37067	1.10452
Change in cost subsidies for rideshare	1.74854e-02	5.51e-02
Change in guaranteed ride home for rideshare	1.10699	3.80121
Change in other incentive for rideshare	1.638e-02	5.014e-02
Change in HOV lane availability	0.77807	2.61252
Change in last commute distance	-8.33293e-03	-0.76300
Change in household size	0.27921	2.18631

[1] "DA" means 'always drive alone'
[2] "RS" means 'rideshare at least once in last 2 weeks'

(Table 1, continued)

Independent Variable	Estimated Coefficient	t-Statistic
Log[Prob(RS → DA)/Prob(DA → DA)]		
Alternative specific constant	-2.57412	-6.33754
Wave1 reserved parking for rideshare	0.67226	1.83600
Wave1 cost subsidies for rideshare	-0 12810	-0.25398
Wave1 guaranteed ride home for rideshare	-0.43526	-0.73320
Wave1 other incentive for rideshare	0.28686	0.59837
Wave1 HOV lane available	0.70308	2.31456
Wave1 last commute distance (miles)	-4.03358e-03	-0.31449
Wave1 household size (number of people)	7.68584e-02	0.75632
Change in reserved parking for rideshare	0.50102	1.37607
Change in cost subsidies for rideshare	0.57746	1 64132
Change in guaranteed ride home for rideshare	-0.38832	-1.10025
Change in other incentive for rideshare	0.20816	0.60681
Change in HOV lane availability	0.80686	2 41079
Change in last commute distance	-4.94631e-02	-2 71005
Change in household size	2.33403e-02	0 17802
Auxiliary statistics	At Convergence	Initial
Log likelihood	-790.89	-1079.9
Number of observations	779	
Percent correctly predicted	60.976	

Table 2. Policy simulations ignoring attrition

Giving everyone access to a guaranteed ride home

	% Change	Std. Error	Baseline
RS[1] → RS	21.43712	8 08602	2.80998e+05
DA[2] → RS	68.47014	12.81228	2.15717e+05
RS → DA	-37.61644	6.00235	1.54883e+05
DA → DA	-16.15972	5.80547	9.26241e+05

Giving all freeway users access to high-occupancy vehicle lanes

	% Change	Std. Error	Baseline
RS → RS	12.79049	6.53451	2.80998e+05
DA → RS	15.05573	8.19401	2.15717e+05
RS → DA	12.89004	10.24839	1.54883e+05
DA → DA	-9.54215	2.53246	9.26241e+05

[1] "RS" means 'rideshare at least once in last 2 weeks'.
[2] "DA" means 'always drive alone'.

Table 3. Binomial logit attrition model

Dependent Variable	Count	Percent
In Both Waves	1107	59.97
Attrited	739	40.03
Independent Variables[1]	**Estimated Coefficient**	**t-Statistic**
Annual household income<=$75,000	-0.20233	-1.81388
High school graduate	-0 90640	-2 08486
Some college, but no degree	-1 03234	-2.48198
College degree, including graduate	-0 96502	-2 30214
Older than 24 and younger than 35	-0.40301	-1.95426
Older than 34 and younger than 45	-0.31492	-1.52823
Older than 44 and younger than 55	-0.46445	-2.08844
Older than 54 and younger than 65	-0 47694	-1.80652
Production/manufacturing	0.85561	3 86404
Sales	0.61101	2.83996
Other occupation	0.60538	2 30767
Survey received at work site	-0.25986	-2.37420
Always lived in Southern Ca.	0 29458	2.77893
Considered moving next year	0.29522	2.52530
Non-white	0 47706	3.42080
Arrived at work between 7:00 and 9:00	-0.13672	-1.17095
Years lived at present address (years)	-1.99842e-02	-2.20902
Reserved parking for rideshare	0.30232	2.54872
Household owned vehicles<=3	0.33727	2 00087
Home owner	-0.17925	-1.50927
Always rideshare in last two weeks	-0 69192	-1.71807
Always rideshare and household income<=$75,000	0.57605	1 44266
Always rideshare and moving next year	0.75718	1.80477
Always rideshare and having kids under 16	0 43372	1.20571
Sometime rideshare in last two weeks	1.12516	2 52191
Sometime rideshare and college degree	-0 73210	-2.57659
Sometime rideshare and age>24and<35	-0.74236	-2.53687
Sometime rideshare and household vehicles<=3	-0.73703	-1.74108
Sometime rideshare and having kids under 16	0.49162	1 83682
Constant	0.65782	1 35897
Auxiliary statistics	**At Convergence**	**Initial**
Log likelihood	-1164	-1279 5
Number of observations	1846	
Percent correctly predicted	64.626	

[1] All dummies except for years lived at present address.

We now assume that this attrition model is accurate and use it to implement the multiply-imputed WESMLE described in the second section of this chapter. Recall that we need to draw the multiple imputations from the attrition model to reflect all the uncertainty in the estimated attrition model. Since we only need to predict attrition probabilities, this uncertainty is all due to the uncertainty in the attrition model parameter estimates, which asymptotically follow a multivariate normal distribution. Therefore, we make a random draw from this estimated multivariate normal sampling distribution, and then use each such draw to calculate one set of attrition probabilities according to the binomial logit probability function.

Table 4 gives the resulting multiple imputations choice model estimates and standard errors calculated according to equations (2) and (3). Table 5 gives the multiply-imputed policy simulations along with their standard errors. These policy simulations are derived from our model by assuming that each respondent represents ω_i observationally equivalent people in the population, where ω_i is the inverse sampling probability (weight) for the i[th] respondent. Therefore, the choice probability P_{ij} is the proportion of these people who choose discrete alternative j. The population prediction for the number of people choosing alternative j is then given by:

$$D_j = \sum_i \omega_i P_{ij}(\theta, x_i). \tag{9}$$

Table 4. Dynamic mode-choice model using multiple imputations

Independent Variable	Estimated Coefficient	t-Statistic
log[Prob(RS → RS)/Prob(DA → DA)]		
Alternative specific constant	-3.57419	-7.96097
Wave1 reserved parking for rideshare	0.60114	1.80491
Wave1 cost subsidies for rideshare	-0.31677	-0.77887
Wave1 guaranteed ride home for rideshare	1.07399	2.31534
Wave1 other incentive for rideshare	1.49544	3 35440
Wave1 HOV lane available	0.71817	2.56257
Wave1 last commute distance (miles)	3.76844e-02	4.49140
Wave1 household size (number of people)	0.14102	1 43273
Change in reserved parking for rideshare	3.97260e-02	0.11392
Change in cost subsidies for rideshare	0.24751	0.77558
Change in guaranteed ride home for rideshare	0.54366	1 86915
Change in other incentive for rideshare	0.70037	1.96800
Change in HOV lane availability	0.82880	2.66336
Change in last commute distance	-5.76262e-03	-0.87952
Change in household size	0 10892	0 89631

(Table 4, continued)

Independent Variable	Estimated Coefficient	t-Statistic
$\text{Log}[\text{Prob}(DA \rightarrow RS)/\text{Prob}(DA \rightarrow DA)]$		
Alternative specific constant	-2.61501	-6.33719
Wave1 reserved parking for rideshare	0.22452	0.58268
Wave1 cost subsidies for rideshare	-0.39645	-0.83566
Wave1 guaranteed ride home for rideshare	1.44138	2.71527
Wave1 other incentive for rideshare	-0.34688	-0.65624
Wave1 HOV lane available	-0.16155	-0 47804
Wave1 last commute distance (miles)	1.98905e-02	1.99657
Wave1 household size (number of people)	0 13974	1.38162
Change in reserved parking for rideshare	0.35672	0.94212
Change in cost subsidies for rideshare	7.63908e-02	0.22020
Change in guaranteed ride home for rideshare	1.13001	3 48130
Change in other incentive for rideshare	3.89724e-02	0.10611
Change in HOV lane availability	0 91167	2.75916
Change in last commute distance	-9 40193e-03	-0.73885
Change in household size	0.23266	1.64069
$\text{Log}[\text{Prob}(RS \rightarrow DA)/\text{Prob}(DA \rightarrow DA)]$		
Alternative specific constant	-2.58580	-5 74700
Wave1 reserved parking for rideshare	0.81852	1.97322
Wave1 cost subsidies for rideshare	-0 15863	-0.28383
Wave1 guaranteed ride home for rideshare	-0.53282	-0.81679
Wave1 other incentive for rideshare	0.32568	0.64161
Wave1 HOV lane available	0.84614	2.57378
Wave1 last commute distance (miles)	6 35757e-03	0 48766
Wave1 household size (number of people)	7.29724e-02	0.63168
Change in reserved parking for rideshare	0.88354	2.15201
Change in cost subsidies for rideshare	0.51384	1.33210
Change in guaranteed ride home for rideshare	-0.57599	-1 45041
Change in other incentive for rideshare	-8 69461e-02	-0.22912
Change in HOV lane availability	0.88045	2.50771
Change in last commute distance	-3 94063e-02	-2.07785
Change in household size	5.61468e-02	0.37541

Our estimates are derived by replacing the unknown parameters θ by their estimates, $\hat{\theta}$, from Table 4. Policy simulations are carried out by comparing the estimates of D_j for different values of the policy variables in x_i.

There are two sources of error in our estimates of D_j: the estimation errors in the parameters θ and the sampling weights ω_i. Conditional on the sampling weights, the variance in $\hat{D} = \left(D_1(\hat{\theta}), D_2(\hat{\theta}), D_3(\hat{\theta}), D_4(\hat{\theta})\right)$ can be estimated by:

$$\hat{\Omega} = \hat{D}_\theta' \hat{V} \hat{D}_\theta , \tag{10}$$

Table 5. Policy simulations using multiple imputations

Giving everyone access to a guaranteed ride home

	% Change	Std. Error	Baseline
RS[1] $\rightarrow$ RS	18.21683	8.79526	7.60048e+05
DA[2] $\rightarrow$ RS	70.49838	14.19468	5.22866e+05
RS $\rightarrow$ DA	-44.80742	6.53906	4.22541e+05
DA $\rightarrow$ DA	-14.19958	6.56034	2.23272e+06

Giving all freeway users access to high-occupancy vehicle lanes

	% Change	Std. Error	Baseline
RS $\rightarrow$ RS	12.14453	6.56288	7.60048e+05
DA $\rightarrow$ RS	17.96968	9.61923	5.22866e+05
RS $\rightarrow$ DA	12.53924	10.53031	4.22541e+05
DA $\rightarrow$ DA	-10.71509	2.71159	2.23272e+06

[1] "RS" means 'rideshare at least once in last 2 weeks'.
[2] "DA" means 'always drive alone'.

where $\hat{D}_\theta$ is the matrix of first derivatives of D with respect to θ evaluated at $\hat{\theta}$ and $\hat{V}$ is a consistent estimator of the covariance of $\hat{\theta}$ (for more details, see Chow, 1983, pp. 182-183). The covariance due to the estimation error in the sampling weights could be handled using the same techniques, but here it is more convenient to use multiple imputations since these can be computed as part of the multiple imputation choice model estimators. For each weight vector, which is simply the inverse of the sampling probability multiplied by 1 minus the imputed attrition probability, we compute $\hat{D}$ and its covariance estimator $\hat{\Omega}$. The final estimate of D is given by:

$$\overline{D} = \sum_{i=1}^{m} \hat{D}_i / m, \tag{11}$$

where m is the number of imputed weight vectors and $\hat{D}_i$ is the estimator for the i[th] weight vector. If $\overline{\Omega}$ is the corresponding average of the covariance estimates $\hat{\Omega}_i$ and

$$S = \sum_{i=1}^{m} \left(\hat{\Delta}_i - \overline{\Delta}\right)\left(\hat{\Delta}_i - \overline{\Delta}\right)' \Big/ (m-1) \tag{12}$$

is an estimate of the covariance among the m estimates for each weight vector, then

$$\Psi = \overline{\Omega} + (1+m^{-1})S \tag{13}$$

is the estimate of the total covariance of $\overline{D}$.

Comparison between these multiple imputations results (Tables 4 and 5) and the results ignoring attrition (Tables 1 and 2) suggest that attrition is not a serious problem for these models and data. This suggestion is reinforced by calculating the Hausman test statistic given in equation (8), which is not significant. We can also examine the effects of ignoring the estimation uncertainty in the attrition weights by simply computing the WESMLE and policy forecasts without drawing any multiple imputations. The resulting policy forecasts are given in Table 6. The results are very similar to those in Table 5, except that the standard errors have dropped by approximately 6%.

Table 6. Policy simulations using WESMLE with single weight

Giving everyone access to a guaranteed ride home

	% Change	Std. Error	Baseline
RS[1] → RS	20.62864	8.34177	8.05749e+05
DA[2] → RS	69.01154	13.33462	6.18642e+05
RS → DA	-44.74986	5.79344	4.21397e+05
DA → DA	-15.29072	6.16246	2.64589e+06

Giving all freeway users access to high-occupancy vehicle lanes

	% Change	Std. Error	Baseline
RS → RS	10.45871	6.64034	8.05749e+05
DA → RS	19.46045	9.22377	6.18642e+05
RS → DA	13.48626	11.39300	4.21397e+05
DA → DA	-9.88298	2.67457	2.64589e+06

[1] "RS" means 'rideshare at least once in last 2 weeks'.
[2] "DA" means 'always drive alone'.

It is important to stress that our finding that attrition did not matter in this example crucially depends on the particular model and policy simulations we examined. Given the large number of insignificant coefficients in our mode choice model, it is likely that a smaller model would be more sensitive to the presence of non-ignorable attrition.

CONCLUSIONS

Combining the WESMLE and Rubin's Multiple Imputations methodology provides a simple but general procedure for consistent estimation and forecasting when there is non-ignorable panel attrition. Although full maximum likelihood estimation is generally more efficient, the methodology explored in this chapter is computationally much simpler. The methods used here can also be extended to simultaneously cope with choice-based panel sampling and non-ignorable missing data.

Although the methods proposed in this chapter are a useful addition to the tools available to correct for the effects of panel attrition, it should be stressed that no *ex-post* econometric technique can substitute for minimizing attrition in the first place. If wave-to-wave attrition is not minimized, then the panel rapidly becomes useless for long-run dynamic analysis.

Acknowledgements. This research was sponsored by the University of California Transportation Center, with funding from the U.S Department of Transportation and the California State Department of Transportation (Caltrans). Additional computer funding was provided by the University of California Research Unit in Mathematical Behavioral Sciences. The authors wish to thank Seyoung Kim, Tom Golob, and Carole Uhlaner for their assistance None of these people and institutions are responsible for any errors or the views expressed herein.

REFERENCES

BROWNSTONE, D. (1991) Multiple Imputations for Linear Regression Models Technical Report MBS 91-37, Research Unit in Mathematical Behavioral Sciences, University of California, Irvine, California.

BROWNSTONE, D. and GOLOB, T.F. (1992) The effectiveness of ridesharing incentives. Discrete-choice models of commuting in Southern California. *Regional Science and Urban Economics*, **22**, 5-24.

CHOW, G.C. (1983) *Econometrics*. McGraw-Hill, New York, New York.

HAUSMAN, J.A (1978) Specification tests in econometrics *Econometrica*, **46**, 1251-1272.

IMBENS, G. (1992) An efficient method of moments estimator for discrete choice models with choice-based sampling. *Econometrica*, **60**, 1187-1214.

MANSKI, C.F. and LERMAN. S (1977) The estimation of choice probabilities from choice-based samples *Econometrica*, **45**, 1977-1988

RUBIN, D.B. (1986) Statistical matching using file concatenation with adjusted weights and multiple imputations *Journal of Business and Economic Statistics*, 4, 87-94.

RUBIN, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys.* John Wiley and Sons, New York, New York.

UHLANER, C.J. and KIM. S. (1993) Designing and Implementing a Panel Study of Commuter Behavior: Lessons for Future Research. Working Paper 93-2, Institute of Transportation Studies, University of California, Irvine, California.