

Using Census Match Data to Evaluate Models of Survey Nonresponse

John Dixon

Bureau of Labor Statistics, Room 1950

2 Massachusetts Ave., NE

Washington, DC 20212-0001

Abstract

Several methods are used to model nonresponse in surveys: panel information, item nonresponse, last 5 percent, and area characteristics. This study compares several methods for modeling nonresponse in the Current Population Survey (CPS) using match data from the Decennial Census as a criterion. Recommendations based on the comparisons and limitations resulting from imperfect matching will be discussed.

Introduction

Studying nonresponse to household surveys is difficult because of a lack of information about nonrespondents. For panel surveys information can be borrowed from other panels. Survey households may also be matched with other sources, usually administrative data (registers) or censuses.

For a single administration of a survey, information can be modeled based on characteristics of those interviewed early and late in the interview process. The lateness of response (for example, the last 5 percent) can be used, since if the effort to collect the data had ended earlier, they would have been nonrespondents (Bates and Crieghton, 2000; Chiu, Riddick, and Hardy, 2001). The nonresponse to items can also be used as a surrogate for nonresponse (Dixon, 2002; Loosveldt, Pickery, and Billiet, 2002).

Data Sources

A key source of data in this study resulted from matching Census long-form data to Current Population Survey (CPS) cases. Therefore, information obtained from the Census could be used to describe nonresponse cases in the CPS. Data from the CPS was selected for February through May, 2000 to cover the response time frame for the 2000 Census long form¹ (there were 212,914 enumerated persons with interviews or refusals in this time period, noncontact was not analyzed in this paper).

Details about the CPS can be found in Technical Paper 63. The CPS is the primary source of information on the labor force characteristics of the

U.S. population. Similar estimates can be generated from the Census. However, many methodological differences may contribute to differences between the CPS and Census;

- Reference period (CPS: asks about the week containing the 2nd Tuesday of the month, Census: asks about last week, but over a several month period).
- The CPS consists of 8 separate interviews spread out over a 16 month period using a complex sample rotation design. The Census long form was done once.
- Collection mode (CPS: personal visit on 1st and 5th interviews, other interviews done predominately by telephone; Census: self-administered done mostly by mail; 72 percent, drop off form, 18.8 percent; and the rest mostly by personal interview).
- Interviewers (CPS interviewers are much more experienced).
- Instrument (Census paper form, CPS computer-assisted interview).
- Questions (CPS asks about active search for work, self employment, owning a business, multiple jobs, retirees); Census is more general and asks fewer questions about labor-force status).
- Collection period: CPS for 10 days, Census for over a month (as long as 7 months).

Methods

The matching process failed to match about 10 percent of the CPS household members using the Census long form. The match was less successful for those who refused the CPS interview (no match for 25 percent of refusers).

The variables used to model nonresponse were adapted from Groves and Couper (1998), and Dixon (2001). A model with 17 predictors and 72 interactions was examined and reduced to a model with 8 predictors and 5 interactions. The adjusted pseudo r-square went from .23 to .20. While the goodness of fit statistics indicated there were other terms which should be added to the model, this model represented a trade-off between complexity and fit.

¹ Census Day was April 1, 2000.

Draft- Opinions expressed are of the author and not BLS

Unweighted data were used since the frame of analysis was the interviewed persons, and no inference to a national sample was intended. Similarly, no adjustment was made for sample design for the same reasons. The variances are for the chosen sample, not for national estimates.

Two methods for modeling nonresponse based on the current survey respondents were used. The last 5 percent of the respondents was used to represent potential nonrespondents (some noncontact, some refusal), and item refusal was used as a surrogate for unit refusal.

Two sources of information on nonrespondents are also used in the models: panel information from nonrespondents who had responded in previous months, and information from the Census long form matched to the CPS.

The advantage of the panel data is the completeness of the match. The disadvantage is it has no useful information about households which never respond. The advantage of the Census match is that it can provide information about some of those who never respond to the survey. However, a disadvantage is that the success of matching is related to nonresponse. The Census can be used to examine the deficiency of the survey model, and the survey can be used to examine the impact of matching for the Census.

A multinomial logistic model was used to test the hypothesis that the relationship between household and personal characteristics used to predict nonresponse are consistent for the panel data and the Census data (proportional odds). The dependent variable was the source of the data (Census, Panel, or interview). Logistic models were also used to examine the effect of matching using only the CPS panel data. These models used the source of the data as the dependent variable and included "match status" as one of the predictors. The effect of match status was also examined with Breslow-Day homogeneity of odds-ratio tests.

Logistic models were used on the matched data to evaluate the difference between the nonrespondents who had information from panel responses and those who never responded to the CPS but had information from the Census long form. These models only used the refusers, so the estimated coefficients would describe the difference between refusers who only had data from the Census match and those who had data from the other CPS panels. The dependent variable was a binary indicator for whether the information on refusers came from the Census or the CPS panels.

A logistic model was used to contrast the employment status for those who responded to the survey to those who refused the survey based on information from other panels or the Census.

Results

Prediction of refusal

The logistic models in Table 1 (for complete tables, contact the author, tables are abbreviated due to space limits) show 8 models with one predictor, and one model with eight predictors. The odds ratio can be interpreted as the probability of refusal for one group compared to another (e.g., married compared to non-married respondents).

Households with a child present (Kid) were more likely to refuse (OR: 5.770), as were older households (Age; OR:1.002) and households with members in school (School; OR:1.225).

Households that were less likely to refuse were Hispanic households (Hispanic; OR:0.392), households with a married respondent (Married; OR:.894), larger households (Number; OR:0.761), households with relatives present (Relatives; OR:0.599), and households with a White respondent (White; OR:0.581).

The moderating effect of the other variables can be seen in the difference in the estimates between the single predictor models and the multivariate model. While households with a married respondent were less likely to refuse (Married; OR:.894), when adjusted for the other variables they were more likely to refuse (OR:1.426). This was the only effect to reverse direction. The variables which increased the likelihood of refusal after adjusting for other variables were "child present" (Kid; OR:11.061), school attendance (School; OR:1.572), and White respondent (White; OR:0.715). Larger households had a reduced likelihood of refusal after adjusting for other variables (Number; OR: 0.705).

Interaction effects

All of the variables were involved in interactions except "Married." Table 2 shows the 5 models involving pairs of variables and their interaction, and one model with 8 variables and 5 interactions. "Relatives present" increased the likelihood of refusal for older households and White households. Larger households (Number) increased the likelihood of refusal for households with school attendance (School) or young children present (Kid), but reduced the likelihood for Hispanic households (Hispanic).

Match status

Match status was evaluated in two different sets of logistic models. The first set of models looked at the moderating effect of match status in a model of refusals predicted by a number of household and personal characteristics (the same 17 variables used in other models of nonresponse) in Table 3.

A logistic model predicting refusal was compared to a model that included match status as a variable. Where the coefficients differed by more than two standard errors there may be a moderating effect. These variables would share a relationship with refusal which is related to match status. "School attendance," "small children in the household," and the "size of the household" were moderated by match status by reducing the effect.

A logistic model predicting refusal which included match status as an interaction term for each of the predictors was used to see if the effect of match status was differential for any of the predictors (Table 4). "School attendance" and "small children present" which had moderating effects, also had interactions. "Size of the household" didn't interact significantly. All but one of the interactions indicated higher refusals for those matched with the Census (Age, "small children present", "size of the household", "home ownership", and "school attendance"). "Relatives present" resulted in lower refusals for matched households.

The Breslow-Day homogeneity of odds-ratios tests (Table 11) showed the last 5 percent was least impacted by match status. The tables which formed the basis for the test were "employment status" by "refusal prediction" stratified by "match status".

Census and CPS Panel data for refusers

A logistic model using only refusers who matched the Census was used to compare the difference between those who had CPS panel data and those who only had Census data (Table 5). An indicator for Census/CPS was used as the dependent variable. Separate models for seventeen variables which had been found related to refusals were used as independent variables. An additional model was used with all the variables as simultaneous predictors to assess their unique relationship. Hispanic members were more likely to be in the Census only (4.2771), but only when adjusted for the other variables, as were homeowners (1.0288). Refusers from multiple unit structures (MUL) and larger households (NUM) were more likely to be in the Census only, and never respond to the survey. "Relatives present" were less likely to be in the

Census only, as were Male, Black and White refusers.

Last 5 percent

A logistic model was used to examine how well the last 5 percent of the respondents could be used as a substitute for nonresponse (Table 6). The same variables used to predict refusal were put in a model where an indicator for the last 5 percent was used instead of the refusal variable. In addition, a measure of the number of attempted contacts was used to indicate the amount of effort to get an interview. The model fit relatively poorly (Max-rescaled R-square of .04 compared to .20 for the model of refusal). The strongest variable was the number of attempted contacts.

Item refusal

A logistic model was used to examine how well item refusals could be used as a substitute for unit refusal (Table 7). The same model as used in the last 5 percent was used. The dependent variable was an indicator for whether there were any item refusals or not. The model fit poorly (Max-rescaled R-square of .03). The strongest variables were the number of attempted contacts, household size, and age of the respondent.

Panel refusal

A logistic model was used to examine how well refusal to other panels of the CPS could be used to substitute for refusals in the CPS panel which was matched to the Census (Table 8). The model fit poorly (Max-rescaled R-Square 0.0440). The strongest variables were household size, number of attempted contacts, and relatives present.

Labor force - Unemployment

Since the category "employed" is less sensitive to measurement error relative to the categories "unemployed" and "not in the labor force" (Biemer and Bushery, 1999) it will be used as an indicator of labor force status. This will more clearly show effects of nonresponse, since "unemployed" may add more measurement error (Palumbo & Siegel, 2004).

Table 9a shows the agreement between the Census and the CPS. The Kappa coefficient (0.8148) indicates moderate agreement. This is as good a match between surveys with different time frames, questions, and collection procedures as we are likely to achieve. Even with the same respondents, but different times, the CPS achieved a Kappa of 0.8706 (Table 9c). The agreement between the Census and the Panel estimates was 0.7737 (Table 9b).

Those who refused the CPS had higher employment as measured by the Census (Table 9e);

68.70 versus 59.33. Part of the difference could be accounted for by the difference in matching, CPS measures for the matched were higher than for the nonmatched; 63.68 versus 60.22 (Table 9d). Tables 9f-9n show the various combinations of source of refusal and source of employment. Refusals had a higher percentage of employed relative to completed interviews for all refusal sources except item refusal.

The three refusal propensity measures were compared in predicting employment using three single predictor logistic models and a three predictor logistic model (Table 10). Panel refusal and the last 5 percent were most similar in their separate coefficients. The multivariate model shows the last 5 percent is redundant with the other propensity measures, since it becomes non-significant.

Discussion

The seminal work of Groves and Couper (1998) matched CPS households with the 1990 Census. The current work examined a person-level match. There were several differences in findings, which isn't unexpected when going to a different level of analysis.

Groves and Couper found that single person households resulted in less cooperation, while the current study similarly found that larger households were more likely to cooperate, as were married householders. Groves and Couper found that younger and older households were more likely to cooperate, while this study found a slight trend toward not cooperating for older household members. The age effect interacted with whether there were other relatives present; younger members with other relatives in the household were more likely to cooperate. Both studies found that Hispanics were more likely to cooperate. This study found that Hispanics in larger households were even more likely to cooperate.

Employment based on the Census was 59.33 for those interviewed in the CPS compared to 68.70 for those who refused (Table 9e). The combined was 59.46, showing a very slight bias (underestimating by .13 percent). The actual bias would need to account for the sample design with weights and complex variance estimation. The Census employment measure and the CPS refusal are the standards used to compare the other measures. In a study which

matched the United Kingdom census with 6 surveys, Beerten and Freeth (2004) found that where the household reference person was unemployed he or she was less likely to respond to the survey, which was similar to the current study. While the U.S. Census measure of employment was lower than the CPS and Panel measures, it didn't matter in terms of assessing the different methods of estimating bias. The last 5 percent propensity worked much better than the item refusal propensity. If this finding can be replicated with other surveys, it would be encouraging as a method of assessing potential bias. The poor showing of the item propensity could be due to either a poor propensity estimate, or a confounding of the mechanisms which produce item nonresponse. Dixon (2002) showed that item nonresponse was related to subsequent unit nonresponse, and to lower estimates of unemployment, which this study also found. This suggests that item nonresponse may be sufficiently different from unit nonresponse that it can't be used to detect bias (at least in the CPS). Item refusal may be related to employment status and propensity to respond. Both item nonresponse and the last 5 percent are sensitive to survey procedures, since the interviewer may not press for responses to refused items, or may encourage item refusal as a trade-off for unit response. The last 5 percent may be sensitive to the effort interviewers put into contacting reluctant and hard-to-contact households earlier in the interview period. The propensity measure based on other panel nonresponse worked well, comparable to the last 5 percent. For surveys with panel data this may work to check on the last 5 percent method.

Limitations

Additional work needs to be invested in studying noncontact. The relationship between personal characteristics and household and interview characteristics could be modeled with multilevel models (Dixon and Tucker, 2000; Fraboni, Rosina, Orsini, and Baldazzi, 2002). Additional methods of estimating bias (e.g., benchmarking) would be useful to evaluate.

The item nonresponse models need to be further studied to evaluate why they didn't perform as expected. Contrasting who didn't respond to items with those who refused the entire survey could be enlightening.

References:

- Bates, N. and Crieghton, K., “The Last Five Percent: What Can We Learn from Difficult Interviews?”, In Proceedings of the Annual Meetings of the American Statistical Association, August 13-17, 2000. <http://www.fcsm.gov/committees/ihsng/asa2000proceedings.pdf>
- Beerten, R., and Freeth, S., “Exploring Survey Nonresponse in the UK: The Census-Survey Nonresponse Link Study”, Paper presented at the 2004 American Association for Public Opinion Research, Phoenix, Arizona.
- Biemer, P., and Bushery, J., "Estimating the Error in the Labor Force Data Using Markov Latent Class Analysis", Proceedings of the FCSM Conference, 1999. <http://www.fcsm.gov/99papers/biemer.pdf>
- Chiu, P., Riddick, H. and Hardy, A., “A Comparison of Characteristics Between Late/Difficult and Non-late/Difficult Interviews in the National Health Interview Survey”, In Proceedings of the Annual Meetings of the American Statistical Association, August 5-9, 2001. <http://www.fcsm.gov/committees/ihsng/chiupasa2000.pdf>
- Dixon, J. and Tucker, C., Modeling Household and Interviewer Nonresponse Rates from Household and Regional Characteristics. A paper presented at the International Workshop on Household Survey Nonresponse, Budapest, September, 2000. <http://www.fcsm.gov/committees/ihsng/buda5.pdf>
- Dixon, J., “Relationship Between Household Nonresponse, Demographics and the Unemployment Rate in the Current Population Survey”, In Proceedings of the Annual Meetings of the American Statistical Association, August 5-9, 2001. <http://www.fcsm.gov/committees/ihsng/jdixon2001.pdf>
- Dixon, J., “The Effect of Item and Unit Nonresponse on Estimates of Labor Force Participation”, In Proceeding of the Annual Meetings of the ASA, August 11-15, 2002. <http://www.fcsm.gov/committees/ihsng/asa02a.pdf>
- Fraboni, R., Rosina, A., Orsini, S., and Baldazzi, B., "A Multilevel Approach to the Analysis of Refusal Risk in the Italian Multipurpose Survey", paper presented at the International Conference on Improving Surveys, August 2002, Copenhagen, Denmark.
- Groves, R. and Couper, M., Nonresponse in Household Interview Surveys, 1998, New York: Wiley.
- Loosveldt, G., Pickery, J., and Billiet, J., "Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey", Journal of Official Statistics, v.18.4, 2002.
- Palumbo, T. and Siegel, P., “Accuracy of Data for Employment Status as Measured by the CPS-Census 2000 Match”, Final Report for the Census 2000 Evaluation B.7, May 4, 2004.
- U.S. Department of Labor, Bureau of Labor Statistics. *Current Population Survey Technical Paper 63 - Design and Methodology*. Washington, D.C. 2000.

Appendix A

Table 8 - Panel refusal

Parameter	DF	Estimate	StdError	Chi-Square	Pr > ChiSq
Max-rescaled R-Square 0.0440					
Intercept	1	-1.2581	0.0818	236.4939	<.0001
Age	1	-0.0111	0.00130	72.7446	<.0001
Hispanic (hsp)	1	-0.1147	0.0940	1.4897	0.2223
Young Child(kid)	1	0.3535	0.0698	25.6124	<.0001
Married	1	0.1799	0.0332	29.4150	<.0001
Size of HH(num)	1	-0.2236	0.0121	339.4368	<.0001
Relatives(rel)	1	-1.0681	0.0776	189.6323	<.0001
School(sch)	1	-0.1157	0.0923	1.5728	0.2098
White(wht)	1	-0.4365	0.0565	59.7288	<.0001
Age*Rel	1	0.0137	0.00145	89.0694	<.0001
Hsp*Num	1	-0.0381	0.0230	2.7329	0.0983
Kid*Num	1	0.1595	0.0175	83.1802	<.0001
Num*Sch	1	0.1306	0.0237	30.4414	<.0001
Rel*Wht	1	0.2589	0.0649	15.9166	<.0001
# contacts(cnt)	1	0.1562	0.0109	204.1607	<.0001

Table 9a-Employment indicators from Census and CPS

Percent	CPS Employed	CPS Other
Census Employed	57.16	3.08
Census Other	5.69	34.07

Kappa 0.8148

Table 9b Employment from Census and Panel

Percent	Panel Employment	Panel Other
Census Employment	57.26	2.97
Census Other	7.66	32.11

Kappa 0.7737

Table 9c CPS Employment and Panel Employment

Percent	Panel Employment	Panel Other
CPS Employment	61.52	1.93
CPS Other	4.00	32.55

Kappa 0.8706

Table 9d
Match Status and CPS Employment

Row Pct	CPS Employed	CPS Other
Not matched	60.22	39.78
Matched	63.68	36.32

Statistic DF Value Prob
Chi-Square 1 52.0551 <.0001

Table 9e
Refusal and Census Measure of Employment

Column Percent	Interview	Refused	Overall
Census Employed	59.33	68.70	59.46
Census Other	40.67	31.30	40.54

Statistic DF Value Prob
Mantel-Haenszel Chi-Square 1 12.9690 0.0003

Table 9f
Refusal and CPS Panel Measure of Employment

	Interview	Refused	Overall
Panel Employed	65.58	70.88	65.63
Panel Other	34.42	29.12	34.37

Statistic DF Value Prob
Mantel-Haenszel Chi-Square 1 19.4393 0.0001

Table 9g
Last 5 percent Refusal and CPS Measure of Employment

	Interview	Refused	Overall
CPS Employed	63.17	71.56	63.41
CPS Other	36.83	28.44	36.59

Statistic DF Value Prob
Mantel-Haenszel Chi-Square 1 175.9550 0.0001

Table 9h
Item Refusal and CPS Measure of Employment

	Interview	Refused	Overall
CPS Employed	67.42	42.36	63.41
CPS Other	32.58	57.64	36.59

Statistic DF Value Prob
Mantel-Haenszel Chi-Square 1 6928.2819 0.0001

Table 9i
Last 5 percent Refusal and Census Measure of Employment

	Interview	Refused	Overall
Census Employed	59.31	65.42	59.47
Census Other	40.69	34.58	40.53

Statistic DF Value Prob
Mantel-Haenszel Chi-Square 1 6.0314 0.0141

Table 9j
Item Refusal and Census Measure of Employment

	Interview	Refused	Overall		
Census Employed	63.33	39.67	59.47		
Census Other	36.67	60.33	40.53		
Statistic		DF	Value	Prob	
Mantel-Haenszel Chi-Square		1	503.4172	0.0001	

Table 9k
Item Refusal and Panel Measure of Employment

	Interview	Refused	Overall		
Panel Employed	70.06	43.61	65.63		
Panel Other	29.94	56.39	34.37		
Statistic		DF	Value	Prob	
Mantel-Haenszel Chi-Square		1	5235.0142	0.0001	

Table 9l
Panel Refusal and Census Measure of Employment

	Interview	Refused	Overall		
Census Employed	58.98	71.89	59.47		
Census Other	41.02	28.11	40.53		
Statistic		DF	Value	Prob	
Mantel-Haenszel Chi-Square		1	29.7282	0.0001	

Table 9m
Panel Refusal and Panel Measure of Employment

	Interview	Refused	Overall		
Panel Employed	65.19	75.95	65.63		
Panel Other	34.81	24.05	34.37		
Statistic		DF	Value	Prob	
Mantel-Haenszel Chi-Square		1	203.7530	0.0001	

Table 9n
Panel Refusal and CPS Measure of Employment

	Interview	Refused	Overall		
CPS Employed	62.92	77.24	63.41		
CPS Other	37.08	22.76	36.59		
Statistic		DF	Value	Prob	
Mantel-Haenszel Chi-Square		1	572.7623	0.0001	

Table 10 - Models Predicting Employment from Refusal Propensity Scores

Parameter	DF	Multivariate model				3 Single Predictor Models				
		Estimate	StdErr	Chi-Sq	Pr>ChiSq	Estimate	StdError	Chi-Sq	Pr>ChiSq	
Intercept	1	0.7460	0.0153	2373.2292	<.0001	0.2612	0.0824	10.058	0.0015	
Last 5 percent	1	-0.1092	0.0888	1.5106	0.2191	-1.9904	0.0374	2837.8	<.0001	
Item refusal	1	-1.9877	0.0374	2821.4579	<.0001	0.5759	0.0733	61.684	<.0001	
Panel refusal	1	0.5227	0.0806	42.0896	<.0001					
Max-rescaled R-Square		0.1713								

Table 11 - Breslow-Day tests of Homogeneity of Odds-Ratios stratified by Match Status

	5%		Item		CPS	
	Chi-sq	Prob.	Chi-sq	Prob.	Chi-sq	Prob.
CPS employment	0.1241	.7254	171.7514	.0001	.	.
CPS panel employment	4.4477	.0350	140.8503	.0001	91.9046	.0001