

CHARACTERISTICS OF NONRESPONDENTS TO THE CURRENT POPULATION SURVEY (CPS) AND CONSUMER EXPENDITURE INTERVIEW SURVEY (CEIS)¹

Brian A. Kojetin

Bureau of Labor Statistics, 2 Massachusetts Ave. N.E., Room 4915, Washington, D.C. 20212

KEY WORDS: Nonresponse, Current Population Survey, Consumer Expenditure Survey

I. Introduction

Although survey nonresponse has been a topic of research for many years, only recently has systematic attention been given to developing a theoretical framework for understanding nonresponse. Groves, Cialdini, and Couper (1992) have described a theoretical model of survey participation that integrates socio-demographic and survey design factors with psychological characteristics of the respondent and interviewer along with aspects of the interaction between the respondent and the interviewer. Specifically, they identified categories of factors reflecting the 1) social context, such as political, social, economic, and neighborhood characteristics; 2) respondent characteristics including socio-demographic characteristics, experience with surveys, knowledge of the topic, etc.; 3) survey design attributes such as burden, topic, mode of administration, and respondent selection; 4) characteristics of the interviewer including socio-demographic characteristics, interviewing experience, and expectations; and 5) the interaction that occurs between the respondent and the interviewer which results in a decision to cooperate or refuse to participate in the survey. This model for the study of nonresponse elucidates a variety of factors on several levels of analysis that can influence a person's cooperation with a survey request and also recognizes that people may have good reasons for not participating in a survey.

In the present investigation the characteristics of respondents and nonrespondents to the Current Population Survey (CPS) and Consumer Expenditure Interview Survey (CEIS) are compared utilizing a sample that was matched to 1990 decennial Census records. The CPS and the CEIS are economic surveys sponsored by the Bureau of Labor Statistics (BLS) with data collection by the Census Bureau. These surveys provide vital economic data about employment and spending patterns of U. S. households. These two surveys represent a subset of the seven surveys that

were matched to the 1990 decennial census records (See Groves & Couper, 1992a). Researchers examining nonresponse with these census-survey match data sets have focused chiefly on cooperation with the survey request (Groves & Couper, 1993a, 1993b). However, some recent attention has also focused on contact rates (Couper & Groves, 1993c), but these analyses included only household-level variables.

The Census-survey match data provides the opportunity to make comparisons between the respondents and nonrespondents on these two BLS surveys on a variety of Census household and block level variables. From a practical standpoint, it would be useful to know the absolute and relative contribution of factors reflecting these different levels of analysis to the prediction of survey nonresponse because some factors are more easily observed, measured or obtained than others are. Therefore, it may be more cost efficient to utilize some of these more accessible factors in making post-survey nonresponse adjustments. For example, some characteristics of the neighborhood may be obtainable from census records or other sources, and some characteristics of the housing unit may be easily observed externally on a visit and could subsequently be utilized for statistical adjustment for nonresponse. Because there may be important differences as well as similarities between refusals and noncontacts, these two groups were also compared separately to survey respondents. Finally, preliminary multivariate models that utilized all available information were also examined to see how well survey nonresponse could be predicted from all of the available short form Census variables.

II. Method

Design

To study the characteristics of nonrespondents to these two BLS surveys, respondent and nonrespondent households were matched to their decennial census records. Through comparisons of respondents and nonrespondents on the variables from the decennial census, some aspects of the social context and some respondent characteristics are examined. Because these census variables also covered several different levels of analysis, an examination of the influence of

¹The views expressed in the paper of those of the author and do not necessarily represent those of the Bureau of Labor Statistics

factors reflecting each level of analysis on predicting nonresponse was made.

Census Matching

A sample of respondent cases and all nonrespondent cases from both surveys were selected from a three-month period around census day (April 1, 1990). Address level matching was made from survey sample control information and Census Bureau resources. For the CPS there was a total eligible sample of 2684 with 1445 completed interviews and 1239 nonresponse cases. For the CEIS there was a total eligible sample of 1802 with 844 interviewed cases and 958 nonresponding cases. The overall household level match rate was 93.4% for the CPS and 96.4% for the CEIS. For further details of the survey-census match operation, see Couper and Groves, (1992a).

Analysis Plan

For this initial set of analyses, the main goal has been to obtain descriptive information and to follow a data driven or exploratory strategy for model building. In other words, these analyses were not conducted to test specific theoretical hypotheses about the nature of nonrespondents to BLS surveys, but were done to assess how well survey nonrespondents could be predicted using all of the available information from the decennial Census short form. Future work will address creating more parsimonious models and testing specific theoretical predictions.

In creating models of nonresponse, it may be important to distinguish two different types of census variables that exist at the household level (and are also aggregated at the block level). Specifically, we can distinguish between characteristics of the *housing unit* such as type of structure, number of rooms, value or rent, etc., and characteristics of the *people* occupying the housing unit, such as their ages, marital status, ethnicity, etc.. Similarly, at the block level, we can distinguish between the housing characteristics of the block and the composition of people living there. It may be important to distinguish between these variables for some practical as well as theoretical considerations. We may expect that the housing unit characteristics will be more stable over time than the particular occupants will be. Furthermore, most aspects of the housing unit are more easily observable or available than are characteristics of all of the occupants.

Census variables were identified that reflected different levels of analysis (general area, neighborhood, and household) as well as variables that primarily represented the characteristics of the housing units or structures compared to the characteristics of the occupants. Specifically, the available short-form

variables were classified as referring to 1) the general area where the potential respondent lived, such as the designation of Metropolitan Statistical area; 2) the kind of structures in the immediate area (Census block), such as the percentage of single family homes; 3) the kinds of occupants in the immediate area, such as the percentage of married couples with children; 4) the structure of the housing unit, such as how many separate living units there are; and 5) the occupants of the housing unit, such as the number of people in the household.

First, univariate analyses were conducted to examine whether each Census short-form variable individually significantly discriminated respondents from nonrespondents. These analyses were either logistic regressions for continuous Census variables (most of the block level variables) or crosstabulations for the categorical Census variables. Second, each category of variables was entered as a group in a logistic regression. This analysis provided information on the how well each category group was able to predict survey nonresponse. Third, all of the variables were forced into a logistic regression analysis to maximally predict survey response given all of the available Census variables. All of the analyses were conducted using SUDAAN which calculates standard error estimates and statistical tests using Taylor Series approximation, reflecting the stratification and clustering of the survey designs. Data are weighted for all analyses to reflect different probabilities of selection in the original studies and the survey-census match sample.

III. Results

III.A CPS

Area and Block Characteristics

As can be seen in the first four columns of Table 1, nonrespondents were more likely than respondents to be located in more dense urban areas with higher percentage of the block devoted to multi-unit dwellings and lower percentage of housing units on the block being owner occupied, or single family units. Nonrespondents are also more likely to be on blocks that have a higher percentage of minorities, single persons, households consisting of unrelated individuals, and group quarters and a smaller percentage of married couples with children. In addition, nonrespondents are more likely to live on blocks with higher average monthly rental cost and house values.

Household Level Characteristics

Nonrespondents on the CPS were more likely than respondents to come from single person households who rent housing units with fewer rooms, in multi-unit structures. Several variables appear to represent

partially the effects of single person households on response. Specifically, households with no adult men, households with no adult women, the composition of the household, the size of the household, the number of adults in the household, and marital status all show strong differences between respondents and nonrespondents often (though not necessarily entirely) because they reflect the greater propensity of single person households to be survey nonrespondents. Some of the housing unit characteristics may also go along with being a single person living alone because these people are probably also more likely to rent, live in a multi-unit structure, and have fewer rooms than two or more people living together.

III.B CEIS

Area and Block Characteristics

There were few area or block characteristics that reliably distinguished Type A nonrespondents from respondents to the CEIS. As can be seen in the last three columns of Table 1, nonrespondents were more likely to be located in more dense urban areas, and were more likely than respondents to be on blocks that have a higher percentage of single persons, with higher average monthly rental cost. They were also less likely to live on blocks with single parent families, or a larger percentage of children less than 5 than or household members less than 20.

Household Level Characteristics

In a similar manner to the CPS, Type A nonrespondents on the CEIS were more likely than respondents to come from single person households. Nonrespondents were correspondingly less likely to be married with children or to have children less than 5 years old in the household, but there were few other significant differences between respondents and nonrespondents.

III.C Multivariate Analyses

A set of multivariate logistic regression analyses were conducted by entering each group of variables separately to examine their ability to predict survey nonrespondents. In addition, all of the variables were forced into a logistic regression analysis to maximally distinguish respondents from refusals or noncontacts given all of the available Census variables. The results of these analyses can be seen in Table 2. More differences between respondents and nonrespondents on the Census variables were observed for the CPS at the univariate level than on the CEIS and, accordingly, we were able to predict nonresponse on the CPS better using this information. For both the CPS and CEIS the category of variables that best predicted survey nonresponse involved characteristics of persons living in the household, followed by either the housing unit

structure (CPS) or structures in the Census Block (CEIS).

III.D Comparing Respondents to Refusals and Noncontacts

Additional analyses were conducted to determine whether refusals and noncontacts differed from survey respondents on any of the Census short form variables that were not reflected in the previous analyses comparing respondents and nonrespondents. In a similar manner to the previous set of analyses, refusals and noncontacts were contrasted with respondents separately on each Census variable². Then, each group of variables was entered into a logistic regression model to optimally predict survey response or type of nonresponse.

CPS

Refusals were more likely than respondents to come from blocks with a lesser percentage of persons over 65 years of age and a greater percentage of children less than 5 years old. Refusals were also more likely to be Black and non-Hispanic. Noncontacts were more likely to occur in blocks with a higher percentage of boarded-up units and with households where all members are less than 35 years of age.

As can be seen in Table 2, separating refusals and noncontacts lead to better predictions than treating them as a single group. Furthermore, it appears that the noncontacts differed the most from survey respondents on these Census variables.

CEIS

Similar analyses were conducted to compare respondents with refusals and noncontacts. Refusals were less likely than respondents to be Hispanic and were more likely to come from households with fewer adults.

There were very few noncontacts in the dataset for the CEIS, and, therefore, the earlier description of Type A nonrespondents largely characterized the differences between the refusals and respondents. Noncontacts were more prevalent in areas with higher crime rates, and higher percentages of 2 to 9 unit structures, non-family households, minorities and persons over 65 years of age. Noncontacts were also more likely to occur on blocks with lower percentages of single unit housing structures, owner-occupied housing units, and households with married couples with children. Noncontacts were also less likely to live in single unit structures and were more likely to be

² Tables of descriptive means and percentages for refusals and noncontacts are not included because of space limitations. Complete tables are available from the author.

renters. Noncontacts were less likely to be married and more likely to have never been married.

As can be seen in Table 2, the pattern of findings for refusals and noncontacts on the CEIS is similar to that of the CPS, with both refusals and noncontacts predicted better separately than they were taken together.

IV. Conclusions

The present analyses revealed significant differences between respondents and nonrespondents in both the CPS and CEIS surveys and some unique differences among respondents, refusals, and noncontacts. An examination of multivariate logistic regressions showed that the available census variables were able to account for a fairly substantial proportion of variance between respondents and nonrespondents, especially when refusals and noncontacts are examined separately. Further analyses are being conducted to create more parsimonious multivariate models of nonresponse.

V. References

- Couper, M. P. & Groves, R. M. (1993c). *Household-Level Effects on Survey Participation*. Paper presented at the Fourth International Workshop on Household Survey Nonresponse, Bath, England.
- DeMaris, A. (1992). *Logit Modeling: Practical Applications*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-086. Newbury Park, CA: Sage.
- Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56, 475-495.
- Groves, R. M. & Couper, M. P. (1992a). *Survey-Census Match procedures*. Washington, DC: Bureau of the Census (unpublished memorandum).
- Groves, R. M. & Couper, M. P. (1993a). *Unit Nonresponse in Demographic Surveys*. Paper presented at the Bureau of the Census' Annual Research Conference, Arlington, Virginia.
- Groves, R. M. & Couper, M. P. (1993b). *Multivariate Analysis of Nonresponse in Personal Visit Surveys*. Paper presented at the Joint Statistical Meetings, San Francisco, California.

Table 1. Characteristics of Respondents and Type A Nonrespondents to the CPS and CEIS

Variable	CPS			CEIS		
	Respondents	Nonrespondents	χ^2	Respondents	Nonrespondents	χ^2
General Area						
County level Crime Rates	5639 (174)	5895 (146)	ns	5484 (99)	5587 (155)	ns
Urbanicity			**			*
Central City	15.32 (2.00)	24.07 (2.88)		14.49 (1.15)	20.41(1.99)	
Urban, not C.C.	20.31 (2.14)	22.87 (2.68)		21.74 (1.34)	23.51 (1.87)	
Other MSA	35.34 (2.68)	27.42 (2.34)		31.92 (1.59)	26.75 (2.00)	
Other	29.02 (2.32)	25.64 (2.15)		31.86 (1.72)	29.33 (2.06)	
Population density	3091 (222)	4507 (414)	**	3344 (164)	3961 (358)	*
Structures in Block						
% single detached units	62.86 (2.00)	49.46 (1.75)	**	62.29 (1.46)	59.40 (1.61)	ns
% multi-unit: 2-9 units	14.14 (1.14)	19.58 (1.11)	**	14.12 (.83)	14.18 (.95)	ns
% multi-unit: 10-49 units	7.41 (.95)	13.29 (1.04)	**	6.83 (.67)	8.47 (.69)	+
% multi-unit: 50+ units	2.96 (.64)	5.80 (.77)	**	4.24 (.60)	4.95 (.60)	ns
% vacant	7.25 (.44)	8.19 (.38)	+	7.12 (.37)	6.82 (.36)	ns
% boarded up	.19 (.06)	.34 (.06)	ns	.35 (.07)	.25 (.06)	ns
Mean rental cost	467.50 (15.3)	509.80 (16.6)	*	429.70 (10.1)	491.78 (13.4)	**
Mean house values	115,835 (4849)	133,880 (7294)	*	110,443 (5123)	128,368 (5635)	ns
Mean number of rooms	5.45 (.07)	5.04 (.06)	**	5.38 (.05)	5.44 (.05)	ns
Persons in Block						
% occupied units	2.64 (.03)	2.48 (.02)	**	2.67 (.03)	2.60 (.02)	*
% owner occupied	67.25 (1.62)	56.76 (1.53)	**	65.48 (1.20)	65.88(1.24)	ns
% persons in group qtrs.	.25 (.08)	.74 (.18)	**	.56 (.13)	.66 (.18)	ns
% single persons	24.01 (.93)	28.94 (.69)	**	23.59 (.69)	25.51 (.63)	*
% non-family members	5.35 (.34)	6.29 (.28)	*	4.97 (.34)	5.55 (.27)	ns
% married couples w/kids	33.28 (.96)	28.08 (.71)	**	32.79 (.77)	32.16 (.66)	ns

Table 1 continued

Variable	CPS			CEIS		
	Respondents	Nonrespondents	Sig.	Respondents	Nonrespondents	Sig.
% single parent families	11.36 (.54)	12.71 (.50)	+	12.84 (.45)	11.43 (.38)	*
% minorities	21.26 (1.57)	26.53 (1.50)	**	23.32 (1.22)	23.25 (1.20)	ns
% children < 5 years	6.80 (.22)	6.89 (.17)	ns	7.09 (.18)	6.56 (.16)	*
% people > 65 years	13.95 (.65)	13.57 (.51)	ns	14.78 (.60)	13.76 (.51)	ns
% people < 20 years	26.30 (.56)	24.54 (.41)	**	26.91 (.48)	25.47 (.38)	*
persons per room	.50 (.01)	.51 (.01)	ns	.51 (.01)	.49 (.01)	**
Predominant Race			ns			*
White, non Hisp.	82.38 (2.10)	80.30 (1.74)		80.27 (1.54)	81.27 (1.47)	
Black, non Hisp.	8.72 (1.53)	9.73 (1.29)		10.52 (1.15)	9.94 (1.15)	
Other, non Hisp.	.45 (.26)	1.09 (.45)		.29 (.15)	.63 (.23)	
Hispanic, all races	4.46 (1.14)	3.97 (.84)		5.19 (.87)	2.31 (.57)	
No one group pred	4.00 (1.12)	4.91 (.82)		3.73 (.71)	5.85 (.88)	
Housing Unit Structures						
Number of Rooms			**			ns
1-3 rooms	13.10 (1.32)	24.60 (1.73)		13.96 (1.37)	13.94 (1.30)	
4-5 rooms	35.68 (1.83)	38.20 (1.65)		37.72 (1.99)	36.86 (1.88)	
6-7 rooms	34.41 (1.67)	25.55 (1.49)		33.02 (1.89)	33.60 (1.73)	
8 or more rooms	16.81 (1.53)	11.65 (1.13)		15.30 (1.29)	15.59 (1.27)	
Units in Structure			**			ns
Mobile Home	5.94 (1.09)	4.52 (.85)		7.19 (1.14)	4.76 (.95)	
Single fam. home	69.51 (2.25)	55.55 (2.02)		69.32 (1.79)	66.80 (1.70)	
2-9 unit structure	13.93 (1.55)	19.03 (1.59)		11.75 (1.18)	14.37 (1.23)	
10-49 unit struct.	5.85 (1.08)	12.39 (1.25)		6.77 (.93)	8.00 (.95)	
50 + unit struct.	4.80 (.87)	7.34 (1.21)		4.34 (.72)	5.61 (.81)	
Other	.68 (.32)	1.18 (.39)		.63 (.32)	.46 (.22)	
Monthly Rent			ns			ns
0-262	23.20 (3.17)	19.23 (2.35)		26.54 (3.03)	21.40 (2.57)	
262-412	27.54 (3.39)	28.40 (2.82)		30.05 (3.23)	28.13 (3.18)	
412-575	23.53 (3.25)	21.70 (2.29)		23.04 (2.81)	23.17 (2.81)	
575+	25.74 (3.53)	30.67 (3.36)		20.37 (2.62)	27.29 (3.06)	
House Value			ns			+
0-47,500	23.98 (2.31)	19.72 (2.37)		29.16 (2.41)	22.29 (2.20)	
47,500-85,000	24.75 (2.14)	25.29 (2.55)		26.11 (2.16)	25.23 (2.01)	
85,000-137,500	20.89 (1.90)	19.32 (2.05)		19.08 (1.99)	18.73 (1.83)	
137,500+	30.38 (2.53)	35.67 (3.10)		25.66 (2.06)	33.76 (2.53)	
Housing Unit Persons						
all people > 64 years	16.12 (1.20)	12.10 (1.16)	*	16.71 (1.40)	16.29 (1.50)	ns
all people < 35 years	22.30 (1.34)	23.86 (1.58)	ns	21.78 (1.62)	18.88 (1.49)	ns
Children < 5 years	15.14 (1.15)	9.80 (1.00)	**	17.58 (1.52)	9.60 (1.04)	**
no adult males	23.49 (1.26)	37.40 (1.72)	**	23.16 (1.64)	29.04 (1.68)	**
no adult females	14.57 (1.09)	25.79 (1.51)	**	14.01 (1.33)	19.53 (1.43)	**
Race/Ethnicity			ns			ns
Hispanic	6.05 (.97)	5.34 (.86)		6.74 (1.04)	4.56 (.80)	
White, non Hisp.	79.79 (1.82)	75.90 (1.56)		79.20 (1.58)	80.46 (1.44)	
Black, non Hisp.	10.47 (1.40)	13.92 (1.42)		11.24 (1.13)	11.28 (1.17)	
Other, non Hisp.	3.69 (.64)	4.84 (.75)		2.83 (.61)	3.70 (.67)	

Table 1 continued

Variable	CPS			CEIS		
	Respondents	Nonrespondents	Sig.	Respondents	Nonrespondents	Sig.
Household Composition			**			**
Single person	23.26 (1.45)	40.79 (1.85)		21.78 (1.61)	29.72 (1.70)	
Male w/child	.36 (.17)	.43 (.29)		.60 (.29)	1.08 (.38)	
Female w/child	3.52 (.59)	5.71 (.80)		4.12 (.77)	5.28 (.83)	
Spouses w/child	26.33 (1.53)	16.12 (1.36)		28.25 (1.81)	18.83 (1.43)	
Spouses w/o child	22.90 (1.38)	19.57 (1.37)		20.09 (1.62)	22.55 (1.64)	
Other nuc. family	7.42 (.78)	5.43 (.74)		11.10 (1.22)	11.51 (1.25)	
Other related	7.29 (.91)	4.92 (.80)		7.83 (1.02)	5.12 (.87)	
Two nonrelated	8.93 (.94)	7.02 (.90)		6.24 (.90)	5.90 (.93)	
Household Size			**			**
One	22.27 (1.39)	38.20 (1.78)		20.74 (1.54)	27.80 (1.61)	
Two	32.68 (1.56)	30.10 (1.56)		31.05 (1.85)	32.72 (1.78)	
Three	17.97 (1.16)	15.35 (1.18)		17.82 (1.48)	15.39 (1.34)	
Four	14.52 (1.15)	9.62 (1.01)		16.76 (1.39)	15.47 (1.31)	
Five or more	12.55 (1.12)	6.73 (.85)		13.63 (1.37)	8.62 (1.03)	
Number of Adults	1.93 (.03)	1.68 (.03)	**	1.97 (.03)	1.90 (.03)	ns
Marital Status			**			ns
Married	60.08 (1.66)	46.33 (1.86)		60.08 (1.90)	55.57 (1.83)	
Wid., Div., Sep.	27.00 (1.47)	31.23 (1.42)		28.21 (1.69)	31.62 (1.66)	
Never Married	12.92 (1.13)	22.44 (1.55)		11.71 (1.27)	12.81 (1.31)	
Tenure			**			ns
Owner	68.01 (2.09)	54.86 (1.85)		65.74 (1.85)	66.85 (1.79)	
Renter	30.22 (2.09)	42.45 (1.84)		32.20 (1.81)	31.00 (1.72)	
Other	1.77 (.39)	2.69 (.59)		2.06 (.54)	2.15 (.54)	

Note: Values in parentheses are standard errors. ** $p < .01$ * $p < .05$ + $p < .10$

Table 2. Logistic Regression Models
(Values shown are R_L^2)³

Variables Entered	CPS			CEIS		
	Nonresponse Rate	Refusal Rate	Noncontact Rate	Nonresponse Rate	Refusal Rate	Noncontact Rate
General Area	.025*	.023	.028*	.016*	.013*	.029
Structures in Census Block	.132*	.121*	.143*	.143*	.150*	.102*
Persons in Census Block	.033*	.033	.036	.018	.019	.042
Housing Unit Structure	.208*	.191*	.222*	.122*	.117*	.129*
Housing Unit Persons	.275*	.248*	.291*	.183*	.173*	.281*
Total	.329*	.391*	.448*	.243*	.330*	.436*

* $p < .05$

³ Note: See DeMaris (1992)

$$R_L^2 = \frac{-2 \log L_0 - (-2 \log L_1)}{-2 \log L_0}$$