Estimation with Response Error and Non-response: Food Stamp Participation in the SIPP

Christopher R. Bollinger and Martin H. David

University of Kentucky and University of Wisconsin – Madison

Error in survey data originates from failure to contact the sample and from false answers to verifiable questions. These errors may be systematic and associated with uncooperative or unreliable respondents. Zabel (1995) models attrition in the *Survey of Income and Program Participation* and finds systematic demographic and design effects. Bollinger and David (1997) model response error and identify correlations to income per capita. In this analysis we link missing interviews in a panel and response error through a trivariate probit analysis. Robustness of the correlation between attrition and response error is examined by comparing variants of the model. The joint model of response error and attrition becomes the first stage of a pseudo-likelihood estimate of a model of Food Stamp participation. The model is significantly different from naive probit on the survey data.

Acknowledgement 1 *The authors gratefully acknowledge support from the National Science Foundation, grants SBR-9511935 and SBR-952174. We thank the Census Bureau for indispensable assistance and particularly Kent Marquis. Participants in a variety of seminars contributed substantially to our understanding of the problem. The work could not have been completed without the assistance of Joseph Hendrickson. The views expressed here are those of the authors, and do not represent views of the Census Bureau. The authors are solely responsible for errors and will be grateful for comments.*

.

1 Introduction

Participation in the Food Stamp program by eligible households rose twenty percent from 1988 to 1993 and declined substantially after 1995. Policy makers were astounded by these fluctuations, which demonstrate the importance of understanding behavior that leads eligible households to apply for benefits. The long history of scholarly modeling of participation in transfer programs relies on household survey data (Keane and Moffitt 1998, Moffitt 1992). Unfortunately, those data fail to reveal all participants in the Food Stamp Program since survey responses are known to have errors.

Food Stamp Program participation is measured in a number of surveys. The *Survey of Income and Program Participation (SIPP)* panel is the most accurate, as it was designed to measure the use of transfer programs by U.S. households. Nevertheless, *SIPP* underestimates Food Stamp Program recipiency, because some respondents fail to report benefits received (Bollinger and David 1997, Marquis and Moore 1990). Estimates of models of participation clearly needs to include response errors.

Previous work by Bollinger and David (1997) found that response error in the *SIPP* was related to demographic characteristics. Estimates of models of Food Stamp

participation adjusted for response error are significantly different from estimates which fail to account for response error. The response error models are estimated on validated responses to the first interview of the *SIPP*. Models of Food Stamp participation are estimated using data from the fourth interview because it elicits the detailed asset information needed to measure eligibility. How these estimates are affected by changes in the sample due to attrition has not been examined prior to this paper.

Zabel (1995) and Rendtel and Büchel (1994) clearly establish that attrition is determined both by the characteristics of the respondent and the characteristics of the survey methodology. Zabel's (1998) estimate of a dynamic model of attrition reveals nearly monotonic declines in attrition probabilities. (A similar result holds for the *German Socio-economic Panel*.) He also finds the number of item imputations predicts subsequent attrition. He finds comparable relationships for the *Panel Survey of Income Dynamics*. This suggests that respondents first fail to answer some questions and subsequently refuse to answer any questions. Further, a change in interviewer is also found to be positively related to attrition. One explanation for the latter finding is that respondents learn to trust an interviewer; those with a high level of distrust react to a strange interviewer by refusing to be interviewed. Again, the same finding occurs in the *German Socio-economic Panel* (Rendtel 1990).

We hypothesize that a latent variable – propensity to cooperate – determines both response error and missed interviews. Cooperative respondents, with high positive propensity to cooperate, attempt to provide correct information to the enumerator and make themselves available for interviews. noncooperators, with negative values for propensity to cooperate, may be unavailable for interview. Alternatively, they may be interviewed and fail to respond truthfully, particularly to cognitively difficult or sensitive questions. The results below demonstrate a positive relationship between failure to report Food Stamp participation and missing interviews during the first year of the 1984 *SIPP* panel. Further, we find that the improved predictions of response error have an important impact on estimation of a model of Food Stamp participation.

Section 2 presents a model of Food Stamp participation which allows for response error in the survey report of household participation. Section 2 also presents a trivariate model of response error and interview nonresponse. The latent variable structure recognizes that the unobserved propensity to cooperate variable enters into both the propensity to commit errors in reporting Food Stamp receipt and the propensity to miss interviews in the panel. Section 3 briefly describes the data. Section 4 presents estimates of the joint models of response error and interview nonresponse. Section 5 uses the results from Section 4 to estimate models of Food Stamp participation. Assumptions concerning the model of response are tested. Results reveal the necessity

to adjust for response error and nonresponse in estimating behavioral parameters.

2 Models of Food Stamp Participation and Response Error

2.1 Food Stamp Participation and Response Error

The Bollinger and David (1997) model of response error leads to estimation of Food Stamp Program participation using a pseudo-likelihood function. (We call this model the *simple response error* model.) The function incorporates conditional expectations of propensities to respond incorrectly. The model of participation at time t is formulated on household i's true Food Stamp participation indicator, FS^*_{it}.

The probit model of Food Stamp participation for households who are eligible is given by

$$FS^*_{it} = \begin{cases} 1 & if \quad X_{it,F}\beta_F + \varepsilon_{it,F} > 0 \\ 0 & otherwise \end{cases} \quad i = \{asset\text{-}eligible\}. \tag{1}$$

where $\varepsilon_{it,F} \sim N(0,1)$.

Data for the model of Food Stamp participation come from the fourth interview (wave) of the 1984 SIPP panel. The sample is limited to households headed by a married couple who are asset-eligible for the Food Stamp Program. Asset-eligible families have little to live on in the absence of current income sources. The Food Stamps Program also requires income-eligibility – gross and net income less than established thresholds. The choice to become eligible may entail reducing some income

source; therefore, income is endogenous and co-determined with Food Stamp participation. These data are described more completely below and in table 7. Regressors in the participation model can be grouped into three categories:

1. **Earning capacity of the household**: *Hwagehigh* and *Hwagelow* measure the conditional expectation of the husband's wage. *Wwagehigh* and *Wwagelow* measure the conditional expectation of the wife's wage. These regressors are used to avoid endogeneity of earnings and Food Stamp participation. See page 23 for additional discussion.

2. **Demographics:** *Fpovmw4* is the poverty threshold for the household based on official poverty thresholds that depend largely on the number of adults and children. *Kidlt18* measures the number of children in the household. *Disabledh* indicates that the head of the household is disabled. *SMSA* indicates households living in a standard metropolitan statistical area.

3. **Assets:** *Lassets* measures liquid assets counted in the eligibility test. *Hhtheq* measures equity in the ownership of the residence. *Othwlt* measures any other assets held.

The true participation variable is not measured in survey data. Only the household's report of participation, FS_{it}, is available. An indicator for an error of omission (a false negative) is denoted by $O_{it} = 1$. An indicator for an error of commission (a false positive report) is denoted by $C_{it} = 1$. The relationship between observed and true Food Stamp participation is given by the identity:

$$FS_{it} \equiv (1 - O_{it}) FS^*_{it} + C_{it} (1 - FS^*_{it}) . \tag{2}$$

Bayes' rule implies

$$\begin{aligned} \Pr [FS_{it} = 1 \,|\, X_{it}] &= (1 - \Pr [O_{it} = 1 | \, X_{it}] - \Pr [C_{it} = 1 | \, X_{it}]) \, F (X_{it,F} \beta_F) \\ &\quad + \Pr [C_{it} = 1 \,|\, X_{it}] , \end{aligned} \tag{3}$$

where $F(.)$ is the CDF of the standard normal distribution. This relationship can be expressed as

$$\Pr [FS_{it} = 1 | X_{it}] = (1 - p_{it} - q_{it}) \, F (X_{it,F} \beta_F) + q_{it} \tag{4}$$

where p_{it} and q_{it} are the probability of the household committing errors of omission and commission respectively.

Hausman et al. (1998) examine a model similar to the *simple response error* model. They assume $p_{it} = p$ and $q_{it} = q$. Bollinger and David (1997) establish that the assumption is not supported in the *SIPP*.

2.2 Response Error and nonresponse

Bollinger and David (1997) estimate probit models for p_{it} and q_{it} using a subsample of the *SIPP* which was matched to administrative records for three states. This subsample contains both the true state of the household, FS^*_{it}, and the household response, FS_{it}. Sample measurements permit calculation of the indicators O_{it} and C_{it}, and subsequent estimation of probit models of the response errors:

$$O_{it} = \begin{cases} 1 & if \quad X_{it,o}\beta_o + \varepsilon_{it,o} > 0 \\ 0 & otherwise \end{cases} \quad \text{if } FS^*_{it} = 1 \tag{5}$$

and

$$C_{it} = \begin{cases} 1 & if \quad X_{it,c}\beta_c + \varepsilon_{it,c} > 0 \\ 0 & otherwise \end{cases} \quad \text{if } FS^*_{it} = 0. \tag{6}$$

The two probits partition the population. The vectors $X_{i1,o}$ and $X_{i1,c}$ contain one regressor, per capita earnings, in our application of the model (see page 18). The error terms $\varepsilon_{i1,o}$ and $\varepsilon_{i1,c}$ are unobservable and probit estimation implies that the distributions of the two errors are standard normal distributions.

These response error models fail to account for nonresponse, an obvious deficiency. The sample of households interviewed in the fourth interview (wave) where asset data is provided is a choice-based sub-sample. The text table that follows underscores the increasing selectivity of the available data. The probability of missing individuals at the first contact is 4.9% (column labelled 0). Subsequent attrition is shown in the

remaining columns. By the fourth contact attrition has cumulated to 16.9%. Additional nonresponse comes from others who miss the wave 4 interview, but complete later interviews.

Contact	0	1	2	3	4	5	6	7	8	Total
Percent attriting (individuals)	4.9	5.9	6.1	5.4	4.8	4.4	3.7	2.6	0*	37.8

*By definition, attritors must miss the two interviews prior to the last interview for which they are eligible. (Source: Bollinger-David (1996, Table 2))

The *cooperator* hypothesis put forth in the introduction suggests that behavior associated with missing interviews is similar to behavior associated with response error in interviews granted. Hence the *simple response error* models estimated in Bollinger and David (1997) ignore a potentially important observable characteristic: the interview nonresponse pattern of the household over the entire panel. This omission is particularly important when the timing of the available data is considered. Recall that the models of response error can only be estimated on data derived from the first interview. Models of Food Stamp participation can only be estimated on data derived from the fourth. Since the fourth interview occurs nearly a year after the first interview, the relationship between response error and subsequent interview nonresponse is important.

Recognizing the inadequacy of the *simple response error* models in Bollinger and David (1997), we broaden the scope of the model to include measures of missed interviews. The latent variable, propensity to cooperate, measures the household's

willingness and ability to provide accurate data at each interview and to continue to provide interviews.

The multivariate *cooperator* model for response error and interview nonresponse incorporates three indicators as dependent variables: O_{i1}, C_{i1}, A_{iT}. The variable $A_{iT} = 1$ indicates some interview nonresponse for the i^{th} household over a time window $1 \leq t \leq T$. Individual interview nonresponse is aggregated to the household, assuring comparable units of analysis in every dimension of the model. nonresponse is aggregated over a time window that begins in the first wave of interviewing, since every household includes someone interviewed in wave 1.

A probit model is applied to interview nonresponse. Formally,

$$A_{iT} = \begin{cases} 1 & if \quad X_{iT,A}\beta_A + \varepsilon_{iT,A} > 0 \\ 0 & otherwise \end{cases} \quad T = 4, 8 \ . \tag{7}$$

Probit estimation entails the assumption that $\varepsilon_{iT,A}$ is normally distributed.

Recapitulating, each probit (eqs. 1,5,6, and 7) requires errors drawn from a standard normal distribution. The multivariate model that is applied to the four indicator variables is presented in the following table:

Variables					Parameters				
Behavior	Indicator	Conditioning	Random	Universe	Slopes	Covariance			
True FS	FS^*_{it}	$X_{it,F}$	$\varepsilon_{\cdot it,F}$	All i,t	β_F	1	0	0	0
Nonresponse	A_{iT}	$X_{iT,A}$	$\varepsilon_{\cdot iT,A}$	All i	β_A	0	1	ρ_o	ρ_c
Omission	O_{i1}	$X_{i1,o}$	$\varepsilon_{\cdot i1,o}$	$FS^*_{i1}=1$	β_o	0	ρ_o	1	ρ_{oc}
Commission	C_{i1}	$X_{i1,c}$	$\varepsilon_{\cdot i1,c}$	$FS^*_{i1}=0$	β_c	0	ρ_c	ρ_{oc}	1

ρ_{oc} represents the unidentified correlation between errors of omission and commission.

The *cooperator* model is defined by the last three rows of the above table. It joins the models of omission, commission and nonresponse. The *cooperator* model tests **Hypothesis 1:** $\rho_o > 0$. We interpret a positive correlation between $\varepsilon_{\cdot iT,A}$ and $\varepsilon_{\cdot i1,o}$ to mean that latent *cooperativeness* induces individuals to provide interviews and accurate responses. Individuals with negative values of the latent *cooperativeness* variable will miss interviews and provide inaccurate responses. Technically, it also would suggest that $\rho_c > 0$ and $\rho_{oc} > 0$. However, errors of commission are extremely rare in a single cross section; significant results for the correlation between $\varepsilon_A$and ε_c are not expected. We cannot simultaneously observe errors of omission and commission, so the parameter ρ_{oc} is not estimable in this context.

The *cooperator* model of response error and nonresponse is a trivariate probit model with parameters $\beta_A, \beta_o, \beta_c, \rho_o$, and ρ_c. Cross-sectional data from the first

contact with the sample generate coefficient estimates of the *cooperator* model by MLE. Only a subsample of the observations is validated by administrative records. The log-likelihood of the sample data is actually composed of three terms. The first term is a bivariate probit on the subsample of households where administrative records indicate true receipt of food stamps ($FS^*_{it} = 1$). The second term is a bivariate probit on the subsample of observations where administrative records indicate no receipt of food stamps ($FS^*_{it} = 0$). The final term is a univariate probit for the nonresponse model on the remaining households where no administrative records are available:

$$\begin{aligned}
\mathcal{L}_1 \;=\; & \sum_{FS^*_{i1}=1} \{A_{iT} \cdot O_{i1} \cdot PO_{11} + A_{iT} \cdot (1 - O_{i1}) \cdot PO_{10} \\
& + (1 - A_{iT}) \cdot O_{i1} \cdot PO_{01} + (1 - A_{iT}) \cdot (1 - O_{i1}) \cdot PO_{00}\} \\
& + \sum_{FS^*_{i1}=0} \{A_{iT} \cdot C_{i1} \cdot PC_{11} + A_{iT} \cdot (1 - C_{i1}) \cdot PC_{10} \\
& + (1 - A_{iT}) \cdot C_{i1} \cdot PC_{01} + (1 - A_{iT}) \cdot (1 - C_{i1}) \cdot PC_{00}\} \\
& + \sum_{FS^*_{i1}=missing} A_{iT} \cdot F\left(X_{iT,A}\beta_A\right) + (1 - A_{iT}) \cdot F\left(-X_{iT,A}\beta_A\right). \quad (8)
\end{aligned}$$

The terms $PO_{r,s}$ denote the log probability of observing $A_{iT} = r$ and $O_{i1} = s$, and terms $PC_{r,s}$ denote the log probability of observing $A_{iT} = r$ and $C_{i1} = s$ ($r, s =$ 0,1). For example, two of the corresponding integrals of the normal density function are given by

$$PO_{11} = \ln \int_{-X_{i1,o}\beta_o}^{\infty} \int_{-X_{iT,A}\beta_A}^{\infty} f(\varepsilon_A, \varepsilon_o; \rho_o) d\varepsilon_A d\varepsilon_o \quad (9)$$

$$PC_{11} = \ln \int_{-X_{i1,c}\beta_c}^{\infty} \int_{-X_{iT,A}\beta_A}^{\infty} f(\varepsilon_A, \varepsilon_c; \rho_o) d\varepsilon_A d\varepsilon_c \quad (10)$$

where $f(.)$ is the pdf of the standard normal distribution. Thus identification and estimation of (ρ_o, β_o) hinges on the first term and identification and estimation of (ρ_c, β_c) hinges on the second term. The parameters in the nonresponse model β_A are found in all three terms. Specification of conditioning variables derives from economic concepts for modeling O_{i1}, C_{i1}. Specification of conditioning variables for A_{iT} derives from the analyses of Zabel (1995).

Estimation of the parameters of the participation model (eq. 1) will be undertaken in a second step. The likelihood function for the participation model is given by

$$\begin{aligned} \mathcal{L}_2 = & \sum_{i=eligible} FS_{it} \cdot \ln\{((1 - p_{it} - q_{it}) F(X_{it,F}\beta_F) + q_{it})\} \quad (11) \\ & + (1 - FS_{it}) \cdot \ln\{((1 - p_{it} - q_{it}) F(-X_{it,F}\beta_F) + p_{it})\}. \end{aligned}$$

The results from the estimation of the *cooperator* model are used to construct predicted values $\widehat{p}_{it}$ and $\widehat{q}_{it}$ on the wave 4 asset-eligible sample. These predictions are used to form a pseudo-likelihood function which is maximized with respect to the parameter β_F. Pseudo-maximum likelihood has a long history and desirable asymptotic properties (See Gourieroux et al. 1984). Estimation requires three assumptions:

Assumption 1: Models of response error apply to the United States population, even though they are estimated on a subsample of the SIPP.

Assumption 2: The coefficients in the *cooperator* model are time-invariant.

Assumption 3: The covariances between ε_F and $(\varepsilon_A, \varepsilon_o, \varepsilon_c)$ are zero.

Assumption 1 is necessary, since the matched validation data which allow for observation of the true variable FS^*_{it} are only available for Pennsylvania, Florida and Wisconsin subsamples of the SIPP. Bollinger and David (1997) discuss this issue and provide evidence that this subsample is not markedly different in observable characteristics from the SIPP sample as a whole. Assumption 2 is required since estimation of the participation model must use wave 4 data where detailed asset information is available to construct a sample of eligible households.

The covariance matrix above embodies assumption 3. It implies the model of Food Stamp participation and the *cooperator* model are separable and can be independently estimated. We view assumption 3 as most problematic, because the same event that precipitates attrition may well change Food Stamp participation.

3 Data

The data all derive from the 1984 panel of the *Survey of Income and Program Participation*. The 1984 SIPP panel consists of 8 interviews covering a 32-month pe-

riod. The first interviews were conducted in October 1983. Each interview, termed a "wave," asks a standard set of questions concerning sources of income, participation in government transfer programs and labor force activity. Each wave also contains questions on special topics. Different topics are covered at each wave. Of importance for the models of Food Stamp participation is the supplemental information on asset holdings obtained at wave 4. These data define the Food Stamp eligibles and appear as conditioning variables in the Food Stamp participation models.

3.1 Response error

The data used for estimation of the joint model of response error and nonresponse come from a special subsample of the SIPP for which state administrative records of Food Stamp Program participation are available. Florida, Pennsylvania and Wisconsin provided the Census Bureau with administrative records of participants in the Food Stamp Program. These records were matched by the Census Bureau to the SIPP data for waves 1 and 2. The match was based on name, social security number, address, and demographic information. These data are referred to as the validation data. Any discrepancy between reported Food Stamp participation and the administrative record is considered response error. Bollinger and David (1997) discuss the assumption that administrative records are correct. Details of the match process can be found in Marquis and Moore (1990).

Table 1 presents descriptive statistics for the variables used in estimation of the *cooperator* model. The under-reporting of Food Stamp participation in the survey data can be clearly seen by comparing the first two rows. Table 2 shows the extent of error in reporting receipt of food stamps in household questionnaires relative to administrative records. The counts pertain to the reference month prior to the interview. Net underreporting is smaller than errors of omission because some households falsely declare receipt of food stamps.

Tables 3 and 4 show the gross correlation of errors of omission and commission to a measure of missed interviews, *anymiss* (see below). Interview nonresponse and response error are related. Table 3 demonstrates that overall, 29% of participating households have a missing interview, but 36% of participating households with an error of omission have a missing interview. Conversely, 11% of participating households have an error of omission, but 15% of participating households with a missing interview have an error of omission. Table 4 demonstrates that overall, 28% of nonparticipating households have a missing interview, but 43% of nonparticipating households with an error of commission have a missing interview. Similarly, 0.3% of nonparticipating households have an error of commission, but 0.4% of nonparticipating households with a missing interview have an error of commission. The differences shown are more systematically analyzed in the *cooperator* model.

3.2 Nonresponse

Two measures of interview nonresponse are used in this study. Each is evaluated for the time windows $T = 4, 8$. Variants of these four measures are created by including or excluding out-of-scope households. Table 1 displays means for the eight variables generated by selecting a measure, window, and variant. The more inclusive measure corresponds to the indicators *anymiss* and *anymiss4*. These indicators equal 1 if any member of the household at the first interview misses any subsequent interview. *Anymiss* is measured on the time window that includes all waves for which a member of the household was eligible for interview. *Anymiss4* refers to interviews missed in the first 15 months (waves 2 through 4). The second measure of missing interviews corresponds to the indicators *anyatt* and *anyatt4*. These indicators equal 1 if any member of the household missed the last two interviews for which that person was eligible. This construction assures that at least two consecutive noninterviews are recorded. Individuals who miss only the last interview for which they were eligible are not counted in *anyatt* or *anyatt4* (but would be counted in *anymiss* and *anymiss4)*.

The *cooperator* model of noninterviews relates to respondent behavior. The most pertinent measure of interview nonresponse includes cases where individuals in a household take action to avoid giving an interview – refusing to give an interview,

for example. When a person moves out of scope of the sampling frame, the meaning of nonresponse becomes ambiguous. An individual is deemed out of scope by the Census Bureau for many reasons. Moving out of the United States or to remote areas, death, entering an institutional living arrangement, or entering the military are all classified as out of scope. Some of these actions constitute choices to be non-cooperative. Others may be related to a household crisis and may indicate inability to provide accurate answers. Some are simply random acts unrelated to non-cooperation. Because information is incomplete and reasons for being out of scope are aggregated, it is not clear that out-of-scope individuals should be classified as missing interviews. The four variables described above exclude all out-of-scope cases. The variables *anymissu, anyattu, anymiss4u* and *anyatt4u* include out-of-scope cases.

Relationships among the various measures of interview nonresponse should be noted. Considering the measures aggregated over the entire panel, approximately two-thirds of the households with any individual missing an interview ($anymiss = 1$) had an individual who left the sample ($anyatt = 1$). When nonresponse is aggregated over the second through the fourth waves, forty percent of households with any missed interviews include an individual that leaves the sample. Results are similar when out-of-scope cases are included in the measure of missed interviews.

A_{iT} depends on household membership at a point in time. Individuals enter and leave households throughout the panel. The point in time that is relevant to the *cooperator* model is $t = 1$; $t = 4$ is relevant to the participation model. A household that contained noncooperators at interview 1, may no longer contain those individuals at interview 4. For example, a household under stress from marital problems may contain a noncooperator who leaves the household due to divorce. The responses prior to the divorce will be less accurate than responses after resolution of the crisis.

Our preferred measure of nonresponse is *anymiss4*. (The missing interviews measure includes all households that contain attritors. The corresponding attrition measure excludes some households that miss interviews.)

3.3 Regressors: cooperator model

Regressors can be considered in four groups.

1. **Demographic variables for the head of the household**. *Headage* is the age of the head of the household, *headsex* indicates female headed households, *headwhite* indicates that the head of the household is white, *headms* indicates households headed by a married couple, and *headed* measures the level of education in years.

2. **Financial variables.** *Pcfinc* is the total income from all sources for

the adults in the primary family of the household divided by the number of individuals in that family. In most cases, the primary family is the entire household. The attrition literature suggests that total income or *pcfinc* should appear in the model for nonresponse (Zabel 1995). *Pcearn* is the employment earnings for all adults in the household divided by the number of individuals in the household. Bollinger and David (1997) suggest that *pcearn* should appear in the models for response error for Food Stamp participation. Last, *homeown* indicates that the household owns the housing unit in which they reside.

3. **Family characteristics.** *Fkidlt18* is the number of children in the primary family under age 18. *Fkidgt18* is the number of children in the primary family over age 17. *Oldest* is the age of the oldest adult child in the primary family minus 17.

4. **Persons other than the primary family.** *Nonfam* is the number of people in the household not in the primary family. *Nonfinc* is the total income from all sources of the nonfamily members of the household.

Table 1 presents descriptive statistics for all variables in the nonresponse model on the 19,856 households used to estimate the noninterview indicator, A_{iT}. Variables

are listed in the order in which they appear in the *cooperator* submodel.

4 Estimates

4.1 Cooperator Model

The *simple response error* model used in Bollinger and David (1997) is nested within the *cooperator* model.[1] Our preferred specification for the *cooperator* model uses the variable *anymiss4* to measure nonresponse. The rationale is that the measure is close in time to when response errors were measured. Furthermore, attrition, which entails missing consecutive interviews prior to the termination of data collection, is a narrow measure of uncooperativeness that will not register persons with temporary problems who miss interviews and freely give interviews at a later time.

Table 5 presents the estimated *cooperator* model when interview nonresponse is measured for $T = 4$. This window is particularly relevant for two reasons. First, wave four data are used to estimate a model of Food Stamp participation. Interview nonresponse prior to and including wave 4 is the relevant behavior for sample selection.

[1]Bollinger and David (1997) found that household per capita earnings was significant in prediction of both errors of omission and commission. With commission it is nearly a tautology: the likelihood of incorrectly reporting participation in Food Stamps for high-income households would be nearly zero. One would expect that errors of commission would stem from two possible sources: program confusion or interviewer bias. In each case, the relationship between likelihood of response error and household income would be negative. Errors of commission were found to be positively related to household income. In this case one would suspect that errors of omission are due to either stigma or threat. Although response error may have been weakly related to other variables, the previous work suggests that per capita household earnings are the most important determinant of response probability.

Second, the time window is less than a year after the measurements of response error. One would expect the strongest relationship between response error and interview nonresponse would manifest itself in adjacent periods. That is, *noncooperators* are more likely to commit errors and begin missing interviews early in the panel.

Estimates of ρ_o are the main focus of the model as they test the *cooperator* hypothesis. If ρ_o is significant, then estimation of the model of participation must incorporate the *cooperator* model. Table 5 reveals estimates of ρ_o that vary in significance. When *anymiss4* is used for the measure of nonresponse, the estimate is 0.412. The one-tailed test of the *cooperator* hypothesis ($H_{null} : \rho_0 \leq 0$, $H_{Accept} : \rho_o > 0$) gives a test statistic of 2.16, which supports the *cooperator* hypothesis at all conventional levels. When *anymiss4u* is used, the estimate is 0.339. The one tailed test gives a test statistic of 1.88, which is significant at the 95% (one-tailed) level, again supporting the *cooperator* hypothesis. Although the sign is correct, the estimates of ρ_o are not significantly different from zero when *Anyatt4 and Anyatt4u* (two or more consecutively missed interviews) are used.

Other dependent variables and $T = 8$ have been investigated. Response error can be measured for the screening question. An error in the screener induces response error for recipiency of food stamps any time during the 4-month reference period. Table 6 presents estimates of the correlation coefficient ρ_o for 12 alternative specifica-

tions of the *cooperator* model. The slope coefficients in these models are qualitatively similar to those in Table 5 (available on request). Estimates of ρ_o show varying support for the *cooperator* hypothesis. The correlation is strongest for the model fit for $T = 4$ and error in the screening question. It is not surprising that the relationship between response error and nonresponse attenuates when interview nonresponse is measured over a longer time horizon. Interview nonresponse occurring close in time is most important in predicting response error. The generally insignificant results for $T = 8$ need further investigation on other samples.

4.2 Food Stamp Participation

The sample used for estimation of the Food Stamp participation model, called the primary sample, derives from the fourth interview of the 1984 panel of the SIPP. Married-couple households in which both husband and wife were of working age (18–64) are included. Households whose assets exceed the threshold for eligibility are excluded. Asset eligibility can only be established at the fourth and seventh waves, when questions are asked about asset holdings and debt. Table 7 presents descriptive statistics for each of the variables used in estimation of the participation model.

4.2.1 Regressors of principal interest

The model of participation focuses upon the effect assets have on the decision to participate. Hence, three variables which sum to net worth are included in the analysis: *Lassets*, *Hhtheq*, and *Othwlt*. The variable *Lassets* contains liquid assets that determine eligibility for the Food Stamp Program, including checking accounts, savings accounts, stocks, and bonds, less associated debt. The variable *Hhtheq* is the self-reported value of the house owned by the householder less the amount of outstanding mortgages. The final variable *Othwlt* contains net wealth from automobiles, businesses, second homes, and other sources that are not counted in the other two categories.

Earnings capacity for the husband and wife, *Hwage* and *Wwage,* was predicted from standard human capital models (Bollinger and David 1997). The models predict wage rates from schooling, experience, and family life cycle. The models employ standard corrections for selectivity. The predicted values are independent of the earner's current level of work activity, thereby avoiding the simultaneity bias that would be associated with earnings as a regressor. The coefficient estimated describes a response to earnings capacity (Garfinkel and Haveman 1977; Morgan *et al.* 1962).

Earnings capacity for the husband and wife and the asset variables *Hhtheq* and *Othwlt* are splined. The spline allows a nonlinear response and assures that responses

estimated for the lower part of the distribution of these four variables are not affected by outliers in the upper tail of the distribution.

4.2.2 Estimates

The likelihood function for the observed report of Food Stamp participation is a function of the probabilities of errors of omission and commission (eq. 4). Using the model of joint response error and interview nonresponse estimated in the previous section, a pseudo-likelihood function for FS_{it} can be constructed. The probability of omission errors conditional on nonresponse can be derived as

$$p_{it} = \Pr\left[O_{it} = 0 \middle| X_{it}, A_{it} = 1\right] * A_{it} + \Pr\left[O_{it} = 0 \middle| X_{it}, A_{it} = 0\right] * (1 - A_{it})$$

$$= \frac{\Pr\left[O_{it} = 0,\, A_{it} = 1 \,|\, X_{it}\right]}{\Pr\left[A_{it} = 1 \middle| X_{it}\right]} * A_{it} + \frac{\Pr\left[O_{it} = 0,\, A_{it} = 0 \,|\, X_{it}\right]}{\Pr\left[A_{it} = 0 \,|\, X_{it}\right]} * (1 - A_{it})\,, \qquad (12)$$

while commission errors can be derived as

$$q_{it} = \Pr\left[C_{it} = 0 \middle| X_{it}, A_{it} = 1\right] * A_{it} + \Pr\left[C_{it} = 0 \middle| X_{it}, A_{it} = 0\right] * (1 - A_{it})$$

$$= \frac{\Pr\left[C_{it} = 0,\, A_{it} = 1 \,|\, X_{it}\right]}{\Pr\left[A_{it} = 1 \,|\, X_{it}\right]} * A_{it} + \frac{\Pr\left[C_{it} = 0,\, A_{it} = 0 \,|\, X_{it}\right]}{\Pr\left[A_{it} = 0 \,|\, X_{it}\right]} * (1 - A_{it})\,. \qquad (13)$$

Parameters from the Table 5 model *anymiss4* are used to calculate household level predictions for $\hat{p}_{i4}$ and $\hat{q}_{i4}$. The calculation takes the household membership for the month prior to interview at the wave 4 contact. This means that some "split-offs" from wave 1 households now are in separate dwellings. The noninterview behavior

of their "roommates" is aggregated into the nonresponse measure, even though those persons may not have been in the sample at wave 1. The measure A_{i4} results from this construction. Reported standard errors are computed to account for sampling variance from the first stage estimation of the models for p_{it} and q_{it}. See Gourieroux et al. (1984) and Bollinger and David (1997) for detailed explanation.

The primary sample contains households who are clearly not eligible to receive food stamps because they have high incomes (although they are asset-eligible). The primary sample has a much higher mean for $\hat{q}_{i4}$ than the value observed for households known to receive food stamps, $FS^* = 1$. The probit model for errors of omission predicts that households with sufficient income to be ineligible are very likely to withhold information. But those same households are very unlikely to receive food stamps. This tension results in the difference in estimates between the *naive probit* and models incorporating response errors.

Table 8 compares alternative estimates of probits on indicators with errors in samples with missing interviews. The *naive probit* model makes no correction for response error; it is biased because it cannot account for false negative answers. The *simple response error* model, first presented in Bollinger and David (1997), overcorrects for errors in the data because the *cooperator* model confirms **hypothesis 1.** Error-prone households are more likely to miss interviews than correct reporters. The *cooperator*

model uses $\hat{p}_{i4}$ and $\hat{q}_{i4}$ to incorporate behavior relating to response error and missed interviews in the estimation of the participation model. The results are intuitive. Differences between the *naive probit* and the *simple response error* models shrink. Sampling error in the *cooperator* model necessarily is more correctly estimated than in the *naive probit* because the *cooperator model* cannot be ignored.

One might be tempted to compare the three models using a likelihood ratio or other test. However the three participation models are not nested. The likelihood functions used to calculate each one are different in the given values for $\widehat{p}_{it}$ and $\widehat{q}_{it}$. The underlying distributions which give rise to the likelihood function are different and thus standard tests are not applicable. The correct locus for testing the *cooperator* model against the *naive* or *simple response error* models is on the cooperator model (tests shown on page 21). Technically, $\mathcal{L}_1$ (eq. 8) and $\mathcal{L}_2$ (eq. 11) can be added and jointly maximized. The joint model is nested. However, the test for comparing models would again be tests on the parameters in the models of response error and ρ_o, already shown. Most importantly, the pseudo-likelihood method permits these estimates of the *cooperator* model to be used in connection with other samples where survey responses are not validated.

An intuitive measure of the importance of the response error correction is differences in the predictive results of the models. We calculate the predicted probability

of receiving food stamps for each household, varying the level of two important conditioning variables – earning capacity of the head and home equity. Conditioning variables not used as classifiers are set at their sample means. Table 9 presents the mean probability of Food Stamp participation within earning capacity or home equity classes for all three models. The qualitative differences are remarkable. For both classifying variables, the *cooperator* model has a participation rate that drops off markedly faster than the base model, but somewhat slower than the *response error* model. For example, in the first row, it can be seen that the probability of participation at the very low potential wage rate of $5 per hour is nearly the same in the *naive probit* and *cooperator* model, and somewhat lower in the *simple response error* model. However, at the high end of the range (predicted wages over $13.36 per hour increases only *Hwagehighb*), the *cooperator* model predicts only 1.6% of such households will participate, whereas the simple model predicts 1%, and the base model predicts 2.4%.

5 Conclusions

Two important findings emerge from this analysis: response error is correlated with interview nonresponse, and estimates of models of Food Stamp participation should be adjusted for response error while controlling for interview nonresponse. Section 4.1

presents evidence that families with higher than average response error in the first interview also exhibit higher than average interview nonresponse in the first year of the panel. This finding supports **hypothesis 1** and a theory of *cooperativeness*. *Cooperators* reveal themselves by willingness and availability to give interviews throughout the panel and by the accuracy of their data. *Noncooperators* reveal themselves by missed interviews and the inaccuracy of their data.

Estimates of the *cooperator* model can be combined with data from other survey samples measuring Food Stamp participation, so long as assumptions 1-3 can be maintained.

These findings imply strong direction for data collection and estimation. Validation data need to be collected from samples that are coordinated with major panel studies. We do not know enough about circumstances that lead to poor quality response nor do we know which attributes cause response errors to vary systematically. In many areas of economic study, particularly the collection of asset and asset-related income data, aggregate data lead to a presumption that underreporting is substantial. The lapse of resources directed at measuring response error since Ferber (1969a), Ferber et al. (1969b) and Lansing *et al.* (1961) needs to be erased. A second conclusion is that validation data need to be accessible to a large community of users so that appropriate models of error can be scientifically developed.

A second implication is that scientists can no longer presume that nonresponse and response errors are separable. Both confound model estimation. Conventional weighting schemes for nonresponse will often not improve model estimates. In panel data it may be that nonresponse profiles can proxy for some part of response error. The extent to which that proxy assists estimation can only be determined from careful validation of a substantial number of domains of survey measurement. Should our **hypothesis 1** be supported in most studies of validation data, survey design would need to redirect resources that are now directed at reducing nonresponse towards measuring response errors.

Our findings emphasize the need to validate survey data. Morgenstern (1963) called for validation as a regular aspect of data collection. Others have echoed this call. We provide strong evidence that validation can improve estimation of models on probability samples of households.

References

Bollinger, Christopher R., Martin H. David, and Kent Marquis. 1996. Identifying unreliable respondents: Can missed interviews proxy for validation? *Proceedings of the Survey Research Section of the American Statistical Association*, Vol. 2 pp. 675-680.

Bollinger, Christopher R. and Martin H. David. 1997. Modeling discrete choice with response error: Food Stamp Participation. *Journal of the American Statistical Association*, Vol. 92, no. 439, pp. 827-835.

Ferber, R.J. Forsythe, H. W. Guthrie, and E. S. Maynes. 1969a. Val-

idation of a National Survey of Consumer Financial Characteristics: Savings Accounts. *Review of Economics and Statistics*, 51, pp 436-444.

Ferber, R., J. Forsythe, H. W. Guthrie, and E. S. Maynes. 1969b. Validation of a National Survey of Consumer Financial Characteristics: Common Stocks. *Journal of the American Statistical Association*, Vol. 64, pp. 415-432.

Garfinkel, Irwin and Robert H. Haveman. 1977. *Earnings Capacity, Poverty, and Inequality.* New York, NY: Academic Press.

Gourieroux, C., A. Monfort and A. Trognon 1984. Pseudo Maximum Likelihood Methods: Theory. *Econometrica*, vol. 52, pp. 681-700.

Hausman, J. A., Jason Abrevaya, and F. M. Scott-Morton. 1998. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, Vol. 87, no. 2, pp.239-270.

Keane, M. and R. Moffitt. 1998. A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review*, Vol. 39, no. 3, pp 553-589.

Lansing, John B., Gerald P. Ginsberg, and Kaisa Braaten. 1961. *An Investigation of Response Error.* Urbana, IL: Bureau of Economic and Business Research, University of Illinois.

Marquis, Kent and Jeffrey Moore. 1990. Measurement Errors in the SIPP Program Reports. *Proceedings of the Bureau of the Census Annual Research Conference.* Washington, DC: Bureau of the Census, pp. 721-745.

Moffitt, R. 1992. Incentive Effects of the US Welfare System: A Review,.*Journal of Economic Literature.* Vol. 30, no. 1, pp1 - 61.

Morgan, James N., Martin H. David, Wilbur J. Cohen, and Harvey E. Brazer. 1962. *Income and Welfare in the United States.* New York: McGraw-Hill. Table 6-16.

Morgenstern, Oskar. 1963. *On the Accuracy of Economic Observations.* Princeton, NJ: Princeton University Press. Census, 721-745.

Rendtel, Ulrich and Felix Büchel. 1994. Tests for non-ignorable panel attrition and their application on [sic] wage estimates from the *German Socio-economic Panel.* Deutsches Institut für Wirtschaftsforschung (DIW) Discussion Paper

Rendtel, Ulrich. 1990 Participation rates in panel surveys - Influence, confidence, and social selection: Development of participation rates in the *German Socio-economic Panel. Kölner Zeitschrift für Soziologie und Sozialpsychologie. Vol 42,* no. *2,* pp. *280-299.*

Zabel, Jeffrey E. 1995. A comparison of attrition in the *Panel Survey of Income Dynamics* and the SIPP. SIPP Working Paper. Washington D. C.: U. S. Bureau of the Census.

Zabel, Jeffrey E. 1998. An analysis of attrition in the *Panel Survey of Income Dynamics* and the *Survey of Income and Program Participation* with an Application to a Model of Labor Market Behavior, *Journal of Human Resources,* Vol. 33, no. 2, pp. 479-506.

6 Tables

Table 1: Descriptive Statistics: Validation and Non-response Data, (last month prior to interview)

Symbol	Label	Definition	T	N	Mean	Std. Dev.
FS^*	RFS	FSP record: food stamps rec'd		2685	0.067	0.251
FS	QRS	SIPP response: food stamps rec'd		2685	0.062	0.242
	pcearn	Per Capita Household earnings		19856	656.3	836.9
A_T	Anymiss	Any adult misses any interview	8	19856	0.283	0.450
	Anymissu	*Anymiss* plus out of scope cases	8	19856	0.308	0.462
	Anyatt	Any adult misses last two interviews	8	19856	0.182	0.386
	Anyattu	*Anyatt* plus out of scope cases	8	19856	0.206	0.404
	Anymiss4	Anymiss on waves 2-4 only	4	19856	0.164	0.370
	Anymiss4u	*Anymiss4* plus out of scope	4	19856	0.178	0.383
	Anyatt4	Anyatt on waves 2-4 only	4	19856	0.069	0.253
	Anyatt4u	*Anyatt4* plus out of scope	4	19856	0.078	0.268
	pcfinc	Total family income/family size		19856	984	1707
	headage	Age of head of household		19856	47.8	17.5
	Headsex	Gender of head (1=female)		19856	0.316	0.465
	headwhite	Race of head (1 = white)		19856	0.869	0.338
	Headms	Marital Status of head (1=married)		19856	0.590	0.492
	Headed	Education of head		19856	12.1	3.51
	homeown	Ownership of residence (1 if own)		19856	0.643	0.479
	fkidlt18	Number of minor children in family		19856	0.695	1.09
	fkidgt18	Number of adult children in family		19856	0.233	0.588
	Oldest	Age - 17 of oldest adult child, 0 otherwise		19856	4.32	10.28
	Nonfam	Number of people not in primary family		19856	0.074	0.349
	nonfinc	Income of people not in primary family		19856	50.1	299.2

Table 2: Households Receiving Food Stamps

Validation sample: Last Month Prior to Interview, 1983				
Administrative	Survey response			
report	No	Yes	Total	Error rate
No	2496	**8**	2504	0.3%
Yes	**22**	159	181	12.2
Total	2518	167	2685	
	Errors indicated in bold.			

Table 3: Omission Errors and Anymiss

Households receiving food stamps: administrative report					
		Error of Omission			
Anymiss		No	Yes	Total	Error rate
None	0	114	14	128	11%
Some	1	45	8	53	15
Total		159	22	181	29
%		88	12	100	

Table 4: Commission Errors and Anymiss

Households not receiving food stamps: administrative report					
		Error of Commission			
Anymiss		No	Yes	Total	Error rate
None	0	1778	4	1782	0.2%
Some	1	691	3	694	0.4
Total		2469	7	2476	0.3

Table 5: Cooperator Model of Response Error and Missed Interviews

Coefficient		Measure of interview nonresponse			
Symbol	Name	anymiss4u	anymiss4	anyatt4u	anyatt4
	Constant	-0.437^*	-0.567^*	-1.05^*	-1.19^*
		(0.068)	(0.070)	(0.086)	(0.090)
	Headage	-0.0040^*	-0.0047^*	-0.0028^*	-0.0029^*
		(0.00078)	(0.00081)	(0.0010)	(0.0011)
	Headsex	-0.047	-0.015	-0.046	-0.045
		(0.029)	(0.030)	(0.035)	(0.037)
	HWhite	-0.131^*	-0.137^*	-0.043	-0.055
		(0.032)	(0.032)	(0.039)	(0.040)
β_A	Headms	-0.078^*	-0.011	-0.074^*	-0.042
		(0.030)	(0.031)	(0.036)	(0.037)
	Headed	0.0056	0.0074^*	0.0031	0.0050
		(0.0035)	(0.0036)	(0.0044)	(0.0046)
	PCfinc	-0.000093^*	-0.000072^*	-0.000076^*	-0.000042
		(0.000016)	(0.000018)	(0.000021)	(0.000023)
	PCfinc^2	$4.6\,E-9^*$	$3.5\,E-9^*$	$3.9E-9^*$	$2.4E-9$
		$(1.1\,E-9)$	$(1.4\,E-9)$	$(1.5\,E-9)$	$(2.1\,E-9)$
	Homeown	-0.135^*	-0.105^*	-0.114^*	-0.078^*
		(0.026)	(0.026)	(0.033)	(0.034)
	Fkidlt18	-0.067^*	-0.062^*	-0.068^*	-0.061^*
		(0.012)	(0.012)	(0.015)	(0.016)
	Fkidgt18	0.226^*	0.202^*	0.161^*	0.156^*
		(0.029)	(0.029)	(0.036)	(0.038)
	Oldest	0.0011	0.0021	0.0010	0.0010
		(0.0017)	(0.0018)	(0.0022)	(0.0023)
	Nonfam	0.316^*	0.306^*	0.239^*	0.174^*
		(0.034)	(0.034)	(0.039)	(0.042)
	Nonfinc	0.000020	0.000030	0.000020	-0.000022
		(0.000040)	(0.000040)	(0.000042)	(0.000043)

(Table Continued on Next Page)

* indicates significance at the 5% two-tailed test.

Table 5 (Con't): Cooperator Model of Response Error and Missed Interviews						
Coefficient		Measure of interview nonresponse				
Symbol	Name	anymiss4u	anymiss4	anyatt4u	anyatt4	
β_o	Constant	-1.51^*	-1.50^*	-1.52^*		-1.53
		(0.17)	(0.17)	(0.18)		(0.17)
	PCearn	0.0026^*	0.0026^*	0.0026^*		0.0026
		(0.0007)	(0.0006)	(0.00066)		(0.0006)
β_c	Constant	-2.49^*	-2.46^*	-2.48		-2.48
		(0.15)	(0.16)	(4.09)		(4.58)
	PCearn	-0.000093	-0.00070	-0.00073		-0.0007
		(0.0015)	(0.0018)	(0.0042)		(0.01)
ρ_o	Omission ρ	$0.339^\dagger$	$0.412^\dagger$	0.339		0.368
		(0.18)	(0.19)	(0.30)		(0.283)
ρ_c	Commission ρ	0.177	0.213	0.152		0.192
		(1.46)	(1.51)	(2.8)		(3.38)
Mean likelihood		-0.436	-0.417	-0.261		-0.240
Sample Sizes: $N_{nonresponse} = 19856$, $N_{commission} = 2467$ for commission, $N_{omission} = 181$						
*indicates significance at the 5% two tailed test, † indictates significance at 5% level one tailed test						

Table 6: Estimates of ρ_o: Alternative Models

Time window	AnumissU	Anymiss	AnyattU	Anyatt
	A. Last month response error			
Full panel $T = 8$	0.192	0.090	0.097	-0.056
	(0.165)	(0.177)	(0.191)	(0.222)
	B. Screener (4-mo.) response error			
Full panel $T = 8$	$0.289^\dagger$	0.157	0.251	0.094
	(0.143)	(0.154)	(0.158)	(0.186)
15 months, $T = 4$	$0.368^\dagger$	$0.430^\dagger$	$0.407^\dagger$	$0.422^\dagger$
	(0.151)	(0.147)	(0.227)	(0.230)
† indictates significance at 5% level one tailed test				

Table 7: Descriptive Statistics for Primary Sample

Variable	Definition	Mean	Std. Dev.
FS	Survey response: food stamps rec'd	0.098	0.297
HwageA	Predicted Husband's wage below mean	12.17	1.50
HwageB	Predicted Husband's wage above mean	1.26	2.81
WwageA	Predicted Wife's wage below mean	6.80	0.99
WwageB	Predicted Wife's wage above mean	0.66	0.91
Fpovmw4	Poverty threshold	857.7	225.8
Fkidlt18	Number of minor children	1.48	1.29
OthwltA	Other wealth below $75K	4892	15959
OthwltB	Other wealth above $75K	1845	15876
HhtheqA	Home equity below $75K	17085	22253
HhtheqB	Home equity above $75K	1232	8503
Lassets	Liquid assets	408	476
Disabledh	Disabled head of HH	0.13	0.33
SMSA	Live in SMSA	0.44	0.50

N=2316

Table 8: Food Stamp Participation Model (Eligible households, Aug-Nov 1984)

Variable	Naive probit	Simple Response error	Cooperator (anymiss4)
$\bar{p}_4, \bar{q}_4$, (s_p, s_q)	0,0	0.007, 0.261 (0.009,0.195)	0.005, 0.394 (0.001, 0.279)
Constant	1.48*	1.99*	1.86*
	(0.35)	(0.47)	(0.56)
Hwagelow	-0.14*	-0.18*	-0.16*
	(0.03)	(0.04)	(0.07)
Hwagehigh	-0.073*	-0.10*	-0.12*
	(0.022)	(0.03)	(0.04)
Wwagelow	-0.21*	-0.20*	-0.19*
	(0.05)	(0.06)	(0.07)
Wwagehigh	-0.19*	-0.16	-0.19
	(0.09)	(0.13)	(0.12)
Fpovmw4	0.0003	0.0004	0.00027
	(0.0002)	(0.0003)	(0.00025)
Fkidlt18	0.11*	0.12*	0.092
	(0.04)	(0.05)	(0.085)
OthwltA	-4.7 $E-6$	-2.25 $E-6$	-4.0 $E-6$
	(4.6 $E-6$)	(5.97 $E-6$)	(7.1 $E-6$)
OthwltB	3.38 $E-7$	-1.12 $E-6$	-0.96 $E-6$
	(5.42 $E-7$)	(7.24 $E-6$)	(7.6 $E-6$)
HhtheqA	-1.13 $E-5$*	-1.84 $E-5$*	-1.5 $E-5$*
	(.29 $E-5$)	(0.48 $E-5$)	(.72 $E-5$)
HhtheqB	8.77 $E-7$	113 $E-7$	-5.7 $E-7$
	120 $E-7$	(135 $E-7$)	(15.6 $E-7$)
Lassets	-0.0010*	-0.0025*	-0.0018
	(0.00015)	(0.00052)	(0.0030)
Disabledh	1.32*	1.58*	1.56*
	(0.15)	(0.21)	(0.31)
SMSA	0.084*	0.071	0.11
	(0.090)	(0.12)	(0.12)
Loglikelihood	-563.5	-516.9	-518.8
N = 2316			

Table 9: Mean Probability of Food Stamp Participation

Model	Classifier:				
	A. Earning capacity (1984 $/hour):				
	5	7	9	11	13
Naive probit	0.196	0.128	0.078	0.045	0.024
Simple response error	0.191	0.109	0.055	0.025	0.010
Cooperator	0.196	0.120	0.067	0.034	0.016
	B. Home Equity ($1000)				
	None	5	10	15	18
Naive probit	0.048	0.042	0.037	0.033	0.031
Simple response error	0.032	0.026	0.021	0.017	0.015
Cooperator	0.040	0.034	0.029	0.024	0.022