THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

MODELING FOOD STAMP PARTICIPATION IN THE PRESENCE OF REPORTING ERRORS

No. 179

C. R. Bollinger
M. H. David
University of Wisconsin

U. S. Department of Commerce BUREAU OF THE CENSUS

MODELING FOOD STAMP PARTICIPATION
IN THE PRESENCE OF REPORTING ERRORS

Christopher R. Bollinger and Martin H. David[1]

Martin David, University of Wisconsin, 1180 Observatory Dr., Madison, WI 53706

Abstract

KEY WORDS: Response error, SIPP, means-tested transfer income

Validation study of reported program participation by Marquis and Moore (1990) reveals a bias due to net under-reporting of Food Stamp Program participation of 13% (Wave 1-2 data from the 1984 Survey of Income and Program Participation). We extend that analysis conditioning on demographic and economic covariates. The resulting model of over- and under-reporting is included in MLE of the probability of Food Stamp Program participation (David and MacDonald 1992). Survey response and administrative data are aggregated to families in this analysis. Under-reporting is modeled by probit analysis on demographic and economic variables using conceptual insights from cognitive research and economic theory. Aggregation of reports to families eliminates apparent errors associated with discrepancies in identifying the individual certified to receive Food Stamps and multiple reporting of recipiency for a family group. The probability of under-reporting Food Stamp recipiency increases with family income normalized by family size. Other results from the response error analysis are less stable across interviews, and have greater sampling errors. Women are better reporters than men, and married couples are more reliable than single householders. Estimating the model of Food Stamp participation using the models of over- and under-reporting results in substantial shifts of the probit coefficients. Estimates of several coefficients of particular conceptual interest in modeling Food Stamp participation are more than fifty percent larger in the model conditioned on probabilities of response errors.

[1] The authors would like to thank Kent Marquis and Jeff Moore for their generous help in obtaining access to the data and getting started with the analysis. We would also like to thank the staff at Bureau of the Census Housing and Household Economic Statistics and at Census Survey Methods Research for their help and support.

Section 1: Introduction

Many times in economics and other social sciences, cross sectional or panel surveys attempt to measure the participation of individuals in various activities, groups and government programs. One such study is the 1984 Survey of Income and Program Participation (hereafter SIPP). SIPP asks the members of 20000 households about their participation in various government transfer programs such as AFDC, Unemployment Insurance and Food Stamps. Many economists and other social scientists have utilized this data to explore participation in these programs. Although seldom mentioned, it is well understood that these studies are limited by the quality of the data. In particular, errors in the reporting of participation in a program can severely misrepresent the true state of the world and make inference based upon these data faulty.

This paper extends the path breaking validation study of program participation reported in SIPP by Marquis and Moore (1990, hereafter MM). MM used a record check study to examine the response error in a subsample of the 1984 Survey of Income and Program Participation data for eight government transfer programs: Aid For Dependent Children (AFDC), Food Stamps (FS), Unemployment Insurance (UI), Workers' Compensation (WORK), Federal Civil Service Retirement (CSRET), Old Age Survivors Disability Insurance (OASDI), Supplemental Security Income (SSI), and Veterans' Pensions (VET'S). MM found that some programs had only trivial net reporting errors (OASDI has a 1% net over-reporting error), while others had significant reporting errors (AFDC has a 39% net under-reporting). Possibly more significantly, MM found that not all response errors were under reports. In fact, while OASDI had negligible net reporting error, there were still 5.7% under-reporting and 1.5% over-reporting. (These are probabilities of false negatives and false positives, which need to be weighted by the proportion recorded as receiving or not receiving OASDI in administrative records to arrive at net bias.)

The analysis below focuses on the Food Stamp Program. A primary reason was our interest in examining the effect of reporting errors on estimates of a model of FS "take -up", or participation rates (David and MacDonald 1992). Similar models have many practical and policy implications (Cf. Martini 1992). The effect of response error on parameter estimates for this class of models is of great interest. A second reason for studying this program is that MM studied response errors of individuals. Individuals are the pertinent unit of analysis in many cases, but the Food Stamp program applies to an aggregate of individuals. The correct unit of analysis is the household in most cases (Martini 1992, 7). The aggregation of data into household units eliminates apparent errors at the individual level. Failure to identify the individual certified to receive Food Stamps is of no consequence, if another member of the family erroneously responds that she is certified (two compensatory errors). Multiple reports of recipiency are also unimportant.[2]

[2] In some cases two individuals can be certified to receive FS; the survey design also permits reporting of certification by several individuals.

In section two of this paper we describe the validation data collected by Marquis and Moore and analyzed in both their work and ours. We discuss the relevance of the sample universe to the SIPP universe, other data sources, and to the US population as a whole. Aspects of the SIPP questionnaire and how its structure may have led to particular patterns of response error are examined. We summarize some basic results about Food Stamp response error from the MM study, and give comparable results for our aggregated data.

In section three of this paper, a conceptual model of response errors is presented. We argue that gender, mutual status, and normalized family income can play a role in the accuracy of family reports of Food Stamp Program participation.

In section four, the results from estimating several specifications of the response error model presented in section three are reported. Hypotheses regarding the concepts presented in section three are tested, and some explanations for the results are interpreted.

In section five, the theory for utilizing the estimates of section four to correct the Food Stamp participation model is presented. The participation model of David and MacDonald (1992) is then re-estimated, correcting for response error. The corrected results are compared to the original estimates and the impact of failing to correct for measurement error is discussed.

Section II: The Data

Sources of data

The primary purpose of this study was to gain an understanding of the extent and structure of measurement error in the reporting of Food Stamp Program participation. Data necessarily are collected from two sources. The "primary" data set is the Survey of Income and Program Participation (SIPP) 1984 Panel. Validating data derive from administrative records of three states: Florida, Pennsylvania and Wisconsin.[3] All participants in the FS Program during a 12-month interval were included in the set of validating records matched to SIPP records.

The universe of states providing validating data was determined by three constraints:

(1) The state had to contain a large number of individuals who were sampled by SIPP.

(2) State records maintained for the FS program (and other programs not discussed here) had to be centralized, computerized, accessible, complete, and of high quality.

[3] Data were collected for a fourth state. Unfortunately, those data were apparently not extracted correctly and proved to be unusable.

(3) The agency responsible for administering the FS program had to be willing to release identifiable individual-level data to the Census Bureau for purposes of research (Documentation of the project from Marquis and Moore, U.S. Bureau of the Census).

The time window covered by validating data is June 1983 through May 1984. These months correspond to the survey months of the first two waves of SIPP interviewing (Marquis and Moore 1990). Each interview collects retrospective data on a four-month reference period; in addition some information is obtained for the date of interview. Four independent samples, called rotation groups, were interviewed in successive months. As a consequence interviewing occurs in every month, and every respondent has the task of recalling events in the four prior months.[4] Our analysis focuses primarily on responses pertaining to the last reference month before the interview month for waves one and two.

Matching and errors

Validating data were matched to the primary data by Social Security Number, name (first, last and middle initial), house number, street name, apartment number, city, zip code, gender and date of birth. The match leads to two outcomes:

- Administrative reports about FS exist and were linked to SIPP individuals.

- No administrative records exist for the SIPP individual and non-participation is assumed as the default for variables derived from administrative reports in the combined data.

An individual receiving FS benefits any time during the 1-year period coincident with the first two waves of SIPP was matched to the corresponding SIPP household member. However, a few individuals claiming receipt of FS are not reported as receiving FS in the 1-year window for which FS records were examined; those individuals have no administrative reports. This phenomenon can occur because of individual response error or a matching error.

The combined data provide the capability for estimating response error pertaining to questions about FS participation. It is used to distinguish six categories of error.

[4] Thus the interview for wave one, rotation group one, was conducted during October 1983; the reference period represents June through September of 1983. The interview month for wave one, rotation group two, was conducted in November 1983; the corresponding reference period represents July through October 1983, and so forth.

Source of error:

Type of error	Respondent	Administrative records
	(1)	(2)
Commission	Over-reporting recipiency with probability p_i, Errors in payment reported	Administrative errors identified in quality control samples
	(3)	(4)
Omission	Under-reporting recipiency with probability q_i, Failure to report amounts	Incomplete records
	(5)	(6)
Matching errors	Failure to provide valid Social Security number (SSN) refusals, transpositions, etc.	Failure to include record in matching file

Matching errors can not be isolated. Errors of type (5) result in exclusion of the SIPP respondent from analysis when the SSN was refused.[5] Remaining matching errors (5 and 6) are improbable, but possible (Scheuren and Oh, 1981). Overreporting can occur because individuals move into the sample and their FS experience is recorded outside the States from which administrative records were sampled. In fact, this situation occurs only in wave 1 in our analysis, and its importance is questionable since the rate of overreporting is greater in wave 2 (where in-movers are excluded from our statistics) than in wave 1 (where in-movers can not be identified). In our analysis we assume no matching error, and no administrative error (2, 4 or 6). Because administrative errors are carefully reviewed through the FS quality control program, and administrators are penalized for inaccuracies, these assumptions are plausible. However, it is clear that the assumptions give an upper bound to estimates of over-reporting and an unknown bias to the errors of payment

Unit of analysis and sample

The monthly data from waves 1 and 2 are tabulated by Marquis and Moore (1990) at an individual level. However, since Food Stamps are awarded to groups of individuals who share

[5] When an erroneous number was provided, a small probability exists that the number corresponds to someone with the same name and birth date. In that case apparent errors are an artifact.

cooking facilities, the correct unit of analysis is most often the Census household (Martini 1992, 7).[6]

The sample studied here is determined by response to the first SIPP interview. The data were clustered by household. The wave one sample consists of adults responding in wave one.[7] These persons comprise 2685 household units. The wave two sample includes adults who were interviewed at the same residence in both waves one and two less those households gaining or losing persons through change of residence. The wave 2 sample individuals comprise 2564 family units.

The choice of sample for wave 2 is determined by the design of the validation study and the design of the SIPP questionnaire. No record data would be available for persons moving out of the participating states. (Movers are represented in wave one data for four months, and for a sample that is not attritted.) Examining validating records for eight reference months was essential to understanding propensity of some individuals to deny the receipt of FS categorically, and to estimate conditioning effects. Beginning in wave two, the screening question for FS was conditioned on answers to the screening question in the prior wave. That conditioning would not apply to persons entering the sample at wave two.

The validation sample is not a representative sample of either the U.S. population or the SIPP respondents. In Table 1 demographic characteristics of the validation sample are compared to the characteristics of the take up sample studied in section five. Although some differences are notable, none seem to indicate substantial differences between the samples.

Insofar as the model of FS participation includes demographic characteristics associated with the selection biases of the validation sample, the above differences in sample characteristics are of less concern. Only second-order interactions (i.e. race interacting with state) entail potential biases.

Questionnaire and data structure

The structure of the questionnaire itself is of interest. In wave one, each adult member of the family was asked whether they were the authorized recipient of food stamps at any time during the preceding four months. Respondents answering "yes" were then queried about recipiency and payments received in each of those four months. Respondents answering "no" to the screening question were not asked any additional questions about food stamps.

In wave two, the majority of respondents who answered "no" to the screening question during the wave one interview respond to the same sequence of questions about food stamp as in wave one: first a screener, then detailed questions if the screener is answered affirmatively. Respondents who

[6] It is impossible to have two Food Stamp units in a household, and the FS unit may exclude an unrelated individual who does not share in meal preparation and eating. Such cases are rare in the validation sample. It is more important to our results that aggregation over the household allows us to study consensual unions and other unusual living arrangements where unrelated adults are aggregate into a single FS unit.

[7] Individuals less than 18 years of age were excluded since they are not asked about FS recipiency.

answered " yes" to the screener during wave one are treated differently. Respondents were asked first to confirm their food stamp status during wave one; respondents then were asked the screening and detail questions about Food Stamps with respect to the reference period for wave two.

Types of error and probabilities of error

An overview of the quality of FS reporting in wave 1 is provided by Table 2. Nearly all data from the matched sample is consistent, but that is because FS are reported in administrative records for only 7.7% of families (209/2685). About three-quarters of families with reported FS-(154/209), correctly claim to receive payments for the same duration as is reported. A few respondents inconsistently claim program benefits though none is reported, 0.28% (7/2476); this is termed the probability of over-report. 13.9% (29/209) of respondents fail to claim any FS during the 4-month reference period when administrative records report payments. While errors in the duration of benefits characterize an additional 12.4% (26/209) of families, positive and negative errors roughly balance. The error of "report and no claim", an error of omission, is clearly a problem. The table clearly indicates an excess of errors of omission on the part of respondents over errors of commission, leading to a net downward bias in SIPP-based estimates of FS recipiency.

Table 3 displays the data used in modeling response error in Section III. The rows are determined by respondents' answers to the screening question and detail provided for the month prior to the interview; information reported from administrative records determines the columns. Labels from A-I refer to the interior of the table and assist in understanding assumptions underlying estimated error rates. The column labelled "Yes" includes all families in which FS were paid in the month prior to the interview. The proportion correctly claiming receipt of FS is 1 - q = .88 (159/181); q = .12 (22/181) constitute errors of omission. The row containing these correct responses requires "yes" answers to both the screening and the detail questions. Errors of commission arise when reports show no payment or no record is matched to persons giving affirmative answers to both questions. This type of error has an extremely small probability, because only 6.7% (181/2685) of the sample receive FS and errors of commission relate to the remaining population p = .0032 (8/2504). This error rate is problematic. Families in cell G may be truthfully reporting participation, but matching error inhibited linkage to the appropriate administrative report. If so, both probability of over-reporting and probability of under-reporting should be reduced.[8] Both cells G and H are treated as errors in the subsequent analysis, giving an upper bound to estimates of errors of commission.

The structure of the response error is common to both waves. First, the predominant source of errors is inconsistent responses to the screening question. 88% (35/39) of errors of omission in waves one and two did not report receiving FS at all. A similar pattern holds for errors of

[8] If matching is the source of error, an appropriate measure of over-reporting would exclude the estimated number of mismatches in calculating the probability of over-reporting, i.e. (8-3)/2504. The estimated number of mismatches reporting FS should also be subtracted from the apparent number of under-reports, since a survey report should be matched to those cases. This would make a modest reduction in the probability of under-reporting (22-3)/181.

commission. Families over-reporting food stamp participation in the month prior to the interview did not receive food stamps at any time during the wave. Second, some people are good reporters and some people are bad reporters. This manifests itself in a number of ways. In wave one, of the people who correctly report food stamp recipiency in the month prior to the interview, 41% report the amount of food stamp payment correctly; 67% report the amount within $10. In wave 2, 49% correctly report the level and 66% are within $10.

Reporting accuracy manifests itself over time. In Table 4 we report the error rates for the wave two sample broken down by wave one reporting accuracy. Most striking is that if no error was committed in wave one, the probability of committing an error of omission in wave two is very low: only 5% of the families who correctly reported FS participation in wave one and only 1.7% of the families who correctly reported no FS participation in wave one committed an error of omission in wave two. The full wave two sample displayed an 11% rate for errors of omission. Although sample sizes are small, we find that famines who committed an error of commission in wave one never committed an error of omission in wave two and families who committed an error of omission in wave one never committed an error of commission in wave two. Also, 27% of the families who committed an error of omission in wave one also committed an error of omission in wave two. Similarly, 25% of the families but committed an error of commission in wave one also committed an error of commission in wave two.

Table 5 compares reporting error at the individual level (reported in MM) to errors calculated at the household level. The rate of under-reporting for the household unit of analysis is significantly smaller than for individuals. Two factors contribute to the difference. Confusion as to who is certified to receive Food Stamp occurs. This leads to both one over-report and one under-report. Since the number of individuals not receiving Food Stamps is large, a small change in the number of over-reports is insignificant, but a change in the number of under-reports has a large effect upon that rate. The second major source of difference is the few cases where two people are reported as the official recipients for Food Stamps but the SIPP data recorded only one of them as certified. This leads to an under-report which does not exist at the family level. Again, despite infrequent occurrences, the small number of households receiving food stamps makes the number significant.

Table 5 also indicates that differences in error rates between wave one and wave two are inconsequential. Cognitive theories provide some additional insight

Section III: Modeling Response error

Cognitive theories of response error

Several theories relate response error to cognitive processes. Two theories are of particular interest here: retrieval and teaming theories. Error in retrieval of information includes simple forgetting, telescoping, and program name confusion. Barriers to recall can be created by the threat posed in revealing behavior that is not generally acceptable (Sudman and Bradburn 1974). Cognitive research indicates that association of the desired event with the personal history of the responding individual may assist recall.

Evidence presented in MM suggest that simple retrieval problems are not responsible for much response error in reporting FS. They did not explore the issue of threat directly, which we explore below. Connections between personal history and receipt of FS are established by the reporting required for SIPP, and the experimental studies done in preparing the SIPP questionnaire establish that the order of questioning in the present form, which touches on events that may trigger participation in various Federal income support programs, is superior to the order in which questions about recipiency are integrated with questions about amounts (Kulka 1984).

Learning theories of response error suggest that repeated interviews can increase or decrease response error. When the error arises from unfamiliar terms, e.g. "AFDC' rather than "welfare", the respondent may learn the new vocabulary and reduce errors attributable to name confusion. However, the respondent may also learn the conditioning pattern in the questionnaire. This could lead to an increase in response error on the part of uncooperative respondents. That is, the respondent learns that a "yes" answer to the FS screening question triggers a longer set of questions. To avoid a lengthy interview, the respondent answers "no" to the screening question, following a "yes" report. Again, MM's analysis fails to demonstrate significant conditioning of this kind.

Attributes of error-prone households

The question we attempt to answer here is whether there are any characteristics of the household or members of the household that are associated with response error. Plausible hypotheses can be forwarded that errors of omission are associated with: marital status, education, gender of the reporter or recipient, age, race, or income level. The chain of reasoning relates to the interviewing process, to threat, and to saliency.

One important process which can lead to response errors associated with demographic characteristics is the structure of the interview. Because both husband and wife are interviewed when the householder is married, two opportunities arise to elicit FS recipiency, while the single person is only asked once; thus omission errors should be smaller for households including married couples than households including only non-married persons.[9]

Using FS may be stigmatized in some communities. This gives rise to three hypotheses. Omission errors are likely to increase with income, as FS use is more prevalent and accepted among very low income households than among middle-income households. Fraudulent use of FS leads to the same hypothesis - Households earning income that is not reported to the FS administrators will be understandably reluctant to admit to participation in the program; their income is necessarily higher than the income of households with little income or no members in the labor market.

Differences in the ethnicity of the respondent and the interviewer may heighten threatening aspects of reporting FS participation. As most interviewers are white, this effect, in the aggregate, is plausibly to increase omission error by non-whites relative to whites.[10]

Threat may also be more important for older persons whose self-esteem may be more damaged by accepting means-tested income support than is the case for younger persons who regard the support as unconnected to personal failures.

Another aspect is the respondent's awareness of the details of the food stamp program. As most shopping for food is done by women, it is likely that FS are handled by women and they will report more completely than men in a household with several adults. It is plausible that persons with more education are better acquainted with the name of the FS program, reducing omission error.

The notation p_v q_i has already been introduced to describe the probability that the i^{th} household makes an erroneous report. p_i applies to households that do not receive FS and describes the probability of errors of commission (false positive); q_i applies to household that receive FS and describes the probability of errors of omission (false negatives). We assume a single index model for the probabilities of response error. That is, the probability of a response error, conditional upon observable characteristics Z_i, is represented by:

$$p_i = F(Z_i ß_o) \quad \text{and} \quad q_i = F(Z_i ß_1).$$

The model is fully derived below. The probabilities vary because individuals are heterogenous in their motivation, willingness, and ability to respond truthfully to the interviewing process. Circumstances that can be objectively measured may influence contribute to this variation, as hypothesized above.

Realizations of errors are denoted by the indicator variables:

---

[9] However, the opposite effect could be observed: Married individuals could evidence more errors of omission if confusion in the family about who should report is large. This would not affect error rates at the household level of analysis.

[10] Because race of enumerator was not linked to the validation data, it was not possible to test for enumerator effects directly.

y_{i1} $= 1$ Iff FS recipiency is claimed and is not reported
$= 0$ Otherwise

y_{io} $= 1$ Iff FS recipiency is not claimed and is reported
$= 0$ Otherwise

A standard probit model can then be used to model p_i and q_i and gives rise to a particular specification for the function F(.):

Response error Model

(1) $Y^*_{i0} = Z_i\, ß_0 + u_{i0} \quad u_{i0} \sim \text{iid } N(0,1)$

(2) $Y^*_{i1} = Z_i\, ß_1 + u_{i1} \quad u_{i1} \sim \text{iid } N(0,1)$

(3) $Y^*_{i0} = 1$ If $y^*_{i0} > 0$
$= 0$ Otherwise.

(4) $Y^*_{i1} = 1$ If $y^*_{i1} > 0$
$= 0$ Otherwise.

The variable y_{i0} is only observed if the respondent is a food stamp recipient, and y_{i0}, is only observed if the respondent is not a food stamp recipient. Hence, the indicator y_{i0} is equal to one if an error of omission has occurred; y^*_{i1} equals one if an error of commission has occurred. This model allows us to estimate $ß_0$, $ß_1$ using maximum likelihood estimation. This then allows us to simulate the probability of response error given the true status for other samples, including members of the SIPP sample who were not included in the validation sample.

The take-up model, that is the model of the probability of FS participation, is specified with an error conditioned on the truefood stamp status. This specification will provide a correction that incorporates the impact of measurement error on estimates of model parameters. Consider the following food stamp participation model.

Food Stamp Participation Model

(5) $y_{if} = X_i\, ß_i + u_{if} \quad u_{if} \sim \text{iid } N(0,1)$

(6) $\text{FStrue}_i = 1$ If $y^*_{if} > 0$
$= 0$ Otherwise.

This standard probit model of program participation has been estimated in many contexts. However, the researcher only observes the reported food stamp participation, FSrep. If FSrep is generated by the response error model, then the following is true by definition:

(7) $\Pr[FSrep = 1 \mid X] = \Pr[y_{i0} = 0 \mid X_i] \Pr[FStrue = 1 \mid X_i].$
$+\Pr[y_{il} = 1 \mid X_i] \Pr[FStrue = 0 \mid X_i].$

Equation (7) can be estimated if the probability of false positives and the probability of false negatives are known or estimable. (Below we refer to the vector of these estimated probabilities as the vector error rates.) The simplest model assumes that p and q are independent of household characteristics. Then

(8) $F(FSrep \mid X_i; ß_f, p, q \qquad \{[1 - p - q] F(X_i ß_f) + p\}^{*}$
$\{[1 - p - q] [1 - F(X_i ß)] + q\}$

$(1 - p - q) F(X_i ß_f) + p$

Thus, estimation of $ß_f$ requires modeling the probability of error given the true status in constructing the likelihood function. For that reason, we estimate the model (1) - (4).

One additional point should be noted. The take-up model is estimated on data from Wave 4, where a balance sheet of assets and debts is measured.. It is not possible to determine eligibility in waves 1 and 2. The eligible population in wave 4 will differ from the eligible population in waves 1 and 2, where validity was checked. Hence it is not possible to use validity data directly in estimating the take-up model, even for the sub-sample for which validity is determined.

Estimated models of response error

Table 6 displays several specifications for models of errors of omission. Model A1 is the most concise and summarizes interactions for gender and marital status shown in Model A2. Both models are estimated on Wave 1. Model A3 pools data from new entrants to FS in wave 2 with Wave 1, excluding persons subject to the additional confirmatory questions for prior claimants. We prefer to use the evidence from the Wave 1 sample, as it is not confounded by attrition and sample selection entailed for wave 2 cases.

The results range from weak to definitive. Differences in probabilities of errors of omission may exist among single men, single women and married couples (Models A1, A2), but the number of parameters required to identify this interaction is too large to afford a finding of significant effects. Caution is in order given the absence of such effects in the pooled wave 1 and wave 2 estimates (Model A3). Surprisingly, the probability of error is smaller when a proxy interview is taken. As noted in section two above, other evidence suggests that there are good reporters and bad reporters (or cooperators and noncooperators). Hence, this may simply be another example of this phenomena. Most importantly, of course, the hypothesis that proxy interviews are a source of reporting error must be rejected. The strongest finding is that larger per capita incomes lead to an increase in the probability of reporting error.

This may be due to stigma effects, lack of familiarity with the program, or the threat associated with program fraud.

Because errors of commission are rare, reliable modeling of relationships to personal characteristics is impossible. Model B1, in table 6, describes a significant decline in errors of commission as per capita income rises. The finding may simply mirror the absence of eligibility for FS among higher income persons, and the lack of effort required to give the truthful negative answer for those persons.

Section IV- Modeling participation rates

We estimate participation from a model to predict the impact of extending eligibility to some who are currently asset-ineligible. The assumptions underlying our modeling approach are explained below.

Our reduced-form model of the participation decision differs from past work in a number of respects:

a. The sample on which participation is modeled includes only asset-eligible units.

b. The sample includes only couples where both partners are of working age, between eighteen and fifty-nine years old.

c. The choice is related to economic constraints on the household, rather than economic outcomes. Expected wages are used as a regressor and actual earnings are disregarded. Below we refer to the estimate of the expected wage as 'earnings capacity', since the model of the wage conditions on conceptually and empirically important observable factors that condition the wage level of individuals with differing characteristics.

d. The model conditions on all components of net wealth. This allows the household to respond to the whole of its resources. The choice to participate balances program benefits against the costs (stigma and transaction) of participation. Because wealth is a resource that can be used to compensate for temporary income drops, it seems reasonable to expect that the rate at which couples participate is reduced by the presence of wealth. Because the counted and uncounted components of wealth are treated separately in our model, we can gain some insight into the impact of program rules on behavior.

e. The response to both expected earnings (earnings capacity) and wealth is modeled using splines to capture nonlinearities.

f. Different responses to earnings capacity are permitted for husband and wife. Assuming that the worker with the higher earnings capacity is the first to participate in the labor force, we model separate effects for the earnings capacity of the higher-paid worker and the capacity of the lesser-paid worker.

Earnings Capacity

People with the same earnings capacity include some slackers and some high achievers. It is of more interest to know the response of persons with equal earnings capacities to an increase in capacity (that might be effected by training or policies affecting labor demand) than to know that achievers will work their way out of eligibility and slackers will not. Using expected wage rates to measure earnings capacity provides a way of scaling demographic characteristics by their economic importance.[11]

Earnings capacity is represented by four variables:

a. Whichever is smaller: the wage of the higher-earning worker or the mean wage of higher earners (in Table 7 this variable is labeled "wage of high earner, below mean");

b. The amount of the wage of the higher-earning worker in excess of the mean wage, which will be zero for those with below-average wages ("wage of high earner, above mean");

c. Whichever is smaller: the wage of the lower-earning worker or the mean wage ("wage of low earner, below mean"); and

d. The amount of the wage of the lower-earning worker in excess of the mean wage, which will be zero for those with below-average wages ("wage of low earner, above mean")

For this sample there was no woman whose earnings capacity exceeded her husband's. As a result the earnings capacity of the husband is identical to the earnings capacity of the higher earner.

Our expectation is that increased earnings capacity inhibits participation. In part this is due to the increased likelihood that high-capacity earners will find the search for work relatively short. We also expect that the earnings capacity of the higher-endowed worker will be more powerful in reducing participation.

Measures of Wealth

Five variables relate to net wealth. The average net wealth of households in the sample for which participation rates are modeled is over $25,000 - even though the sample includes only asset-eligible couples. The variables tested pertain to countable assets, home equity, and equity in

---

[11] In a variant of the model, we add some characteristics used in modeling earnings capacity to determine whether the economic scaling is appropriate or incomplete. Permitting interactions between expected wage and its determinants did not produce significant improvement in the likelihood of the takeup model and resulted in negligible changes in the coefficients on the expected wage level.

self-employment enterprises. Countable assets over $1500 ($3000 for households containing an aged or disabled person) preclude participation. Modeling on households whose countable assets are below this limit is appropriate, since conditioning variables can have no effect on participation by couples whose countable assets exceed the threshold; their participation rates are identically equal to zero.[12]

No administrative regulation links home equity to asset eligibility. Income in-kind from the home is not estimated and counted; the value of home equity and debt are not considered in the asset test. Our hypothesis is that participation rates should fall as the value of home equity rises. Two arguments suggest this hypothesis. If the transaction costs of obtaining a second mortgage or other credit are less than those associated with applying for Food Stamps, then couples who anticipate a spell of eligibility of short duration may prefer to take out loans rather than apply for FS benefits. The second argument relates to stigma and to past use of Food Stamps. We suspect that homeowners are less likely to live in communities where Food Stamp use is prevalent; thus information costs of applying for FS are greater than for non-owners. In addition, we hypothesize that the proportion of couples who have ever used FS falls as home equity increases also increasing the transaction cost of an application. Lastly, reluctance to use means-tested programs is also likely to be higher among established homeowners.

The effect of equity in a business is not as clear as the effect of home equity. Income derived from the business will be counted and can result in exceeding the threshold of income eligibility. This definitional connection to program participation implies that equity in the business should have a more attenuating effect on participation than does home equity. However, business equity does not establish the owner in middle class living patterns, while home ownership is likely to do so. On those grounds, business equity might not inhibit participation as much as does home equity. The direction of a net effect cannot be predicted. Furthermore, any measured effect will be less precise than the measured effect of home equity because fewer people own businesses than own homes.

The amount of countable assets is linked to countable income, just as the amount of business equity is linked to countable income. The extent of the likely relationship is easier to predict because yields on financial assets are less variable than yields from self-employment. If the expected yield on financial assets were 12 percent per year, 1 percent of the value of countable assets would be added to countable income during the month for which FS assistance is desired. At the threshold of eligibility, $1500, $15 would be added to countable income each month.[13] In other words, the

---

[12] Measurement errors in wealth imply that some couples are excluded because they overreport countable assets or because we were unable to document actual values of automobiles. Other couples should be excluded from this population because they underreport assets on the survey and should be excluded the asset-eligible population.

[13] All couples with countable assets above $1500 were generally excluded from the probit computation. An exception is that families containing an adult, other than the husband and wife, who is over fifty-nine years of age and families that contain a disabled person of any age are entitled to a $3000 asset threshold. Eligibility simulation included this exception to
the general rule.

definitional tie between holding assets and income eligibility is likely to be extremely small. For that reason we would not expect countable wealth to reduce participation significantly for definitional reasons. However, the transaction cost of applying for benefits and the anticipation of receipts that could cause assets to exceed the $1500 level could easily inhibit couples from applying. For example, assume that the couple has $1300 in countable assets. Assume also that the probability that either earner receives an automatic deposit of S400 within the month is 0.5. The expectation is that the couple will not meet the asset-eligibility test, and that the effort spent in applying will be wasted. We hypothesize that the anticipation of additional countable assets will inhibit participation more that home equity does.

Results of the model of participation are presented in Table 7. Table 8 shows the mean value of variables over the sample of 2699 couples who were asset-eligible under 1984 law. Model A (Table 7) was our conceptually preferred model.

Section V: Adjusting models of participation rates for response error

In table 7 we report estimates for the model of food stamp participation. The baseline model, with no correction for measurement error, is reported in column A. The four remaining models consider four possible estimates for the misclassification rates. Column B uses the simple marginal response error rates estimated from the full Wave one validation sample to correct the likelihood function for the measurement error. Column C uses the marginal response error rates for married couples in the Wave one Validation sample to affect the correction. Columns D and E use two different vector response error rates. Column D conditions only upon household per capita income, and corresponds to the vector response effort rates from models A4 and B1 of Table 6. Column E conditions the omission effort rate (q_i) on per capita income as well as gender, marital status, interview status and a gender-marital status interaction term and corresponds to the vector response error in columns A2 and B1 of Table 6.

As noted in the section above, the strongest and most stable correlate with response error is the per capita income. When this aspect is included in the participation model two very interesting results occur. First, the predicted wage variable of the higher earner (below the mean) increases in absolute magnitude (both over the uncorrected model as well as over the model corrected with marginal misclassification rates). Hence, corrected for misclassification, food stamp participation is more sensitive to changes in predicted wage. In addition to the change in the coefficient on predicted wage, the coefficients on countable assets and home equity also rise dramatically. The coefficient on countable assets more than doubles, while the coefficient on home equity increases by over 50%. This implies that food stamp participation is much more sensitive to counted assets, as well as other assets than uncorrected models would suggest. This has important policy implications concerning changes in the asset test.

The best perspective on the differences between the coefficients in the models is to assume that Model E represents the correct, unbiased estimate of the coefficient vector ($ß_f$). Then results of other models can be presented relative to model E (i.e. the ratio of $ß_{km}$ to $ß_{kE}$ where k indexes explanatory variables and m indexes models A, B, and D). These ratios are presented in Figure 1, for those conditioning variables that are highly significant in Model E. What stands out in the Figure is that the relative bias of response attributable to home equity and countable assets in Model A is large. The coefficient on countable assets in Model A is less than 0.4 of Model E.

A second finding is that Model B fails to eliminate much of the relative bias in the coefficients. Thus observation of expected vector error rates over a population can do little to correct the coefficients of a model that depends on heterogeneous characteristics of the population. In contrast, differences between Model D and Model E are less than the difference between Model B and Model E, except for the spline on the wage of the higher earner. This provocative, though by no means definitive, outcome suggests that a crude model of vector effort rates is an appropriate strategy for removing bias from typical estimators of behavioral models based on survey data.

The differential in relative bias for different characteristics is a sobering finding. In conjunction with the finding that Model B does little to correct for the bias of survey -based estimators, we conclude that a far larger proportion of resources used in collecting sample survey data must go towards validation and modeling vector error rates. Sampling errors are small as compared to the bias arising from reporting errors that are not exceedingly large. Yet we conclude that the response to countable assets more than doubles from estimates based on survey data alone.

Finally, we note that embedding validation sample in the survey design from SIPP was essential to understanding sources of bias. Had the primary data been collected using a different questionnaire and design than SIPP it would have been difficult to locate the source of error in the screening question and to identify the fact that some respondents report with high accuracy, while others do not.

Table 1
Comparison of Key Characteristics in Samples Used
1984 SIPP Panel'[a]

	Wave 1[b] SIPP Validation	Full Sample	Wave 4[c] Asset Eligible[d]
Sample Size (Number of Households)	2685	8169	2644
Pct. Claiming FS	6.2	3.3	9.6
Income (Monthly)	$ 2214 (3390)	$2890 (2020)	$1910 (1350)
Pct. Married	61%	100%	100%
Age of H.H head	48.9 (18)	40.0 (10)	37.3 (11)
Pct female head	28	NA	NA
Pct. black head	7.4	9.6	15.9

[a] Only persons aged 18 to 59 are included in the samples of households studied. Younger persons were not asked FS questions.

[b] The validation sample includes both married and single persons. Only persons residing in Florida, Pennsylvania, and Wisconsin PSU's are included. The sample is therefore not representative of the US population.

[c] The behavioral model for FS participation is estimated for couples of working age, i.e., households whose head is married with spouse present and where both spouses are aged 18 to 59. This sample represents the corresponding US population.

[d] Countable assets less than $1500 ($3000 in households containing persons over 59 years of age or disabled persons).

Table 2
Errors in Reporting Food Stamp Recipiency
over the four-month reference period for Wave 1
(1984 SIPP Validation Sample)

	SIPP Claims			Administrative Reports
Validation Sample (households)	2685			2685
No Receipt of FS	2469			2476
False positive claims		7		
FS received in reference Period	180			209
False negative claims		29		
False claims, total		36		
Consistent positive claims	180			
Duration Consistent		154		
Duration not Consistent		26		
Claimed Months > reported			12	
Claimed Months < reported			14	

Probability (defined over reference period)

False positive	7/2476 = 0.00283
False negative	29/209 = 0.139

TABLE 3
Errors in Reporting Food Stamp Recipiency
Month Prior to Interview Wave 1
(1984 SIPP Validation Sample - Florida, Pennsylvania, Wisconsin)

SIPP Questionnaire Family member certified to receive FS (screener)?	No Match	Administrative data Matched: Food Stamps received? No	 Yes	Total
No	2422 A	56 B	20 c	2498
Yes Received FS last month?				
No	3 D	15 E	2 F	20
Yes	3 G	5 H	159 I	167
Total	2428 J	76 K	181 L	2685

Type of error	Commission	Omission
rate	p = 8/2504 = 0.0032	q = 22/181 = 0.12
formula	p = (G + H)/(J + K)	q = (C + F)/L

Explanation of table cells

A Consistency assumed, possibly over-stated due to matching errors.
B Consistent.
c False negative; error in response to screening question (omission).
D Possible match failure; additional matched could reveal false negatives.
E Consistent.
F False negative; error in response to question about last month (omission).
G False positive assumed; possible match failure included.
H False positive; errors in response to question about last month (commission).
I Consistent.

Table 4
Errors in Wave Two
By Accuracy of Wave One Report
(1984 SIPP Validation Sample, Month Prior to Interview)

	Errors in Wave Two Report				
Accuracy of	Errors of Commission		Errors of Omission		Total
Wave One Report	Percent	Number	Percent	Number	Family Units
Correctly Reported No FS Participation	0.38%	9	0.17%	4	2394
Correctly Reported FS Participation	1.4%	2	5%	7	141
Commission	25%	2	0	0	8
Omission	0%	0	27.2%	6	22

** Subsample too small for statistical inference to be drawn.

Table 5
Comparison of Individual Level Validation
with Household Level Validation
(1984 SIPP Validation Sample, Month Prior to Interview)

	Errors of Commission		Errors of Omission	
	Percent	Number	Percent	Number
Individual Level Data (Marquis and Moore)				
Wave One	.36%	18	22%	41
Wave Two	.44%	22	19%	31
Household Level Data (Bollinger and David) Full Sample				
Wave One	.32%	8	12%	22
Wave Two	.54%	13	11%	17
Married Couples				
Wave One	.08%	1	11%	5
Wave Two	0%	0	3%	1

Table 6
Estimates of Vector Error Rate (1984 SIPP Validation Sample)

	Model				
	Models of Omission			Model of Commission	
Variable	A1	A2	A3	A4	B1
Intercept	-0.269 (0.624)	0284 (0.674)	-0.432 (0.548)	-1.450 (0.180)	-1.875 (0.231)
Family Per Capita Income (in 000's)	1.124* (0.452)	1.119* (0.453)	1.136* (0.405)	0.934* (0.412)	*-1.653 (0.549)
Gender		-0.779 (0.353)	-0.357 (0.323)		
Marital Status		-0.717 (0.494)	-0.559 (0.466)		
Gender* Marital		1.060 (0.629)	0.673 (0.595)		
Single Male	0.576 (0.408)				
Single Female	-0.205 (0.313)				
Interview Status	-1.101 (0.551)	-1.108 (0.557)	-0.656 (0.418)		
Log Likelihood	-59.86	-59.71	-70.43-	64.55	46.06
Sample Size	181	181	197	181	2504

Dependent Variable for Models A1-A4 is indicator for false negative; Dependent Variable for Model B1 is indicator for false positive. Models A1, A2, A4 and B1 are estimated using data from the last month before the interview for wave one of the 1984 SIPP panel. Model A3 uses pooled data from the last month before the interview in wave one and those individual who are new entrants to the food stamp program in wave two (using response for last month before the interview).

* indicates that coefficient is significant at the 5% level.

Table 7
Probit Models of Food Stamp Participation Models Corrected for Response Error
(1984 SIPP panel, Wave 4, Asset Eligible Sample)

	A	B	C	D	E
Intercept	1.341* (0.349)	1.764* (0.416)	1.615* (0.405)	1.984* (0.475)	1.804* (0.432)
Wage of High Earner, Below Mean	-0.142* (0.033)	-0.155* (0.039)	-0.142* (0.036)	-0.185* (0.046)	-0.181 (0.043)
Wage of High Earner, Above Mean	-0.075* (0.022)	-0.092* (0.028)	-0.083* (0.025)	-0.092* (0.031)	-0.077 (0.027)
Wage of Low Earner, Below Mean	-0.197* (0.046)	-0.218* (0.054)	-0.220* (0.051)	-0.210* (0.060)	-0.198 (0.056)
Wage of Low-Earner, Above Mean	0.204* (0.089)	-0.190- (0.107)	0.200-* (0.097)	0.157 (0.128)	-0.159 (0.122)
Poverty Threshold Level (000's)	0.421* (0.189)	0.375 (0.225)	0.349 (0.225)	0.408 (0.254)	0.375 (0.234)
Number of Kids Less than 18	0.085 (0.036)	0.128* (0.043)	0.113* (0.043)	0.134* (0.051)	0.126* (0.045)
Other Wealth, Below 75000 (0000's)	-0.049 (0.045)	-0.031- (0.053)	0.046 (0.049)	-0.018 (0.060)	-0.015 (0.058)
Other Wealth, Above 75000 (0000's)	0.004 (0.054)	-0.004 (0.067)	0.003 (0.057)	-0.011 (0.073)	-0.009 (0.069)
Home Equity, Below 75000 (0000's)	-0.114* (0.029)	-0.148* (0.040)	-0.127* (0.035)	-0.189* (0.051)	-0.180* (0.048)
Home Equity, Above 75000 (0000's)	0.003 (0.120)	0.045 (0.130)	0.016 (0.126)	0.118 (0.136)	0.113 (0.131)
Counted Assets (in 000's)	-0.908* (0.133)	-1.971* (0.340)	-1.169* (0.318)	-2.572* (0.646)	-2.519 (0.621)

Head Disabled	1.319*	1.481*	1.432*	1.572*	1.450*
	(0.146)	(0.189)	(0.175)	(0.214)	(0.188)
Live in SMSA	0.103	0.075	0.086	0.065	0.064
	(0.089)	(0.105)	(0.098)	(0.121)	(0.114)
Log Likelihood	-572.89	-569.888	-574.633	-528.899	-534.797

* indicates that coefficient is significant at the 5% level.

Table 8
Descriptive Statistics for Variables
Used in Estimating Food Stamp Participation Models

Variable	Mean	Std. Dev.	Minimum	Maximum
Wage of High Earner, Below Mean	12.09	1.59	4.94	13.37
Wage of High Earner, Above Mean	1.27	2.92	0	36.63
Wage of Low Earner, Below Mean	6.76	1.04	0.17	7.39
Wage of Low Earner, Above Mean	0.62	0.89	0	5.55
Poverty Threshold	871.	250.	452	1936
Number of Kids Less than 18	1.44	1.32	0	12
Other Wealth, Below 75000	4920	16188	-103398	75000
Other Wealth, Above 75000	2036	18539	0	480853
Home Equity, Below 75000	16503	22258	-15456	75000
Home Equity, Above 75000	1240	8533	0	121000
Counted Assets	461	680	-9	16897
Live in SMSA	0.44	0.49	0	1
Head Disabled	0.13	0.33	0	1

Notes To Accompany Table 7

All Models were estimated on the Asset Eligible Sample (n = 2699).
A: Model A is the uncorrected Probit Model of Food Stamp Program Participation
B: Model B uses the marginal reporting effort rates: p = 0.0032 (over-reports), q = 0.1215 (under-reports)
C: Model C uses the marginal reporting error rates for married couples: p = 0.0008, q = 0.1111
D: Model D uses the vector reporting effort rates, conditional on only family per Capita Income (Model A4 in table 6 for under-reports, Model B1 in table 6 for over-reports)
E: Model E uses the vector reporting error rates: Models A2 and B1 from table 6.

AVAILABLE UPON REQUEST IN HARD COPY FORM OR DISK FORM.

Figure 1: RATIOS OF COEFFICIENT ESTIMATES TO MODEL E

The coefficients considered in Figure 1 are only those which were significant at the 5% level for model E. We numbered them in the same order as listed in tables 6 and 7 above. Hence:

1 = Constant
2 = Wage of High Earner, Below Mean
3 = Wage of High Earner, Above Mean
4 = Wage of Low Earner, Below Mean
7 = Number of Kids Less than 18
10 = Home Equity, Below $75000
12 = Counted Assets
13 = Head Disabled

The first bar (labeled Model A) represents the ratio of the particular coefficient estimated in Model A to the coefficient estimate from Model E. The second bar (labeled Model B) is the ratio of the estimate from Model B to the estimate from Model E, and the third bar is the ratio of the estimate from Model D to the estimate from Model E. As noted in the text, this can be interpreted as the relative bias attributable to response error.

References

Aborn, Murray (ed.) 1991. "Special Issue on Cognitive and Survey Measurement" Applied Cognitive Psychology 5(3): 1734.

David, Martin H. and Maurice MacDonald, 1992. "Inflation, Asset Testing, and Food Stamp eligibility: Better Tests Without Increased Program Cost?" Madison, WI: Institute for Research on Poverty Discussion Paper 971-92.

Kulka, Richard, 1984. Documentation of the Income Survey Development Program Research Triangle Park NC: Research Triangle Institute.

Marquis, Kent and Jeffrey Moore, 1990. "Measurement Errors in SIPP Program Reports." SIPP working paper 9008, Washington DC: Us Bureau of the Census/ Center for Survey Methods Research.

Marfiffl, Alberto, 1992. "Participation in the Food Stamp Program: A Multivariate Analysis" Current Perspectives on Food Stamp Participation. Alexandria VA: USDA/Food and Nutrition Service/Office of Analysis and Evaluation.

Scheuren, Fritz and Lock Oh, 1980. "Fiddling Around with Nonmatches and Mismatches" Measuring the Impact on Family and Income Statistics of Reporting Differences Between The Current Population Survey and Administrative Records: Report #11. Washington DC: US. Department of Health, Education and Welfare, Social Security Administration, Office of Policy Research and Office of Research and Statistics. SSA publication number 13-11750.

Sudman, Seymour and Norman Bradburn, 1974. Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine.

Appendix

This appendix contains a brief a discussion of the details of modeling and estimation for the food stamp participation model, controlling for the reporting errors. The basic Food Stamp Participation model, described in equations (5) and (6) in the paper is a standard probit model. This implies that

(9) $\quad Pr[FStrue_i = 1 \mid X_i] = F(X_i\ \beta_f)$.

The function F() is the standard normal cumulative density function. If the true Food Stamp participation status were known (i.e. if $FStrue_i$ is observed), standard maximum likelihood estimation would consistently estimate the unknown parameter vector β_f. However, in this case only the reported food stamp status $FSrep_i$ is observed. There may be errors in the reporting of food stamp status.

To model the reporting errors in the food stamp variable, we make the following set of assumptions:

(A1) $\quad Pr[FStrue_i = 1 \mid X_i\ Z_i = F(X_i\ \beta_f)$

(A2) $\quad Pr[FSrep_i = 0 \mid X_i\ Z_i\ FStrue_i = 1] = F(Z_i\ \beta_1)$

(A3) $\quad Pr[FSrep_i = 1 \mid X_i\ Z_i\ FStrue_i = 0] = F(Z_i\ \beta_0)$.

Assumption (A1) implies that the true food stamp participation model is a probit model and is dependent on the set of variables X_i. The variables Z_i are independent of $FStrue_i$ once we have conditioned on X_i. (This does not rule out the same variable appearing in both X_i and Z_i.) Assumption (A2) implies that the probability of errors of omission ($FSrep_i = 0$ when $FStrue_i = 1$) also follows a probit model and is dependent on the set of variables Z_i but independent of the set X_i after conditioning on Z_i. Assumption (A3) establishes a similar model for errors of commission.

The above assumptions (and the models described in the body of the paper) allow simple calculation of the probability of observing $FSrep_i = 1$ conditional upon the set of variables X_i (the food stamp participation correlates) and Z_i (the reporting error correlates):

(10) $\quad Pr[FSrep_i = 1 \mid X_i\ Z_i] = Pr[FSrep_i = 1 \mid FStrue_i = 1, X_i\ Z_i] * Pr[FStrue_i$

$= 1 \mid X_i\ Z_i]$ + Pr[FStruej = 1 1 X, Zj

+ Pr[FSrep, = 1 1 FStrue, = 0, Y,,, Zj ' Pr[FStruej = 0 1 Yv Zj.

The above equality follows from Bayes rule. Using assumptions (Al), (A2), and (A3), we can write:

F(Y, ff) + F(Z, (1 - F(X, Of))
(12) Prf%rep, = 1 1 Y,, Zj (1 - F(Z, 01) - F(Z, F(X, P,) + F(Z, O,)

In the paper, pi represents the conditional probability of errors of commission (represented in assumption (A3)) and ch represents the conditional probability of errors of ommission.

The first step in estimating Pf is to utilize the validation data to estimate Po and P,. (See Section III and Table 6 for the results of US estimation.) C)Once the parameters for the models of omission and commission have been estimated they are used to construct the log likelihood function for estimating Of.

(11) Prf%repi = 1 1 X, Zj (1 - F(Z,

This can be rewritten as:

(12) $Pr[FSrep_i = 1 \mid X_i, Z_i] = (1\text{-}F(Z_i\, ß_1) - F(Z_i\, ß_0)) * F(X_i\, ß_f) + F(Z_i\, ß_0)$

In the paper, p_i represents the conditional probability of errors of commission (represented in assumption (A3)) and q_i represents the conditional probability of errors of omission.

The first step in estimating $ß_f$ is to utilize the validation data to estimate $ß_0$ and $ß_1$. (See Section III and Table 6 for the results of this estimation.) Once the parameters for the models of omission and commission have been estimated they are used to construct the log likelihood function for estimating $ß_f$.

That log likelihood function is:

(13) $L_i - FSrep_i * \log(Pr[FSrep_i = 1 \mid X_i, Z_i]) + (1 - FSrep_i) * \log(1 - Pr[FSrep_i = 1 \mid X_i, Z_i])$

Equation (12) then gives the expression for the probabilities on the right had side. The parameter $ß_f$ can be estimated using any standard maximum likelihood procedure. We used the prepackaged Gauss procedure.

However, the variance matrix of the estimates produced by prepackaged procedures is not correct. It fails to account for the sampling from the first stage estimates of $ß_0$ and $ß_1$. The prober form for the variance matrix can be found analytically and easily estimated. However, some additional assumptions are necessary:

> (A4) It is assumed that the samples used to estimate the reporting error models are independent of the sample used to estimate the food stamp model. We justify this obviously strong assumption by noting that the two samples were taken more than 1 year apart.
>
> (A5) Let N_0 and N_1 be the sample sizes utilized in estimating $ß_0$ and $ß_1$ respectively. Let N_f be the sample size utilized in estimating $ß_f$. Then Lim $N_f/N_0 = K_0$ and the Lim $N_f/N_1 = K_1$. (This assumption is also needed to prove consistency of $ß_f$.)

It should also be noted the samples used to estimate $ß_0$ and $ß_1$ are constructed to be independent of each other.

It is also convenient to define some standard notation. Let W be the expectation of the second derivative with respect to $ß_f$ twice of the Likelihood function (13) evaluated at the true parameters ($ß_0$, $ß_1$, and $ß_f$) (this is also known as the Fisher Information matrix). Let V_0 and V_1 be the asymptotic variance matrices of the estimates of $ß_0$ and $ß_1$ respectively. Let U_0 be the expected value of the second derivative with respect to $ß_f$ and then $ß_0$ of the likelihood function (13) evaluated at the true parameters. Let U_1 be defined similarly, relative to $ß_1$.

It can then be shown by utilizing standard linear approximation theorems (a.k.a. delta method) that the asymptotic variance of $ß_f$ is:

(14) $\text{Variance}(ß_f) = W^{-1} (W + K_0 * U_0 * V_0 * U_0' + K_1 * U_1 * V_1 * U_1') W^{-1}$

Note that if V_0 and V_1 are zero matrices (i.e. if $ß_0$ and $ß_1$ were known and not estimated) the above expression would simplify to the inverse of W which is the standard maximum likelihood result. The terms W and U_j are easily estimated (in fact, the pre-packaged procedure estimates W^{-1}, and reports it as the variance matrix for the estimate). Estimates for the terms V_0 and V_1 are reported along with the estimates of $ß_0$ and $ß_1$ from the first stage.