

Title: "Non response in rotating panel surveys: analysis on Argentina's labor force survey".

Authors: Claudio Comari* , Augusto E. Hozzowski**

Abstract

The Permanent Household Survey (“Encuesta Permanente de Hogares”, EPH in its Spanish acronym) is the official labor force survey of Argentina. The EPH’s sample is divided into four different rotating panels, like other surveys of its kind. The selected households remain for six quarters in the sample under the rotation scheme known as 2-2-2. That is: two quarters into the sample, two quarters out and again two quarters into the sample, so each rotating panel is included to be interviewed in four times.

In this paper we evaluate the effects of the use of rotating panels on the *non response* and the *unemployment rate*, using a set of data of several quarters of the EPH. In order to do so, the information was organized according to the number of times each group of households was included to be interviewed.

The observation of a higher *non response rate* in the panels included for the first time than the observed in those included several times was an unexpected result; as we show in this paper, the *non response* decreases for each panel throughout the successive participations into the sample.

Keywords: *Rotating panels, attrition, rotation group bias, labor force survey*

* Instituto Nacional de Estadística y Censos INDEC - Argentina.

** Instituto Nacional de Estadística y Censos INDEC - Argentina.

Introduction

This paper presents the main results obtained by INDEC (the National Institute of Statistics and Censuses) in the analysis of the EPH survey conducted to establish the impact of two of the traditional issues faced by the specialized literature about the use of rotating panel in probabilistic sample design: the issue of *rotation group bias on estimators* and the issue of *attrition of the sample* according to the increase of the number of waves or participations.

The subjects associated with the non response are relevant for INDEC because of its trend to increase; it is closely related to the decrease of the unemployment rate and the consequent reduction of the possibility to establish contact between the interviewers and the residents in the selected dwellings.

The first section of this article presents the general characteristics of the EPH sample design and some rotation schemes used in labor force surveys in several countries. In the second section, we analyze the effect of the rotation group bias on estimators and its impact on the EPH. The third part is about the issue of attrition, in which an unexpected result is found, given the international experience: in the EPH's rotating panels the non response decreases for each group along the successive participations into the sample.

1. Sample design of the Permanent Household Survey (EPH) in Argentina

EPH samples are based on a probabilistic sample of 'areas'. The design is similar to the one used for most of the labor force surveys:

- *Estimation Domains*: every 'urban agglomeration' with more than 100,000 inhabitants ('agglomerations'). *Independent* samples in each of these 'agglomerations'.
- *Sample design in each agglomeration*: two-stage and stratified design.
- *Primary sampling units (PSU)*: areas that in general constitute enumeration districts (geographical delimitations designed for the Population Census). On average, an enumeration district covers an area with 350 dwellings; it is a geographical delimitation designed for the census proceedings. In some cases, the areas are divisions of enumeration districts or groups of enumeration districts to guarantee a minimum number of dwellings inside the areas.
- *Number of areas selected in each agglomeration*: the number varies between 60 and 120 areas, except for Greater Buenos Aires (~12,692,000 inhabitants) where approximately 520 PSU are selected.
- *PSU selection method*: stratified sampling, *pps* inside each stratum.
- *Secondary sampling units (SSU)*: private dwellings.

- *Number of dwellings selected in each SSU*: it varies between 8 and 12 dwellings.
- *Dwelling selection method inside each selected PSU¹ until 2013*: selection based on systematic sampling of dwellings. The list of dwellings is classified geographically before the selection takes place.
- *Dwelling selection method inside each selected PSU since 2013 and subsequent*: selection based on systematic sampling of two compact segments of five dwellings each one. The list of segments is classified geographically before the selection takes place.

As with other surveys of its kind, one of the most important characteristics in EPH's sampling since its redesign in 2003, has been the use of samples composed by four different rotating panels for each quarter. According to Binder & Hidirolou (1988)² "...This method of sampling is also referred to as 'sampling on successive occasions with partial replacement of units' (Yates, 1949; Patterson, 1950) and 'sampling for a time series' (Hansen, Hurwitz and Madow, 1953)".

The selected dwellings remain in the sample for one and half year under the rotation scheme known as 2-2-2. That is: two quarters into the sample, two quarters out and again two quarters, so each rotating panel is included in the sample to be interviewed in four times. With the 2-2-2 scheme, each quarter shares 50% of the dwellings with the following quarter and in this way with the same two quarters in the following year.

There are several schemes to do the sequence of interviews: on a quarterly basis, Brazil and Mexico use the scheme named 1-2-(5), where each dwelling is interviewed successively by 5 quarters and then they are definitely excluded from the sample; the United Kingdoms uses the scheme 1-(2)-5 by mobile quarters; in Chile, France and Spain the 1-2-(6) scheme is used, so each dwelling is interviewed by 6 successive quarters before being definitely excluded from the sample. There are other schemes used on a monthly basis like the ones used by Canada, the United States of America and Australia.

The optimal design of the sample depends, in general, on the parameter of interest³ and therefore the use of rotating panels must be in accordance with the research goals, e.g. : "(...) if one is primarily interested in the total consumption over a number of periods, it is less likely that reinterviews are attractive than if one is primarily interested in period-to-period changes in consumption"⁴.

¹ Comari, C., Hoszowski A. (2013). Changes in the selection of dwellings in the labor force survey of Argentina: A Simulation. *Joint Statistical Meetings 2013 Proceedings*. Survey Research Methods Section, p.p. 2626 - 2636.

² Binder D., Hidirolou A. (1988): Sampling in Time. *Handbook of Statistics*. Vol. 6. P. R. Krishnaiah and C. R. Rao, eds. p.p 187.

³ Cochran G. (1977). *Sampling Techniques*. Wiley.

⁴ Nijman T. et al. (1991). The efficiency of rotating-panel designs. *Journal of Econometrics* 49 p.p. 374. North-Holland

The use of rotating panels is highly valued within the academic field because it allows the creation of *quasi-longitudinal* studies due to the possibility of matching records from microdata files belonging to different periods of observation.

From the point of view of national statistical offices, the use of rotating panel provides certain advantages:

- it stabilizes the estimates because in a same observation period it shares several common units with the previous period;
- it reduces the demand over the master sample without compromising the allowance of units for other surveys;
- it therefore reduces the costs associated with the setting-up and update of the master sample;
- it helps making the logistics work easier by the repetition of units;
- it allows the study of flows and transitions, at an individual or aggregate level;
- it allows to better screen elusive populations or events, among other advantages.

As it was already mentioned, the rotation schemes vary and each one involves its own advantages and disadvantages.

2. The rotation group bias issue

The effects of *rotation group bias* are largely discussed by Bailer (1975)⁵, Solon (1986)⁶, Waite, Huggins and Mack (1998)⁷ and Erkens (2012)⁸, among other authors.

For several surveys, the effects are so important that the statistical offices have special procedures for its correction. That is the case of the Current Population Survey of the United States in which, to reduce the effect of the rotation group bias,

⁵ Bailer, B.(1975). The Effects of Rotation Group Bias on Estimates from Panel Surveys. *Journal of the American Statistical Association*, Vol. 70, No. 349, pp. 23-30.

⁶ Solon, G.(1986). Effects of Rotation Group Bias on Estimation of Unemployment. *Journal of Business and Economic Statistics*, 4(1), p.p. 105-109.

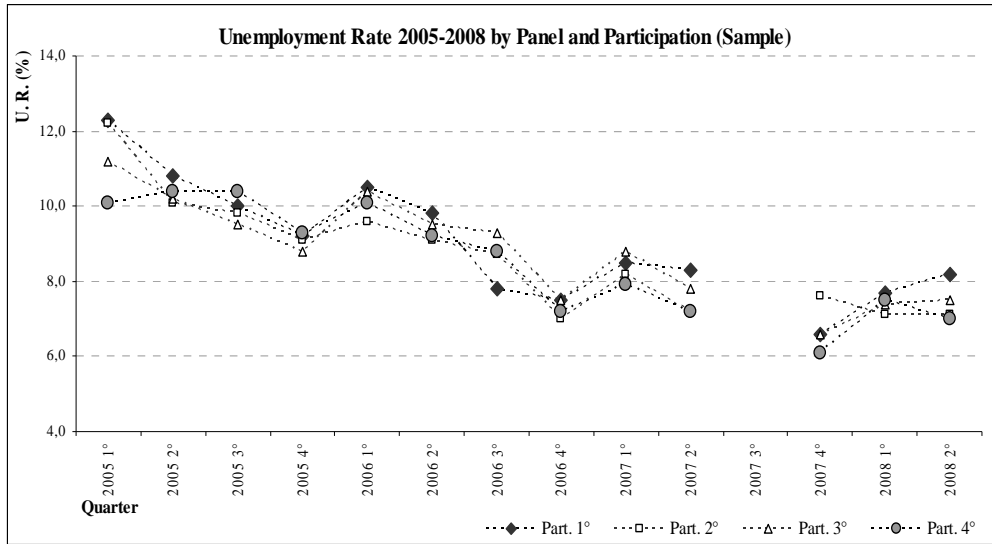
⁷ Waite, P., Huggins V. and Mack S.(1998). Assessment of Efforts to Reduce Nonresponse Bias: 1996 Survey of Income and Program Participation (SIPP). *Nonresponse in survey research. Proceedings of the Eighth International Workshop on Household Survey Nonresponse 24-26 September 1997*. Zentrum für Umfragen, Methoden und Analysen (ZUMA). Mannheim.

⁸ Erkens, G. (2012). Changes in Panel Bias in the U.S. Current Population Survey and its Effects on Labor Force Estimates. *Joint Statistical Meetings 2012 Proceedings*. Survey Research Methods Section, p.p. 4220 - 4232.

a special treatment for the calculation of the final estimations is used (see Bailar; 1975)⁹.

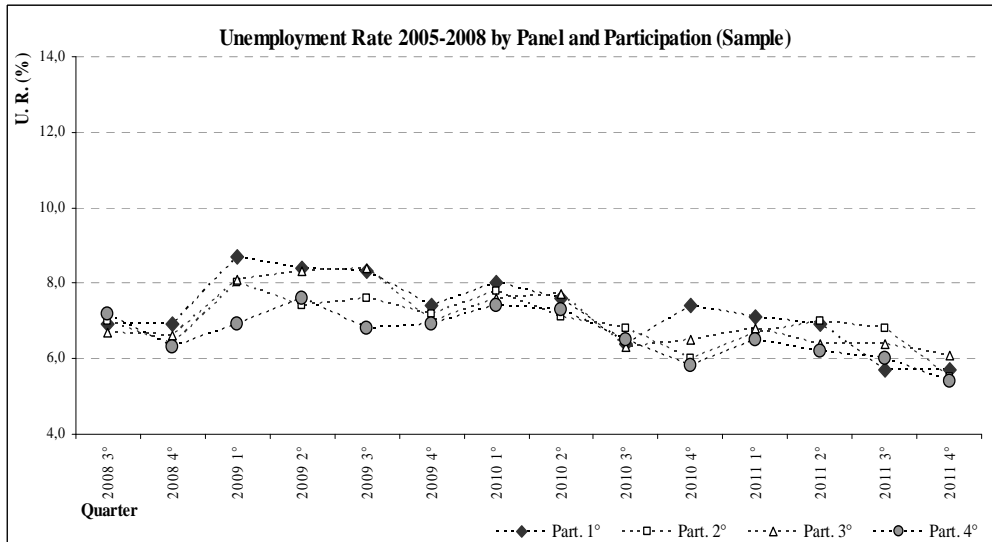
The *unemployment rate* is the most important estimator derived from the EPH, so it is important to know how much it is exposed to that *rotation group bias*, if it exists. The following graphics (1 and 2) present the sample values results obtained for each quarter between 2005 and 2011¹⁰ in accordance with the number of participations of the panel.

Graphic 1



Source: EPH – INDEC. First quarter 2005 until second quarter 2008.

Graphic 2



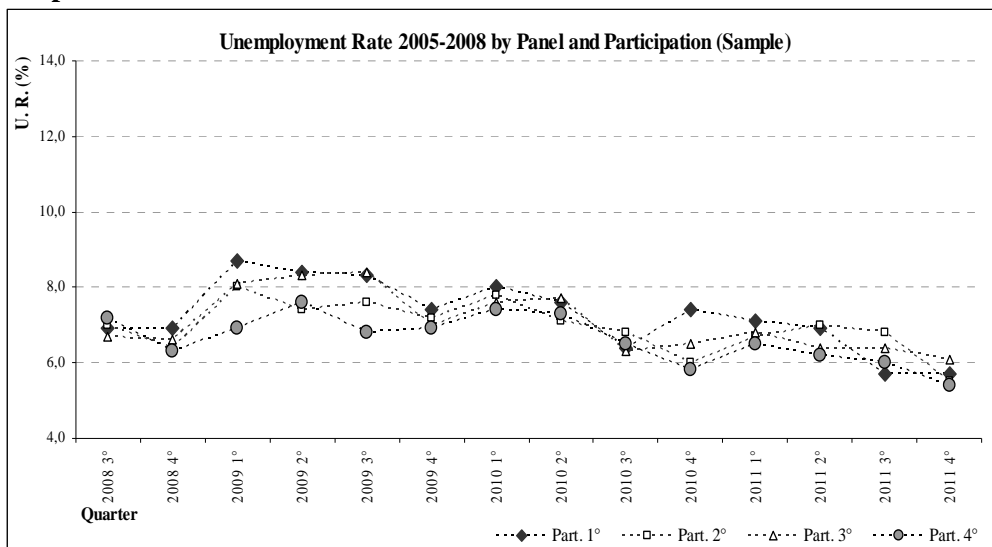
Source: EPH – INDEC. Third quarter 2008 until fourth quarter 2011.

In the following graphics we will see the results obtained in relation to the *unemployment rate* after weighting by the design factor (graphics 3 and 4).

⁹ Bailar, B. (1975). Op. cit., p.p. 25.

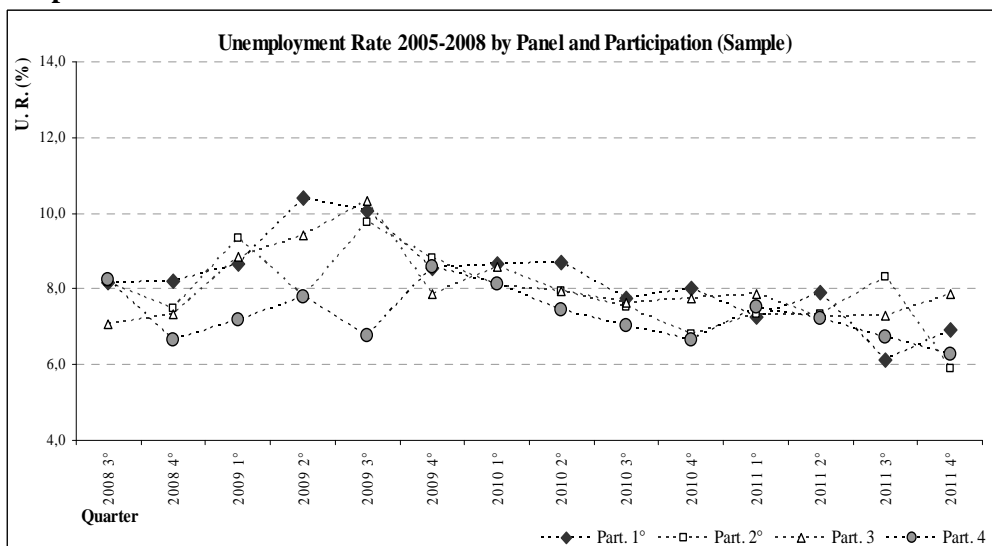
¹⁰ The 2007 third quarter is not included because the collection was not completed due to a strike of interviewers in several *agglomerations*.

Graphic 3



Source: EPH – INDEC. First quarter 2005 until second quarter 2008.

Graphic 4



Source: EPH – INDEC. Third quarter 2008 until fourth quarter 2011.

It is not possible to find the presence of the *rotation group bias* in these graphics. To improve this analysis, a contingency table between participation and number of rank of unemployment rates observed in each quarter is created.

Being the rotation group with lower unemployment rate classified as *Rank 1* and the group with higher rate as *Rank 4*, the table of the rotation groups according to the number of participation and rank of unemployment rate observed is the following:

Table 1: Distribution of unemployment rates according to participation and rank. Percentages by row.

Participation	Rank			
	1 %	2 %	3 %	4 %
1	14.29	21.43	25.00	39.29
2	21.43	32.14	28.57	17.86
3	25.00	21.43	25.00	28.57

4	39.29	25.00	21.43	14.29
---	-------	-------	-------	-------

This table shows a strong association between rank and number of participation. A higher rate is more frequently observed in the first participation and a lower rate is often observed in the 4th time of the interview.

That is also confirmed by the evaluation of the totality of the rates through the Skillings-Mack's test, which gave a result of 5.357 with P-value = 0.1474, a significant value for this type of evaluation. Due to this reason, it is important to evaluate the relevance of its impact over the final unemployment rate, measuring for each rotating group how much it can overestimate in the first participation respect to the mean of the other three groups. The statistics of this factor are:

	P25	P50	P75
Increase	-.1076	1.507	3.334

The median increase is 1.51%, value that can be considered minimum.

In summary: there are more probabilities to obtain a higher unemployment rate by the groups that participate for the first time and a higher probability to obtain the lowest rate when the group is observed by the 4th time, although these trends are not systematic. It also could be said that the impact over the rates is extremely weak.

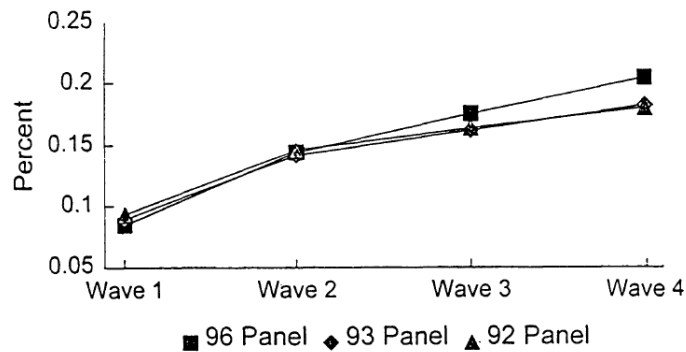
The result showed allows us to think that the EPH is almost free of *rotation group bias*. The explanation may lay in the differences in the sample performance found in the comparison between its kinds of surveys.

3. The problem of the attrition in the rotation panels

Another important issue in the studies of sampling by rotation panels is the loss of observations along the periods.

For example, in a study about SIPP, the following graphic about the growth of the non response is presented:

Figure 2: Household sample loss rates, Wave 1, 2, 3 and 4



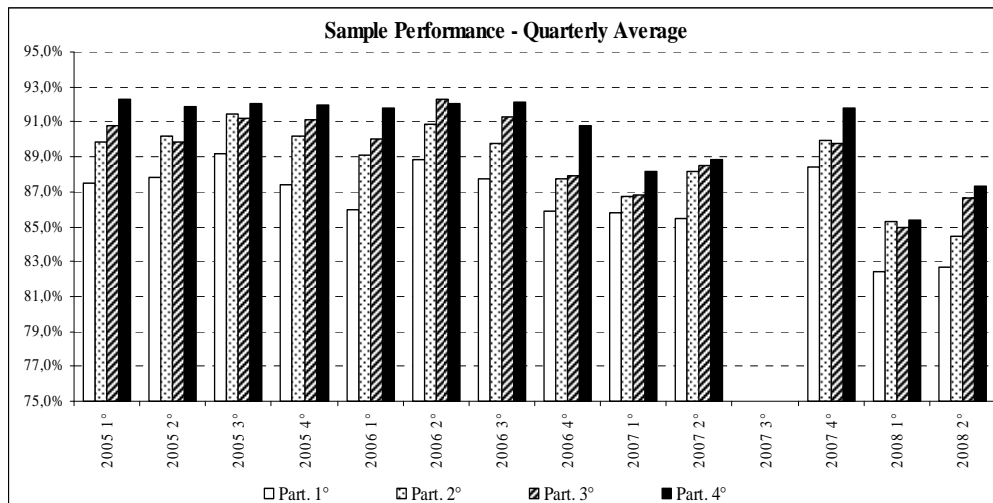
Source: Waite,P., Huggins V. and Mack S.(1998). *Assessment of Efforts to Reduce Nonresponse Bias: 1996 Survey of Income and Program Participation (SIPP)*. Nonresponse in survey research. Proceedings of the Eighth International Workshop

on Household Survey Nonresponse 24-26 September 1997. Zentrum für Umfragen, Methoden und Analysen (ZUMA), p.p. 26.

The non response increases within the group as the number of participations rises. This phenomenon is normally expected, but in the case of EPH this is not true.

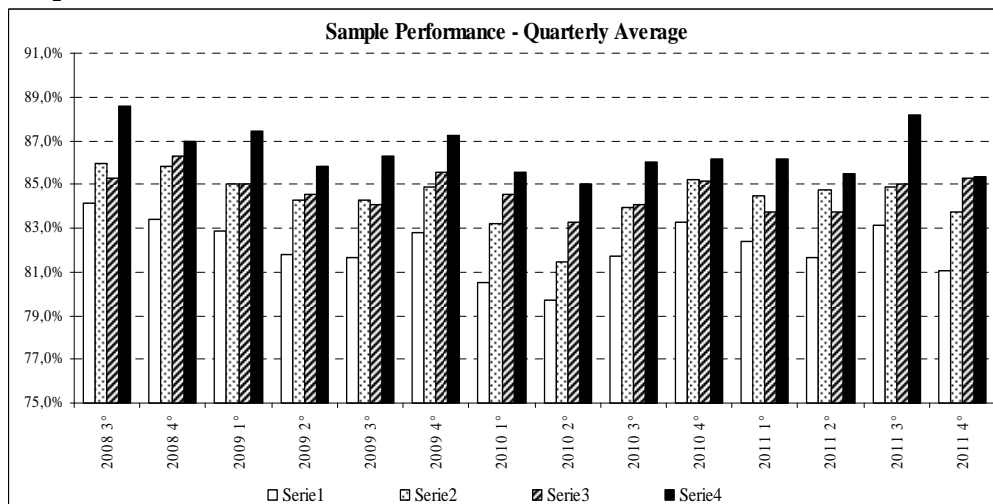
In the following graphics we show the evolution of the percentage of responses obtained from 2005 until 2011, according to the number of participations of each panel expressed in quarterly averages.

Graphic 5



Source: EPH – INDEC. First quarter 2005 until second quarter 2008.

Graphic 6



Source: EPH – INDEC. Third quarter 2008 until fourth quarter 2011.

In the second graphic, the following effects may be identified: the decrease in the general response mentioned in the introduction and the increase of the percentage of the response according to the number of times that the panel is included in the sample.

This last observation can be the explanation of the weak incidence of the *rotation group bias*; while in those studies that show *rotation group bias* the rotating panels

are subjects to losses (attrition), in the EPH the rotating panels obtain a better performance along its participations.

The final reasons remain unknown; although our hypotheses towards cultural factors may be valid. Perhaps, over time, the interviewers have obtained useful additional information about the households that allows a successful interview there where before they failed. Another hypothesis may be that the households that are more elusive in the earlier turns change their willingness and become cooperatives. We don't have enough evidence for a satisfactory explanation, only the certainty that the trends are different among the surveys.

Conclusions

The results obtained in the examination of EPH's microdata files belonging to 2005 until 2011 show us that:

- There is a relation between the number of participations and the level of unemployment rate;
- There is an extremely weak incidence of *rotation group bias* over the main estimators obtained from EPH;
- The sample is not mainly affected by attrition but instead shows a better performance as the number of participations increases;
- The weak incidence of *bias* is possibly due to the trend of a better performance according to the increase in the number of participations.

Bibliography

Bailar B.(1975). The Effects of Rotation Group Bias on Estimates from Panel Surveys. *Journal of the American Statistical Association* , Vol. 70, No. 349. p.p. 23-30.

Binder D., Hidiroglou A. (1988): Sampling in Time. *Handbook of Statistics*. Vol. 6. P. R. Krishnaiah and C. R. Rao, eds. p.p. 187-211.

Cochran G. (1977). *Sampling Techniques*. Wiley.

Comari C., Hoszowski A. (2013). Changes in the selection of dwellings in the labor force survey of Argentina: A Simulation. *Joint Statistical Meetings 2013 Proceedings*. Survey Research Methods Section. p.p. 2626-2636.

Erkens G. (2012). Changes in Panel Bias in the U.S. Current Population Survey and its Effects on Labor Force Estimates. *Joint Statistical Meetings 2012 Proceedings*. Survey Research Methods Section. p.p. 4220-4232.

Nijman T. et al. (1991). The efficiency of rotating-panel designs. *Journal of Econometrics*. North-Holland. 49. p.p. 373-399.

Solon G.(1986). Effects of Rotation Group Bias on Estimation of Unemployment. *Journal of Business and Economic Statistics*, 4(1). p.p. 105-109.

Waite P., Huggins V. and Mack S.(1998). *Assessment of Efforts to Reduce Nonresponse Bias: 1996 Survey of Income and Program Participation (SIPP)*. Nonresponse in survey research. Proceedings of the Eighth International Workshop on Household Survey Nonresponse 24-26 September 1997. Zentrum für Umfragen, Methoden und Analysen (ZUMA). Mannheim. p.p. 23-44.