Centre for International Studies on Economic Growth

TOR VERGATA

CEIS Tor Vergata
RESEARCH PAPER SERIES

Working Paper No. 99 March 2007

A sample selection model for unit and item nonresponse in cross-sectional surveys

Giuseppe De Luca Franco Peracchi

CEIS Tor Vergata - Research Paper Series, Vol.33 , No. 99 , March 2007

A sample selection model for unit and item nonresponse in cross-sectional surveys*

Giuseppe De Luca and Franco Peracchi
University of Rome "Tor Vergata"

This version: November 2006

Abstract

We consider a general sample selection model where unit and item nonresponse simultaneously affect a regression relationship of interest, and both types of nonresponse are potentially correlated. We estimate both parametric and semiparametric specifications of the model. The parametric specification assumes that the errors in the latent regression equations follow a trivariate Gaussian distribution. The semiparametric specification avoids distributional assumptions about the underlying regression errors. In our empirical application, we estimate Engel curves for consumption expenditure using data from the first wave of SHARE (Survey on Health, Aging and Retirement in Europe).

Keywords: Unit nonresponse, item nonresponse, cross-sectional surveys, sample selection models, Engel curves.

JEL classification: C14, C31, C34, D12

*We thank Chuck Manski, Frank Vella and seminar participants at Northwestern and UCL for helpful comments. Financial support for this research was provided through the European Community's Program 'Quality of Life' under the contract No. QLK6-CT-2002-002426 (AMANDA). This paper is based on data from the early Release 1 of SHARE 2004, which is preliminary and may contain errors that will be corrected in later releases. The SHARE data collection has been primarily funded by the European Commission through the 5th framework program (project QLK6-CT-2001-00360 in the thematic program 'Quality of Life'). Additional funding came from the US National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064). Data collection in Austria (through the Austrian Science Fund, FWF) and Switzerland (through BBW/OFES/UFES) was nationally funded.

1 Introduction

Nonresponse is a very important source of nonsampling errors in sample surveys. A distinction is usually made between two forms of nonresponse: unit and item nonresponse. Unit nonresponse occurs when eligible sample units fail to participate to a survey because of failure to establish a contact, or explicit refusal to cooperate. Item nonresponse occurs instead when responding units do not provide useful answers to particular items of the questionnaire.

The relevance of distinguishing between unit and item nonresponse is twofold. First, data users can improve model specification, because different information is usually available for studying the two types of nonresponse. In fact, the information available to study unit nonresponse is usually confined to the information obtained from the sampling frame or the data collection process, whereas the additional information collected during the interview can be used to study item nonresponse. Second, understanding the different types of error generated by unit and item nonresponse plays a key rule at the survey design stage, where resources have to be allocated efficiently to reduce nonresponse errors. For instance, improving incentive schemes and follow-up procedures can help reduce unit nonresponse, while reducing the complexity of the questionnaire can help reduce item nonresponse.

For panel surveys, one can also distinguish a particular form of unit nonresponse, namely sample attrition. This occurs when a responding unit in one wave of the panel drops out in a subsequent wave. In this paper, we are mainly concerned with problems of nonresponse in cross-sectional surveys or, equivalently, in the first wave of panel surveys. Response rates in the first wave of a panel are typically much lower than in subsequent waves. For example, the overall household nonresponse rate in the first wave of the European Community Household Panel, a large longitudinal survey of the European population, is about 30 percent, whereas the overall household attrition rate in the next two waves is about 10 percent (Eurostat 1997). Despite its importance, however, response rate in the first wave has received little attention relative to panel attrition, largely because of the lack of information on unit nonrespondents.

One crucial issue in studying both unit and item nonresponse is establishing whether or not the mechanism generating missing observations is random. Following Rubin (1976), we distinguish between three missing data mechanisms: missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR). A mechanism is MCAR if missingness does not depend on the values of the variables in the data matrix. A mechanism is MAR if, after conditioning on a set of observed covariates, there is no relation between missingness and the observed outcomes. A mechanism is NMAR if missingness and the observed outcomes are related even after conditioning on the set of observed covariates. When mechanisms underlying (unit or item) nonresponse are NMAR, ignoring nonresponse errors or relying on the MAR assumption may lead to invalid inference about population parameters of interest.

An important strategy in order to reduce nonresponse errors consists of planning preventive measures to cope with nonresponse at the survey design stage. Well-designed surveys aim to reduce unit nonresponse rates by choosing the most appropriate fieldwork period, interview mode, interviewer training, follow-up procedures and incentive schemes. Other aspects of the questionnaire design, like length of the interview, wording of the questions and their reference period, are more likely to affect item nonresponse rates. Empirical studies by Groves and Couper (1998), Groves *et al.* (2002), O'Muircheartaigh and Campanelli (1999) and Riphahn and Serfling (2002), show that all these aspects of survey design are crucial to explain the response rates achieved in sample surveys.

Unfortunately, despite the preventive measures adopted for minimizing nonresponse errors, response rates are rarely close to 100 percent. This explains why most of the survey nonresponse literature focuses on the development of statistical methods for ex-post adjustments of nonresponse errors (see Lessler and Kalsbeek 1992, and Little and Rubin 2002). Weighting adjustment methods, which involve the assignment of weights to sample respondents in order to compensate for their systematic differences relative to nonrespondents, have been traditionally used to deal with problems of unit nonresponse, whereas imputation procedures, which aim to fill in missing values to produce a complete dataset, have been traditionally used to deal with problems of item nonresponse. Although ex-post adjustment techniques have reached a high level of sophistication, such methods commonly assume that the missing data mechanism is MAR, and they do not generally allow compensating simultaneously for errors due to unit and item nonresponse.

This paper differs from previous studies in two respects. First, problems of selectivity due to unit and item nonresponse are analyzed jointly. Second, missing data mechanisms underlying the different types of nonresponse are allowed to be NMAR. In particular, we analyze a general sample selection model where unit and item nonresponse can jointly affect a regression relationship of interest, and the two types of nonresponse can be correlated. Attention focuses on two alternative specifications of the model, one parametric and the other semiparametric. In the parametric specification, errors in the two selection equations (one for unit and one for item nonresponse) and in the equation for the outcome of interest are assumed to follow a trivariate Gaussian distribution.

In the semiparametric specification, we avoid distributional assumptions about the errors in the three equations. After discussing issues related to identification and estimation of the two kinds of model, we provide an empirical application by using data from the first wave of SHARE (Survey on Health, Aging and Retirement in Europe), a survey conducted in 2004 across eleven European countries. The aim of this analysis is to investigate the potential selectivity associated with unit and item nonresponse in the estimation of Engel curves for food consumption at home and total nondurable consumption.

The remainder of the paper is organized as follows. Section 2 formalizes the motivation of this study, and presents a general framework to analyze problems of unit and item nonresponse. Sections 2.1 and 2.2 consider problems of identification and estimation of the parametric and semiparametric model respectively. Section 3 presents our data. Section 4 presents our empirical results. Finally, Section 5 summarizes our main findings and offers some conclusion.

2 The statistical model

In what follows, we are interested in estimating the conditional mean function of a random outcome by using data from a survey. Initially, a set of n units is drawn at random from the population of interest. Nonresponse may then select the sample at two stages. First, unit nonresponse may reduce the sample size to $n_1 < n$ responding units. Second, item nonresponse may further reduce the number of usable observations to $n_2 < n_1$. This loss of observations causes an efficiency loss relative to the ideal situation of complete response. This efficiency loss needs not be the main concern, however, because lack of independence between the missing data mechanism and the outcome of interest may also generate selectivity in the observed sample and may lead to biased estimates of the population parameters.

To formalize the statistical problem, we consider a sequential framework where individuals first decide whether to participate to the survey. Given participation, they then decide whether to answer a specific item of the questionnaire. Thus, the indicator of unit response, Y_1, is always observed, the indicator of item response, Y_2, is only observed for the units that agree to participate to the survey, and the response process is completely described by two elements: the probability of unit nonresponse, $\pi_0 = \Pr\{Y_1 = 0\}$, and the probability of item nonresponse conditional on unit response, $\pi_{0|1} = \Pr\{Y_2 = 0 \,|\, Y_1 = 1\}$. Our objective is to obtain consistent estimates of the mean function of the outcome of interest Y_3 (conditional on covariates) allowing for selectivity generated by unit and item nonresponse.

By the law of iterated expectations[1]

$$\begin{aligned} \mathrm{E}(Y_3) = \pi_0\, \mathrm{E}(Y_3 \,|\, Y_1 = 0) + (1-\pi_0)[\pi_{0|1}\, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0) + \\ + (1-\pi_{0|1})\, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)]. \end{aligned} \tag{1}$$

The sampling process identifies π_0, $\pi_{0|1}$ and $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)$ but not $\mathrm{E}(Y_3 \,|\, Y_1 = 0)$ and $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0)$. Hence, in the absence of additional information, $\mathrm{E}(Y_3)$ is not identifiable.

The assumption that both unit and item nonresponse are MAR is convenient because it implies that $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1) = \mathrm{E}(Y_3)$, which enables one to directly exploit the information contained in the subsample of fully responding units. If this assumption does not hold and estimates of $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)$ are used to estimate $\mathrm{E}(Y_3)$, then the overall nonresponse bias is

$$\begin{aligned} \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1) - \mathrm{E}(Y_3) = \pi_0[\mathrm{E}(Y_3 \,|\, Y_1 = 1) - \mathrm{E}(Y_3 \,|\, Y_1 = 0)] + \\ + \pi_{0|1}[\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1) - \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0)]. \end{aligned} \tag{2}$$

Thus, the overall nonresponse bias depends on two separate components, respectively proportional to the probability of unit nonresponse and the probability of item nonresponse conditional on unit response. The overall nonresponse bias is zero when $\pi_0 = \pi_{0|1} = 0$ (neither unit nor item nonresponse), or $\mathrm{E}(Y_3 \,|\, Y_1 = 1) = \mathrm{E}(Y_3 \,|\, Y_1 = 0)$ and $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1) = \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0)$ (both unit and item nonresponse are MAR). Equation (2) makes it clear that there is also a third case in which the overall nonresponse bias is zero, namely when the bias components due to unit and item nonresponse have opposite sign and offset each other.

A simple way of allowing for differential selectivity effects of unit and item nonresponse is to adopt the following straightforward generalization of the classical sample selection model of Heckman (1979)

$$Y_j^* = \mu_j + U_j, \qquad j = 1,2,3, \tag{3}$$

$$Y_1 = 1\{Y_1^* \geq 0\}, \tag{4}$$

$$Y_2 = 1\{Y_2^* \geq 0\}, \qquad \text{if } Y_1 = 1, \tag{5}$$

$$Y_3 = Y_3^*, \qquad \text{if } Y_1 Y_2 = 1, \tag{6}$$

where the Y_j^*, $j = 1,2,3$, are latent continuous random variables representing, respectively, the propensity to participate to the survey, the propensity to answer to the item of interest, and the outcome variable in the uncensored sample, and the U_j are regression errors with zero mean. The μ_j are assumed to depend linearly on a k_j-vector of fully observable exogenous variables X_j, that is,

[1] In the following, explicit conditioning on covariates is suppressed to simplify notation.

$\mu_j = \alpha_j + \beta_j^\top X_j$, $j = 1, 2, 3$, where α_j and β_j are unknown parameters to be estimated. The latent variables Y_j^* are related to their observed counterparts Y_j through the observation rules (4)–(6), where $1\{A\}$ is the indicator function of the event A.

The primary interest of the analysis is estimation of the parameters in μ_3 from the sub-sample of fully observed units, for which

$$\mathrm{E}(Y_3^* \,|\, Y_1 Y_2 = 1) = \mu_3 + \sigma_3 \, \mathrm{E}(U_3 \,|\, U_1 > -\mu_1, U_2 > -\mu_2). \tag{7}$$

If any of the two nonresponse mechanisms is NMAR, then the conditional expectation on the right hand side of (7) is different from zero, and traditional estimation methods, such as ordinary least squares, lead to inconsistent estimates of the parameters of interest. Consistent estimation can be based on simple generalizations of the classical Heckman two-step procedure. Parametric and semiparametric versions of this procedure are presented in Sections 2.1 and 2.2 respectively. In both cases, point identification of $\mathrm{E}(Y_3)$ is achieved by restricting the shape of the joint distribution of (Y_1, Y_2, Y_3).

Although unavoidable if one seeks point identification, assumptions about the shape of the joint distribution of (Y_1, Y_2, Y_3) may not be entirely convincing, especially if they cannot be tested. When the lack of credible assumptions about the response process prevents point identification of $\mathrm{E}(Y_3)$, partial identification in the sense of Horowitz and Manski (1998) and Manski (2003) would still be possible if one could bound $\mathrm{E}(Y_3 \,|\, Y_1 = 0)$ and $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0)$ in (1). Specifically, suppose that $\mathrm{E}(Y_3 \,|\, Y_1 = 0)$ can take any value in the interval $[a_0,\, a_1]$, whereas $\mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 0)$ can take any value in the interval $[b_0,\, b_1]$. Then, $\mathrm{E}(Y_3)$ must necessarily lie in the interval $H = [c_0, c_1]$, where

$$c_0 = \pi_0 a_0 + (1 - \pi_0)[\pi_{0|1} b_0 + (1 - \pi_{0|1}) \, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)],$$

$$c_1 = \pi_0 a_1 + (1 - \pi_0)[\pi_{0|1} b_1 + (1 - \pi_{0|1}) \, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)],$$

whose width is equal to

$$c_1 - c_0 = \pi_0 (a_1 - a_0) + (1 - \pi_0) \pi_{0|1} (b_1 - b_0).$$

When Y_3 is non-negative, as in the empirical example in Section 4, it is natural to choose $a_0 = b_0 = 0$. The resulting bounds of the identification region for $\mathrm{E}(Y_3)$ are

$$c_0 = (1 - \pi_0)(1 - \pi_{0|1}) \, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)],$$

$$c_1 = \pi_0 a_1 + (1 - \pi_0)[\pi_{0|1} b_1 + (1 - \pi_{0|1}) \, \mathrm{E}(Y_3 \,|\, Y_1 = 1, Y_2 = 1)],$$

and the width of the identification region is $c_1 - c_0 = \pi_0 a_1 + (1 - \pi_0) \pi_{0|1} b_1$. Even in this case, however, credible values for a_1 and b_1 are not easily found (unless when Y_3 is a 0-1 indicator, in which case it is natural to set $a_1 = b_1 = 1$), and so we do not pursue this approach further.

2.1 Parametric estimation

Our parametric framework assumes that the latent regression errors follow a trivariate Gaussian distribution with zero mean and correlation matrix

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}.$$

We also normalize the variances of U_1 and U_2 to one in order to identify the coefficients of the binary response equations.

In this parametric setting, the vectors $\beta = (\beta_1, \beta_2, \beta_3)$ and $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ of regression parameters can be estimated consistently through the two-step procedure proposed by Poirier (1980) and further developed by Ham (1982). Here, we slightly modify their procedure in order to account for partial observability of Y_2. The procedure exploits the fact that, under our set of assumptions, we have the explicit representation

$$\mathrm{E}(U_3 \,|\, U_1 > -\mu_1, U_2 > -\mu_2) = \rho_{13}\lambda_1(\theta) + \rho_{23}\lambda_2(\theta), \tag{8}$$

where $\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \rho_{12})$, the $\lambda_j(\theta)$ are bias correction terms given by

$$\lambda_1(\theta) = \frac{\phi(\mu_1)\,\Phi(\sigma^{-1}(\mu_2 - \rho_{12}\mu_1))}{\Phi_2(\mu_1, \mu_2; \rho_{12})},$$
$$\lambda_2(\theta) = \frac{\phi(\mu_2)\,\Phi(\sigma^{-1}(\mu_1 - \rho_{12}\mu_2))}{\Phi_2(\mu_1, \mu_2; \rho_{12})},$$

$\sigma = \sqrt{1 - \rho_{12}^2}$, and $\phi(\cdot)$, $\Phi(\cdot)$ and $\Phi_2(\cdot, \cdot; \rho)$ denote, respectively, the density of the standardized Gaussian distribution, its distribution function and the distribution function of the bivariate Gaussian distribution with zero means, unit variances and correlation coefficient ρ. The basic idea of the two-step procedure is to obtain consistent estimates of the bias correction terms in (8), and then use them as additional explanatory variables in an otherwise standard OLS regression.

In the first step of the procedure, we consider a bivariate probit model with sample selection for (Y_1, Y_2), and estimate the parameter θ by maximum likelihood (ML). Identifiability of θ requires imposing at least one exclusion restriction on the two sets of exogenous covariates X_1 and X_2. Subject to the identifiability restrictions, the log-likelihood for a random sample of n units is of the form

$$L(\theta) = \sum_{i=1}^{n} \left[Y_{i1}Y_{i2} \ln \pi_{i11}(\theta) + Y_{i1}(1 - Y_{i2}) \ln \pi_{i10}(\theta) + (1 - Y_{i1}) \ln \pi_{i0}(\theta)\right], \tag{9}$$

where we conventionally set $Y_{i2} = 0$ whenever $Y_{i1} = 0$ and, dropping the suffix i for simplicity,

$$\begin{aligned}
\pi_{11}(\theta) &= \Pr\{Y_1 = 1, Y_2 = 1\} = \Phi_2(\mu_1, \mu_2; \rho_{12}), \\
\pi_{10}(\theta) &= \Pr\{Y_1 = 1, Y_2 = 0\} = \Phi(\mu_1) - \Phi_2(\mu_1, \mu_2; \rho_{12}), \\
\pi_0(\theta) &= \Pr\{Y_1 = 0\} = 1 - \Phi(\mu_1).
\end{aligned}$$

A ML estimator $\hat{\theta}$ maximizes (9) over the parameter space $\Theta = \Re^{k_1+k_2+2} \times (-1, 1)$. This estimator is asymptotically normal under general conditions, and is consistent if the bivariate probit model is correctly specified. Within this model, the hypothesis of conditional independence between unit and item nonresponse can be tested either through a Wald test on the significance of ρ_{12}, or through a likelihood ratio test that compares the maximized values of the log-likelihood in (9) with the sum of the log-likelihoods of two simple probit models, one for Y_1 and one for Y_2 given $Y_1 = 1$.

In the second step of the procedure, estimates $\hat{\lambda}_j = \lambda_j(\hat{\theta})$, $j = 1, 2$, of the bias correction terms in (8) are used as additional predictors in the augmented regression model

$$Y_3 = \alpha_3 + \beta_3^\top X_3 + \sigma_3\rho_{13}\hat{\lambda}_1 + \sigma_3\rho_{23}\hat{\lambda}_2 + \epsilon_3, \tag{10}$$

where $\epsilon_3 = U_3 - \sigma_3\rho_{13}\hat{\lambda}_1 - \sigma_3\rho_{23}\hat{\lambda}_2$ is a heteroscedastic regression error with zero conditional mean. The vector of parameters $\gamma = (\alpha_3, \beta_3, \sigma_3\rho_{13}, \sigma_3\rho_{23})$ may be estimated consistently by OLS, but inference must take into account the heteroscedasticity induced by censoring and the additional sampling variability induced by the use of the generated regressors $\hat{\lambda}_1$ and $\hat{\lambda}_2$ instead of λ_1 and λ_2. Ham (1982) provides consistent estimators of the σ_3 and the asymptotic covariance matrix of the OLS estimator of γ. Alternatively, standard errors may be obtained via the nonparametric bootstrap.

Although implementing this two-step procedure is relatively straightforward, a major concern is identifiability of the parameters in model (10). The identification problem is closely related to that arising in the classical Heckman two-step procedure (see Vella 1998 and Puhani 2000 for an extensive discussion). Parameters of the second estimation step may in principle be identified through the nonlinearity of the inverse Mills ratio. However, since the inverse Mills ratio is linear over a wide range of its argument, identification obtained through the nonlinearity of the inverse Mills ratio is often weak and exclusion restrictions are typically imposed to assist identification in the second estimation step.[2] The above considerations also hold for sample selection models with

[2] Leung and Yu (1996) show that the quasi-linearity of the inverse Mills ratio causes essentially a problem of collinearity with the other covariates of the second estimation step, which in turn leads to inflated standard errors and unreliable estimates.

two censoring equations. Although larger values of the correlation coefficient ρ_{12} increase slightly the nonlinearity of the bivariate Mills ratio, the function is still linear for wide ranges of the two indexes μ_1 and μ_2. Exclusion restrictions (that is, variables which are included in X_1 and X_2 but excluded from X_3) become then crucial to guarantee identifiability of the parameters in the second estimation step.

As suggested by Fitzgerald *et al.* (1998) and Nicoletti and Peracchi (2005), features of the data collection process and socio-demographic characteristics of the interviewers can be promising candidates for this set of exclusion restrictions. Because these variables are external to the subjects under investigation and are not under their control, one should expect them to be irrelevant in explaining the outcome of interest. On the other hand, results of several data validation studies have shown that these variables are typically important predictors of both unit and item response.

2.2 Semiparametric estimation

Parametric estimators of sample selection models are known to be sensitive to incorrect specification of the model. During the last twenty-five years, a large body of literature has been concerned with finding semiparametric procedures for consistent estimation in the presence of various forms of misspecification.

In this section, we focus on model (3)–(6) and consider semiparametric estimation procedures that are robust to departures from the assumption of Gaussian errors. As noted by Vella (1998), semiparametric estimation of sample selection models involves two main difficulties. First, one cannot invoke distributional assumptions to estimate parameters of the two binary response models. Second, one cannot use distributional relationships to find an analytical expression for the bias correction term in equation (7). In this case, the conditional expectation for the outcome variable of interest can be written as the partially linear model

$$\mathrm{E}(Y_3 \,|\, Y_1 Y_2 = 1) = \mu_3 + g(\mu_1, \mu_2), \qquad i = 1, \ldots, n_2, \tag{11}$$

where $\mu_3 = \alpha_3 + \beta_3^\top X_3$ is the linear part of the model, and g is an unknown function of the two indexes $\mu_1 = \alpha_1 + \beta_1^\top X_1$ and $\mu_2 = \alpha_2 + \beta_2^\top X_2$. Notice that model (11) maintains a double index structure. Although restrictive, this structure is useful to avoid the curse of dimensionality problem.

Under regularity conditions, semiparametric estimation of model (3)–(6) can again be carried out through a two-step procedure. In the first step, we adapt the semi-nonparametric (SNP) estimator of Gallant and Nychka (1987) to the bivariate binary-choice model with sample selection to obtain consistent estimators of μ_1 and μ_2.[3] A similar approach has been used, among others, by Gabler *et al.* (1993), Gerfin (1996), Melenberg and van Soest (1996), and Stewart (2004) for SNP estimation of univariate binary-choice and ordered-choice models. In the second step, the remaining parameters of the partially linear model in (11) are instead estimated by the semiparametric approach of Robinson (1988).

Alternative approaches for semiparametric estimation of sample selection models with multiple selection rules are the semiparametric least square (SLS) estimator of Ichimura and Lee (1991) and the semiparametric two-step estimator of Das, Newey and Vella (2003). The SLS estimator is likely to be less efficient of our semiparametric two-step procedure because it only focuses on the sub-sample of fully observed units for simultaneous estimation of the parameters in the selection equations and in the equation for the outcome of interest. In addition, this estimator is more computational demanding because it requires to compute kernel regressions at each iteration of the estimation process. The semiparametric two-step procedure of Das, Newey and Vella (2003) is instead an attractive alternative because it allows for several generalizations (such as the presence of endogenous regressors). However, their estimator requires independence of the error terms in the selection equations, which we do not want to impose a priori.

Before describing our estimation procedure in detail, it is important to mention the conditions under which this semiparametric model is identified. First, identification of μ_3 from equation (11) requires identification of the two indexes μ_1 and μ_2. As argued by Newey (1999), consistency of $\hat{\mu}_1$ and $\hat{\mu}_2$ guarantees identification of the underlying indexes, but different consistent estimators may correspond to different identifying assumptions. Identification conditions necessary for consistent estimation of μ_1 and μ_2 are provided in the following section. Second, it is important to notice that the intercept coefficient α_3 is absorbed into the unknown function g and is not separately identified. Third, as shown by Robinson (1988), identification of the slope coefficients β_3 requires that X_1 and X_2 are not linear combinations of X_3. As for the parametric case, exclusion restrictions are then necessary to guarantee identifiability of the parameters in the second estimation step.

The bivariate binary-choice model with sample selection consists of equations (4) and (5). The log-likelihood function for a random sample of n observations has the same form as (9) with $\pi_{11}(\theta)$,

[3] Parameters of the two response equations could also be estimated by the semiparametric maximum likelihood (SML) approach of Lee (1995), which generalizes to sequential choice models the approach originally proposed for binary-choice models by Klein and Spady (1993). Like other semiparametric estimators based on kernel density estimaton, the SML estimator is likely to be very computational demanding since kernel regression needs to be conducted at each iteration of the maximization process. In particular, because of both the large sample size and the large number of covariates used in our empirical application, the implementation of this estimator would be very time consuming.

$\pi_{10}(\theta)$ and $\pi_0(\theta)$ replaced by

$$\begin{aligned}
\pi_{11}(\delta) &= \Pr\{Y_1 = 1, Y_2 = 1\} = 1 - F_1(-\mu_1) - F_2(-\mu_2) + F(-\mu_1, -\mu_2), \\
\pi_{10}(\delta) &= \Pr\{Y_1 = 1, Y_2 = 0\} = F_2(-\mu_2) - F(-\mu_1, -\mu_2), \\
\pi_0(\delta) &= \Pr\{Y_1 = 0\} = F_1(-\mu_1),
\end{aligned}$$

where $\delta = (\alpha_1, \alpha_2, \beta_1, \beta_2)$, F_j is the unknown marginal distribution function of the latent regression error U_j, $j = 1, 2$, and F is the unknown joint distribution function of (U_1, U_2).

Following Gallant and Nychka (1987), we approximate the unknown joint density f of the latent regression errors by an Hermite polynomial expansion of the form

$$f^*(u_1, u_2) = \frac{1}{\psi_K} \tau_K(u_1, u_2)^2 \, \phi(u_1) \, \phi(u_2), \tag{12}$$

where $\tau_K(u_1, u_2) = \sum_{h,k=0}^K \tau_{hk} u_1^h u_2^k$ is a polynomial of order K in u_1 and u_2, and

$$\psi_K = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau_K(u_1, u_2)^2 \phi(u_1) \phi(u_2) \, du_1 \, du_2$$

is a normalization factor to ensure that f^* is a proper density. The class of densities that can be approximated by this polynomial expansion includes densities with arbitrary skewness and kurtosis, but excludes violently oscillatory densities or densities with too fat or too thin tails (see Gallant and Nychka 1987, p. 369).

Since the polynomial expansion in (12) is invariant to multiplication of $\tau = (\tau_{00}, \tau_{01}, \ldots, \tau_{KK})$ by a scalar, some normalization is needed. After setting $\tau_{00} = 1$, expanding the square of the polynomial in (12) and rearranging terms gives

$$f^*(u_1, u_2) = \frac{1}{\psi_K} \left[\sum_{h,k=0}^{2K} \tau^*_{hk} u_1^h u_2^k \right] \phi(u_1) \, \phi(u_2),$$

with $\tau^*_{hk} = \sum_{r=a_h}^{b_h} \sum_{s=a_k}^{b_k} \tau_{rs} \tau_{h-r,k-s}$, where $a_h = \max(0, h-K)$, $a_k = \max(0, k-K)$, $b_h = \min(h, K)$, and $b_k = \min(k, K)$. Integrating $f^*(u_1, u_2)$ alternatively with respect to u_2 and u_1 gives the following approximations to the marginal densities f_1 and f_2

$$f_1^*(u_1) = \frac{1}{\psi_K} \left[\sum_{h,k=0}^{2K} \tau^*_{hk} \, m_k \, u_1^h \right] \phi(u_1) = \frac{1}{\psi_K} \left[\sum_{h=0}^{2K} \gamma_{1h} \, u_1^h \right] \phi(u_1), \tag{13}$$

$$f_2^*(u_2) = \frac{1}{\psi_K} \left[\sum_{h,k=0}^{2K} \tau^*_{hk} \, m_h \, u_2^k \right] \phi(u_2) = \frac{1}{\psi_K} \left[\sum_{k=0}^{2K} \gamma_{2k} \, u_2^k \right] \phi(u_2), \tag{14}$$

where m_h and m_k are the hth and kth central moments of the standardized Gaussian distribution, $\gamma_{1h} = \sum_{k=0}^{2K} \tau^*_{hk} m_k$, $\gamma_{2k} = \sum_{h=0}^{2K} \tau^*_{hk} m_h$, and $\psi_K = \sum_{h=0}^{2K} \gamma_{1h} m_h = \sum_{k=0}^{2K} \gamma_{2k} m_k$. As for the bivariate density function, γ_{10} and γ_{20} are normalized to one by imposing that $\tau_{i0} = \tau_{0j} = 0$ for all $i, j = 1, \ldots, K$. Thus, if $\gamma_{1h} = 0$ for all $h \geq 1$, then $\psi_K = 1$ and the approximation f_1^* coincides with the standard normal density. Similarly, the approximation f_2^* coincides with the standard normal density when $\gamma_{2k} = 0$ for all $k \geq 1$. Thus, Wald tests for the joint significance of these two sets of parameters provide tests for Gaussianity of the marginal distributions of U_1 and U_2.

Semiparametric identification of the model requires imposing some location restriction either on the distributions of the latent regression errors U_1 and U_2, or on the systematic part of the model. Gabler *et al.* (1993) impose restriction on the τ_{hk} parameters to guarantee that the error term in their model has zero mean. Melenberg and van Soest (1996) argue that forcing the error terms to have mean zero is too cumbersome and propose instead the simpler approach of normalizing the intercept coefficients α_1 and α_2 to be equal to their parametric estimates.

Subject to these identifiability restrictions, integrating the joint density (12) gives the following approximation to the joint distribution function F

$$\begin{aligned} F^*(u_1, u_2) = \Phi(u_1)\Phi(u_2) &+ \frac{1}{\psi_K} A_1^*(u_1, u_2)\phi(u_1)\phi(u_2) \\ &- \frac{1}{\psi_K} A_2^*(u_2)\Phi(u_1)\phi(u_2) - \frac{1}{\psi_K} A_3^*(u_1)\phi(u_1)\Phi(u_2). \end{aligned}$$

Similarly, integrating the marginal densities (13) and (14) gives the following approximations to the marginal distribution functions F_1 and F_2,

$$\begin{aligned} F_1^*(u_1) &= \Phi(u_1) - \frac{1}{\psi_K} A_3^*(u_1)\phi(u_1), \\ F_2^*(u_2) &= \Phi(u_2) - \frac{1}{\psi_K} A_2^*(u_2)\phi(u_2), \end{aligned}$$

where

$$\begin{aligned} A_1^*(u_1, u_2) &= \sum_{h,k=0}^{2K} \tau^*_{hk} A_h(u_1) A_k(u_2), \\ A_2^*(u_2) &= \sum_{h,k=0}^{2K} \tau^*_{hk} m_h A_k(u_2) = \sum_{k=0}^{2K} \gamma_{2k} A_k(u_2), \\ A_3^*(u_1) &= \sum_{h,k=0}^{2K} \tau^*_{hk} m_k A_h(u_1) = \sum_{h=0}^{2K} \gamma_{1h} A_h(u_1), \end{aligned}$$

with $A_0(u_j) = 0$, $A_1(u_j) = 1$, and $A_r(u_j) = (r-1)A_{r-2}(u_j) + u_j^{r-1}$, $j = 1, 2$. Notice that, the SNP approximations to the marginal distribution functions F_1 and F_2 have the Gaussian distribution function as the leading term and differ from this by the product of the Gaussian density function and a polynomial of order $(2K-1)$. The approximation to the joint distribution function F has instead the product of two Gaussian distribution functions as the leading term and differs from this by a complicated function of u_1 and u_2. Thus, the approximations to marginal distribution functions nest the univariate Gaussian distribution function, while the approximation to the joint distribution function only nests the bivariate Gaussian distribution with zero correlation coefficient.

Semiparametric estimators of β_1 and β_2 can then be obtained by maximizing a pseudo log-likelihood function in which the unknown distribution functions F, F_1 and F_2 are replaced by their approximations F^*, F_1^* and F_2^*. As shown by Gallant and Nychka (1987), the resulting pseudo-ML estimator of θ^* is $\sqrt{n}$-consistent provided that the degree K of the polynomial increases with the sample size.[4] In practice, for a given sample size, the value of K may be selected either through a sequence of likelihood ratio tests, or by model selection criteria like the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).

2.2.1 Second step

Given consistent estimates of the two indexes μ_1 and μ_2, the parameters in μ_3 can be estimated by the semiparametric approach of Robinson (1988). Specifically, the partially linear model (11) implies that

$$Y_3 - \mathrm{E}(Y_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1) = \beta_3^\top [X_3 - \mathrm{E}(X_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1)] + \epsilon_3, \tag{15}$$

where $\mathrm{E}(\epsilon_3|\mu_1, \mu_2, Y_1Y_2 = 1) = 0$. After replacing μ_1 and μ_2 with their estimates from the first step and the unknown conditional expectations in (15) with their nonparametric estimates, the slope coefficients β_3 can be estimated by OLS with no intercept. Robinson (1988) shows that, under mild regularity conditions, this estimator is $\sqrt{n}$-consistent and asymptotically normal. Like for the parametric two-step procedure, computation of the standard errors needs to take into account the heteroscedasticity induced by censoring and the additional sampling variability induced by the use of the generated regressors $\hat{\mu}_1$ and $\hat{\mu}_2$. In our empirical application, this is done via the nonparametric bootstrap.

Finally, the nonlinear function g can be estimated nonparametrically by the residual component

$$\hat{g}(\mu_1, \mu_2) = \hat{\mathrm{E}}(Y_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1) - \hat{\beta}_3^\top \, \hat{\mathrm{E}}(X_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1), \tag{16}$$

[4] Notice that, Gallant and Nycka (1987) do not provide asymptotic normality results. Results on the asymptotic normality of the SNP estimator are given in Gabler *et al.* (1993).

where $\hat{\mathrm{E}}(Y_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1)$ and $\hat{\mathrm{E}}(X_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1)$ denote respectively the nonparametric estimates of $\mathrm{E}(Y_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1)$ and $\mathrm{E}(X_3 \,|\, \mu_1, \mu_2, Y_1Y_2 = 1)$.[5] Notice that, since the rate of convergence of $\hat{g}$ depends on the rate of convergence of the nonparametric estimators in (16), this estimator is not $\sqrt{n}$-consistent in general.

3 Data

Our empirical application uses data from the Survey on Health, Ageing and Retirement in Europe (SHARE), a standardized multi-purpose household survey designed to investigate several aspects of the elderly population in Europe.

Survey respondents are asked about their household expenditures on three sub-categories of consumption (food expenditure at home, food expenditure outside home, and phone expenditure) and on total nondurable consumption. Because of the large fractions of zeros on food expenditure outside home and phone expenditure, we only consider food expenditure at home and total nondurable expenditure. Although the measure of primary interest for many economic studies is total nondurable expenditure, recent data validation studies by Browning *et al.* (2002), Battistin *et al.* (2003) and Winter (2004) show that information collected through sub-categories of consumption expenditure is usually more accurate than that collected through a "one-shot" question on total nondurable expenditure. In addition, food expenditure at home is typically an important component of total nondurable expenditure, and is of direct interest in itself.

3.1 Country coverage and sampling design

The first wave of SHARE, conducted in 2004, covered 15,544 households and 22,431 individuals in eleven European countries (Austria, Belgium, Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden and Switzerland).

In each country, the target population consists of all people living in residential households who have at least 50 years of age, plus their (possibly younger) partners. The target population is further restricted by excluding people who currently do not reside at the sampled address, or died before the starting of the field period, or are unable to speak the specific language of the national questionnaire, or are physically or mentally unable to participate to the survey.

All national samples are selected through probability sampling, but sampling procedures are

[5] Alternatively, the nonlinear component of the model can be estimated by a nonparametric regression of $Y_3 - \hat{\beta}_3^\top X_3$ on $\hat{\mu}_1$ and $\hat{\mu}_2$.

not completely standardized across countries. We distinguish between two groups of countries depending on the nature of the sampling frame adopted. In one group of countries (Denmark, Germany, Italy, Netherlands, Spain, and Sweden), the sampling frame is a population register containing information at least on years of age and gender of the sampled units (in Germany only information on broadly defined age class and gender is available). In the other group of countries, the sampling frame is either a telephone register (Austria, Belgium, Greece and Switzerland) or a register of dwellings (France), and contains no information on the background characteristics of the sampled units. In these countries, age-eligibility was assessed through a preliminary screening phase in the field. However, because of nonresponse during the screening phase, it was not possible to determine the eligibility status of about 15 percent of the gross sample. For this second group of countries, the analysis of unit nonresponse is therefore complicated by the lack of sampling frame information and unknown eligibility of a fraction of the gross sample. To avoid these problems, we only consider the countries of the first group for which sampling frame information on years of age and gender is available (Denmark, Italy, Netherlands, Spain, and Sweden).

Table 1 provides the number of eligible households, the unweighted household response rate, and the main sub-components of the household nonresponse rate (that is, noncontact rate, refusal rate and other non-interview rate) by country.[6] The household response rate ranges from a minimum of 47 percent in Sweden to a maximum of 62 percent in the Netherlands, and is equal to 55 percent on average. Focusing attention on the reasons for nonresponse, refusal to participate to the survey is the main reason (35 percent), although in some countries a non negligible fraction of nonresponse is also due to noncontact (13 percent in Spain) and other non-interview reasons (5 percent in Sweden).

Conditional on unit response, SHARE also experienced non-negligible amounts of missing data for open-ended questions on income, assets and consumption expenditures. Item response rates for the two consumption expenditure items of interest are reported in Table 2.[7] The cross-country average of the item response rates is equal to 85 percent for food expenditure at home and 81 percent for total nondurable expenditure, again with considerable variation across countries. The lowest item response rates are in Spain (78 and 77 percent respectively), while the highest are in Sweden (93 and 90 percent respectively).

Although response rates obtained in the first wave of SHARE do not differ considerably from

[6] For each country, the unweighted household response rate is computed as the fraction of eligible households with at least one interviewed person. Further details on the computations of these outcome rates are given in Börsch-Supan and Jürges (2005).

[7] For each consumption expenditure question, the item response rate is compiuted as the fraction of eligible respondents with a "Don't know" or "Refusal" answer.

those obtained by other comparable European surveys, the results of Tables 1 and 2 indicate that unit and item nonresponse may be two important sources of nonsampling errors.

3.2 Outliers in consumption expenditure

A preliminary analysis of PPP-adjusted consumption expenditure data reveals clearly the presence of outliers in the tails of the empirical distribution of these variables. This is a typical problem with data collected through retrospective and open-ended questions. On the one hand, there are households who report zero or very low expenditures. Although these may be plausible values for some consumption categories (like food outside home and phone), they are highly suspicious for food expenditure at home and total nondurable expenditure. On the other hand, we also find extremely high expenditure values which are presumably due to interviewer's typing errors.

To deal with these problems, we symmetrically trim 1 percent of the observations from each tail of the two empirical distributions.[8] Summary statistics (that is, number of nonmissing observations, mean, standard deviation, minimum and maximum) of the two PPP-adjusted and trimmed consumption expenditure distributions are shown in Table 3.

3.3 Choice of predictors

In SHARE, predictors of unit nonresponse can be obtained by exploiting the information coming from the sampling frame, the survey agencies and the fieldwork. By matching these three sources of data, we are able to get information on background characteristics of the selected household member (like years of age and gender), interviewers' characteristics (like years of age, gender and years of education), number and timing of calls made (total number of calls and indicators for calls made in the evening and the week-end), and length of the fieldwork (measured by the number of days between the first and the last call).

For the sub-sample of responding households, the additional information collected during the interview can be used to study nonresponse on specific items of the questionnaire. The multi-disciplinary nature of the SHARE data offers the unique opportunity of assessing whether item nonresponse on consumption expenditure questions is related to different types of economic and health variables, while controlling for features of the data collection process and background characteristics of respondents and interviewers.

[8] A sensitiveness analysis with 1.5 and 2 percent of trimming on each tail does not lead to qualitative different results. Thus, for efficiency reasons, we only present results with 1 percent of trimming.

Since consumption expenditure questions are asked to the household member who is most knowledgeable about housing matters (the "household respondent" or HR), a set of variables related to socio-demographic characteristics, cognitive abilities, and health conditions of the HR has been included as predictors of item response. Our set of socio-demographic variables includes years of age (entering as a quadratic term), gender, years of education, current job situation, marital status, household size, and indicators for having children aged less than 6 years and for living in small cities. Cognitive abilities are measured through the scores obtained in the mathematical, orientation in time, and delayed recall tests performed during the cognitive function (CF) module of the SHARE interview. The set of health variables includes instead indicators for less than good self-perceived health, for at least one ADL limitation, and for self-reported problems in managing money.[9]

One problem with estimating Engel curves using the SHARE data is that household income is typically affected by item nonresponse, measurement errors and outliers. To deal with the first problem, we use the imputed gross annual household income provided in the SHARE public release database, but include an indicator for imputed values.[10] To reduce the impact of measurement errors and outliers, we do not use income directly but use instead a set of indicators for income quartiles. There are three reasons for using imputed household income. First, considering the overall process leading to missing data on consumption and household income is complicated. This is because the latter is obtained by aggregating 27 different sources of income (19 collected at the individual level and 8 collected at the household level). Second, nonresponse on household income does not seem to be as problematic as nonresponse on consumption expenditure. As shown in Table 4, household income is fully observed for 47 percent of the households that agree to participate to the survey, while at least one income component is missing for the remaining 53 percent. Of these, 27 percent have only one missing component, 15 percent have two missing components, 7 percent have three missing components, and only 4 percent have more than three missing components. Accordingly, for households with missing income components, the fraction

[9] A description of these cognitive ability tests and health measures can be found in Börsch-Supan *et al.* (2005).

[10] In SHARE, gross household income is obtained by aggregating different sources of income collected both at the individual and the household level. Missing data are imputed separately for each income source by using conditional hot-deck and regression imputations. The first method is used to impute missing data on the amount variables, while the second is used to imputed missing data on the frequency variables (e.g. the number of months in which the respondent has received a payment). For the imputation of amounts, the set of covariates consists of unfolding-bracket intervals and country dummies. If no unfolding-bracket information is available, a richer set of covariates is used (typically, age, gender, education and country). For the imputation of frequencies, the set of covariates consists instead of age, gender, quartiles of the imputed amount, and country dummies. Multiple imputations are then provided by five independent replications of the one-step imputation procedure. Here, for simplicity reasons, we only use the first of these independent replications.

of total household income that is imputed is equal to 37 percent on average, but for half of them this fraction does not exceed 20 percent. Third, initial nonrespondents to questions about income amounts are asked a sequence of unfolding-bracket questions, namely whether the amount is larger than, smaller than, or about equal to a given threshold. This sequence provides valuable categorical information that is exploited in the imputation of household income. By design, however, no sequence of unfolding-bracket questions is used to collect additional information on nonrespondents to consumption expenditure questions.[11]

To control for differences in the interview process, we use a set of measures of the cognitive burden of the interview. These variables include the length of the HR interview, and a set of indicators for partial proxy interviews, full proxy interviews, and interviews not conducted at the respondent's home.

As a measure of the interviewers' computer skill, we also include the length of the Interviewer (IV) module. The IV module contains a set of close-ended questions on the background characteristics of the interviewers and the conditions of the interview process. Since this module is only completed by the interviewer without involving the respondent, its length provides a proxy measure of the interviewers' computer skill.[12]

Definitions and summary statistics (number of nonmissing observations, mean and standard deviation) of the predictors used in our empirical application are given in Table 5.

4 Empirical results

This section presents the results obtained by using the first wave of SHARE to investigate whether selectivity associated with unit and item nonresponse may bias the estimation of Engel curves for household consumption expenditure.

4.1 Parametric estimates

We estimate five alternative models. Model 1 is a standard linear model estimated for the fully responding units without accounting for selectivity generated by nonresponse. Model 2a is a classical sample selection model estimated for the unit respondents and only accounts for selectivity generated by item nonresponse. Model 3a is a classical sample selection model estimated for the

[11] Extensive discussions on the definition of household income, item response rates, unfolding-bracket questions, and imputation procedures adopted in SHARE are given in Börsch-Supan and Jürges (2005).

[12] Estimates of a regression model reveals that this module tends to last longer for interviewers with higher age and lower education. Results are omitted to save space.

full sample with a single indicator ($D = Y_1 Y_2$) for unit and item response. Model 4a is a generalized sample selection model which accounts for selectivity generated by unit and item nonresponse, but assumes that errors in the unit and the item response equations are independent. Finaly, Model 5a is a generalized sample selection model which accounts for selectivity generated by unit and item nonresponse, and does not impose independence of the error terms in the two response equations. To control for possible effects due to "mismatch" between the interviewer and the interviewee, we also present the results obtained by introducing interaction terms in the age, gender and schooling level of the interviewer and the interviewee (Models 2b, 3b, 4b, and 5b). Parametric estimates of these models are provided in Tables 6–13.

All estimated models share two common features. First, given the high comparability of the SHARE data, we pool data from the various countries and introduce country indicators plus their interactions with income quartiles to capture unobserved heterogeneity across countries. Pooling the data allows us to increase efficiency of estimation and helps reducing problems of collinearity due to the limited within-country variability of some variables (like the characteristics of the fieldwork and the interviewers).[13] Second, identifiability of the model parameters is achieved by imposing a common set of exclusion restrictions. As mentioned in Section 2.1, our exclusion restrictions are based on characteristics of the fieldwork, the interview process and the interviewers. In particular, characteristics of the fieldwork are used to predict unit nonresponse, characteristics of the interview process are used to predict item nonresponse, and socio-demographic characteristics of the interviewers are used to predict both. Thus, if we distinguish between sampling frame or fieldwork information (Z), characteristics of the household or the HR (V), characteristics of the interviewer or the interview process (W), and country dummies (D), then the sets of predictors in each of the three model equations are

$$
\begin{aligned}
X_1 &= (Z, W, D), \\
X_2 &= (V, W, W^*, D), \\
X_3 &= (V, D).
\end{aligned}
$$

where W^* are characteristics of the interviewer or interview process which are only relevant for the item response equation. The use of this large set of exclusion restrictions should protect against problems of collinearity, especially in the second estimation step.

Estimates of the response probability for Model 3 are presented in Table 6, while estimates of the probability of participation for Models 4 and 5 are presented in Table 7. After dropping

[13] As the reference country, we always take Sweden (the country with the largest number of observations).

a few cases with missing data on the covariates, the sample size consists of 12,537 households, of which 6,746 (53.8 percent) participated to the survey. The reference category for the unit response equations is always a Swedish male aged 55, who received 4 call attempts during fieldwork lasting one month, and was approached by a male interviewer of the same age and schooling level.

Other things being equal, we find that the relationship between the probability of participation and the age of the sampled household member has an inverted-U shape, with a maximum at 62 years of age. Women are less likely to participate than men, but the differences are not strongly significant. The interviewer's gender does not seem to matter, whereas the interviewer's age is positively related to unit response. The interviewer's education, the total number of calls and the length of the fieldwork are negatively related to unit response. These negative relations may simply reflect the strategy of increasing the number of calls (specially those in the evening) and switching to more experienced interviewers when there are difficulties in reaching contact and gaining respondents' cooperation. As for the interaction between characteristics of the interviewee and the interviewer, we find that the age interactions are significant at the 1 percent level, while the gender interaction is not. While the AIC tends to select the less parsimonious models with interactions, the BIC tends to select the more parsimonious ones without interactions.

Estimates of the probability of item response for food expenditure at home and total nondurable expenditure are presented in Tables 8 and 9 respectively. The reference category for the item response equations is now a Swedish male worker aged 55, with 13 years of education, living in a couple without children, residing in a big city, with good cognitive abilities and health conditions, in the second income quartile, with income not missing, interviewed by a male interviewer of the same age and schooling level, and with the interview conducted at the respondent home, without proxy, and lasting one hour.

Other thing being equal, we find that the probability of item response tends to fall with the age of the HR and with the presence of children aged less than 6 years. Living in a small city, being employed, being single, or being more educated are negatively related to item response, but the estimated effects are only weakly statistically significant. Even after controlling for the respondent's background characteristics (like age and education), the scores obtained on the cognitive ability tests are positively related to item response, while other health measures are not. The positive and significant coefficients on the third and forth income quartiles suggest that item nonresponse may lead to selection of households with higher income. Furthermore, the negative coefficient on the dummy for income imputations suggests that nonresponse to income questions is positively related to nonresponse to consumption expenditure questions. Among characteristics of the interview process and the interviewers, we find that allowing the interviewee to be assisted by a proxy respondent, conducting the interview at the respondent home, and using more experienced interviewers (that is, interviewers with better computer skill) are positively related to item response. Interaction terms do not seem to be important predictors of the probability of item response. For food consumption at home, both AIC and BIC tend to select models without interactions. For total nondurable consumption, AIC tends to select models with interactions, while BIC tends to select models without interactions.[14] Estimates of the correlation coefficient ρ_{12} are very close to zero, and the corresponding likelihood ratio tests never reject conditional independence between unit and item nonresponse. Accordingly, the differences between the estimated coefficients of the probit model (Models 4a and 4b) and the bivariate probit model (Models 5a and 5b) are not statistically significant.

Finally, Tables 10–13 present estimates of the Engel curves for food expenditure at home and total nondurable expenditure. For both consumption expenditure items, the selection biases associated to unit and item nonresponse have opposite sign and therefore partly offset each other: the first (unit nonresponse) is positive, the second (item nonresponse) is negative.

Estimates of the model parameters can be used to estimate the relative total nonresponse bias for the reference category (corresponding to $X = x$)

$$\mathrm{RB}_{\mathrm{total}}(x) = \frac{\mathrm{E}(U_3 \,|\, U_1 > -\mu_1(x), U_2 > -\mu_2(x))}{\mu_3(x)},$$

where $\mu_j(x) = \alpha_j + \beta_j^\top x$, $j = 1, 2, 3$. For our parametric model, this is just the sum of the relative biases due to unit and item nonresponse. The relative unit nonresponse bias for the refence category ranges between 8 percent and 10 percent for food expenditure at home, and between 14 percent and 18 percent for total nondurable expenditure. The relative item nonresponse bias ranges instead between -2 percent and -3 percent for food expenditure at home, and between -3 percent and -5 percent for total nondurable expenditure. For food expenditure at home, the coefficients on the bias correction terms λ_{unit} and λ_{item} are not statistically significant, and estimates of the five models are not very different. We conclude that unit and item nonresponse appear to be purely random. For total nondurable expenditure, the coefficients on the bias correction terms are statistically significant at the 1 percent level. Therefore, neither unit nor item nonresponse errors are ignorable, and only estimates of Models 4 and 5 are consistent.

[14] In this case, only interaction terms in age are statistically significant at the 5 percent level.

As mentioned in Section 2.2, the assumption of Gaussian errors is critical for our parametric estimates of sample selection models. Parametric estimators are indeed inconsistent if this distributional assumption is not valid. Conditional moment tests provide a simple way of testing for this assumption.[15] Because errors of the unit and the item response equations are independent under the null hypothesis (see Tables 8 and 9), tests for Gaussianity can be conducted separately by testing for Gaussianity in two simple probit models. Following Pagan and Vella (1989), these conditional moment tests can be computed via artificial regressions in which the sample third and forth moments are regressed against an intercept and the score obtained from the probit model. Under the null, the coefficients on the intercepts should be equal to zero. Thus, a test for the joint significance of these coefficients is a test for Gaussianity in the probit model. The independence of the error terms in the two response equations is also useful to test Gaussianity in the outcome equation. In this case, a straightforward generalization of the RESET-like test proposed by Pagan and Vella (1989) consists of augmenting the second estimation step with the additional variables $\zeta_{Uj} = \hat{\mu}_1^j \, \hat{\lambda}_1$, $\zeta_{Ij} = \hat{\mu}_2^j \, \hat{\lambda}_2$ and $\zeta_{rs} = \zeta_{Ur} \, \zeta_{Is}$ (with $j = 1, 2, 3$ and $r, s = 0, 1, 2, 3$), and then testing their joint significance.[16] Tables 14 and 15 focus on Model 5a and provide tests for the assumption of Gaussian errors in the two response equations and in the outcome equation respectively. Overall, our results suggest that the Gaussian assumption is strongly rejected for the unit response equation, but not for the other two equations. As a consequence, parametric estimators may be inconsistent.

4.2 Semiparametric estimates

In this section, we focus on Model 5a and presents estimates of the semiparametric two-step procedure that are robust to departures from the assumption of Gaussian errors.

In the first step of the procedure, parameters of the unit and item response equations are estimated jointly by the SNP estimator discussed in Section **??** for two choices of K (the degree of the Hermite polynomial used for approximating the bivariate density of the error terms), namely $K = 3$ and $K = 4$. The model specifications underlying these different choices are then compared through a likelihood ratio test, AIC, and BIC. For all model selection criteria, the preferred specification has $K = 3$ for both food expenditure at home and total nondurable expenditure. Parametric and SNP estimates of these models are presented in Tables 16–19. Notice that, because of the different scale, estimated coefficients of the SNP model and the bivariate probit model with sample selection are not directly comparable. Thus we compare ratios of the estimated coefficients, obtained

[15] For an extensive discussion of conditional moment tests see Newey (1985) and Pagan and Vella (1989).

[16] Here, $\hat{\lambda}_1$ and $\hat{\lambda}_2$ denote the classical inverse Mills ratios evaluated at $\hat{\mu}_1$ and $\hat{\mu}_2$ respectively.

after dividing by the absolute value of the coefficient for the length of the fieldwork (*lfield*) in the unit response equation and the basolute value of the coefficient on the length of the IV module (*ivlength*) in the item response equation (and then dividing again by 100). The standard errors of these normalized coefficients are computed using the delta method.

For the unit response equation, we find significant differences between the SNP and the parametric estimates. The main differences occur for interviewers' characteristics, features of the fieldwork, and country dummies. The assumption of Gaussian error in the unit response equation is again rejected at the 1 percent level by a Wald test on the joint significance of the γ_{1h} parameters in the approximation (13). The semiparametric estimate of the marginal density function exhibits positive skewness and lower kurtosis than a standard normal density. Furthermore, the density plot in Figure 1 also reveals the presence of multiple modes.

For the item response equation, we find that Gaussianity is still rejected for food expenditure at home but not for total nondurable expenditure. The marginal densities underlying the two item response equations appear to have a similar shape. They are both platikurtic, exhibit negative skewness, and have a secondary mode in the lower tail of the distribution (see Figure 1). For the item response equation, however, departures from the assumption of Gaussian errors appear to be less harmful. Once the different scale in taken into account, the differences between the parametric and the SNP estimates in Table 19 are small.

In the second step of the procedure, estimates of the two indexes $\hat{\mu}_1$ and $\hat{\mu}_2$ are used to estimate a partially linear model for the outcome variables of interest. Following Robinson (1988), the unknown conditional expectations in (15) are estimated nonparametrically by Nadaraya-Watson bivariate kernel regression estimators, where the bivariate kernel is the product of two univariate bias reducing kernels with the same bandwidth $h_n = O(n^{-1/p})$. In computing the nonparametric estimates, we also trim observations for which the denominator of the Nadaraya-Watson estimator is smaller than a threshold $b_n = O(n^{-1/r})$. In the application, we experiment with different values of p and r. After replacing the unknown conditional expectations in (15) with their nonparametric estimates, the vector β_3 of slope parameters is estimated by standard OLS with no intercept.[17] Standard errors of the OLS estimator are instead computed by the nonparametric bootstrap with 50 replications.

Semiparametric estimates of the partially linear model for food expenditure at home and total nondurable expenditure are presented in Tables 20 and 21 respectively. To explore sensitiveness of

[17] As mentioned in Section 2.2, the intercept coefficient is absorbed in the nonlinear function g and is not separately identified.

Robinson's estimator with respect to choice of the bandwidth parameter h_n and the threshold b_n, estimation is carried out for various alternative combinations of p and r.[18] In particular, results are presented for $p = 5$ and $r = 21$ (Model A), $p = 6$ and $r = 13$ (Model B), $p = 7$ and $r = 10$ (Model C). Parametric estimates are also reported to facilitate comparisons.

According to our estimates, the Robinson's estimator is not very sensitive to the choice of the bandwidth h_n and the threshold b_n. The only exception is Model A in Table 20. For this model, the low value of the bandwidth and the high value of threshold lead to imprecise estimates. For food expenditure at home, we find that the estimated coefficients of the partially linear are very close to their parametric counterparts. This suggests that parametric estimates are only marginally affected by departures from Gaussianity. For total nondurable expenditure, instead, the differences between the parametric and the semiparametric estimates are somewhat larger, especially for the coefficients on the income quartiles and their interactions with the country dummies.

5 Conclusions

In this paper we investigate problems of selectivity generated by unit and item nonresponse in cross-sectional surveys, or equivalently, in the first wave of a panel survey.

We first analyze a general sample selection model in which unit and item nonresponse can simultaneously affect a regression relationship of interest through NMAR missing data mechanisms. Issues concerning identification and estimation have been considered for two alternative specifications of this model. In the parametric specification, errors in the two selection equations and in the equation for the outcome of interest are assumed to follow a trivariate Gaussian distribution, and model parameters are estimated by the parametric two-step procedure proposed by Poirier (1980). In the semiparametric specification, error terms are assumed to follow an unknown distribution, and model parameters are estimated by a semiparametric two-step procedure which involves a generalization of the SNP estimator proposed by Gallant and Nychka (1987) in the first step, and the semiparametric estimator of Robinson (1988) in the second step.

We then use data from the first wave of SHARE to investigate whether selectivity associated with unit and item nonresponse may bias the estimation of Engel curve for food expenditure at home and total nondurable expenditure. Overall, the amount of bias generated by unit and item nonresponse does not seem to be ignorable. According to our estimates, the relative unit nonresponse bias for

[18] Combinations of p and r are selected to satisfy conditions imposed on the choice of the bandwidth parameter and the trimming factor (see Robinson 1988, Theorem 1).

the refence category ranges between 8 percent and 10 percent for food expenditure at home, and between 14 percent and 18 percent for total nondurable expenditure. The relative item nonresponse bias ranges instead between -2 percent and -3 percent for food expenditure at home, and between -3 percent and -5 percent for total nondurable expenditure. According to several specifications of our sample selection models, unit and item nonresponse errors appear to be purely random for food expenditure at home, whereas they are not ignorable for total nondurable expenditure.

Diagnostic tests based on the conditional moment framework of Pagan and Vella (1989) and on the SNP framework do not support the assumption of Gaussian errors, specially in the unit response equation. Nevertheless, estimates of our semiparametric two-step procedure do not lead to very qualitative different results.

References

Battistin E., R. Miniaci, and G. Weber (2003), "What Do We Learn from Recall Consumption Data?", *Journal of Human Resources*, 38: 354–385.

Börsch-Supan A., A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber (2005), *Health, Ageing and Retirement in Europe - First Results from the Survey of Health, Ageing and Retirement in Europe*, MEA, Mannheim.

Börsch-Supan A., and H. Jürges (2005), *The Survey of Health, Ageing and Retirement in Europe - Methodology*, MEA, Mannheim.

Browning M., T.F. Crossley, and G. Weber (2002), "Asking Consumption Questions in General Purpose Survey", *Economic Journal*, 113: 540–567.

Das M., W.K. Newey, and F. Vella (2003), "Nonoparametric Estimation of Sample Selection Models", *Review of Economic Studies*, 70, 33–58.

Eurostat (1997), "Response rates for the first three waves of the ECHP", PAN 92/97, Luxembourg.

Fitzgerald J., P. Gottschalk, and R. Moffitt (1998), "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics", *Journal of Human Resources*, 33: 251–299.

Gabler S., F. Laisney, and M. Lechner (1993), "Semiparametric Estimation of Binary-Choice Models with an Application to Labor-Force Participation", *Journal of Business and Economic Statistics*, 11: 61–80.

Gallant A.R., and D.W. Nychka (1987), "Semi-Nonparametric Maximum Likelihood Estimation", *Econometrica*, 55: 363–390.

Gerfin M. (1996), "Parametric and Semi-Parametric Estimation of Binary Response Model of Labor Market Participation", *Journal of Applied Econometrics*, 11: 321–339.

Groves R.M., and M.P. Couper (1998), *Nonresponse in Household Interview Surveys*, Wiley, New York.

Groves R.M., D.A. Dillman, J.L. Eltinge, and R.J.A. Little (2002), *Survey Nonresponse*, Wiley, New York.

Ham J.C. (1982), "Estimation of a Labour Supply Model with Censoring due to Unemployment and Underemployment", *Review of Economic Studies*, 49: 335–354.

Heckman J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47: 153–161.

Horowitz J., and C.F. Manski (1998), "Censoring of Outcomes and Regressors due to Survey Nonresponse: Identification and Estimation using Weights and Imputations", *Journal of Econometrics*, 84: 37–58.

Ichimura H., and L.F. Lee (1991), "Semiparametric Least Square of Multiple Index Models: Single Equation Estimation", in W.A. Barnett, J. Powell, and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge.

Klein R., and R. Spady (1993),"An Efficient Semiparametric Estimator for the Binary Response Model", *Econometrica*, 61: 387–421.

Lee L.F. (1995), "Semiparametric Maximum Likelihood Estimation of Polychotomous and Sequential Choice Models", *Journal of Econometrics*, 65: 381–428.

Lessler J.T., and W.D. Kalsbeek (1992), *Nonresponse Errors in Surveys*, Wiley, New York.

Leung S.F., and S. Yu (1996), "On the Choice between Sample Selection and Two-Part Models", *Journal of Econometrics*, 72: 197–229.

Little R.J.A., and D.B. Rubin (2002), *Statistical Analysis with Missing Data* (2nd ed.), Wiley, New York.

Manski C.F. (2003), *Partial Identification of Probability Distribution*, Springer, New York.

Melenberg B., and A. van Soest (1996), "Measuring the Costs of Children: Parametric and Semiparametric Estimators", *Statistica Neerlandica*, 50: 171–192.

Nicoletti C., and F. Peracchi (2005), "Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel", *Journal of the Royal Statistical Society*, *Series A*, 168: 763–81.

Newey W. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica*, 53: 1047–1070.

Newey W. (1999), "Two Step Series Estimation of Sample Selection Models", unpublished.

O'Muircheartaigh C., and P. Campanelli (1999), "A Multilevel Exploration of the Role of Interviewers in Survey Nonresponse", *Journal of the Royal Statistical Society*, *Series A*, 162: 437–446.

Pagan A., and F. Vella (1989), "Diagnostic Test for Models Based on Individual Data: A Survey", *Journal of Applied Econometrics*, 4: 29–59.

Poirier D. (1980), "Partial Observability in Bivariate Probit Models", *Journal of Econometrics*, 12: 209–217.

Puhani P. A. (2000), "The Heckman Correction for Sample Selection and its Critique", *Journal of Economic Surveys*, 14: 53–68.

Riphahn R. T., and O. Serfling (2002), "Item Nonresponse on Income and Wealth Questions", Discussion Paper No. 573, IZA Bonn.

Robinson P.M. (1988), "Root-N-Consistent Semiparametric Regression", *Econometrica*, 56: 931–954.

Rubin D.B. (1976), "Inference and Missing Data", *Biometrika*, 63: 581–592.

Stewart M.B. (2004), "Semi-Nonparametric Estimation of Extended Ordered Probit Models", *The Stata Journal*, 4: 27–39.

Vella F. (1998), "Estimating Models with Sample Selection Bias: A Survey", *Journal of Human Resources*, 33: 127–169.

Winter J. (2004), "Response Bias in Survey-Based Measures of Household Consumption", *Economics Bulletin*, 3: 1–12.

Table 1: Unweighted household response rates.

Country	Eligible	Response rate	Noncontact rate	Refusal rate	Other noninterview rate
Denmark	1742	.61	.09	.29	.01
Italy	2505	.54	.08	.36	.02
Netherlands	2509	.62	.05	.32	.01
Spain	2619	.50	.13	.36	.01
Sweden	3956	.47	.06	.42	.05
Total	13331	.55	.08	.35	.02

Table 2: Unweighted item response rates for consumption expenditure questions.

Country	Eligible	Food at home	Nondurable consumption
Denmark	1178	.81	.79
Italy	1376	.85	.84
Netherlands	1559	.89	.77
Spain	1341	.78	.77
Sweden	1850	.93	.90
Total	7304	.85	.81

Table 3: Summary statistics for consumption expenditure questions. Yearly amounts expressed in 100 Euro. Empirical distributions trimmed symmetrically by 2 percent.

Variable	Obs.	Mean	Std.	Min	Max
Food at home	6067	49.6	42.4	.9	640
Nondurable consumption	5780	118.0	86.9	11.7	1280

Table 4: Summary statistics for the fraction of imputed household income by missing income components.

Missing components	Eligible	Fraction of imputed income Mean	P_{25}	P_{50}	P_{75}
0	3411				
1	1943	.20	.00	.02	.29
2	1091	.47	.07	.42	.88
3	527	.59	.18	.68	1.00
4 or more	332	.72	.47	.88	1.00
Total	7304	.37	.01	.20	.76

Table 5: Definitions and summary statistics for the main covariates. U denotes the unit response equation, I denotes the item response equation, C denotes the consumption equation.

Variable	Obs.	Mean	Std.	Description	Equation
gs_fem	13114	.54	.50	Female sampled hh member	U
gs_age	13114	65.0	10.7	Age of the sampled hh member	U
iv_fem	13100	.75	.43	Female interviewer	U, I
iv_age	13100	49.4	11.9	Interviewer age	U, I
iv_yedu	12804	13.1	2.95	Interviewer years of education	U, I
tot_call	13114	3.95	5.36	Total number of call attempts	U
call_ev	13114	.51	.50	Dummy for calls made in the evening	U
call_we	13114	.25	.43	Dummy for calls made in the week-end	U
lfield	13114	42.0	47.4	Length of fieldwork (days)	U
hr_fem	7087	.55	.50	HR female	I, C
hr_age	7076	64.9	10.5	HR age	I, C
hr_yedu	7060	9.30	4.51	HR years of education	I, C
hr_nowork	7041	.66	.47	No paid work in the last 4 weeks	I, C
single	7065	.33	.47	HR lives as single	I, C
hsize	7087	2.15	1.05	Household size	I, C
children	7032	.02	.14	Dummy for children aged less than 6 years	I, C
s_city	6930	.22	.41	Household lives in a small city	I, C
math	7031	.23	1.17	Score on mathematical test (-2–2)	I, C
recall	6999	-1.73	2.04	Score on delayed recall test (-5–5)	I, C
orient	7043	.76	.65	Score on orientation in time test (-3–1)	I, C
hr_health	7050	.39	.49	Less than good self-perceived health	I, C
hr_adl	7048	.10	.30	At least one ADL limitation	I, C
hr_pmm	7087	.04	.19	Self-reported problems in managing money	I, C
income_q1	7084	.25	.43	1st quartile gross annual HH income	I, C
income_q2	7084	.25	.43	2nd quartile gross annual HH income	I, C
income_q3	7084	.25	.43	3rd quartile gross annual HH income	I, C
income_q4	7084	.25	.43	4th quartile gross annual HH income	I, C
inc_mis	7087	.53	.50	Gross annual income missing	I, C
f_proxy	6958	.02	.12	Full proxy interview (CO module)	I
p_proxy	6958	.07	.26	Partial proxy interview (CO module)	I
int_out	6956	.04	.20	Interview not conducted at the respondent home	I
int_length	6996	72.6	27.0	Length of the HR interview (min.)	I
iv_length	7072	1.89	1.59	Length of the IV module (min.)	I

Table 6: Parametric estimates of the response probability for the model with a single indicator for both unit and item response (* denotes a p-value between 1% and 5%, ** denotes a p-value below 1%). $T_1(5)$ is a χ^2-test for the interaction terms and has 5 degree of freedom (df), $T_2(4)$ is a χ^2-test for the age interaction terms and has 4 df. Sample size $n = 12537$.

	Food at home		Nondurable expenditure	
Variable	Model 3a	Model 3b	Model 3a	Model 3b
gs_fem	-.0348	-.0307	-.0695 **	-.0596
gs_agec	.0038	.0018	.0011	-.0029
gs_agec^2	-.0005 **	-.0005 **	-.0004 **	-.0003 *
iv_fem	.0147	.0193	.0075	.0169
iv_agec	.0053 **	.0089 **	.0060 **	.0106 **
iv_agec^2	.0001	.0001	.0002 *	.0002
iv_yeduc	-.0181 **	-.0182 **	-.0129 **	-.0130 **
tot_call	-.0198 **	-.0195 **	-.0208 **	-.0205 **
call_ev	-.0626 *	-.0632 *	-.0967 **	-.0969 **
call_we	.0708 *	.0697 *	.0975 **	.0964 **
lfield	-.0039 **	-.0039 **	-.0037 **	-.0037 **
DK	-.0578	-.0554	-.0637	-.0617
ES	-.3586 **	-.3556 **	-.3373 **	-.3338 **
IT	-.2611 **	-.2590 **	-.2401 **	-.2375 **
NL	.0057	.0059	-.1424 **	-.1421 **
gsiv_sex$_{ff}$		-.0041		-.0112
gsiv_age$_{11}$		-.0003		-.0003
gsiv_age$_{12}$		.0000		.0000
gsiv_age$_{21}$		-.0000		-.0000
gsiv_age$_{22}$		-.0000		-.0000
_cons	.1825 **	.1965 **	.1795 **	.1915 **
$k_1 + k_2$	15	20	15	20
$T_1(5)$		16.06 **		21.67 **
$T_2(4)$		16.06 **		21.63 **
AIC	16776.22	16770.15	16754.58	16742.82
BIC	16895.21	16926.32	16873.56	16898.99

Table 7: Parametric estimates for unit response. Information criteria for Models 5a and 5b are not reported because they are estimated jointly with the models presented in Tables 8 and 9. Sample size $n = 12537$.

			Food at home		Nondurable expenditure	
Variable	Model 4a	Model 4b	Model 5a	Model 5b	Model 5a	Model 5b
gs_fem	-.0456 *	-.0348	-.0457 *	-.0348	-.0455 *	-.0345
gs_agec	.0053 *	.0013	.0053 *	.0013	.0053 *	.0013
gs_agec^2	-.0004 **	-.0003 **	-.0004 **	-.0003 **	-.0004 **	-.0003 **
iv_fem	.0278	.0378	.0278	.0378	.0275	.0375
iv_agec	.0048 **	.0076 **	.0048 **	.0076 **	.0048 **	.0076 **
iv_agec^2	.0002 **	.0002	.0002 **	.0002	.0002 **	.0002
iv_yeduc	-.0155 **	-.0156 **	-.0155 **	-.0156 **	-.0156 **	-.0157 **
tot_call	-.0220 **	-.0217 **	-.0220 **	-.0217 **	-.0219 **	-.0216 **
call_ev	-.0726 **	-.0729 **	-.0727 **	-.0729 **	-.0785 **	-.0784 **
call_we	.0661 *	.0646 *	.0661 *	.0646 *	.0688 *	.0672 *
lfield	-.0042 **	-.0042 **	-.0042 **	-.0042 **	-.0042 **	-.0042 **
DK	.1749 **	.1768 **	.1749 **	.1768 **	.1737 **	.1757 **
ES	-.2109 **	-.2066 **	-.2109 **	-.2066 **	-.2111 **	-.2071 **
IT	-.1506 **	-.1481 **	-.1506 **	-.1481 **	-.1511 **	-.1487 **
NL	.1103 **	.1106 **	.1103 **	.1106 **	.1098 **	.1102 **
gsiv_sex$_{ff}$		-.0131		-.0132		-.0132
gsiv_age$_{11}$		-.0002		-.0002		-.0002
gsiv_age$_{12}$		.0000		.0000		.0000
gsiv_age$_{21}$		-.0000		-.0000		-.0000
gsiv_age$_{22}$		-.0000		-.0000		-.0000
_cons	.2138 **	.2296 **	.2139 **	.2296 **	.2166 **	.2319 **
k_1	15	20	15	20	15	20
$T_1(5)$		14.98 *		14.97 *		14.71 *
$T_2(4)$		14.91 **		14.91 **		14.65 **
AIC	16638.97	16633.97				
BIC	16757.95	16790.13				

Table 8: Parametric estimates for item response on food expenditure at home. Interaction terms between characteristics of the interviewee and the interviewer, and between country dummies and income quartiles are not reported to save space. $T_1(18)$ is a χ^2-test for all interaction terms and has 18 df. $T_2(4)$ is a χ^2-test for the age interaction terms and have 4 df. $T_3(12)$ is a χ^2-test for the interaction terms between country dummies and income quartiles and have 12 df. Sample size $n_1 = 6746$.

Variable	Model 4a	Model 4b	Model 5a	Model 5b
hr_fem	.1333 **	.1925 *	.1335 **	.1925 *
hr_agec	-.0080	-.0060	-.0080	-.0060
hr_agec^2	.0000	-.0001	.0000	-.0001
hr_yeduc	-.0078	-.0067	-.0077	-.0067
hr_nowork	.1504 *	.1547 *	.1502 *	.1547 *
single	-.1262 *	-.1209 *	-.1263 *	-.1209 *
hsize	.0040	.0080	.0040	.0080
children	-.4626 **	-.4624 **	-.4627 **	-.4624 **
s_city	-.0381	-.0425	-.0381	-.0425
math	.0442	.0439	.0442	.0439
recall	.0265 *	.0259	.0265 *	.0259
orient	.1493 **	.1522 **	.1493 **	.1522 **
hr_health	.0086	.0111	.0086	.0111
hr_adl	-.0902	-.0882	-.0902	-.0882
hr_pmm	-.0827	-.0909	-.0827	-.0909
income_q$_1$	-.2757	-.2810	-.2759	-.2810
income_q$_3$	.3522 *	.3593 *	.3520 *	.3593 *
income_q$_4$	.3073 *	.3093 *	.3073 *	.3093 *
inc_mis	-.6863 **	-.6841 **	-.6862 **	-.6841 **
f_proxy	-.0261	-.0388	-.0257	-.0388
p_proxy	.2608 **	.2530 **	.2607 **	.2530 **
int_out	-.1251	-.1155	-.1247	-.1155
int_length	.0021 *	.0021 *	.0021 *	.0021 *
iv_length	-.0358 **	-.0361 **	-.0357 **	-.0361 **
iv_fem	-.0224	.0200	-.0225	.0200
iv_agec	.0011	.0062	.0011	.0062
iv_agec^2	-.0003 *	-.0002	-.0003 *	-.0002
iv_yeduc	-.0141	-.0244 *	-.0141	-.0244 *
DK	-.5862 **	-.5603 **	-.5893 **	-.5602 **
ES	-.2413	-.2406	-.2421	-.2405
IT	-.3593 **	-.3445 *	-.3602 **	-.3445 *
NL	-.2842 *	-.2756	-.2864	-.2755
_cons	1.7941 **	1.7663 **	1.8020 **	1.7661 **
k_2	44	50	44	50
$T_1(18)$		27.52		27.44
$T_2(4)$		6.35		6.24
$T_3(12)$		18.75		18.75
AIC	4593.70	4596.58		
BIC	4900.45	4944.23		
$k_1 + k_2$	59	70	59	70
ρ_{12}			-.01	.00
AIC	21232.66	21230.54	21234.67	21232.55
BIC	21648.48	21721.34	21695.73	21775.41

Table 9: Parametric estimates for item response on total nondurable expenditure. Interaction terms between characteristics of the interviewee and the interviewer, and between country dummies and income quartiles are not reported to save space. Sample size $n_1 = 6746$.

Variable	Model 4a	Model 4b	Model 5a	Model 5b
hr_fem	-.1052 *	-.1072	-.1003 *	-.1029
hr_agec	-.0167 **	-.0210 **	-.0172 **	-.0211 **
hr_agec^2	.0003	.0004	.0003	.0004 *
hr_yeduc	-.0074	-.0062	-.0072	-.0060
hr_nowork	.1046	.1120 *	.1014	.1090
single	-.0678	-.0573	-.0676	-.0575
hsize	-.0476 *	-.0420	-.0475 *	-.0420
children	-.5721 **	-.5828 **	-.5665 **	-.5780 **
s_city	-.1122 *	-.1163 *	-.1121 *	-.1162 *
math	.0108	.0101	.0105	.0099
recall	.0397 **	.0385 **	.0392 **	.0382 **
orient	.0622	.0651	.0622	.0652
hr_health	-.0040	-.0052	-.0057	-.0068
hr_adl	.0315	.0334	.0314	.0333
hr_pmm	-.0350	-.0432	-.0338	-.0422
income_q_1	-.2532	-.2490	-.2527	-.2486
income_q_3	.3004 *	.3070 *	.2947 *	.3019 *
income_q_4	.1274	.1355	.1253	.1335
inc_mis	-.6926 **	-.6922 **	-.6835 **	-.6841 **
f_proxy	-.0831	-.0724	-.0752	-.0649
p_proxy	.1953 *	.1819 *	.1918 *	.1787 *
int_out	-.2624 **	-.2580 **	-.2512 *	-.2479 *
int_length	.0008	.0008	.0008	.0008
iv_length	-.0376 **	-.0367 **	-.0370 **	-.0363 **
iv_fem	-.0175	-.0161	-.0187	-.0174
iv_agec	.0030	.0111 **	.0025	.0104 **
iv_agec^2	-.0001	.0000	-.0001	-.0000
iv_yeduc	.0002	-.0113	.0019	-.0095
DK	-.3933 **	-.3676 **	-.4467 **	-.4179 **
ES	-.1935	-.1944	-.2051	-.2049
IT	-.1643	-.1497	-.1800	-.1644
NL	-.4985 **	-.4856 **	-.5341 **	-.5188 **
_cons	1.8890 **	1.8951 **	2.0233 **	2.0190 **
k_2	44	50	44	50
$T_1(18)$		34.63 *		7.93
$T_2(4)$		12.76 *		5.13
$T_3(12)$		20.16		5.13
AIC	5574.07	5569.58		
BIC	5880.83	5917.23		
$k_1 + k_2$	59	70	59	70
ρ_{12}			-.18	-.17
AIC	22213.04	22203.54	22213.68	22204.41
BIC	22628.86	22694.34	22674.74	22747.27

Table 10: Parametric estimates of the Engel curves for food expenditure at home based on nonresponse equations without interaction terms. Standard errors in Models 4a and 5a are computed via the nonparametric bootstrap with 50 replications. $T_1(2)$ is a χ^2-test for selectivity due to nonresponse and has 2 df. Sample size $n_2 = 5884$.

Variable	Model 1	Model 2a	Model 3a	Model 4a	Model 5a
hr_fem	-1.1364	-1.5956	-1.2487	-1.7185	-1.7196
hr_agec	-.1121	-.0909	-.0959	-.0717	-.0716
hr_agec^2	-.0012	-.0009	-.0026	-.0019	-.0019
hr_yeduc	.7653 **	.7896 **	.7629 **	.7877 **	.7877 **
hr_nowork	1.0883	.5142	1.1892	.5992	.5994
single	-7.2677 **	-6.6825 **	-7.3069 **	-6.6983 **	-6.6977 **
hsize	8.8260 **	8.7781 **	8.8380 **	8.7909 **	8.7910 **
children	-1.9551	1.0404	-2.0474	1.0278	1.0298
s_city	-2.5328 *	-2.4035	-2.4627	-2.3471	-2.3469
math	.6823	.5307	.6977	.5421	.5420
recall	-.1512	-.2619	-.1369	-.2541	-.2541
orient	-.7090	-1.5683	-.7224	-1.6053	-1.6060
hr_health	-.6586	-.7203	-.5994	-.6631	-.6628
hr_adl	2.9356	3.1783	2.9400	3.1840	3.1841
hr_pmm	2.4839	3.1970	2.3632	3.1045	3.1046
income_q_1	-.9156	.2129	-.9330	.2300	.2317
income_q_3	-.6944	-1.5305	-.7102	-1.5631	-1.5629
income_q_4	2.2267	1.5194	2.2825	1.5456	1.5458
inc_mis	-.9352	1.7769	-1.0460	1.7543	1.7556
DK	1.4696	3.6433	2.2812	5.1239	5.1349
ES	35.3607 **	36.5089 **	35.0494 **	36.8200 **	36.8228 **
IT	21.7575 **	23.2918 **	21.7532 **	23.7196 **	23.7231 **
NL	17.6481 **	18.3924 **	18.6041 **	19.5787 **	19.5865 **
DK*inc_q_1	-.2981	-.9952	-.2972	-1.0124	-1.0137
DK*inc_q_3	-.1133	.6048	-.0945	.6410	.6406
DK*inc_q_4	-.6427	.0543	-.7214	.0167	.0162
ES*inc_q_1	-7.3420	-7.4043	-7.3697	-7.4248 *	-7.4260 *
ES*inc_q_3	6.3409	7.1329	6.3946	7.1834	7.1832
ES*inc_q_4	7.5541	8.0258	7.5664	8.0584	8.0583
IT*inc_q_1	-5.9353	-7.3404	-5.8787	-7.3226	-7.3242
IT*inc_q_3	9.0998 *	8.9653 *	9.2012 *	9.0475 *	9.0476 *
IT*inc_q_4	6.7618	7.8005	6.7969	7.8780	7.8787
NL*inc_q_1	4.0141	3.5407	4.0296	3.5387	3.5377
NL*inc_q_3	-2.1439	-2.2282	-2.1367	-2.2418	-2.2426
NL*inc_q_4	-5.4750	-5.1941	-5.5891	-5.2914	-5.2919
λ		-14.6377	5.3126		
λ_{unit}				4.9501	4.9775
λ_{item}				-15.0370	-15.0747
_cons	41.8552 **	43.4959 **	37.4091 **	39.5074 **	39.4856 **
k_3	35	36	36	37	37
$T_1(2)$				5.98	5.69
RB$_{\text{unit}}$(0)				.084	.084
RB$_{\text{item}}$(0)		-.028		-.032	-.032
RB$_{\text{total}}$(0)		-.028	.097	.052	.053

Table 11: Parametric estimates of the Engel curves for food expenditure at home based on nonresponse equations with interaction terms. Standard errors in Models 4b and 5b are computed via the nonparametric bootstrap with 50 replications. Sample size $n_2 = 5884$.

Variable	Model 1	Model 2b	Model 3b	Model 4b	Model 5b
hr_fem	-1.1364	-1.5287	-1.2651	-1.6867	-1.6866
hr_agec	-.1121	-.0921	-.0955	-.0706	-.0706
hr_agec^2	-.0012	-.0010	-.0027	-.0022	-.0022
hr_yeduc	.7653 **	.7860 **	.7612 **	.7825 **	.7825 **
hr_nowork	1.0883	.5818	1.1992	.6666	.6667
single	-7.2677 **	-6.7544 **	-7.3228 **	-6.7662 **	-6.7662 **
hsize	8.8260 **	8.7866 **	8.8309 **	8.7946 **	8.7946 **
children	-1.9551	.6214	-2.0015	.7326	.7324
s_city	-2.5328 *	-2.4251	-2.4494	-2.3507	-2.3507
math	.6823	.5508	.7010	.5636	.5636
recall	-.1512	-.2471	-.1332	-.2388	-.2388
orient	-.7090	-1.4601	-.7378	-1.5344	-1.5343
hr_health	-.6586	-.7185	-.5820	-.6426	-.6426
hr_adl	2.9356	3.1396	2.9370	3.1473	3.1472
hr_pmm	2.4839	3.1063	2.3699	3.0238	3.0238
income_q_1	-.9156	.0810	-.9464	.1064	.1063
income_q_3	-.6944	-1.4267	-.7209	-1.4983	-1.4982
income_q_4	2.2267	1.6147	2.2894	1.6254	1.6255
inc_mis	-.9352	1.4081	-1.0623	1.4410	1.4408
DK	1.4696	3.3430	2.3768	5.1520	5.1516
ES	35.3607 **	36.3667 **	35.0199 **	36.7719 **	36.7718 **
IT	21.7575 **	23.0784 **	21.7360 **	23.6034 **	23.6032 **
NL	17.6481 **	18.2815 **	18.6919 **	19.6843 **	19.6841 **
DK*inc_q_1	-.2981	-.9255	-.2957	-.9496	-.9495
DK*inc_q_3	-.1133	.5258	-.0653	.6214	.6213
DK*inc_q_4	-.6427	-.0429	-.6706	-.0061	-.0062
ES*inc_q_1	-7.3420	-7.4352	-7.3396	-7.4189 *	-7.4188 *
ES*inc_q_3	6.3409	6.9892	6.3938	7.0600	7.0600
ES*inc_q_4	7.5541	7.9401	7.5407	7.9761	7.9761
IT*inc_q_1	-5.9353	-7.1584	-5.8558	-7.1413	-7.1412
IT*inc_q_3	9.0998 *	8.9945 *	9.2463 *	9.1261 *	9.1261 *
IT*inc_q_4	6.7618	7.6678	6.8273	7.8170	7.8169
NL*inc_q_1	4.0141	3.5737	4.0626	3.5991	3.5991
NL*inc_q_3	-2.1439	-2.2046	-2.0793	-2.1678	-2.1678
NL*inc_q_4	-5.4750	-5.2365	-5.5343	-5.2846	-5.2846
λ		-12.7211	6.0155		
λ_{unit}				5.9456	5.9449
λ_{item}				-13.5365	-13.5345
_cons	41.8552 **	43.2940 **	36.8428 **	38.5631 **	38.5636 **
k_3	35	36	36	37	37
$T_1(2)$				7.55 *	7.37 *
$\text{RB}_{\text{unit}}(0)$				.101	.101
$\text{RB}_{\text{item}}(0)$		-.026		-.031	-.031
$\text{RB}_{\text{total}}(0)$		-.026	.111	.071	.071

Table 12: Parametric estimates of the Engel curves for total nondurable expenditure based on non-response equations without interaction terms. Standard errors in Models 4a and 5a are computed via the nonparametric bootstrap with 50 replications. Sample size $n_2 = 5605$.

Variable	Model 1	Model 2a	Model 3a	Model 4a	Model 5a
hr_fem	-.7727	1.6296	-1.4725	1.2447	1.3118
hr_agec	-.3307	-.0153	-.3010	.0788	.1104
hr_agec^2	.0025	-.0016	-.0019	-.0064	-.0074
hr_yeduc	3.3039 **	3.4492 **	3.2920 **	3.4391 **	3.4447 **
hr_nowork	-10.0798 **	-12.3649 **	-9.6969 **	-11.9999 **	-12.0779 **
single	-14.2161 **	-12.3453 **	-14.3206 **	-12.3993 **	-12.2824 **
hsize	14.7489 **	15.7047 **	14.8135 **	15.8150 **	15.8830 **
children	-8.2815	8.9799	-8.4939	9.2547	10.2786
s_city	-5.6530 *	-3.3541	-5.4347 *	-3.0532	-2.8825
math	2.5932 *	2.4846 *	2.6474 *	2.5526 *	2.5522 *
recall	.1895	-.6340	.2329	-.6150	-.6589
orient	.0409	-1.6808	-.0270	-1.8118	-1.9321
hr_health	-5.9249 *	-5.9525 *	-5.7362 *	-5.7209 *	-5.6757 *
hr_adl	3.6007	2.5357	3.6105	2.4797	2.4127
hr_pmm	4.7020	6.2745	4.1387	5.6755	5.7338
income_q_1	-.2739	4.9975	-.2547	5.1827	5.5696
income_q_3	10.4333 *	6.1075	10.3681 *	5.8856	5.6763
income_q_4	27.8826 **	25.8034 **	28.2009 **	25.9994 **	25.9275 **
inc_mis	2.8630	17.3704 **	2.4785	17.4286 **	18.2682 **
DK	-6.8578	-.5680	-3.6691	5.9322	7.3659
ES	42.6597 **	46.6014 **	41.9278 **	48.0775 **	48.5169 **
IT	24.4017 **	27.6063 **	24.6381 **	29.3261 **	29.8148 **
NL	28.9011 **	37.5871 **	30.6362 **	42.8853 **	44.2894 **
DK*inc_q_1	1.7820	3.1949	1.6628	3.1119	3.1675
DK*inc_q_3	-19.7524 *	-12.6673	-19.8051 *	-12.4277	-12.0484
DK*inc_q_4	-19.4311 *	-15.9772	-19.9518 *	-16.3004 *	-16.1718 *
ES*inc_q_1	-13.8668	-15.8732	-14.0797	-16.1169 *	-16.3252 *
ES*inc_q_3	24.1875 **	23.1896 *	24.2600 **	23.1822	23.0922
ES*inc_q_4	18.9351 *	16.0216	18.8499 *	15.8520	15.6502
IT*inc_q_1	-7.4403	-13.1667	-7.2445	-13.1485	-13.5422
IT*inc_q_3	9.0314	9.2429	9.4296	9.7015	9.7262
IT*inc_q_4	3.0246	5.6670	3.0168	5.8433	6.0031
NL*inc_q_1	11.8391	7.3424	11.9651	7.3200	6.9687
NL*inc_q_3	2.0529	5.6044	2.3050	5.9826	6.0877
NL*inc_q_4	8.8766	12.6336	8.4170	12.3711	12.5065
λ		-63.0836 **	19.2001 *		
λ_{unit}				21.6801 **	24.8778 **
λ_{item}				-65.1366 **	-71.0698 **
_cons	116.6306 **	120.0218 **	100.2615 **	102.4764 **	100.0399 **
k_3	35	36	36	37	37
$T_1(2)$				20.36 **	20.03 **
$\text{RB}_{\text{unit}}(0)$				.141	.167
$\text{RB}_{\text{item}}(0)$		-.036		-.044	-.047
$\text{RB}_{\text{total}}(0)$		-.036	.132	.097	.121

Table 13: Parametric estimates of the Engel curves for total nondurable expenditure based on nonresponse equations with interaction terms. Standard errors in Models 4b and 5b are computed via the nonparametric bootstrap with 50 replications. Sample size $n_2 = 5605$.

Variable	Model 1	Model 2b	Model 3b	Model 4b	Model 5b
hr_fem	-.7727	.8765	-1.5512	.5709	.6742
hr_agec	-.3307	-.1146	-.3055	-.0018	.0269
hr_agec^2	.0025	-.0003	-.0021	-.0056	-.0064
hr_yeduc	3.3039 **	3.3958 **	3.2862 **	3.3862 **	3.3920 **
hr_nowork	-10.0798 **	-11.6548 **	-9.6712 **	-11.3812 **	-11.4908 **
single	-14.2161 **	-12.9931 **	-14.3721 **	-12.9890 **	-12.8694 **
hsize	14.7489 **	15.3834 **	14.7839 **	15.5265 **	15.5930 **
children	-8.2815	3.5631	-8.3004	4.9842	6.1163
s_city	-5.6530 *	-4.1206	-5.4017 *	-3.6555	-3.4851
math	2.5932 *	2.5195 *	2.6556 *	2.5968 **	2.5943 **
recall	.1895	-.3606	.2438	-.3746	-.4223
orient	.0409	-1.0990	-.0902	-1.3732	-1.4964
hr_health	-5.9249 *	-5.9622 *	-5.6774 *	-5.6770 *	-5.6447 *
hr_adl	3.6007	2.8973	3.5907	2.7644	2.6941
hr_pmm	4.7020	5.6703	4.2066	5.1645	5.2339
income_q_1	-.2739	3.3463	-.3614	3.7244	4.1042
income_q_3	10.4333 *	7.5110	10.3192 *	6.9911 *	6.7488 *
income_q_4	27.8826 **	26.4218 **	28.2645 **	26.5260 **	26.4257 **
inc_mis	2.8630	12.6397 **	2.4355	13.4285 *	14.3369 *
DK	-6.8578	-2.5595	-3.4338	4.8946	6.0373
ES	42.6597 **	45.3297 **	41.9383 **	47.2239 **	47.6174 **
IT	24.4017 **	26.5685 **	24.5745 **	28.5506 **	28.9481 **
NL	28.9011 **	34.8035 **	30.6293 **	40.9827 **	42.1463 **
DK*inc_q_1	1.7820	2.7266	1.7521	2.7887	2.8742
DK*inc_q_3	-19.7524 *	-14.9368	-19.6639 *	-14.1312 *	-13.6933
DK*inc_q_4	-19.4311 *	-17.1190	-19.7419 *	-17.0282 **	-16.8396 *
ES*inc_q_1	-13.8668	-15.3680	-13.9090	-15.5658	-15.7501
ES*inc_q_3	24.1875 **	23.5004 *	24.1584 **	23.3872	23.2893
ES*inc_q_4	18.9351 *	16.8852	18.7042 *	16.4933	16.2839
IT*inc_q_1	-7.4403	-11.3729	-7.0675	-11.5117	-11.8993
IT*inc_q_3	9.0314	9.2021	9.6209	9.8780	9.9142
IT*inc_q_4	3.0246	4.9662	3.1124	5.4345	5.6338
NL*inc_q_1	11.8391	8.6761	12.1595	8.5845	8.2506
NL*inc_q_3	2.0529	4.3778	2.5828	5.1873	5.3498
NL*inc_q_4	8.8766	11.3758	8.6652	11.5167	11.7196
λ		-42.8240 *	21.2050 **		
λ_{unit}				24.1104 **	26.2256 **
λ_{item}				-48.3078 *	-54.6603 *
_cons	116.6306 **	118.9383 **	98.6568 **	99.6900 **	98.1585 **
k_3	35	36	36	37	37
$T_1(2)$				17.42 **	17.14 **
$\text{RB}_{\text{unit}}(0)$				.159	.177
$\text{RB}_{\text{item}}(0)$		-.025		-.033	-.036
$\text{RB}_{\text{total}}(0)$		-.025	.146	.126	.141

Table 14: Conditional moment tests for Gaussianity in the unit and the item response equations. Standard errors are computed via the nonparametric bootstrap with 50 replications. η is the generalized residual from the probit model.

			Food at home exp.		Nondurable exp.	
Equation	Moment	df	Test	p-value	Test	p-value
Unit	$E(\mu_1^2\,\eta)$	1	3.6907	.0002	3.6907	.0002
	$E(\mu_1^3\,\eta)$	1	-1.7369	.0824	-1.7369	.0824
	All	2	13.6837	.0011	13.6837	.0011
Item	$E(\mu_2^2\,\eta)$	1	-1.9013	.0573	-1.7716	.0765
	$E(\mu_2^3\,\eta)$	1	-2.2057	.0274	-1.7418	.0815
	All	2	5.0272	.0810	3.1687	.2051

Table 15: Reset-like tests for Gaussianity in the outcome equation. Standard errors are computed via the nonparametric bootstrap with 50 replications. The definition of ζ is provided in the text.

			Food at home exp.		Nondurable exp.	
Equation	Variable	df	Test	p-value	Test	p-value
Outcome	ζ_{U1}	1	-1.1072	.2682	.1451	.8846
	ζ_{U2}	1	-1.0714	.2840	.0993	.9209
	ζ_{U3}	1	-.9683	.3329	.0769	.9387
	ζ_1	1	-.7432	.4573	.9821	.3260
	ζ_2	1	.9618	.3362	-.8073	.4195
	ζ_3	1	-1.1293	.2588	.4335	.6647
	ζ_{00}	1	.9582	.3379	.4876	.6258
	ζ_{01}	1	.7423	.4579	-.9839	.3252
	ζ_{02}	1	-.9594	.3374	.8094	.4183
	ζ_{03}	1	1.1270	.2597	-.4374	.6618
	ζ_{10}	1	.9701	.3320	.4991	.6177
	ζ_{11}	1	.7465	.4554	-.9816	.3263
	ζ_{12}	1	-.9701	.3320	.8104	.4177
	ζ_{13}	1	1.1376	.2553	-.4300	.6672
	ζ_{20}	1	.9244	.3553	.5744	.5657
	ζ_{21}	1	.7612	.4465	-.9754	.3294
	ζ_{22}	1	-.9694	.3324	.7709	.4408
	ζ_{23}	1	1.1126	.2659	-.3827	.7020
	ζ_{30}	1	.7504	.4530	.6970	.4858
	ζ_{31}	1	.7415	.4584	-.9178	.3587
	ζ_{32}	1	-.9032	.3664	.6904	.4899
	ζ_{33}	1	1.0088	.3131	-.3372	.7360
	All	22	18.8646	.6537	25.3018	.2829

Table 16: Parametric and semiparametric estimates for unit response. $W_1(6)$ is a Wald test for the nonlinear combinations of τ_{hk} parameters in the distribution function of the unit response equation, and has 6 df. Sample size $n = 12537$.

	Food at home		Total expenditure	
Variable	Parametric	Semipar.	Parametric	Semipar.
gs_fem	-.0457 *	-.0556	-.0455 *	-.0503
gs_agec	.0053 *	.0103 **	.0053 *	.0103 **
gs_agec^2	-.0004 **	-.0006 **	-.0004 **	-.0006 **
iv_fem	.0278	.0664	.0275	.0686
iv_agec	.0048 **	.0032	.0048 **	.0033
iv_agec^2	.0002 **	.0003 **	.0002 **	.0003 **
iv_yeduc	-.0155 **	-.0164 **	-.0156 **	-.0164 **
tot_call	-.0220 **	-.0703 **	-.0219 **	-.0706 **
call_ev	-.0727 **	-.0949 *	-.0785 **	-.0884 *
call_we	.0661 *	-.0556	.0688 *	-.0498
lfield	-.0042 **	-.0054 **	-.0042 **	-.0054 **
DK	.1749 **	.4646 **	.1737 **	.4627 **
ES	-.2109 **	.0246	-.2111 **	.0260
IT	-.1506 **	.1035	-.1511 **	.1021
NL	.1103 **	.4912 **	.1098 **	.4873 **
_cons	.2139 **		.2166 **	
σ_1		1.94		1.95
Skewness		.23		.22
Kurthosis		2.09		2.11
$W_1(6)$		260.1 **		284.5 **

Table 17: Normalized estimates for unit response. Results based on the normalization $\beta_{\texttt{lfield}} = -.01$. Standard errors computed by the delta method.

	Food at home		Total expenditure	
Variable	Parametric	Semipar.	Parametric	Semipar.
gs_fem	-.1086 *	-.1021	-.1086	-.0930
gs_agec	.0127 *	.0189 **	.0126 *	.0191 **
gs_agec^2	-.0010 **	-.0011 **	-.0009 **	-.0011 **
iv_fem	.0662	.1220	.0656	.1268
iv_agec	.0114 **	.0059	.0115 **	.0061
iv_agec^2	.0006 **	.0005 **	.0006 **	.0005 **
iv_yeduc	-.0370 **	-.0301 **	-.0373 **	-.0303 **
tot_call	-.0524 **	-.1292 **	-.0523 **	-.1306 **
call_ev	-.1730 **	-.1744 *	-.1875 **	-.1633 *
call_we	.1571 *	-.1021	.1642 *	-.0921
DK	.4161 **	.8534 **	.4149 **	.8555 **
ES	-.5017 **	.0453	-.5041 **	.0481
IT	-.3582 **	.1900	-.3609 **	.1888
NL	.2623 **	.9023 **	.2621 **	.9010 **

Table 18: Parametric and semiparametric estimates of the item response equation. $W_2(6)$ is a Wald test for the nonlinear combinations of τ_{hk} parameters in the distribution function of the item response equation, and has 6 df. Interaction terms between country dummies and income quartiles, and τ_{hk} parameters of the SNP estimator are not reported to save space. Sample size $n_1 = 6746$.

	Food at home		Total expenditure	
Variable	Parametric	Semipar.	Parametric	Semipar.
hr_fem	.1335 **	.2627 **	-.1003 *	-.1513
hr_agec	-.0080	-.0154	-.0172 **	-.0296 *
hr_agec^2	.0000	-.0000	.0003	.0004
hr_yeduc	-.0077	-.0196	-.0072	-.0140
hr_nowork	.1502 *	.4147 **	.1014	.2568 *
single	-.1263 *	-.2314 *	-.0676	-.0969
hsize	.0040	.0094	-.0475 *	-.0750
children	-.4627 **	-.9628 **	-.5665 **	-1.5611 **
s_city	-.0381	-.0696	-.1121 *	-.1490
math	.0442	.0764	.0105	.0268
recall	.0265 *	.0475	.0392 **	.0611 **
orient	.1493 **	.3720 **	.0622	.1856 *
hr_health	.0086	-.0282	-.0057	-.0064
hr_adl	-.0902	-.1193	.0314	.0358
hr_pmm	-.0827	-.2405	-.0338	-.0638
income_q$_1$	-.2759	-.5247	-.2527	-.4049
income_q$_3$	.3520 *	.9162 **	.2947 *	.7407 **
income_q$_4$	.3073 *	1.0194 **	.1253	.5159
inc_mis	-.6862 **	-1.3461 **	-.6835 **	-1.2455 **
f_proxy	-.0257	.0162	-.0752	-.3258
p_proxy	.2607 **	.4532 **	.1918 *	.4030 *
int_out	-.1247	-.2554	-.2512 *	-.4626 *
int_length	.0021 *	.0050 **	.0008	.0023
iv_length	-.0357 **	-.0670 **	-.0370 **	-.0642 **
iv_fem	-.0225	.0318	-.0187	.0756
iv_agec	.0011	.0066	.0025	.0061
iv_agec^2	-.0003 *	-.0003	-.0001	-.0001
iv_yeduc	-.0141	-.0203	.0019	.0064
DK	-.5893 **	-.8319 **	-.4467 **	-.4554
ES	-.2421	-.0574	-.2051	.0554
IT	-.3602 **	-.3890	-.1800	-.0376
NL	-.2864	-.2946	-.5341 **	-.6204 **
_cons	1.8020 **		2.0233 **	
ρ_{12}	-.01	.13	-.18	.14
σ_2		1.68		1.74
Skewness		-.69		-.59
Kurthosis		2.77		2.50
$W_2(6)$		20.01 **		8.72

Table 19: Normalized estimates of the item response equation. Results are based on the normalization $\beta_{\texttt{iv\_length}} = -.01$. Standard errors are computed by the delta method. Interaction terms between country dummies and income quartiles are not reported to save space.

	Food at home		Total expenditure	
Variable	Parametric	Semipar.	Parametric	Semipar.
hr_fem	.0373 *	.0392 *	-.0271	-.0236
hr_agec	-.0022	-.0023	-.0046 *	-.0046
hr_agec^2	.0000	-.0000	.0001	.0001
hr_yeduc	-.0022	-.0029	-.0020	-.0022
hr_nowork	.0420	.0619 *	.0274	.0400
single	-.0353	-.0346	-.0183	-.0151
hsize	.0011	.0014	-.0128	-.0117
children	-.1294 *	-.1438 *	-.1530 **	-.2432 *
s_city	-.0107	-.0104	-.0303	-.0232
math	.0124	.0114	.0028	.0042
recall	.0074	.0071	.0106 *	.0095
orient	.0418 *	.0556 *	.0168	.0289
hr_health	.0024	-.0042	-.0015	-.0010
hr_adl	-.0252	-.0178	.0085	.0056
hr_pmm	-.0231	-.0359	-.0091	-.0099
income_q_1	-.0772	-.0784	-.0683	-.0631
income_q_3	.0985	.1368 *	.0796	.1154 *
income_q_4	.0860	.1522 *	.0338	.0804
inc_mis	-.1920 **	-.2010 **	-.1846 **	-.1940 **
f_proxy	-.0072	.0024	-.0203	-.0508
p_proxy	.0729 *	.0677 *	.0518	.0628
int_out	-.0349	-.0382	-.0679 *	-.0721
int_length	.0006 *	.0007 *	.0002	.0004
iv_fem	-.0063	.0047	-.0051	.0118
iv_agec	.0003	.0010	.0007	.0010
iv_agec^2	-.0001	-.0000	-.0000	-.0000
iv_yeduc	-.0039	-.0030	.0005	.0010
DK	-.1649 *	-.1242 *	-.1207 *	-.0710
ES	-.0677	-.0086	-.0554	.0086
IT	-.1008	-.0581	-.0486	-.0059
NL	-.0801	-.0440	-.1443 *	-.0967 *

Table 20: Semiparametric estimates of the partially linear model for food expenditure at home. Bandwidth parameter and trimming factor in semiparametric models are respectively equal to $n_2^{-1/p}$ and $n_2^{-1/r}$. Standard errors are computed via the nonparametric bootstrap with 50 replications.

		Semiparametric		
		Madel A	Model B	Model C
Variable	Parametric	$p=5$, $r=21$	$p=6$, $r=13$	$p=7$, $r=10$
hr_fem	-1.7196	-1.6955	-1.7192	-1.7951
hr_agec	-.0716	-.0718	-.0532	-.0412
hr_agec^2	-.0019	-.0019	-.0023	-.0026
hr_yeduc	.7877 **	.7588 **	.7784 **	.7992 **
hr_nowork	.5994	.5097	.2468	.0339
single	-6.6977 **	-6.7183 **	-6.6826 **	-6.6244 **
hsize	8.7910 **	8.8078 **	8.8046 **	8.8089 **
children	1.0298	1.7756	3.0217	3.2667
s_city	-2.3469	-2.0892	-2.1361	-2.2457
math	.5420	.4601	.4616	.4315
recall	-.2541	-.2572	-.2678	-.2817
orient	-1.6060	-1.0352	-1.5390	-1.7868
hr_health	-.6628	-.8729	-.8161	-.7790
hr_adl	3.1841	3.2253	3.1152	3.0594
hr_pmm	3.1046	3.1885	3.4310	3.2594
income_q$_1$	.2317	-.6798	-.3920	-.3446
income_q$_3$	-1.5629	-2.1956	-2.6150	-2.9673
income_q$_4$	1.5458	.6432	.1069	-.1858
inc_mis	1.7556	1.9156	2.3986	2.6612
DK	5.1349	6.4337	6.1886	6.1523
ES	36.8228 **	37.4021	37.4702 **	37.5368 **
IT	23.7231 **	24.3287	24.4420 **	24.6609 **
NL	19.5865 **	19.5280	19.5082 **	19.8253 **
DK*inc_q$_1$	-1.0137	-.1658	-.1859	-.1466
DK*inc_q$_3$	.6406	.7481	1.4544	1.8804
DK*inc_q$_4$	.0162	1.1986	1.6028	1.9157
ES*inc_q$_1$	-7.4260 *	-5.2459	-5.5954	-5.5325
ES*inc_q$_3$	7.1832	8.4390	8.4521	8.7232
ES*inc_q$_4$	8.0583	9.8780	10.2181	10.2921
IT*inc_q$_1$	-7.3242	-6.5885	-6.7996	-6.7997
IT*inc_q$_3$	9.0476 *	9.5695 *	10.0035 *	10.3845 *
IT*inc_q$_4$	7.8787	8.3366	8.9581	9.3930
NL*inc_q$_1$	3.5377	5.4522	4.9089	4.6435
NL*inc_q$_3$	-2.2426	-.9920	-.9678	-.9374
NL*inc_q$_4$	-5.2919	-3.9029	-3.8517	-3.8474
λ_{unit}	4.9775			
λ_{item}	-15.0747			
_cons	39.4856 **			
n_2	5884	5795	5835	5852

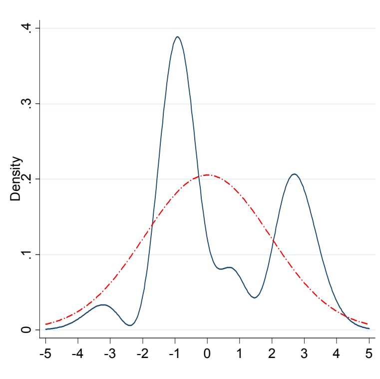
Table 21: Semiparametric estimates of the partially linear model for total nondurable expenditure. Bandwidth parameter and trimming factor in semiparametric models are respectively equal to $n_2^{-1/p}$ and $n_2^{-1/r}$. Standard errors computed via the nonparametric bootstrap with 50 replications.

		Semiparametric		
		Model A	Model B	Model C
Variable	Parametric	$p=5$, $r=21$	$p=6$, $r=13$	$p=7$, $r=10$
hr_fem	1.3118	.9947	1.2516	1.1953
hr_agec	.1104	.1807	.2143	.1847
hr_agec^2	-.0074	-.0058	-.0072	-.0065
hr_yeduc	3.4447 **	3.4959 **	3.5152 **	3.5168 **
hr_nowork	-12.0779 **	-13.7965 **	-13.8466 **	-13.5968 **
single	-12.2824 **	-11.6804 **	-11.8244 **	-11.9894 **
hsize	15.8830 **	16.0205 **	15.8533 **	15.7935 **
children	10.2786	23.1555	23.4829	22.5968
s_city	-2.8825	-3.3996	-3.6721	-3.6607
math	2.5522 *	2.4643 *	2.5447 *	2.4615 *
recall	-.6589	-.8572	-.8233	-.7897
orient	-1.9321	-3.6970	-3.8004	-3.7415
hr_health	-5.6757 *	-6.1505 *	-6.0579 *	-6.1439 *
hr_adl	2.4127	3.0171	2.8381	2.7687
hr_pmm	5.7338	5.5616	5.9330	5.3952
income_q_1	5.5696	6.6534	5.8841	5.2522
income_q_3	5.6763	-.7375	-1.2840	-1.4809
income_q_4	25.9275 **	19.6971	18.8786	18.5371
inc_mis	18.2682 **	21.3783	21.7906	21.3213
DK	7.3659	6.3452	6.3147	5.7556
ES	48.5169 **	44.9698 **	45.8249 **	45.8327 **
IT	29.8148 **	28.6473 **	29.2525 **	29.2918 **
NL	44.2894 **	46.1390 **	46.0851 **	44.9911 **
DK*inc_q_1	3.1675	4.6213	5.4701	5.9248
DK*inc_q_3	-12.0484	-4.3757	-3.8830	-3.8883
DK*inc_q_4	-16.1718 *	-10.5582	-9.6933	-9.4031
ES*inc_q_1	-16.3252 *	-13.2141	-12.5611	-12.3475
ES*inc_q_3	23.0922	37.2361 **	36.6058 **	36.2935 **
ES*inc_q_4	15.6502	25.5702	25.7588	26.0489
IT*inc_q_1	-13.5422	-13.3603	-12.9417	-12.3378
IT*inc_q_3	9.7262	18.4447	18.8368	18.8320
IT*inc_q_4	6.0031	12.4760	12.9065	13.0088
NL*inc_q_1	6.9687	3.8912	6.4949	6.9835
NL*inc_q_3	6.0877	10.4320	11.8479	12.2028
NL*inc_q_4	12.5065	17.0777	18.4888	19.0345
λ_{unit}	24.8778 **			
λ_{item}	-71.0698 **			
_cons	100.0399 **			
n_2	5605	5527	5567	5580

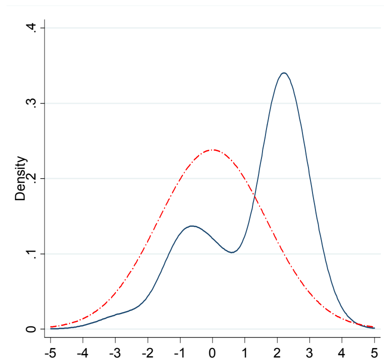
Figure 1: Semiparametric estimates of the error densities.

Food expenditure at home

a) Unit response

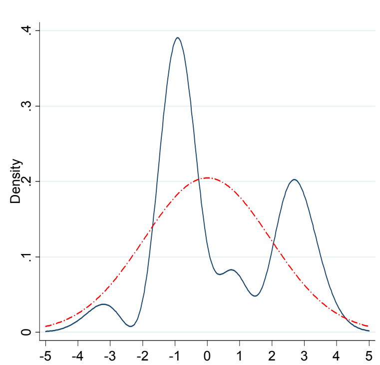


b) Item response



Total nondurable expenditure

c) Unit response



d) Item response